Title

Characterization of Inhibitory Control and Impulsivity Assessments in Healthy Adults Using Factor and Network Modeling

Author

Mattingly, Kamryn

Publication Date

2025

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Characterization of Inhibitory Control and Impulsivity Assessments in Healthy Adults Using Factor and Network Modeling

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Neuroscience

by

Kamryn Mattingly

March 2025

Dissertation Committee:
Dr. Aaron Seitz, Chair
Dr. Lani Bennett
Dr. Edward Zagha

The Dissertation of Kamryn Mattingly is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

I would like to thank my parents, Charlene and Rick Mattingly, for all their love and support while I moved across the country to attend graduate school and earn my PhD. Without them, I wouldn’t have had the opportunity to relocate and start this chapter of my life. I am also grateful for my sister Jane, for always cheering me on and keeping me updated about Mom and Dad in my absence. Thank you for never giving up on me and always believing I can reach my goals, even if it means leaving my hometown.

I am extremely thankful and grateful for my research advisor and committee chair, Dr. Aaron Seitz, for overseeing my research and advising me along the way. Aaron’s patience and empathy have been above and beyond, and I am eternally grateful that he has been a safe and supportive person in my life. While my first impression of Aaron was a mixture of intimidation and curiosity, he helped me to overcome my imposter syndrome and always encouraged me to chase down the answers to my own questions. His mentorship has led to so many new opportunities and ideas and has given me a confidence I may not have discovered otherwise.

To my other committee members and advisors, Dr. Lani Bennett and Dr. Eddie Zagha, who have been on my team in some way or another since my first year of grad school. They have both been core teachers and mentors in my education and have helped in tremendous ways for advancing my understanding of complex subjects in neuroscience. I feel incredibly fortunate to have had the privilege of their mentorship and frequently refer to the notes and papers I’ve collected from my experiences with them.

I would also like to extend deep gratitude for Dr. Megan Peters, for all her support as I struggled in my first year of grad school. Her direct advising over my

research during a lab rotation in my first year gave me a taste of the kind of researcher I hope to be. In rooms that were often dominated by male professors, she always stood tall, raised thought-provoking debates, and held her own no matter what type of situation arose. I will forever be inspired by her confidence and intelligence as I take the next steps in my career after graduate school.

I owe endless thanks and gratitude for all the lab mates, classmates, collaborators, and friends I've made along the way. I'd especially like to express thanks to Radhika, for learning with me through the beginning of the IC project, for all the endless code files, beautiful plots and conference posters. To Anja and Elnaz, I express so much appreciation and gratitude for all the mentorship, and for our analysis team that finally saw the light at the end of the tunnel that all started with some messy data files. Your expertise, patience, curiosity, and ideas will forever inspire me as I take the next steps. To my lab mates and cohort buddies: Morgan, Audrey, Kimia, Krithiga, Sam, Marcello, Sebastian, Becca, and Mohammad: I am forever grateful to have met my second family in my home away from home. All the lab meetings, attempts at journal clubs, social events, and long heart-to-hearts between classes and meetings will always hold space in my heart.

For Krystal, my best friend in Riverside, who was by my side during one of the most challenging periods in my life. Her grace, strength, and confidence taught me how to stand on business, and her caring nature has never left me feeling judged or shamed. She has pushed me to be the best version of myself and continues to be my biggest cheerleader, and I hope to do the same for her as she starts the next chapter of her own education journey.

Dedication

I dedicate this to my son, Theodore. From the time you were an infant, I was a single mom working in a bakery and attending college without a clue about what I wanted to do with my life. I knew it would be a long road, but I wanted to go as far as I could with school so that I could give you the best life possible. Now, you are 10 years old, and I am so incredibly proud and humbled by your beautiful soul, caring heart, and thriving curiosity. I did this all for you and hope to inspire you to achieve anything you can possibly dream of.

ABSTRACT OF THE DISSERTATION

Characterization of Inhibitory Control and Impulsivity Assessments
in Healthy Adults Using Factor and Network Modeling

by

Kamryn Mattingly

Doctor of Philosophy, Graduate Program in Neuroscience
University of California, Riverside, March 2025
Dr. Aaron Seitz, Chairperson

Inhibitory control (IC) is the capacity to interrupt an action in order to reach a specific goal. Impulsivity is the tendency to act rashly despite potentially negative consequences. Conceptually, they imply an inverse relationship, but this has not been consistently found in previous research. IC is measured using performance-based tasks, while impulsivity is generally measured using self-report questionnaires, and this format difference has led to issues in previous studies when comparing directly. In chapter one, this problem is addressed by conducting an Exploratory Factor Analysis (EFA) to identify how performance-based measures of IC and questionnaire scores of impulsivity correlate and group together. We identified four factors across 19 total measures of IC and impulsivity. Three factors consisted of measures with significant loadings from impulsivity assessments, while the fourth factor was showing significant loadings from IC tasks, suggesting IC and impulsivity may be separate constructs driven by separate underlying processes. In chapter two, to explore the relationship structure of IC and impulsivity assessments in a novel way, three network analyses were conducted, using

1) IC measures, 2) impulsivity measures, and 3) both IC and impulsivity measures with data from a healthy adult sample. These analyses revealed sub-networks, or “communities,” that were also largely dominated by assessment type, though some overlap across IC and impulsivity was observed in the full model. Chapter three compares a novel, gamified cognitive task based on the traditional cancellation task, with other traditional IC tasks. We found that UCancellation RT-based metrics significantly predicted TOVA RT variability, suggesting its possible utility as a more appealing alternative to the TOVA in certain cases. Ultimately, the results from this dissertation could help inspire future researchers to remove the redundancy of assessments used to measure IC and impulsivity in both research and clinical settings, while also introducing a novel, gamified measure of IC that may serve as a useful alternative to less-engaging traditional cognitive tasks.

Table of Contents

Lists of Figures

List of Tables

List of Abbreviations

ACC – Anterior Cingulate Cortex

ADHD – Attention Deficit Hyperactivity Disorder

BART – Balloon Analogue Risk Test

BIS/BAS – Behavioral Inhibition System and Behavioral Activation System Scale

BNST – Bed Nucleus Stria Terminalis

CPT – Continuous Performance Task

EFA – Exploratory Factor Analysis

FDR – False Discovery Rate

FT – First Target

IC – Inhibitory Control

IFG – Inferior Frontal Gyrus

IPC – Inferior Parietal Cortex

MS – Milliseconds

MSCS – Multi-Dimensional Self-Control Survey

N.ST – Non-Switch Target

PFC – Prefrontal Cortex

RT – Reaction Time

RTV – Reaction Time Variability

S.D. – Standard Deviation

S.E. – Standard Error

SUPPS-P – Short Form of UPPS-P Impulsivity Scale (Urgency, Premeditation, Perseverance, and Sensation-Seeking)

TOVA – Test of Variable Attention

Background and General Introduction

Inhibitory Control

Inhibitory control (IC) is an important part of basic and higher-level cognitive function and can be defined as the capacity to intentionally withhold an action in favor of a specific goal (Logan, 1997; Schall 2017; Tiego et al. 2018). Several performance-based tasks have been established to capture various measures of inhibitory behaviors (Enkavi, 2019). IC tasks such as these contain many outcome variables related to accuracy, reaction time (RT), and changes in performance across conditions. These metrics are intended to represent specific facets of IC, including attentional control, inhibiting motor responses, and interrupting a prepotent response when presented with a competing stimulus.

Such definitions and terminology of IC have been in ongoing conflict across previous studies, and this issue has intensified as new assessments of IC are introduced into research and clinical settings. This has created a problem where measures taken from common IC tasks are referred to using varying terms that are inconsistently operationalized (Nigg, 2000). For example, some previous studies used the term "response inhibition" while distinct terms were used interchangeably in other studies (e.g., behavioral inhibition, prepotent response inhibition, attention restraint, etc.) (Tiego et al., 2018). This creates confusion regarding what these assessments of IC were designed to measure and how we label such outcome variables.

Table 1. Table comparing terminology used across IC studies, derived from Tiego et al., 2018.

Tiego et al., (2018)	Nigg (2000)	Friedman and Miyake (2004)	Diamond (2013)	Gandolfi et al. (2014)	Stahl et al. (2014)	Kane et al. (2016)
Response Inhibition	Behavioral Inhibition	Prepotent Response Inhibition	Behavioral Inhibition/Behavioral Self-Control	Response Inhibition	Behavioral Inhibition	Attention Restraint
Attentional Inhibition	Interference Control	Resistance to Distracter Interference	Inhibition of Attention/Selective, Focused Attention	Interference Suppression	Stimulus Interference	Attention Constraint

Taking a step back to remember the origin of such IC assessments can provide some clarity that may be overlooked in how some studies have operationalized IC tasks. Commonly used measures of IC have traditionally been developed in association with previously established models of brain circuitry dating back decades. These models have been tested in rodents, primates, and humans and have reliably shown that specific brain regions and neural processes are associated with response inhibition (Schall, 2017; Bechara, 1994). One model attributes IC to the inferior frontal gyrus (IFG), requiring higher-level processes to successfully inhibit an action (Kang et al., 2022). Other previous research challenges this, suggesting that the IFG is merely one participant in the activation of a larger network of brain regions associated with IC, including the pre-supplementary motor area (preSMA) and inferior parietal cortex (IPC), shown in Figure 1 (Hampshire et al., 2010). Many other models of IC have been associated with broader cognitive processes that include working memory, updating, and attention, suggesting that IC is not a single construct, but one facet of overall cognition associated with multiple brain regions (Munakata, 2011; Blain, 2016; Tiego et al., 2018).

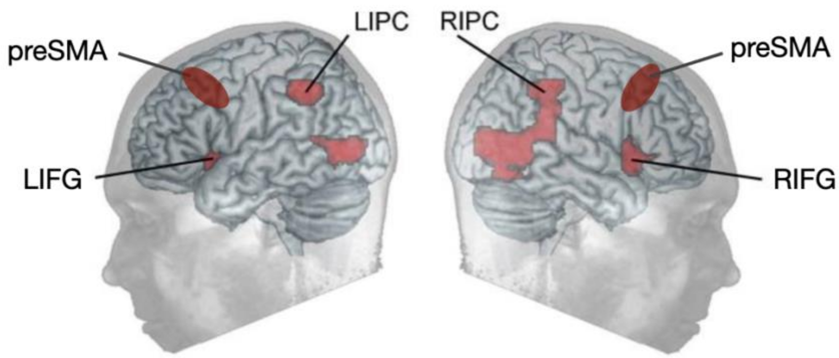


Figure 1. Derived and modified from Hampshire et al., 2010. Regions of the human brain commonly associated with IC, showing both left and right sides.

Another perspective, known as the predicted response-outcome model, describes cognitive control as a process of subsequent decisions and to what extent the feedback of each response guides future decisions (Alexander & Brown, 2019). This model can be demonstrated in the Flanker task, where participants are faced with congruent and incongruent visual stimuli. Participants must consider representations of stimulus and feedback in relation to the current condition and must update their expectation of the outcome in the next condition, as to minimize the likelihood of error. This model is associated with function in the anterior cingulate cortex (ACC), whose function is attributed to conflict detection, and is implicated as a mechanism of updating representations associated with decision making (Alexander & Brown, 2019).

A model of hierarchical cognitive control developed by Koechlin and colleagues implicates sub-regions of the prefrontal cortex (PFC) in cognitive control at four levels: sensory control (premotor cortex), contextual control (posterior lateral PFC), episodic control (anterior lateral PFC), and branching control (polar lateral PFC) (Koechlin, 2007). This model, depicted in Figure 2, suggests that IC is merely one example of cognitive control that is dependent on multiple parameters which have the potential to influence behavior and require the coordination of separate regions of the brain. The different models that implicate brain regions with IC encompass several frontal regions, coupled with motor regions that activate under certain conditions, depending on stimulus and goal, whether it be related to attention, motor response, language selection, emotion, and conflict detection. These models suggest that multiple facets of IC exist and how we interpret distinct metrics from IC tasks should be approached with caution.

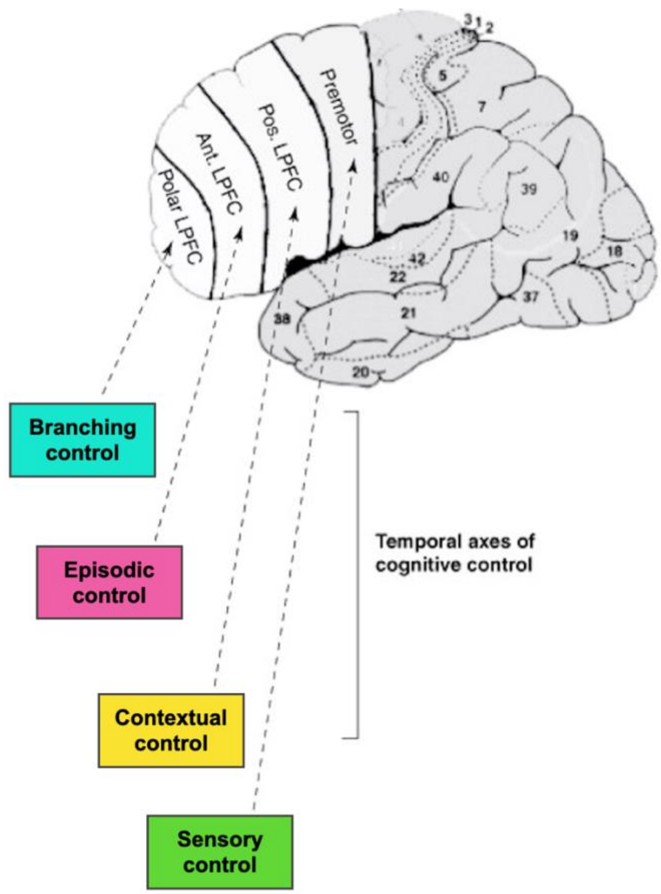


Figure 2. Derived and modified from Koechlin et al., 2007. Regions of the frontal lobe associated with different facets of IC.

While models of IC traditionally put emphasis on measuring how well a person can inhibit a response, the error rates of IC tasks have been used for measuring a separate concept entirely – impulsive behavior (Cyders, 2011; Roberts, 2011; Soutschek, 2020). This practice has grown increasingly common but creates ambiguity regarding what type of underlying trait is captured by error rates from IC measures, and how it might relate to other assessments designed to measure impulsivity directly (Wilbertz, 2014; Cyders, 2011; Enkavi, 2019).

Impulsivity

Impulsivity can formally be defined as hasty behavior that fails to consider long-term consequences in favor of an immediate reward and is a common symptom of clinical diagnoses including Attention Deficit Hyperactivity Disorder (ADHD), addiction and psychopathy (Christ, 2011; Choy, 2022). Impulsivity is generally operationalized using self-report questionnaires that require participants to respond to statements related to impulsive behaviors by deciding how much each statement represents their perception of their own behavior using a Likert-type scale (Gray, 2016; Cyders, 2014, Arnett, 2013).

Such questionnaires usually aim to measure impulsivity across multiple dimensions, or sub-scales. The names for these sub-scales are not in agreement across different commonly used questionnaires, even though these subscale names are often conceptually similar. An example of this can be seen when comparing the names of sub-scales of two separate self-report questionnaires that aim to measure distinct facets of impulsivity. In one questionnaire, known as BIS/BAS (Behavioral Inhibition

System/Behavioral Activation System), there is a sub-category referred to as “Fun-Seeking” (Cyders, 2014). Similarly, there is another questionnaire known as the SUPPS-P Impulsivity Scale that contains a subcategory of “Sensation-Seeking” (Gray, 2016).

The most common approach for determining sub-scales of impulsivity is to conduct a Factor Analysis, which groups variables of interest together based on a correlation or covariance matrix of all variables. In the literature, this is often done based on correlation of items within one questionnaire. This has been demonstrated in both adults and adolescents (Cyders, 2014; Gray, 2016; Watts, 2020; Vervoort, 2019). However, it has been less common for studies to conduct factor analyses that include multiple questionnaires of impulsivity. Some previous studies which include more than one questionnaire of impulsivity have found differing factor structures, which is at least partially attributed to the use of different assessments of impulsivity (MacKillop, 2014; Fino, 2014). Regardless, the factor structures depicted in these two studies do have some overlap, with both containing a factor labeled “reward sensitivity.”

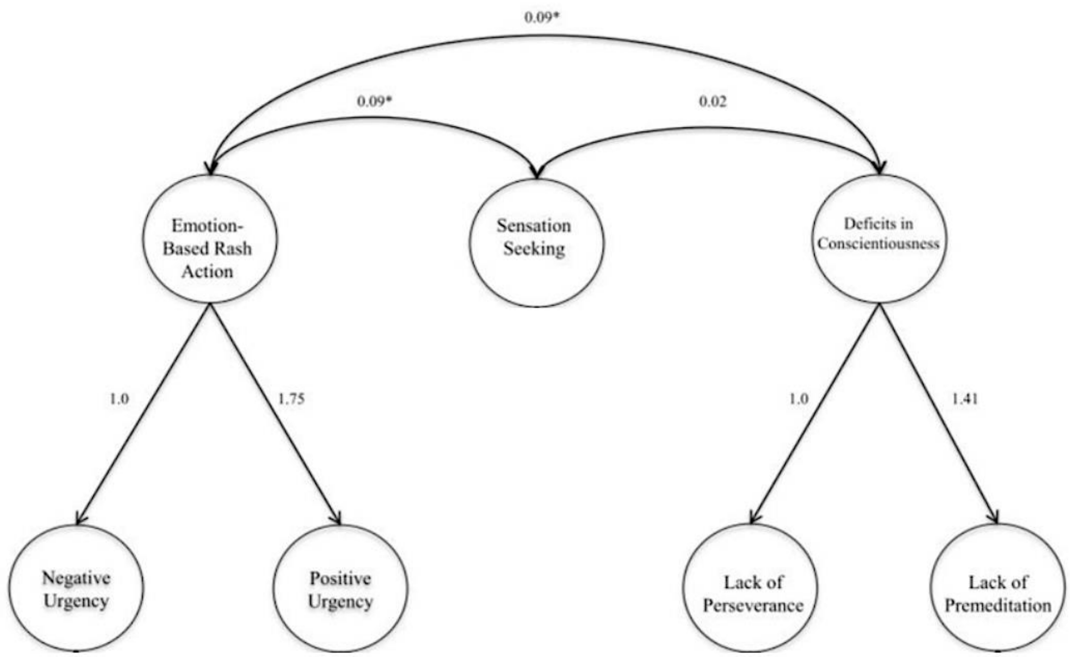


Figure 3. Derived and modified from Cyders et al., 2014. The factor structure produced from a factor analysis on the SUPPS-P Impulsivity Scale implicating distinct facets of impulsivity.

Further, evidence of neural mechanisms of impulsivity have been attributed to the limbic system, which is modulated by dopamine and is associated with reward motivation. The prefrontal regions of the brain, coupled with the insula and striatum, have been associated with impulsivity (Mitchell et al., 2014). Other brain regions that are often implicated in impulsivity are the amygdala and the bed nucleus stria terminalis (BNST), which have also been shown to modulate fear and anxiety, respectively (Owens, 2020). These brain regions have additionally been associated with emotion regulation, a type of IC, that is previously shown to be regulated by sub-regions of the PFC as well as the cingulate cortex (Etkin et al., 2010). This suggests that IC and impulsivity do have some overlap across brain regions, however the coordination of these regions are complex and distinct.

A Clash of Constructs and Formats

While IC and impulsivity both pertain to the tendency to act according to immediate vs future reward, they are typically explored using differing formats (i.e., self-report questionnaire versus performance-based cognitive tasks, respectively). Some studies aim to combine measures of IC and impulsivity, using various methods such as calculating a composite score that combines specific measures taken from self-report questionnaires and performance-based cognitive tasks (Bickel, 2012; Cyders, 2011; Enkavi, 2019).

An example of this is seen in the SUPPS-P Impulsivity Questionnaire that calculates a "Total Impulsivity" rating by summing the scores for all other subscales. Across studies, this practice of creating composite scores of impulsivity and/or IC is not always standardized and frequently neglects to account for differences in assessment format (Logan et al., 1997; Nigg 2000). Additionally, it has been seen in a previous review of self-regulation assessments that self-report questionnaires exhibit better test-retest reliability, while task-based measures of IC have lower reliability (Enkavi, 2019). The implication of more variance across time seen in the IC measures from this study while impulsivity measures remained stable may suggest IC is employed to differing extents depending on context. Conversely, it could suggest that IC may be capturing a state-level effect while impulsivity assessments captured information at the trait-level, or just that IC tasks are less reliable than impulsivity questionnaires (Dang et al., 2020).

While less common, there also exist self-report questionnaires of IC that are structured in a similar format as impulsivity questionnaires, where participants rate themselves according to statements about inhibitory behavior (Nilsen, 2020). Previous studies have shown significant correlations between impulsivity and IC questionnaires but show weaker associations between IC tasks and IC questionnaires (Bickel, 2012; Cyders, 2011; Enkavi, 2019). Additionally, performance-based tasks that aim to measure impulsivity tend to be structured differently than IC tasks in that impulsivity tasks generally involve calculating measures of delay-discounting, which is a metric used to determine a participant’s tendency to prefer a greater reward later as opposed to a smaller reward sooner (Stahl, 2014). The question-and-answer format of impulsivity tasks is argued to be more like that of impulsivity questionnaires than the format of IC tasks, and previous studies have also drawn conflicting conclusions when comparing them directly (Bickel, 2012; Cyders, 2011; Enkavi, 2019; Stahl, 2014).

The Need for Clearer Boundaries

The disconnect between definitions and assessment format poses an issue regarding how measures related to IC and impulsivity are interpreted, especially in instances where such assessments are used for diagnostic purposes, such as ADHD (Arnett, 2013; Kemper, 2018). In many cases, clinicians will diagnose disorders such as ADHD by administering multiple assessments, both performance-based tasks and self-report surveys of impulsivity. The DSM provides a resource of assessments intended to aid in the diagnosis of ADHD, which includes many the assessments described throughout this dissertation (Kemper, 2018).

Despite the issue of terminology, many studies show evidence that separate facets of IC do exist and sometimes correlate with measures of impulsivity (Friedman, 2004; Logan, 1997; Roberts, 2011; Wilbertz, 2014). This has also been exhibited by studies employing

IC tasks in different contexts, such as emotion regulation, working memory, and updating tasks (Munakata, 2011; Fino, 2014; Schall, 2017).

Previous work exploring the use of IC and impulsivity assessments in clinical populations has shown that patients with schizophrenia perform with more variability in RT on IC tasks, and higher averages of impulsivity subscales (Nolan, 2011). Another study exploring IC and impulsivity in an ADHD population revealed more variability in IC tasks compared to healthy controls, while impulsivity subscales were reported as higher on average in the ADHD sample, and similar variance (Roberts, 2011). These two examples show a similar trend of clinical populations may reveal more variability in their IC task performance compared to neurotypical individuals, while differences in impulsivity subscales show robust group differences with clinical populations scoring higher.

While the relationship between IC and impulsivity remains troubled, it remains evident that these assessments are highly informative. Some traditional cognitive tasks of IC that aim to capture metrics related to attention, control, and switching are time-consuming and operate using outdated interfaces. Newer assessments have been introduced and are beginning to establish convergent validity as novel, user-friendly alternatives to traditional tasks (Pahor et al., 2022).

The aim of this dissertation is to establish clearer boundaries between measures from IC and impulsivity assessments by using a variety of statistical methods, while also identifying shared variance among measures that may suggest common underlying processes as they relate to IC, impulsivity, or both. Ultimately, this could help researchers and clinicians make more informed choices about which assessments to use for screening and diagnosis, while avoiding redundancy and extracting information more efficiently. Ideally, this could contribute to saving time and resources without compromising robustness of the data collected.

References

Alexander WH, Brown JW. The Role of the Anterior Cingulate Cortex in Prediction Error and Signaling Surprise. Top Cogn Sci. 2019 Jan;11(1):119-135. doi: 10.1111/tops.12307. Epub 2017 Nov 13. PMID: 29131512.

Arnett AB, Pennington BF, Friend A, et al. The SWAN Captures Variance at the Negative and Positive Ends of the ADHD Symptom Dimension. Journal of Attention Disorders. 2013;17(2):152-162. doi:10.1177/1087054711427399

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition, 50(1-3), 7–15.

Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Gatchalian, K. M., & McClure, S. M. (2012). Are executive function and impulsivity antipodes? A conceptual reconstruction with special reference to addiction. Psychopharmacology, 221(3), 361–387.

Blain, B., Hollard, G., & Pessiglione, M. (2016). Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. Proceedings of the National Academy of Sciences of the United States of America, 113(25), 6967–6972.

Brickenkamp, R., & Zilmer, E. (1998). d2–Test of Attention Hogrefe & Huber Publishers.

Christ, S.E., Kester, L.E., Bodner, K.E. & Miles, J.H. (2011). Evidence for Selective Inhibitory Impairment in Individuals With Autism Spectrum Disorder. Neuropsychology, 25, 690– 701

Choy, O., Raine, A., & Schug, R. (2022). Larger striatal volume is associated with increased adult psychopathy. Journal of Psychiatric Research, 149, 185–193.

Cyders MA, Coskunpinar A. Measurement of constructs using self-report and behavioral lab tasks: is there overlap in nomothetic span and construct representation for impulsivity? Clin Psychol Rev. 2011 Aug;31(6):965-82. doi: 10.1016/j.cpr.2011.06.001. Epub 2011 Jun 13. PMID: 21733491.

Cyders, M. A., Littlefield, A. K., Coffey, S., & Karyadi, K. A. (2014). Examination of a short English version of the UPPS-P Impulsive Behavior Scale. Addictive Behaviors, 39(9), 1372–1376.

Dang J, King KM, Inzlicht M. Why Are Self-Report and Behavioral Measures Weakly Correlated? Trends Cogn Sci. 2020 Apr;24(4):267-269. doi: 10.1016/j.tics.2020.01.007. Epub 2020 Feb 17. PMID: 32160564; PMCID: PMC7977810.

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. 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(12), 5472–5477.

Etkin A, Egner T, Kalisch R. Emotional processing in anterior cingulate and medial prefrontal cortex. Trends Cogn Sci. 2011 Feb;15(2):85-93. doi: 10.1016/j.tics.2010.11.004. Epub 2010 Dec 16. PMID: 21167765; PMCID: PMC3035157.

Fino, E., Melogno, S., Iliceto, P., D'Aliesio, S., Pinto, M. A., Candilera, G., & Sabatello, U. (2014). Executive functions, impulsivity, and inhibitory control in adolescents: A structural equation model. Advances in Cognitive Psychology, 10(2), 32-38. doi:http://dx.doi.org/10.5709/acp-0154-5.

Friedman, N. P., and Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. J. Exp. Psychol. Gen. 133, 101–135. doi: 10.1037/0096-3445.133.1.101

Gray, J. D., Hanna, D., Gillen, A., & Rushe, T. (2016). A closer look at Carver and White's BIS/BAS scales: Factor analysis and age group differences. Personality and Individual Differences, 95, 20–24.

Hampshire A, Chamberlain SR, Monti MM, Duncan J, Owen AM. The role of the right inferior frontal gyrus: inhibition and attentional control. Neuroimage. 2010 Apr 15;50(3):1313-9. doi: 10.1016/j.neuroimage.2009.12.109. Epub 2010 Jan 4. PMID: 20056157; PMCID: PMC2845804.

Kang W, Hernández SP, Rahman MS, Voigt K, Malvaso A. Inhibitory Control Development: A Network Neuroscience Perspective. Front Psychol. 2022 Oct 10;13:651547. doi: 10.3389/fpsyg.2022.651547. PMID: 36300046; PMCID: PMC9588931.

Kemper AR, Maslow GR, Hill S, et al. Attention Deficit Hyperactivity Disorder: Diagnosis and Treatment in Children and Adolescents [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2018 Jan. (Comparative Effectiveness Reviews, No. 203.)

Koechlin E, Summerfield C. An information theoretical approach to prefrontal executive function. Trends Cogn Sci. 2007 Jun;11(6):229-35. doi: 10.1016/j.tics.2007.04.005. Epub 2007 May 1. PMID: 17475536.

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk-taking: The Balloon Analogue Risk Task (BART). Journal of Experimental Psychology: Applied, 8, 75-84.

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and inhibitory control. Psychological Science, 8(1), 60-64. doi:http://dx.doi.org/10.1111/j.1467-9280.1997.tb00545.x

Mackillop, J., Miller, J. D., Fortune, E., Maples, J., Lance, C. E., Campbell, W. K., & Goodie, A. S. (2014). Multidimensional examination of impulsivity in relation to disordered gambling. Experimental and Clinical Psychopharmacology, 22(2), 176–185.

Mitchell MR, Potenza MN. Recent Insights into the Neurobiology of Impulsivity. Curr Addict Rep. 2014 Dec 1;1(4):309-319. doi: 10.1007/s40429-014-0037-4. PMID: 25431750; PMCID: PMC4242429.

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. Trends in Cognitive Sciences, 15(10), 453–459.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. Psychological Bulletin, 126(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220

Nilsen, F. A., Bang, H., Boe, O., Martinsen, Ø. L., Lang-Ree, O. C., & Røysamb, E. (2020). The Multidimensional Self-Control Scale (MSCS): Development and validation. Psychological Assessment, 32(11), 1057–1074.

Nolan, K. A., D'Angelo, D., & Hoptman, M. J. (2011). Self-report and laboratory measures of impulsivity in patients with schizophrenia or schizoaffective disorder and healthy controls. Psychiatry Research, 187(1-2), 301–303.

Owens, M. M., Hyatt, C. S., Gray, J. C., Miller, J. D., Lynam, D. R., Hahn, S., Allgaier, N., Potter, A., & Garavan, H. (2020). Neuroanatomical correlates of impulsive traits in children aged 9 to 10. Journal of Abnormal Psychology, 129(8), 831–844.

Pahor, A., Mester, R.E., Carrillo, A.A. et al. UCancellation: A new mobile measure of selective attention and concentration. Behav Res **54**, 2602–2617 (2022). https://doi.org/10.3758/s13428-021-01765-5

Roberts, W., Fillmore, M. T., & Milich, R. (2011). Linking impulsivity and inhibitory control using manual and oculomotor response inhibition tasks. Acta Psychologica, 138(3),419–428.

Schall, J. D., Palmeri, T. J., & Logan, G. D. (2017). Models of inhibitory control. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 372(1718). https://doi.org/10.1098/rstb.2016.0193

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(2), 850-886. doi:http://dx.doi.org/10.1037/a0033981

Soutschek, A., & Tobler, P. N. (2020). Know your weaknesses: Sophisticated impulsiveness motivates voluntary self-restrictions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 46(9), 1611-1623. doi:http://dx.doi.org/10.1037/xlm0000833

Tiego, J., Testa, R., Bellgrove, M. A., Pantelis, C., & Whittle, S. (2018). A Hierarchical Model of Inhibitory Control. Frontiers in Psychology, 9, 1339.

Vervoort, L., De Caluwé, E., Vandeweghe, L., De Decker, A., Wante, L., Van Beveren, M.-L., Goossens, L., Verbeken, S., Sioen, I., Michels, N., & Braet, C. (2019). Parent-Reported BIS/BAS Scales for Children: Factor Structure and Measurement Invariance Across Age and Gender. Assessment, 26(7), 1282–1295.

Watts, A. L., Smith, G. T., Barch, D. M., & Sher, K. J. (2020). Factor structure, measurement and structural invariance, and external validity of an abbreviated youth version of the UPPS-P Impulsive Behavior Scale. Psychological Assessment, 32(4), 336–347. https://doi.org/10.1037/pas0000791

Wilbertz, T., Deserno, L., Horstmann, A., Neumann, J., Villringer, A., Heinze, H., Schlagenhauf, F. (2014). Response inhibition and its relation to multidimensional impulsivity. NeuroImage, 103,241-248. doi:http://dx.doi.org/10.1016/j.neuroimage.2014.09.021

Chapter 1: Identifying factors of Inhibitory Control and Impulsivity in healthy adults using Exploratory Factor Modeling

The study described in this chapter focused on understanding the relationships between assessments of inhibitory control (IC) and impulsivity. We collected all data during the peak onset of Covid-19 lockdown and administered all assessments online via Zoom. This provided a novel opportunity to explore online data collection that may present an alternative for potential research participants with travel limitations. By including a large number of measures for each construct (IC vs impulsivity), we went a step further beyond a correlation analysis and examined the exploratory factor structure of the measures extracted from these assessments. We found that questionnaire-based measures of impulsivity were largely separate from performance-based measures of IC. These findings added emphasis to the already controversial issue of how IC and impulsivity are defined and operationalized in previous studies.

Considering such assessments are commonly used in the clinical setting for screening and diagnosis of IC-related disorders, such as ADHD, it is important to understand how these measures relate to one another. The process of clinical screening can be costly and time-consuming, and some assessments may be redundant. By taking a closer look at the factor structure that resulted from our EFA, we can begin to have a deeper understanding of which metrics are most informative for each cluster (factor) of assessments. While our data comes from a healthy adult sample, future studies could find benefit from these results for inspiring follow-up studies in clinical populations, especially those with ADHD.

My contributions to this study included overseeing the set-up of the experiment protocol, deciding which assessments to include, processing raw data, conducting the factor analysis, and writing the manuscript in preparation of submitting to a journal for publication.

Title: Identifying factors of Inhibitory Control and Impulsivity in healthy adults using Exploratory Factor Modeling

Authors: Kamryn Mattingly, Radhika Amin, Audrey Carrillo, Anja Pahor, PhD., Jason Reimer, PhD., Susanne Jaeggi, PhD., Aaron Seitz, PhD.

Abstract

Inhibitory control (IC) can be defined as the capacity to resist or interrupt an action in favor of a greater reward, while impulsivity is typically defined as a failure of IC. Most IC assessments are performance-based tasks, while impulsivity is typically measured using self-report questionnaires. Many studies have directly compared measures of IC and impulsivity, but most do so by selecting a small set of measures or by introducing complex composite scores that do not replicate in similar studies. Using data from 270 healthy adults, we conducted an Exploratory Factor Analysis (EFA) to identify how performance-based measures of IC and questionnaire scores of impulsivity correlate and group together. We identified four factors across 19 total measures. Three factors consisted of measures with significant loadings from impulsivity assessments, while the fourth factor was showing significant loadings from IC tasks. These results suggest that IC and impulsivity may be separate constructs, contrary to their definitions, but format limitations may create a confounding effect. Future studies should include IC

questionnaires and performance-based impulsivity tasks to account for this potential confound. Ultimately, these results could help to remove the redundancy of assessments used to measure IC and impulsivity in both research and clinical settings.

1. Introduction

Human behavior can be described across many dimensions and is commonly categorized under the umbrella of "executive function" (EF) (Diamond, 2013). Definitions of EF vary, but largely include processes related to memory, attention, and self-control, also commonly referred to as inhibitory control (IC). IC can be defined as the capacity to intentionally withhold an action in favor of a specific goal (Logan, 1997; Schall 2017; Tiego et al. 2018). Many performance-based tasks have been developed to capture different dimensions of IC such as the Stroop task, Flanker, Antisaccade, and continuous performance tests (CPT) (Enkavi 2019;). These tasks provide outcome variables such as error rate, response time, and differences in performance across conditions. IC tasks are commonly used in research, clinical, and academic settings for their utility in screening participants, assisting in diagnoses of psychiatric conditions, and general testing of someone's cognitive capabilities (Forbes, 1998). Terminology used in the literature to operationalize IC tasks has been inconsistent regarding what certain variables are intended to represent (Tiego et al., 2018). The investigation into measures of IC has also extended across different contexts, such as emotion regulation, working memory, and updating tasks (Munakata, 2011; Fino, 2014; Schall, 2017).

While models of IC often put emphasis on measuring how well a person can inhibit a response, the way in which error rates of IC tasks are interpreted has resulted in a conflict of conceptualization (Cyders, 2011; Roberts, 2011; Soutschek, 2020). Some studies have operationalized such error measures to make inferences about impulsivity, which can formally be defined as urgent behavior that fails to regard long-term consequences in favor of an immediate reward and is a common symptom of psychiatric disorders such as attention deficit hyperactivity disorder (ADHD), addiction and psychopathy (Christ, 2011; Choy, 2022). This practice of using IC tasks as an indicator of impulsivity has grown increasingly common but has created ambiguity regarding what type of underlying trait is captured by error rates from IC measures, and how it might relate to other assessments originally developed to measure impulsivity (Wilbertz, 2014; Cyders, 2011; Enkavi, 2019).

Impulsivity is commonly operationalized using self-report questionnaires that require participants to respond to statements related to impulsive behaviors by deciding how much each statement represents their perception of their own behavior using a Likert-type scale (Gray, 2016; Cyders, 2014, Arnett, 2013). These assessments aim to measure impulsivity, usually across multiple dimensions, or sub-scales. The names for these sub-scales often have some conceptual overlap, but often branch into distinct sub-components that may differ from one assessment to the next. This has been demonstrated when comparing two commonly used impulsivity questionnaires. The BIS/BAS (Behavioral Inhibition System/Behavioral Activation System) questionnaire contains a sub-scale referred to as "Fun-Seeking" (Cyders, 2014). Similarly, the SUPPS-P Impulsivity Scale contains a subscale of "Sensation-Seeking" (Gray, 2016).

One study investigating impulsivity in those with gambling addiction directly compared BIS/BAS and SUPPS-P sub-scales and found a correlation of $r = 0.41$ between Fun-Seeking and Sensation-Seeking although these measures did not correlate significantly with task-based measures of impulsivity (i.e., Balloon Analogue Risk Task (BART) and the Georgia Gambling Task (GGT) (MacKillop et al., 2014; Lejuez, 2002). However, other studies show support that distinct facets of IC do sometimes correlate with impulsivity-related measures, despite format differences (Bickel, 2012; Friedman, 2019; Logan, 1997; Roberts, 2011; Wilbertz, 2014).

Several models of IC have been well-established in past decades and have been tested in humans, primates, and rodents that have reliably shown to associate specific brain regions and neural processes with response inhibition (Schall, 2017; Bechara, 1994). The mechanisms implicated in some models of IC have been associated with broader cognitive processes that include working memory, updating, and attention, suggesting that IC is not a single construct, but rather one dimension involved in a range of cognitive processes (Munakata, 2011; Blain, 2016; Tiego et al., 2018).

A common technique for characterizing questionnaire items into meaningful sub-scales is to conduct a Factor Analysis, which groups variables of interest together based on a correlation or covariance matrix of all survey items. This is typically done for determining how many distinct components, or factors, can be parsed from the overall variance within a single questionnaire (Cyders, 2014; Gray, 2016; Watts, 2020; Vervoort, 2019).

While IC and impulsivity both pertain to the tendency to act according to immediate vs future reward, they are typically explored using differing formats (i.e., performance-based cognitive tasks versus self-report questionnaire, respectively). Some studies aim to combine measures of IC and impulsivity, using various methods such as calculating a composite score that combines specific measures taken from impulsivity questionnaires and performance-based cognitive tasks into a single metric (Bickel, 2012; Cyders, 2011; Enkavi, 2019).

The SUPPS-P Impulsivity Questionnaire provides an example of this, where one can calculate a “Total Impulsivity" rating by summing the scores for all other subscales. Across studies, this practice of creating composite scores of IC and/or impulsivity is not always standardized and may not account for differences in assessment format. This mismatch between assessment format and operationalization of outcome measures has created a lack of repeatability and generalizability across studies of IC and impulsivity.

To better clarify the relationships between measures of IC and impulsivity, we examined a wide set of IC and impulsivity measures in a sample of college students across three universities. We hypothesized that an exploratory factor structure of IC and impulsivity variables, using sub-scales instead of composite scores, would show that some sub-scales of impulsivity load onto factors that include measures of IC, suggesting common underlying mechanisms across specific dimensions of IC and impulsivity.

2. Methods

2.1 Participants

We recruited 511 participants as part of a larger study from three universities in Southern California: UC Riverside, UC Irvine, and California State University San Bernardino. All participants provided informed consent for their participation and were recruited from the SONA research participation system. Due to a high proportion of incomplete datasets (N=219), non-compliance and outliers on cognitive tasks (N=22), only 270 participants had usable data for all assessments included in this analysis. For reaction time-based measures, raw scores that were faster than 200 milliseconds were considered outliers and thus removed. Outliers on other assessments were defined as scores exceeding 4 standard deviations. Included participants ranged in age from 18 to 41 years of age (M = 21, SD = 3.79), and identified as female (N=184), male (n = 84), or other (n = 2). All participants had normal or corrected to normal vision and hearing and were each awarded with SONA research credits following their participation.

2.2 Procedure

All data was collected during online sessions using Zoom video conference software due to Covid-19 restrictions. Participants met one-on-one with a research assistant for administering consent forms and during all data collection. Each participant completed every assessment across 3 sessions, each session lasting no more than one hour. The first session included completion of consent forms, demographics survey, and required software installations. Next, participants were provided with links to complete cognitive tasks and questionnaires. All assessments were counterbalanced across sessions to account for task duration, fatigue effects, and assessment format. Furthermore, each location of data collection administered assessments in 2 different orders, resulting in a total set of 6 task orders to control for potential order effects. Following completion of each session, the research assistant awarded the participant with their SONA credit(s).

2.3 Materials

Behavioral Tasks

AX-CPT. This is a continuous performance cue-probe task which has been shortened to 9 minutes long, compared to a previous use of the task of 90 minutes (Cooper, 2017; Marcora et al., 2009). Participants indicate the target sequence by pressing “E” on the keyboard following the target probe, and for all non-target sequences, “I” is pressed. The target sequence is the letter A (cue) followed by the letter X (probe), with a fixation cross presented between each letter. Other sequences (non-targets) that appear include AY, BX, and BY. There are 180 trials total: 126 AX trials and 18 trials of AY, BX, and BY. The proportion of target trials is .70. The primary variable of interest in this task is the proactive behavioral index (PBI) otherwise referred to as proactive control, calculated by accuracy of (AY-BX)/(AY+BX) (Mäki-Marttunen, 2018). Scores range from -1 to 1 with scores less than 0 indicating reactiveness, while scores greater than 0 represent proactive control.

UCancellation. A timed, tablet-based cancellation test that resembles the D2 (Brickenkamp & Zilmer, 1998), but letters are replaced with pictures of dogs and monkeys (Pahor et al., 2022). 8 items are displayed per row, and every 10 rows has 40 targets. Each row is displayed for 6 seconds. The participant must select the upright dog and the upright monkey separately in single blocks and together in a mixed block. Other dogs and monkeys of different orientations are distractors and considered irrelevant/non-targets.

The goal is to select as many targets and clear as many rows as possible. Score is calculated by number of hits (correctly selected targets) minus number of false alarms.

Antisaccade. The antisaccade task is used to measure inhibition of reflexive eye movements when a visual target is detected on the screen (Everling and Fischer, 1998; Sereno, 1995). Participants are instructed to focus on a central fixation cross, while a yellow square appears on either the left or right side of the fixation cross. Participants are instructed to suppress the reflex of looking at the yellow square, presented for 150 ms, and must instead direct their attention to the opposite side of the fixation cross so they have an opportunity to see the target letter "O" or "Q" presented for 175 ms. A mask of "##" covers the target letter after it's presented, and the participant must indicate which letter was shown by pressing the corresponding key on their keyboard (left, right, or up). This task has a duration of 7 minutes, modified to increase the difficulty by shortening the time between target stimulus and mask (Magnusdottir et al., 2019). There are 18 practice trials and 90 test trials. We used error rate as our variable of interest to represent performance, where lower values indicate better performance and higher values represent more errors.

Flanker. This task displays a horizontal line of five arrows, each arrow pointing left or right. The middle arrow is the target, and the participant is instructed to press the button that matches the direction of the target arrow. The row is spatially jittered to prevent participants from fixating their gaze on the target. Trials can be either congruent, where the direction of the target arrow is consistent with the surrounding arrows, or incongruent, where the direction of the target arrow is inconsistent with the surrounding arrows. This task has 24 practice trials and 60 test trials. Incongruent trials have been shown to lead to slower reaction times, due to the increased cognitive load. Early responses (less than 200 ms) are considered to be too early for inclusion (anticipation error), while responses beyond 3000 ms are considered too slow (inattention error) (Christ et al., 2011). This task was modified from its original design of cartoon fish, each facing either left or right, to a row of arrows. The primary variable of interest is the difference error rate of congruent versus incongruent trials.

TOVA. The Test of Variable Attention (TOVA) is a continuous performance task composed of visual stimuli (Forbes et al., 1998; Greenberg, 1991). In this task, participants indicate with a button press when the target is presented, which is a white square appearing above the central fixation point. Non-target trials occur when the white square is presented below the fixation cross. In half the trials, the target appears frequently (3.5 times for every non-target) and for the other half of trials, the target appears infrequently (once for every 3.5 non-targets). The primary variable of interest is reaction time variability (RTV), defined as the standard deviation of RT for correct responses across all test blocks.

Category Switch Task. In this task, participants must indicate with button presses the answer to questions presented on the screen (Mayr, 2000). Two rules occur during the task: the first is to indicate whether the word presented is something that is living or nonliving. In the second task, participants indicate if the word presented is of something that is small or large. The congruent condition is defined as subsequent trials consisting of the same rule. The incongruent condition corresponds to trials that randomly switch between the two rules. The primary measure of interest for this task is accuracy switch cost, which can be defined as the difference in mean accuracy of congruent versus incongruent trials.

Countermanding (Dogs and Monkeys). This task is a hybrid of Simon and spatial Stroop tasks, and it is also known as the 'Hearts and Flowers' and 'Dots' task (Diamond, 2013). The 'Hearts and Flowers' and 'Dots' tasks require that participants remember 2 rules: stimulus 1 indicates that they should press on the same side of the stimulus and stimulus 2 indicates that they should press on the opposite side of the stimulus (Diamond, 2013). In this version, the participant taps on one of two green buttons in response to the visual stimulus of a dog or a monkey, which appear on the left or right interchangeably. For dogs, the participant taps on the button on the same side of the screen (congruent). For monkeys, the participant taps on the button on the opposite side of the screen (incongruent). On incongruent trials, the participant must inhibit a predisposed response to respond on the same side as the stimulus. The task is self-paced with a 15 second timeout. For this shortened version, there are 3 blocks: 12 congruent trials, 12 incongruent trials, and 48 trials for mixed incongruent and congruent.

Self-report Questionnaires

SUPPS-P Impulsive Behavior Scale – Short. The SUPPS-P Impulsive Behavior scale is a self-report questionnaire that reliably measures five components of impulsive behavior: positive urgency, perseverance (lack of), premeditation (lack of), sensation seeking, and negative urgency (Whiteside & Lynam, 2001). A sixth measure is a total score of impulsivity that sums the scores of all subscales. We included the total score of impulsivity in our descriptive statistics, but removed this for correlational and factor analyses to avoid multicollinearity. Used in this study is the 20-item shortened version of the original 59-item questionnaire that has been validated against the full UPPS-P measure as well as other self-report measures of impulsivity (Cyders et al., 2014).

Behavioral Inhibition/Activation System (BIS/BAS). This self-report assessment measures two dimensions of behavior. The Behavioral Inhibition System (BIS) explores inhibitory responses in situations involving aversive conditions. The Behavioral Activation System (BAS) contains three subscales: Drive, Fun Seeking, and Reward Responsiveness. Drive is defined as unyielding effort towards achieving a goal, Fun Seeking is defined as the motivation and interest in finding reward, and Reward Responsiveness is the tendency of a positive response upon receiving a reward. The primary variables of interest are the three BAS subscales, and the BIS score (Gray, 2016; Vervoort, 2019).

SWAN Scale. This assessment is designed to screen for ADHD in children (Brites, et al., 2015), formally called the "Strengths and Weaknesses of ADHD Symptoms and Normal Behavior." This assessment is traditionally intended as a parent-report survey about a child and contains 18 items. Each item is a statement that is rated on a 4-point scale from "not at all", "just a little", "quite a bit", to "very much." The statements correspond to behaviors that relate to ADHD on two factors: Inattentive Type and Hyperactive Type. The items marked as "not at all" or "just a little" receive a score of 1, while "quite a bit" and "very much" receive scores of 0. Items 1-9 correspond to Inattentive Type while items 10-18 correspond to Hyperactive Type. For each factor, scores are summed. Scores below 6 on either factor suggest that ADHD diagnosis is unlikely. If a participant scores 6 or higher on both factors, they are considered “Combined Type.” This assessment, originally constructed for child and adolescent populations, was modified to be more appropriate for adults. To make this appropriate for adult self-report, as opposed to 3rd-person parent-report, all items were put into 1st person (i.e., “Gives close attention to detail...” is modified to “I give close attention to detail…”). Item 7 was modified from “Keeps track of things necessary for activities (doesn’t lose them)” to “I keep track of things necessary for activities and work” as this includes an adult-appropriate commitment that is more relevant than “activities” alone. Item 13 was modified from “Plays quietly (keeps noise level reasonable)” to “I keep my noise at a reasonable level during a task or activity” to make the item more adult-appropriate. The primary variables of interest are the scores on each of the two subscales: inattentive type and hyperactive type.

2.4 Statistical Analysis

To characterize the factor structure of these IC and impulsivity assessments, we conducted an exploratory factor analysis (EFA) incorporating measures of IC and impulsivity using the psych statistical package in R. First, we calculated a correlation matrix of Pearson correlation coefficients across all variables of interest, and we used this correlation matrix as input for the EFA. To approximate how many factors best fit our model, we produced a Scree plot and ran a parallel analysis, which allows one to better visualize and quantify the number of appropriate factors for a given set of data (Cattell, 1978). Due to the hypothesized correlations of some variables, we applied an oblimin rotation, which removes the constraint of requiring variables to be orthogonal (Russell, 2002). The results of the EFA yield a factor structure containing clusters of our IC and impulsivity measures. The manner in which these scores cluster determined our factors, which were interpreted to represent specific latent traits that each corresponding measure contributes variance towards. Factor loadings less than 0.4 were considered below threshold.

3. Results

3.1 Descriptive Statistics

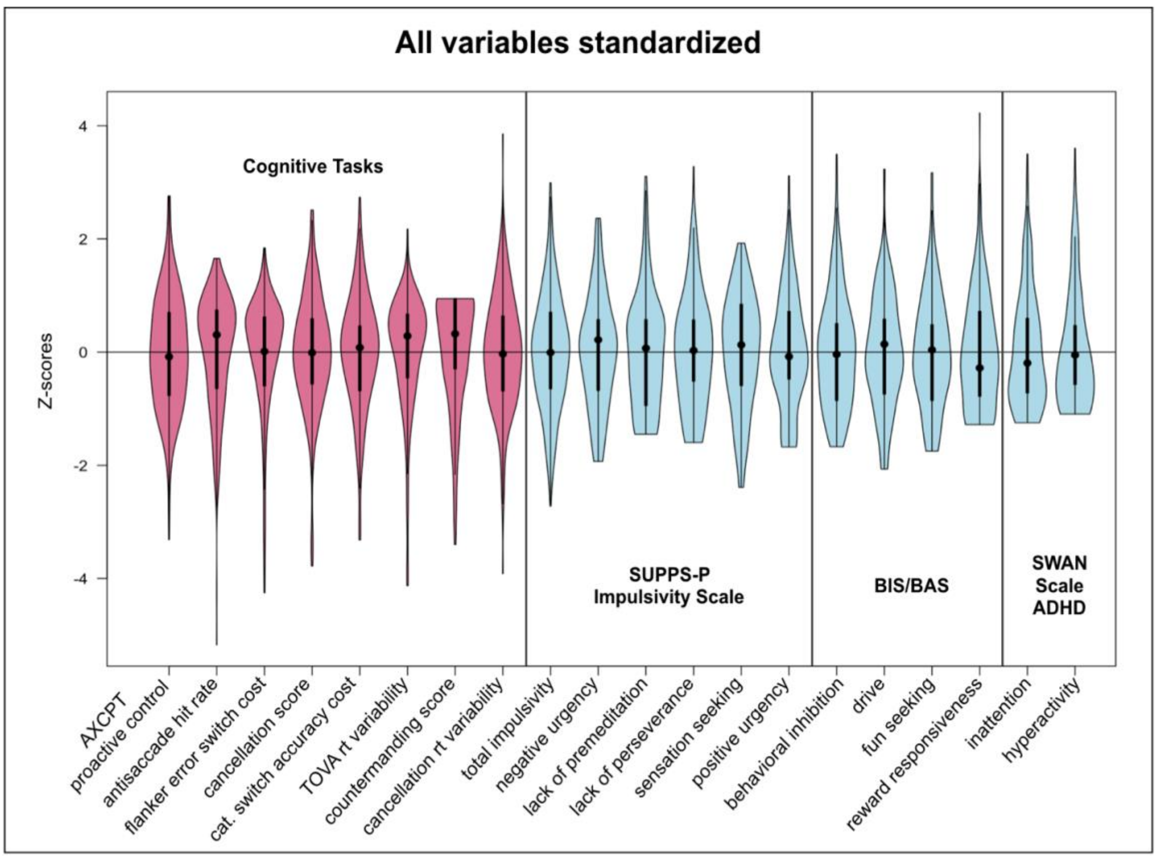


Figure 4. All assessment distributions are represented, with pink distributions for cognitive tasks and blue distributions for impulsivity questionnaires. All variables have been standardized (mean-centered).

Table 2. Descriptive statistics for all cognitive tasks.

Cognitive tasks	mean	S.D	median	Min	max	range	skew	S.E.
AX-CPT proactive control	0.138	0.102	0.130	-0.200	0.420	0.620	0.003	0.006
anti-saccade hit rate	0.757	0.140	0.800	0.033	0.989	0.956	-1.241	0.008
flanker accuracy switch cost	2.836	4.558	2.778	-5.556	22.22	27.778	1.765	0.269
category switch accuracy switch cost	-0.025	0.055	-0.021	-0.208	0.125	0.333	-0.432	0.003
cancel score	282.397	43.237	282.00	119.000	391.0	272.000	-0.568	2.552
cancel RT variability (ms)	-1.022	0.186	-1.028	-1.749	-0.307	1.442	-0.075	0.011
TOVA RT variability (ms)	-106.004	48.749	-92.075	-307.167	0.000	307.167	-1.625	2.878
counter-manding accuracy	96.828	3.355	97.917	85.417	100.0	14.583	-1.163	0.198

Table 3. Descriptive statistics for all impulsivity questionnaire subscales.

Impulsivity questionnaire subscales		mean	S.D.	med	min	max	range	skew	S.E.
SUPPS-P Impuls-ivity Scale	*negative urgency*	9.394	2.793	10.00	4.00	16.00	12.00	0.120	0.165
	positive urgency	8.195	2.505	8.00	4.00	16.00	12.00	0.247	0.148
	lack of pre-meditation	6.864	1.975	7.00	4.00	13.00	9.00	0.473	0.117
	lack of persevere	6.948	1.846	7.00	4.00	13.00	9.00	0.329	0.109
	sensation seeking	10.65	2.781	11.00	4.00	16.00	12.00	-0.33	0.164
BIS/BAS	*behav. inhibition*	13.14	3.677	13.00	7.00	26.00	19.00	0.535	0.217
	drive	8.679	2.264	9.00	4.00	16.00	12.00	0.177	0.134
	fun seeking	7.913	2.237	8.00	4.00	15.00	11.00	0.378	0.132
	reward response	7.557	1.997	7.00	5.00	16.00	11.00	0.728	0.118
SWAN Scale	*In-attention*	2.366	1.895	2.00	0.00	9.00	9.000	0.813	0.112
	hyper-activity	2.098	1.915	2.00	0.00	9.00	9.000	1.016	0.113

3.2 Correlation Analysis

The correlations among all cognitive task variables (Table 4) showed few significant relationships after correcting for multiple comparisons using false discovery rate (FDR) (Benjamini & Hochberg, 1995). Cancellation score revealed significant correlations with cancellation RT variability, antisaccade, and TOVA RT variability (r = 0.285, r = 0.419, r = 0.269, respectively; p < 0.05). Flanker and countermanding also had a significant relationship (r = -0.276, p < 0.05), but neither AX-CPT nor category switch showed any significant correlations with any other cognitive task.

Table 4. Pearson correlations among cognitive tasks. Bold values represent p-value < 0.05. Multiple comparisons were corrected using false discovery rate.

Cognitive task correlations	*AX-CPT*	*anti-saccade*	*flanker*	*cat switch*	*cancel score*	*cancel RT*	*TOVA RT*	*counter-mand*
AX-CPT	1.000							
antisaccade	-0.008	1.000						
flanker	-0.005	0.067	1.000					
category switch	-0.057	0.086	0.013	1.000				
cancellation score	0.109	**0.419**	0.135	-0.003	1.000			
cancellation RT	0.059	0.143	0.144	-0.016	**0.285**	1.000		
TOVA RT	0.022	**0.269**	0.071	0.043	**0.309**	**0.206**	1.000	
counter-manding	0.060	-0.071	**-0.276**	0.088	-0.067	0.026	-0.071	1.000

The correlations among all impulsivity questionnaire subscales (Table 5) showed many significant relationships after correcting for multiple comparisons using FDR. The highest correlations were observed between measures that belong to the same questionnaire, although significant correlations across questionnaires were also observed.

Table 5. Pearson correlation coefficients among self-report impulsivity questionnaires. Bold values represent p-value < 0.05. Multiple comparisons were corrected using false discovery rate. Abbreviations in the top row correspond to the labels in the first column.

Impulsivity questionnaire correlations	*Neg. Urg.*	*L. Pre.*	*L. Per.*	*S.S.*	*Pos.Urg.*	*B.I.*	*Dr.*	*F.S.*	*R.R.*	*In.*	*Hy.*
negative urgency	1.00										
lack of pre-meditation	**0.263**	1.00									
lack of per-severance	0.048	**0.433**	1.00								
sensation seeking	-0.011	0.038	-0.097	1.00							
positive urgency	**0.589**	**0.340**	**0.154**	**0.155**	1.00						
behavior inhibition	**-0.358**	0.030	0.091	**0.228**	-0.112	1.00					
drive	-0.102	0.088	**0.207**	**-0.244**	-0.085	-0.138	1.00				
fun seeking	-0.133	-0.072	0.104	**-0.520**	**-0.184**	-0.064	**0.436**	1.00			
reward respon-siveness	**-0.169**	**0.178**	**0.320**	**-0.167**	-0.057	**0.278**	**0.416**	**0.438**	1.00		
inattention	**0.261**	**0.416**	**0.383**	**-0.156**	**0.193**	**-0.163**	**0.254**	0.081	0.180	1.00	
hyper-activity	**0.241**	**0.316**	**0.279**	0.034	**0.232**	-0.086	0.096	0.022	0.090	**0.50**	1.00

The correlations between cognitive tasks and impulsivity questionnaire subscales (Table 6) revealed only one significant correlation after correcting for multiple comparisons using FDR. The significant relationship was observed between TOVA RT variability and SUPPS-P Lack of Perseverance (r = -0.17, p < 0.05).

Table 6. Pearson correlation coefficients of impulsivity questionnaires and cognitive tasks. Bold values represent p-value < 0.05. Multiple comparisons were corrected using false discovery rate.

Cognitive task vs impulsivity questionnaire correlations	*AX-CPT*	*anti-saccade*	*flanker*	*category switch*	*cancel score*	*cancel RT*	*TOVA RT*	*counter-mand*
negative urgency	0.090	0.040	0.115	-0.093	-0.024	0.019	-0.093	0.030
lack of premeditation	0.067	0.094	0.054	-0.077	0.102	-0.033	-0.019	-0.099
lack of perseverance	0.058	-0.021	-0.043	0.015	0.004	-0.115	**-0.171**	-0.073
sensation seeking	-0.096	0.119	0.053	-0.099	0.002	-0.153	-0.111	-0.098
positive urgency	0.008	0.105	0.071	-0.092	0.058	-0.113	-0.061	-0.032
behavior inhibition	0.041	0.046	-0.038	-0.054	0.027	-0.082	0.052	-0.025
drive	-0.043	0.039	-0.015	0.018	0.092	0.010	0.065	-0.033
fun seeking	0.048	-0.093	-0.073	0.115	0.044	0.095	0.082	-0.074
reward responsiveness	0.029	-0.024	-0.136	0.020	0.095	-0.012	0.045	-0.074
inattention	-0.092	0.002	0.032	0.022	-0.104	0.053	-0.036	-0.088
hyperactivity	0.016	0.048	-0.069	-0.076	-0.090	0.083	-0.078	-0.084

3.3 Exploratory Factor Analysis

We evaluated the shared and unique variance among measures of inhibitory control and impulsivity using exploratory factor analysis (EFA) with an oblique (oblimin) rotation. The factor structure is shown in Figure 5, and all factor loadings are listed in Table 7. The output for the analysis showed four factors to be sufficient, with a mean item complexity of 1.6. The root mean square of residuals (RMSR) was 0.05, and the degrees of freedom (df) corrected RMSR was 0.06. The total number of observations was 270 with a likelihood Chi square value of X^2 = 228.37, p < 0.001. Fit statistics for the model are shown in Tables 8-10.

Factors 1-3 were attributed only to impulsivity questionnaire subscales, but these factors did not significantly correlate, suggesting three distinct constructs related to impulsivity.

While this does not appear to provide insight towards how impulsivity related to IC tasks in this sample, it may provide additional perspective on how the subscales may manifest as distinct components of impulsivity that are instead independent of IC.

Factor 1 included Lack of Premeditation, Lack of perseverance, inattention, and hyperactivity, which were all positively related to one another. Lack of Premeditation and Lack of perseverance both come from the SUPPS-P Impulsivity Questionnaire, which combines both measures into a single higher-level factor referred to as "deficits in conscientiousness" (Watts et al., 2020). Inattention and hyperactivity are the two primary subscales in the SWAN Scale for ADHD screening. Together, the measures within this factor may suggest an underlying construct related to a lack of inhibition and attention, or rather one's own perception of their inhibitory and attentional capabilities.

Factor 2 includes Fun Seeking, Sensation Seeking, and Drive. Fun Seeking and Drive both come from the BIS/BAS questionnaire and are two of the three subscales belonging to the higher-level "behavioral activation system (BAS)" factor of this survey. Sensation Seeking comes from the SUPPS-P Impulsivity Questionnaire, and counterintuitively, was negatively correlated with Fun Seeking. The measures included with this factor suggest a possible construct related to external motivation, such as the wanting of some stimulus (Berridge, 2016).

Factor 3 includes Negative Urgency, Positive Urgency, Reward Responsiveness, and Behavioral Inhibition (BIS). Negative Urgency and Positive Urgency are both subscales of the SUPPS-P Impulsivity Questionnaire and are members of the same higher-level factor referred to as “urgency.” The notion of urgency in this context depends on emotional stimuli, whether positive or negative, to elicit the urgent reaction.

Reward Responsiveness and BIS are both members of the BIS/BAS questionnaire but are characterized as belonging to separate constructs within the structure of this questionnaire. However, they may tap into opposite ends of a similar mechanism—approaching something of value versus the avoidance of aversive stimuli. Further studies would be needed to confirm this, as both Positive Urgency and Reward Responsiveness had relatively high factor loadings on factors 1-3, suggesting they may not belong to one single category or construct, but rather provide information across a range of constructs related to impulsivity.

Factor 4 only included some cognitive task measures: cancellation score, cancellation RT variability, TOVA RT variability, and antisaccade hit rate. Measures that did not load significantly onto any factor were AX-CPT, countermanding score, flanker accuracy switch cost, and category switch accuracy switch cost, and are not included in Figure 5. It is worth noting that each of the “below threshold” measures are all from cognitive tasks.

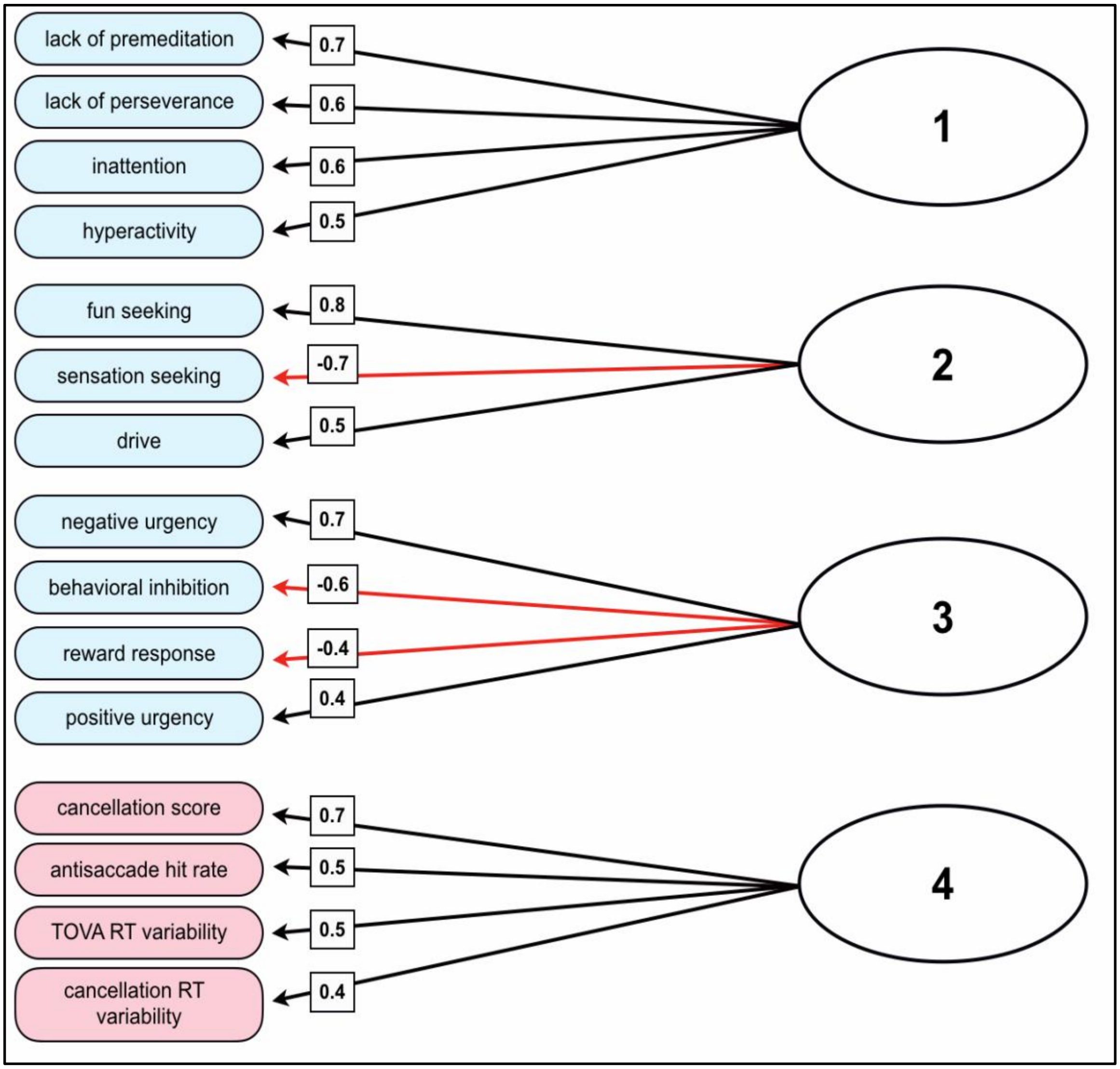


Figure 5. Exploratory factor analysis structure with a 4-factor oblimin rotation. Cognitive tasks are represented in red, while impulsivity questionnaires are labeled blue. Red arrows indicate negative loadings. Factors 1, 2, and 3 are dominated by impulsivity questionnaire subscales, while factor 4 is attributed solely to cognitive task measures. Variables whose factor loadings were all below threshold (< 0.4) are not shown in this diagram (i.e., AX-CPT proactive control, countermanding score, flanker accuracy switch cost and category switch accuracy switch cost).

Table 7. All factor loadings for all variables in the 4-factor oblimin rotation EFA. The right-hand columns indicate the amount of variance explained by the factors, the residual variance of each variable, and the item complexity.

EFA - Oblimin rotation	Factor 1	Factor 2	Factor 3	Factor 4	variance explained	residual variance	item complexity
AX-CPT	0.052	0.039	-0.049	-0.016	0.006	0.994	3.061
antisaccade	-0.107	0.062	0.586	-0.028	0.362	0.638	1.094
flanker	-0.123	0.064	-0.081	0.036	0.030	0.970	2.529
category switch	-0.077	0.073	-0.070	-0.007	0.017	0.983	2.997
cancellation score	0.037	0.056	-0.721	0.028	0.523	0.477	1.020
cancellation RT	0.091	-0.169	0.329	0.199	0.170	0.830	2.407
TOVA RT	0.134	-0.049	0.465	0.044	0.237	0.763	1.208
countermanding	-0.168	-0.032	0.057	-0.112	0.042	0.958	2.099
negative urgency	0.554	0.009	-0.002	-0.563	0.677	0.323	2.000
lack of premeditation	0.676	0.195	-0.054	0.149	0.484	0.516	1.284
lack of perseverance	0.495	0.379	0.092	0.306	0.456	0.544	2.690
sensation seeking	0.238	-0.635	0.004	0.274	0.516	0.484	1.669
positive urgency	0.632	-0.115	-0.039	-0.196	0.489	0.511	1.271
behavior inhibition	-0.045	-0.172	-0.062	0.634	0.424	0.576	1.178
drive	0.013	0.520	-0.051	0.008	0.271	0.729	1.021
fun seeking	-0.203	0.686	-0.017	0.001	0.535	0.465	1.175
reward responsiveness	0.109	0.511	-0.043	0.436	0.489	0.511	2.069
inattention	0.499	0.405	0.084	-0.075	0.396	0.604	2.037
hyperactivity	0.470	0.216	0.076	-0.056	0.264	0.736	1.498

Table 8. Pearson correlations among factors from EFA. None were significant ($p > 0.05$).

Factor correlations	1	2	3	4
1	1			
2	0.08	1		
3	0.10	-0.15	1	
4	0.04	0.00	-0.04	1

Table 9. Variance estimates of EFA for each factor.

Exploratory Factor Analysis output	1	2	3	4
SS loadings	2.00	1.74	1.44	1.30
Proportion Variance	0.11	0.09	0.08	0.07
Cumulative Variance	0.11	0.20	0.27	0.34
Proportion Explained	0.31	0.27	0.22	0.20
Cumulative Proportion	0.31	0.58	0.80	1.00

Table 10. Fit statistics from the 4-factor EFA with oblimin rotation.

EFA fit statistics	
RMSEA	0.069 (90% confidence interval [0.06, 0.08])
BIC	-333.17
TLI factor reliability	0.75
RMSR	0.05
mean item complexity	1.6

4. Discussion

We tested how measures of inhibitory control and impulsivity relate to one another in an exploratory factor analysis framework using a large number of assessments from a healthy adult sample. Contrary to our hypothesis, our analysis showed that measures of IC and impulsivity do not correlate significantly across assessment formats (i.e., questionnaire vs cognitive task), despite conceptual similarities and higher proportions of significant findings from previous studies (Bickel, 2012; Cyders, 2011; Enkavi, 2019). The need for the correction of multiple comparisons in this analysis conflicts with studies that compare two metrics directly, and obtain stronger significance (Logan et al., 1997). When such measures are included in a multiple correlation, and corrections for multiple comparisons are applied, these significant effects may not survive.

Additionally, the results of the EFA revealed a factor structure that clustered predominately in respect to assessment format, rather than conceptual or construct similarities, suggesting format differences may obstruct true underlying associations across assessments of IC and impulsivity. Namely, three out of four factors were dominated by impulsivity questionnaire subscales, while the fourth factor was fully attributed to IC task measures. Using sub-scales instead of calculating composite scores allowed for a deeper understanding of how IC and impulsivity measures related in a healthy adult sample. The results of this analysis suggested that IC may be attributed to a single factor, independent of impulsivity.

Most measures retained in the EFA path diagram in Figure 5 had factor loadings predominately on a single factor, as seen in Table 7. However, reward responsiveness and positive urgency each demonstrated relatively large factor loadings across factors 1-3. Meanwhile, many of the IC task measures, especially the ones that were not retained by the path diagram in Figure 5, showed small to moderate loadings across all factors. This suggests that some IC task measures tap into various dimensions of impulsivity, but do not explain enough variance in the impulsivity questionnaires to reach significance.

We note, however, potential confounds in the interpretation of these results and the meaning of these constructs, as they may relate to the different format of the test instruments used rather than separate constructs entirely. This confound is largely due to the nature of the assessments used—namely that IC assessments are classically cognitive tasks, and impulsivity assessments are typically self-report questionnaires. While self-report questionnaires of IC do exist, they tend to be structured in a similar format as impulsivity questionnaires, where participants rate themselves according to statements about inhibitory behavior (Nilsen, 2020). Previous studies have shown significant correlations between impulsivity questionnaires and IC questionnaires but show weaker associations between IC tasks and IC questionnaires (Bickel, 2012; Cyders, 2011; Enkavi, 2019), again suggesting a disconnect due to format differences.

This disconnect is further inflated by relatively smaller correlations observed among IC tasks when compared to correlations among self-report questionnaire subscales. Additionally, performance-based tasks that aim to measure impulsivity tend to be structured differently than IC tasks in that impulsivity tasks generally involve calculating measures of delay-discounting, which is a metric used to determine a participant's tendency to prefer a greater reward later as opposed to a smaller reward sooner (Stahl, 2014). The question-and-answer format of impulsivity tasks is argued to be more like that of self-report questionnaires than of a traditional cognitive task. Previous studies have also drawn conflicting conclusions when comparing such assessments directly (Bickel, 2012; Cyders, 2011; Enkavi, 2019; Stahl, 2014).

Previous work reviewing self-regulation assessments found that self-report questionnaires exhibited better test-retest reliability when compared to task-based measures of IC (Enkavi, 2019). More variance was seen across time in IC measures while impulsivity questionnaires remained stable, which may suggest IC is employed to different extents depending on context. Conversely, this finding could suggest that IC may be capturing a state-level effect while impulsivity assessments capture information at the trait-level. This could also imply that IC tasks are less reliable across time compared to impulsivity questionnaires.

The mismatch between definitions and assessment format poses an issue regarding how measures related to IC and impulsivity are interpreted, especially in instances where such assessments are used for diagnostic purposes, such as ADHD (Arnett, 2013; Kemper, 2018). In many cases, clinicians will diagnose disorders such as ADHD by administering multiple assessments, both performance-based tasks and self-report questionnaires of impulsivity. The DSM-5 provides a resource of assessments intended to aid in the diagnosis of ADHD, which includes many of the assessments used in this experiment (Kemper, 2018).

Other studies exploring the use of IC and impulsivity assessments in clinical populations have shown that patients with schizophrenia perform with more variability in RT on IC tasks, and higher averages of impulsivity questionnaire subscales (Nolan, 2011). Another study exploring IC and impulsivity in an ADHD population revealed more variability in IC tasks compared to healthy controls, while impulsivity questionnaire subscales were reported as higher on average in the ADHD sample, and similar variance (Roberts, 2011). These two examples show a similar trend of clinical populations may reveal more variability in their IC task performance compared to neurotypical individuals, while differences in impulsivity questionnaire subscales show robust group differences with clinical populations scoring higher. However, a meta-analysis of continuous performance tasks (CPTs) used to aid diagnosis of ADHD found that most studies did not find significant differences in groups of ADHD vs healthy controls, suggesting that CPTs may not provide robust objective measures (Hall et al., 2016).

Other limitations from this study included technical issues related to online data collection. Due to Covid-19 restrictions, all data was collected virtually from each participant’s own device. The variations in settings and devices across participants resulted in larger instances of troubleshooting, and in some cases, failure to collect data. Many assessments used in this study relied on Inquisit/Millisecond software, but unforeseen technical issues prevented data collection across several sessions for many participants. This led to significant data loss and resulted in many participants having incomplete datasets.

Future studies exploring IC and/or impulsivity should pay close attention to what assessments are used, how each measure or sub-scale is operationalized, and if multiple assessments are needed for obtaining the most informative results. Failing to acknowledge inconsistent relationships across IC and impulsivity measures can lead to conflated use of these assessments, especially when cross-format composite scores are introduced. A future study using factor analysis or independent components analysis (ICA) using all questionnaire items rather than subscales would provide additional information to tease apart constructs across questionnaires specifically. Factors that manifest from such analysis could then be compared to cognitive task measures for a potentially more precise basis of comparison across format differences.

In conclusion, our results showed a lack of significant correlation between measures of IC and impulsivity. This finding is counter-intuitive on a conceptual level but suggests a need for future assessments that control for the difference in formats. Establishing consistent operationalization of IC and impulsivity could help to standardize the methodology for which inhibitory and impulsive behaviors are measured in both research and clinical settings.

References

Arnett AB, Pennington BF, Friend A, et al. The SWAN Captures Variance at the Negative and Positive Ends of the ADHD Symptom Dimension. Journal of Attention Disorders. 2013;17(2):152-162. doi:10.1177/1087054711427399

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. Cognition, 50(1-3), 7–15.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple hypothesis testing. J R Stat Soc B 57:289–300

Berridge, K. C., & Robinson, T. E. (2016). Liking, wanting, and the incentive-sensitization theory of addiction. The American Psychologist, 71(8), 670–679.

Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Gatchalian, K. M., & McClure, S. M. (2012). Are executive function and impulsivity antipodes? A conceptual reconstruction with special reference to addiction. Psychopharmacology, 221(3), 361–387.

Blain, B., Hollard, G., & Pessiglione, M. (2016). Neural mechanisms underlying the impact of daylong cognitive work on economic decisions. Proceedings of the National Academy of Sciences of the United States of America, 113(25), 6967–6972.

Brickenkamp, R., & Zilmer, E. (1998). d2–Test of Attention Hogrefe & Huber Publishers.

Brites, C., Salgado-Azoni, C. A., Ferreira, T. L., Lima, R. F., & Ciasca, S. M. (2015). Development and applications of the SWAN rating scale for assessment of attention deficit hyperactivity disorder: a literature review. Brazilian Journal of Medical and Biological Research = Revista Brasileira de Pesquisas Medicas E Biologicas / Sociedade Brasileira de Biofisica... [et Al.], 48(11), 965–972.

CATTELL, R.B. and HORN, J.L. (1978), A CHECK ON THE THEORY OF FLUID AND CRYSTALLIZED INTELLIGENCE WITH DESCRIPTION OF NEW SUBTEST DESIGNS. Journal of Educational Measurement, 15: 139-164. https://doi.org/10.1111/j.1745-3984.1978.tb00065.x

Choy, O., Raine, A., & Schug, R. (2022). Larger striatal volume is associated with increased adult psychopathy. Journal of Psychiatric Research, 149, 185–193.

Christ, S.E., Kester, L.E., Bodner, K.E. & Miles, J.H. (2011). Evidence for Selective Inhibitory Impairment in Individuals With Autism Spectrum Disorder. Neuropsychology, 25, 690–701.

Cooper, S. R., Gonthier, C., Barch, D. M., & Braver, T. S. (2017). The Role of Psychometrics in Individual Differences Research in Cognition: A Case Study of the AX-CPT. Frontiers in Psychology, 8, 1–16. https://doi.org/10.3389/fpsyg.2017.01482

Cyders MA, Coskunpinar A. Measurement of constructs using self-report and behavioral lab tasks: is there overlap in nomothetic span and construct representation for impulsivity? Clin Psychol Rev. 2011 Aug;31(6):965-82. doi: 10.1016/j.cpr.2011.06.001. Epub 2011 Jun 13. PMID: 21733491.

Cyders, M. A., Littlefield, A. K., Coffey, S., & Karyadi, K. A. (2014). Examination of a short English version of the UPPS-P Impulsive Behavior Scale. Addictive Behaviors, 39(9), 1372–1376.

Diamond, A. (2013). Executive Functions. Annual Review of Psychology, 64(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. 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(12), 5472–5477.

Everling, & Fischer. (1998). The antisaccade: A review of basic research and clinical studies. Neuropsychologia, 36(9), 885-899.

Fino, E., Melogno, S., Iliceto, P., D'Aliesio, S., Pinto, M. A., Candilera, G., & Sabatello, U. (2014). Executive functions, impulsivity, and inhibitory control in adolescents: A structural equation model. Advances in Cognitive Psychology, 10(2), 32-38. doi:http://dx.doi.org/10.5709/acp-0154-5

Forbes, G. B. (1998). Clinical utility of the test of variables of attention (TOVA) in the diagnosis of attention-deficit/hyperactivity disorder. Journal of Clinical Psychology, 54(4), 461-476.

Friedman, N. P., & Banich, M. T. (2019). Questionnaires and task-based measures assess different aspects of self-regulation: Both are needed [Review of Questionnaires and task-based measures assess different aspects of self-regulation: Both are needed]. Proceedings of the National Academy of Sciences of the United States of America, 116(49), 24396–24397.

Gray, J. D., Hanna, D., Gillen, A., & Rushe, T. (2016). A closer look at Carver and White's BIS/BAS scales: Factor analysis and age group differences. Personality and Individual Differences, 95, 20–24.

Greenberg, L.M. (1991). T.O.V.A.™ interpretation manual. Minneapolis, MN: Author.

Hall, C. L., Valentine, A. Z., Groom, M. J., Walker, G. M., Sayal, K., Daley, D., & Hollis, C. (2016). The clinical utility of the continuous performance test and objective measures of activity for diagnosing and monitoring ADHD in children: a systematic review. European Child & Adolescent Psychiatry, 25(7), 677–699.

Kang W, Hernández SP, Rahman MS, Voigt K, Malvaso A. Inhibitory Control Development: A Network Neuroscience Perspective. Front Psychol. 2022 Oct 10;13:651547. doi: 10.3389/fpsyg.2022.651547. PMID: 36300046; PMCID: PMC9588931.

Kemper AR, Maslow GR, Hill S, et al. Attention Deficit Hyperactivity Disorder: Diagnosis and Treatment in Children and Adolescents [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2018 Jan. (Comparative Effectiveness Reviews, No. 203.)

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., Strong, D. R., & Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). Journal of Experimental Psychology. Applied, 8(2), 75–84.

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and Inhibitory Control. Psychological Science, 8(1), 60–64.

Mackillop, J., Miller, J. D., Fortune, E., Maples, J., Lance, C. E., Campbell, W. K., & Goodie, A. S. (2014). Multidimensional examination of impulsivity in relation to disordered gambling. Experimental and Clinical Psychopharmacology, 22(2), 176–185.

Magnusdottir, B. B., Faiola, E., Harms, C., Sigurdsson, E., Ettinger, U., & Haraldsson, H. M. (2019). Cognitive Measures and Performance on the Antisaccade Eye Movement Task. Journal of Cognition, 2(1), 1–13. https://doi.org/10.5334/joc.52

Mäki-Marttunen, V., Hagen, T., Aminihajibashi, S., Foldal, M., Stavrinou, M., Halvorsen, J. H., Laeng, B., & Espeseth, T. (2018). Ocular signatures of proactive versus reactive cognitive control in young adults. Cognitive, Affective & Behavioral Neuroscience, 18(5), 1049–1063.

Marcora, S.M, Staiano, W. & Manning, V. (2009). Mental fatigue impairs physical performance in humans. J Appl Physiol, 106: 857–864.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1124–1140.

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. Trends in Cognitive Sciences, 15(10), 453–459.

Nilsen, F. A., Bang, H., Boe, O., Martinsen, Ø. L., Lang-Ree, O. C., & Røysamb, E. (2020). The Multidimensional Self-Control Scale (MSCS): Development and validation. Psychological Assessment, 32(11), 1057–1074.

Nolan, K. A., D'Angelo, D., & Hoptman, M. J. (2011). Self-report and laboratory measures of impulsivity in patients with schizophrenia or schizoaffective disorder and healthy controls. Psychiatry Research, 187(1-2), 301–303.

Pahor, A., Mester, R.E., Carrillo, A.A. et al. UCancellation: A new mobile measure of selective attention and concentration. Behav Res **54**, 2602–2617 (2022). https://doi.org/10.3758/s13428-021-01765-5

Roberts, W., Fillmore, M. T., & Milich, R. (2011). Linking impulsivity and inhibitory control using manual and oculomotor response inhibition tasks. Acta Psychologica, 138(3), 419–428.

Russell DW. In Search of Underlying Dimensions: The Use (and Abuse) of Factor Analysis in Personality and Social Psychology Bulletin. Personality and Social Psychology Bulletin. 2002; 28:1629–1646.

Schall, J. D., Palmeri, T. J., & Logan, G. D. (2017). Models of inhibitory control. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 372(1718). https://doi.org/10.1098/rstb.2016.0193

Sereno, A.B. & Holzman, P.S. (1995). Antisaccades and Smooth Pursuit Eye Movements in Schizophrenia. BIOL PSYCHIATRY, 37, 394-401.

Soutschek, A., & Tobler, P. N. (2020). Know your weaknesses: Sophisticated impulsiveness motivates voluntary self-restrictions. Journal of Experimental Psychology. Learning, Memory, and Cognition. https://doi.org/10.1037/xlm0000833

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(2), 850–886.

Tiego, J., Testa, R., Bellgrove, M. A., Pantelis, C., & Whittle, S. (2018). A Hierarchical Model of Inhibitory Control. Frontiers in Psychology, 9, 1339.

Vervoort, L., De Caluwé, E., Vandeweghe, L., De Decker, A., Wante, L., Van Beveren, M.-L., Goossens, L., Verbeken, S., Sioen, I., Michels, N., & Braet, C. (2019). Parent-Reported BIS/BAS Scales for Children: Factor Structure and Measurement Invariance Across Age and Gender. Assessment, 26(7), 1282–1295.

Watts, A. L., Smith, G. T., Barch, D. M., & Sher, K. J. (2020). Factor structure, measurement and structural invariance, and external validity of an abbreviated youth version of the UPPS-P Impulsive Behavior Scale. Psychological Assessment, 32(4), 336–347.

Whiteside, S. P., & Lynam, D. R. (2001). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. Personality and Individual Differences, 30(4), 669–689.

Wilbertz, T., Deserno, L., Horstmann, A., Neumann, J., Villringer, A., Heinze, H.-J., Boehler, C. N., & Schlagenhauf, F. (2014). Response inhibition and its relation to multidimensional impulsivity. NeuroImage, 103, 241–248.

Chapter 2: Network Model of Inhibitory Control and Impulsivity Assessments

The study described in this chapter includes a subset of data from the same dataset as Chapter 1, with an emphasis on including more measures of IC and impulsivity than our previous analysis. Namely, we included both accuracy-based and reaction time (RT) - based measures for each IC task, and an additional questionnaire subscale of "neuroticism," taken from the Big 5 Personality Scale. To examine the relationships across these measures, three Network Analyses were conducted: IC measures only, impulsivity measures only, and a final model including all IC and impulsivity measures. Each model produced a visually intuitive network diagram with statistics related to how measures clustered together. This approach uses correlation coefficients as the "weights" between measures, which each measure represented as a "node." The findings of this study were consistent with Chapter 1, specifically that measures of IC and impulsivity did not exhibit much overlap, but that distinct facets of IC and impulsivity are detectable in their respective network models. These findings add even more emphasis to the need for careful operationalization when using IC and impulsivity tasks in a research or clinical setting.

My contributions to this study included overseeing the set-up of the experiment protocol, deciding which assessments to include, processing raw data, conducting the network analyses, and writing the manuscript in preparation of submitting to a journal for publication.

Title: Network Model of Inhibitory Control and Impulsivity Assessments

Authors: Kamryn Mattingly, Radhika Amin, Audrey Carrillo, Kevin Rosales, Anja Pahor, Ph.D., Susanne Jaeggi, Ph.D., Jason Reimer, Ph.D., Aaron Seitz, Ph.D.

Abstract

Inhibitory control (IC) is the capacity to interrupt an action in order to reach a specific goal. Impulsivity is the tendency to act rashly despite potentially negative consequences. Conceptually, they imply an inverse relationship, but this has not been consistently found in previous research. IC is measured using performance-based tasks, while impulsivity is generally measured using self-report questionnaires, and this format difference has led to issues in previous studies when comparing directly using methods such as factor analysis. To explore the relationship structure of these assessments in a novel way, we conducted three network analyses, using 1) IC measures, 2) impulsivity measures, and 3) both IC and impulsivity measures with data from 184 healthy adults. The full model containing all measures revealed a moderate density, suggested a commonality across all measures, despite format. Four clusters, or "communities," were identified and demonstrated a mixture of assessment type in two communities, while the remaining two communities were split based on format. Finally, by isolating measures with the highest betweenness centrality estimates, our network analysis revealed the most important "nodes" in our model were from Flanker, TOVA, Antisaccade, UCancellation, SUPPS-P Impulsivity scale (Sensation Seeking) and BIS/BAS Drive. These measures specifically may be most informative for capturing distinct facets of IC and impulsivity.

1. Introduction

Inhibitory control (IC) is commonly defined as the tendency to hold back or modify an action in order to achieve a greater reward, or to avoid a negative consequence (Tiego et al., 2018). A wide variety of previous studies have explored various ways of measuring IC in humans and animals, and a range of assessments have been developed. In humans particularly, IC most commonly measured using tasks that require a participant to respond, or withhold response to, specific stimuli according to specific rules (Munakata, 2011). From there, measures of accuracy and reaction time (RT) can be extracted, and other measures can be calculated from there, such as switch-costs between conditions, or composite scores incorporating both accuracy and RT-based scores.Impulsivity is another domain of behavioral research and is predominantly measured in humans using self-report questionnaires. Impulsivity can be defined as the failure to regulate one's actions when presented with a desirable, usually immediate, reward (Logan, 1997). The questionnaires developed to measure impulsivity have sub-scores that are specific to specific facets of impulsivity, such as drive, positive or negative "urgency," and fun-seeking (Gray, 2016; Whiteside & Lynam, 2001).

By definition alone, IC and impulsivity appear to describe the same behavior in different directions: if someone is impulsive, they have a failure of inhibitory control. However, when comparing measures of IC and impulsivity directly, results are mixed and relationships are weak, inconsistent, and do not easily generalize from one study to the next (Bickel 2012; Tiego et al. 2018). Previous studies have commonly used correlational and factor analyses to investigate measures of IC and impulsivity, but only a small number of measures are chosen, and broad conclusions are drawn that do not replicate onto the next study

(Roberts et al., 2011; Wilbertz et al., 2014). Other studies have gone deeper and compared directly format types and found that scores from cognitive tasks tend to be less stable and more variable, while measures from questionnaires are more reliable (Enkavi, 2019). This format difference and possible lack of reliability may obstruct the ability to identify whether IC and impulsivity are capturing different dimensions of the same behavior.

Network models based on the principles of graph theory provide an alternative way to visualize the relationships across a large number of measures. The network model approach can also help to identify clusters, or “communities” of measures that may be informative towards understanding the common underlying mechanisms driving IC and impulsivity (Burger et al., 2023). The density of a network model can help determine a shared variance across all measures, and these centrality estimates provide insight towards which measures may act as a key link between two other measures that may not be significantly related themselves (Epskamp et al., 2018). Based on previous studies that demonstrate significant relationships between measures of IC and impulsivity, we hypothesized that some measures would correlate, however we also predicted that RT-based and accuracy-based IC measures would “cluster” together, while sub-scores from questionnaires would largely belong to their own clusters. The measures with the highest centrality measure within each “cluster” (community) may serve as the measure capturing the most information regarding distinct facets of IC and impulsivity for screening and diagnostic purposes and could provide a new lens for identifying which measures may be redundant.

2. Methods

2.1 Participants

We recruited 511 participants as part of a larger study. All participants were undergraduate students recruited from UC Riverside, UC Irvine, and California State University San Bernardino. Data collection took place online during the onset of Covid-19 lockdown, and many participants did not complete all sessions. Due to missing data, we had full datasets from N = 272 participants, and after removing outliers, our final sample size for this study was N = 184. Participants reported having normal or corrected-to-normal vision and hearing. Outliers for reaction time-based scores were defined as responses made too quickly (< 200 ms) and for score-based measures, outliers were considered scores that exceeded four standard deviations. Participants included in our final sample were between the ages of 18 and 39 years old (M = 20.6 years; SD = 3.48 years). All participants provided informed consent and received course credit for their participation.

2.2 Materials

Behavioral Tasks

AX-CPT. This is a continuous performance cue-probe task which has been shortened to 9 minutes long, compared to a previous use of the task of 90 minutes (Cooper, 2017; Marcora et al., 2009). Participants indicate the target sequence by pressing “E” on the keyboard following the target probe, and for all non-target sequences, “I” is pressed. The target sequence is the letter A (cue) followed by the letter X (probe), with a fixation cross presented between each letter. Other sequences (non-targets) that appear include AY, BX, and BY. There are 180 trials total: 126 AX trials and 18 trials of AY, BX, and BY. The proportion of target trials is .70. The primary variable of interest in this task is the proactive behavioral index (PBI) otherwise referred to as proactive control, calculated by accuracy of (AY-BX)/(AY+BX) (Mäki-Marttunen, 2018). Scores range from -1 to 1 with scores less than 0 indicating reactiveness, while scores greater than 0 represent proactive control. For this analysis, we calculated proactive control for reaction time and again for error rate (Braver et al., 2009).

UCancellation. A timed, tablet-based cancellation test that resembles the D2 (Brickenkamp & Zilmer, 1998), but letters are replaced with pictures of dogs and monkeys (Pahor et al., 2022). 8 items are displayed per row, and every 10 rows has 40 targets. Each row is displayed for 6 seconds. The participant must select the upright dog and the upright monkey separately in single blocks and together in a mixed block. Other dogs and monkeys of different orientations are distractors and considered irrelevant/non-targets. The goal is to select as many targets and clear as many rows as possible. An outcome measures of interest for this analysis was concentration performance (score) and was calculated by number of hits (correctly selected targets) minus number of false alarms. We also extracted reaction time variability (RTV) for missed blocks and calculated RT switch cost.

Antisaccade. The antisaccade task is used to measure inhibition of reflexive eye movements when a visual target is detected on the screen (Everling and Fischer, 1998; Sereno, 1995). Participants are instructed to focus on a central fixation cross, while a yellow square appears on either the left or right side of the fixation cross. Participants are instructed to suppress the reflex of looking at the yellow square, presented for 150 ms, and must instead direct their attention to the opposite side of the fixation cross so they have an opportunity to see the target letter “O” or “Q” presented for 175 ms. A mask of “##” covers the target letter after it’s presented, and the participant must indicate which letter was shown by pressing the corresponding key on their keyboard (left, right, or up). This task has a duration of 7 minutes, modified to increase the difficulty by shortening the time between target stimulus and mask (Magnusdottir et al., 2019). There are 18 practice trials and 90 test trials. Our measures of interest for this analysis included average RT and error rate.

Flanker. This task displays a horizontal line of five arrows, each arrow pointing left or right. The middle arrow is the target, and the participant is instructed to press the button that matches the direction of the target arrow. The row is spatially jittered to prevent participants from fixating their gaze on the target. Trials can be either congruent, where the direction of the target arrow is consistent with the surrounding arrows, or incongruent, where the direction of the target arrow is inconsistent with the surrounding arrows. This task has 24 practice trials and 60 test trials. Incongruent trials have been shown to lead to slower reaction times, due to the increased cognitive load. Early responses (less than 200 ms) are considered to be too early for inclusion (anticipation error), while responses beyond 3000 ms are considered too slow (inattention error) (Christ et al., 2011). This task was modified from its original design of cartoon fish, each facing either left or right, to a row of arrows. The primary variables of interest were error rate switch cost and RT switch cost.

TOVA. The Test of Variable Attention (TOVA) is a continuous performance task composed of visual stimuli (Forbes et al., 1998; Greenberg, 1991). In this task, participants indicate with a button press when the target is presented, which is a white square appearing above the central fixation point. Non-target trials occur when the white square is presented below the fixation cross. In half the trials, the target appears frequently (3.5 times for every non-target) and for the other half of trials, the target appears infrequently (once for every 3.5 non-targets). The primary variable of interest is reaction time variability (RTV), with a secondary measure of interest as d'prime (Memoria et al., 2018).

Category Switch Task. In this task, participants must indicate with button presses the answer to questions presented on the screen (Mayr, 2000). Two rules occur during the task: the first is to indicate whether the word presented is something that is living or nonliving. In the second task, participants indicate if the word presented is of something that is small or large. The congruent condition is defined as subsequent trials consisting of the same rule. The incongruent condition corresponds to trials that randomly switch between the two rules. The primary measure of interest for this task is accuracy switch cost, which can be defined as the difference in mean accuracy of congruent versus incongruent trials. We also included RT switch cost as a secondary measure of interest.

Countermanding (Dogs and Monkeys). This task is a hybrid of Simon and spatial Stroop tasks, and it is also known as the 'Hearts and Flowers' and 'Dots' task (Diamond, 2013). The 'Hearts and Flowers' and 'Dots' tasks require that participants remember 2 rules: stimulus 1 indicates that they should press on the same side of the stimulus and stimulus 2 indicates that they should press on the opposite side of the stimulus (Diamond, 2013). In this version, the participant taps on one of two green buttons in response to the visual stimulus of a dog or a monkey, which appear on the left or right interchangeably. For dogs, the participant taps on the button on the same side of the screen (congruent). For monkeys, the participant taps on the button on the opposite side of the screen (incongruent). On incongruent trials, the participant must inhibit a predisposed response to respond on the same side as the stimulus. The task is self-paced with a 15 second timeout. For this shortened version, there are 3 blocks: 12 congruent trials, 12 incongruent trials, and 48 trials for mixed incongruent and congruent. The primary measure of interest for this task was mean accuracy.

Self-report Questionnaires

SUPPS-P Impulsive Behavior Scale – Short. The SUPPS-P Impulsive Behavior scale is a self-report questionnaire that reliably measures five components of impulsive behavior: positive urgency, perseverance (lack of), premeditation (lack of), sensation seeking, and negative urgency (Whiteside & Lynam, 2001). A sixth measure is a total score of impulsivity is available by taking the sum of all subscale scores, but this was excluded from our analysis to avoid multicollinearity. We administered the 20-item shortened version of the original 59-item questionnaire that has been validated against the full UPPS-P measure as well as other self-report measures of impulsivity (Cyders et al., 2014).

Behavioral Inhibition/Activation System (BIS/BAS). This self-report assessment measures two dimensions of behavior. The Behavioral Inhibition System (BIS) explores inhibitory responses in situations involving aversive conditions. The Behavioral Activation System (BAS) contains three subscales: Drive, Fun Seeking, and Reward Responsiveness. Drive is defined as unyielding effort towards achieving a goal, Fun

Seeking is defined as the motivation and interest in finding reward, and Reward Responsiveness is the tendency of a positive response upon receiving a reward. The primary variables of interest are the three BAS subscales, and the BIS score (Gray, 2016; Vervoort, 2019).

SWAN Scale. This assessment is designed to screen for ADHD in children (Brites, et al., 2015), formally called the "Strengths and Weaknesses of ADHD Symptoms and Normal Behavior." This assessment is traditionally intended as a parent-report survey about a child and contains 18 items. Each item is a statement that is rated on a 4-point scale from "not at all", "just a little", "quite a bit", to "very much." The statements correspond to behaviors that relate to ADHD on two factors: Inattentive Type and Hyperactive Type. The items marked as "not at all" or "just a little" receive a score of 1, while "quite a bit" and "very much" receive scores of 0. Items 1-9 correspond to Inattentive Type while items 10-18 correspond to Hyperactive Type. For each factor, scores are summed. Scores below 6 on either factor suggest that ADHD diagnosis is unlikely. If a participant scores 6 or higher on both factors, they are considered "Combined Type." This assessment, originally constructed for child and adolescent populations, was modified to be more appropriate for adults. To make this appropriate for adult self-report, as opposed to 3rd-person parent-report, all items were put into 1st person (i.e., "Gives close attention to detail..." is modified to "I give close attention to detail…"). Item 7 was modified from "Keeps track of things necessary for activities (doesn't lose them)" to "I keep track of things necessary for activities and work" as this includes an adult-appropriate commitment that is more relevant than "activities" alone. Item 13 was modified from "Plays quietly (keeps noise level reasonable)" to "I keep my noise at a reasonable level during a task or activity" to make the item more adult-appropriate. The primary variables of interest are the scores on each of the two subscales: inattentive type and hyperactive type.

Mini-Markers (Big 5) Personality Questionnaire. A shortened self-reported assessment to determine personality traits, derived from Goldberg's Unipolar Big-Five Markers (Saucier, 1994). This abbreviated inventory includes 40 markers rather than the original 100 markers, assessing extraversion, agreeableness, conscientiousness, neuroticism, and openness/intellect. Factors from the Mini Markers also provide a valid basis of personality measurement, as such factors have been shown to correlate with the full set of 100 markers. Participants are asked a series of questions and are asked to rank themselves on a 9-point Likert scale for each question. The primary variables of interest for this questionnaire is neuroticism, as this trait is often studied in relation to impulsivity (Whiteside & Lynam, 2001).

2.3 Procedure

The study consisted of three sessions in which participants completed UCancellation, AX-CPT, Antisaccade, Flanker, TOVA, Category Switch Task, and Countermanding tasks on their personal tablets and computers. Participants also completed additional questionnaires as part of a larger study; these questionnaires are not included in this analysis. The assessments were administered via a custom-built app, Recollect the Study, and Inquisit Web via Millisecond Software (version 6.1.0.0), while being monitored by a researcher via Zoom. The interval between sessions was at most two weeks and each session lasted one hour or less. Participants took a 2-minute break between tasks, with the option to skip it. To account for fatigue effects, all assessments were counterbalanced across sessions. The order of assessments was also counterbalanced across data collection sites, and each site implemented two different orders of assessments, alternating with each participant.

2.4 Statistical Analysis

Due to non-normality of some variables, Spearman correlation coefficients were calculated for all measures. Correlations were corrected for multiple comparisons using False Discovery Rate (FDR) (Benjamini & Hochberg, 1995). Three network analyses were calculated and visualized using the *igraph* package version 0.5.1 using R statistical software version 2024.04.1. The first model includes only IC task variables, the second model includes only impulsivity questionnaire subscales, and the third model includes all IC and impulsivity variables. Network density was calculated for each network as the proportion of edge weights out of all possible edge weights. Higher density suggests that all nodes are more highly connected, while lower values suggest a more sparse network architecture (Burger et al., 2023; Epskamp et al., 2018).

Each variable within a network model was represented as a “node” and each correlation coefficient was used as the input for each “edge weight” between nodes. Communities were detected using a “fast greedy algorithm” modularity maximization function, which aims to isolate clusters of highly connected nodes (Newman, 2004). Betweenness centrality estimates were calculated to identify the most “important” or influential nodes within each network.

3. Results

3.1 Descriptive Statistics

Table 11. Descriptive statistics for all variables.

Descriptive Statistics	*mean*	*S.D.*	*median*	*min*	*max*	*range*	*skew*	*S.E.*
age (years)	20.603	3.483	20.000	18.000	39.000	21.000	2.788	0.257
AX-CPT proactive ctrl RT	0.140	0.102	0.134	-0.134	0.419	0.553	0.054	0.007
AX-CPT proactive ctrl errors	0.183	0.571	0.000	-0.949	0.957	1.905	-0.31	0.042
Antisaccade error rate	0.229	0.124	0.200	0.011	0.589	0.578	0.810	0.009
Antisaccade mean RT	650.164	166.483	618.793	447.896	1551.405	1103.509	2.287	12.273
Flanker RT switch cost	-97.701	78.799	-81.031	-441.16	23.798	464.962	-1.89	5.809
Flanker error switch cost	2.868	4.403	2.778	-2.778	22.222	25.000	1.699	0.325
Category switch accuracy switch	-0.024	0.057	-0.021	-0.208	0.125	0.333	-0.45	0.004
Category switch RT switch	339.655	319.233	261.645	-776.91	1977.028	2753.940	1.884	23.534
TOVA RTV	108.069	46.660	93.503	47.817	299.780	251.963	1.516	3.440
TOVA d'prime	4.100	0.633	4.225	2.030	5.078	3.048	-1.03	0.047
Counter-manding accuracy	96.739	3.441	97.917	85.417	100.000	14.583	-1.16	0.254

UCancellation score	279.130	41.072	277.000	137.000	391.000	254.000	-0.01	3.028
UCancellation RTV	262.441	60.757	254.530	120.239	406.048	285.809	0.180	4.479
UCancellation RT switch	298.037	122.921	273.844	39.758	740.739	700.981	0.794	9.062
Negative urgency	9.440	2.715	9.000	4.000	16.000	12.000	0.225	0.200
Lack pre-meditation	6.886	1.971	7.000	4.000	13.000	9.000	0.486	0.145
Lack per-severance	6.929	1.856	7.000	4.000	12.000	8.000	0.322	0.137
Sensation seeking	10.788	2.767	11.000	4.000	16.000	12.000	-0.31	0.204
Positive urgency	8.163	2.510	8.000	4.000	16.000	12.000	0.462	0.185
BIS	13.060	3.475	13.000	7.000	24.000	17.000	0.554	0.256
Drive	8.717	2.282	9.000	4.000	16.000	12.000	0.189	0.168
Fun Seeking	7.918	2.295	8.000	4.000	15.000	11.000	0.299	0.169
Reward Respons-iveness	7.560	2.072	7.000	5.000	16.000	11.000	0.800	0.153
Inattention	2.429	1.929	2.000	0.000	9.000	9.000	0.820	0.142
Hyperactivity	2.228	1.942	2.000	0.000	9.000	9.000	0.987	0.143
Neuroticism	37.484	10.017	38.000	13.000	63.000	50.000	-0.04	0.738

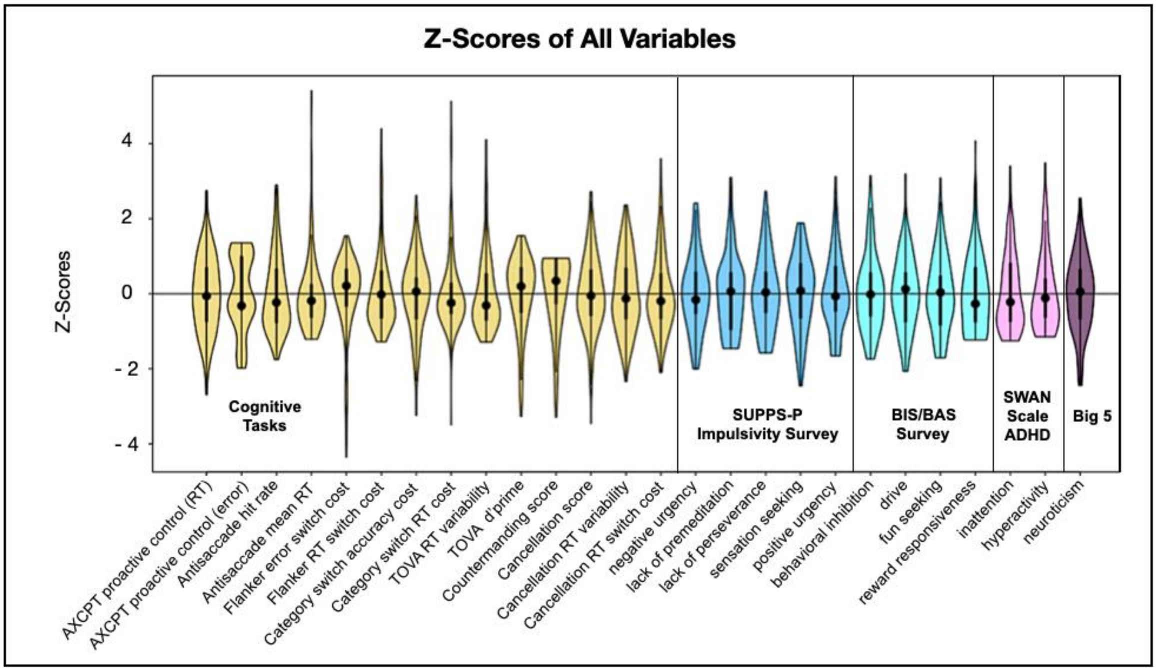


Figure 6. Z-score distributions of all variables.

3.2 Correlation Analysis

To calculate edge weights as input for the network analysis, we conducted a multiple correlation analysis and calculated Spearman correlation coefficients across all 26 variables. The full table is shown in Table 12. Multiple comparisons were corrected using False Discovery Rate (FDR) (Benjamini & Hochberg, 1995).

Table 12. Spearman correlations for all IC and impulsivity variables. Multiple comparisons were corrected using False Discovery Rate (FDR). Bold values represent statistically significant correlations ($p < 0.05$).

Spearman Correlations	AX-CPT proactive ctrl RT	AX-CPT proactive ctrl errors	Anti-saccade error rate	Anti-saccade mean RT	Flanker RT switch cost	Flanker error switch cost	Category switch accuracy switch	Category switch RT switch	TOVA RTV	TOVA d'prime	Counter-manding accuracy	UCancellation score	UCancellation RTV	UCancellation RT switch
AX-CPT proactive ctrl RT	1.000													
AX-CPT proactive ctrl errors	**0.284**	1.000												
Antisaccade error rate	0.049	-0.126	1.000											
Antisaccade mean RT	-0.084	-0.132	**0.476**	1.000										
Flanker RT switch cost	-0.056	0.015	-0.094	-0.068	1.000									
Flanker error switch cost	-0.085	-0.093	-0.152	**-0.202**	-0.106	1.000								
Category switch accuracy switch	-0.130	-0.045	0.023	0.012	0.019	-0.027	1.000							
Category switch RT switch	-0.065	-0.011	**0.217**	**0.241**	-0.067	-0.145	-0.087	1.000						
TOVA RTV	-0.059	-0.177	**0.353**	**0.338**	-0.031	-0.136	0.093	0.173	1.000					
TOVA d'prime	-0.003	0.027	**-0.279**	-0.183	0.075	-0.040	-0.036	-0.096	**-0.521**	1.000				
Countermanding accuracy	0.087	0.098	0.122	0.066	-0.068	-0.192	0.118	0.036	0.061	0.072	1.000			
UCancellation score	0.004	0.045	0.005	-0.068	-0.054	0.103	-0.017	0.018	-0.089	0.046	0.014	1.000		
UCancellation RTV	-0.010	-0.084	**0.297**	**0.280**	-0.050	-0.158	0.068	0.182	**0.331**	-0.168	-0.056	-0.195	1.000	
UCancellation RT switch	0.082	0.013	0.182	**0.216**	-0.034	-0.180	-0.020	0.159	**0.230**	-0.135	-0.094	-0.156	**0.583**	1.000

Spearman Correlations	AX-CPT proactive ctrl RT	AX-CPT proactive ctrl errors	Anti-saccade error rate	Anti-saccade mean RT	Flanker RT switch cost	Flanker error switch cost	Category switch accuracy switch	Category switch RT switch	TOVA RTV	TOVA d'prime	Counter-manding accuracy	UCancellation score	UCancellation RTV	UCancellation RT switch
Negative urgency	0.122	0.120	-0.054	-0.087	0.046	0.110	-0.100	-0.054	0.007	-0.062	0.033	-0.140	-0.088	-0.002
Lack pre-meditation	0.129	0.095	-0.125	-0.166	-0.014	-0.002	-0.080	-0.022	0.012	-0.041	-0.059	0.180	-0.138	-0.059
Lack per-severance	0.136	0.058	0.066	-0.011	-0.011	-0.030	0.010	0.042	0.162	**-0.246**	-0.029	0.115	-0.017	-0.031
Sensation seeking	-0.125	-0.116	-0.193	-0.030	0.138	0.078	-0.043	0.069	0.163	-0.104	-0.088	0.042	0.114	-0.027
Positive urgency	0.031	0.081	-0.107	-0.144	0.083	0.070	-0.071	-0.114	-0.043	0.022	-0.014	0.018	-0.080	-0.070
BIS	0.026	0.014	-0.031	0.123	0.025	-0.061	-0.092	0.020	0.020	-0.038	0.024	-0.078	0.118	0.088
Drive	-0.017	-0.029	-0.059	-0.028	-0.042	-0.066	0.028	-0.145	-0.050	0.075	-0.046	0.021	-0.056	-0.044
Fun Seeking	0.080	0.050	0.079	0.041	-0.056	-0.097	0.115	-0.066	-0.128	0.075	-0.094	-0.004	-0.045	-0.030
Reward Response	0.053	0.037	0.013	0.079	0.039	-0.160	0.021	0.028	-0.034	0.030	-0.097	-0.142	0.036	-0.004
Inattention	-0.047	0.084	-0.054	-0.036	-0.043	0.013	0.005	-0.069	-0.043	0.061	-0.104	-0.053	0.006	0.001
Hyperactivity	0.089	0.065	0.017	-0.101	-0.018	-0.087	-0.044	-0.038	0.039	-0.099	-0.182	-0.128	0.065	-0.042
Neuroticism	0.104	0.034	0.036	-0.039	-0.091	0.184	-0.035	0.008	-0.039	0.003	-0.060	0.060	-0.124	-0.123

Spearman Correlations	Negative urgency	Lack pre-meditation	Lack per-severance	Sensation seeking	Positive urgency	BIS	Drive	Fun Seeking	Reward Response	Inattention	Hyper-activity	Neuroticism
Negative urgency	1.000											
Lack pre-meditation	**0.212**	1.000										
Lack per-severance	0.077	**0.503**	1.000									
Sensation seeking	-0.031	-0.008	-0.106	1.000								
Positive urgency	**0.555**	**0.274**	0.142	0.148	1.000							
BIS	**-0.294**	0.058	0.077	**0.234**	-0.077	1.000						
Drive	-0.102	0.080	**0.214**	**-0.278**	-0.120	-0.162	1.000					
Fun seeking	-0.083	-0.030	0.058	**-0.452**	-0.163	-0.044	**0.459**	1.000				
Reward Response	-0.119	0.182	**0.308**	**-0.199**	-0.071	0.192	**0.437**	**0.417**	1.000			
Inattention	**0.229**	**0.407**	**0.368**	**-0.220**	0.142	-0.141	**0.235**	0.110	0.151	1.000		
Hyperactivity	**0.222**	**0.280**	**0.278**	-0.013	0.171	-0.078	0.097	0.088	0.112	**0.408**	1.000	
Neuroticism	**0.471**	**0.199**	0.069	**-0.272**	**0.303**	**-0.299**	0.047	0.190	0.029	**0.209**	**0.230**	1.000

3.3 Network Analysis

Inhibitory Control Network Model

First, we calculated a network model consisting of only IC measures and observed a moderate density (0.714). The network model is shown in Figure 7. Betweenness centrality estimates were calculated for each variable and indicate the “importance” of a node as it occurs on the same path between two other nodes. Betweenness centrality estimates for this model ranged between 0.00 and 8.938 (M = 3.74, SD = 2.893). All centrality estimates for this model are listed in Table 13.

This model also revealed four communities. The first community contained Antisaccade mean RT & error rate, TOVA RTV & d’prime, and Category Switch RT switch cost. The highest centrality estimate for this community was a tie between Antisaccade mean RT and TOVA RTV (3.938).

The second community contained Flanker error rate & RT switch cost, and Countermanding accuracy. The highest centrality estimate was for Flanker error switch cost (8.562). The third community contained Category Switch accuracy switch cost, and AX-CPT proactive control (RT & error rate). The highest centrality estimate for the third community was Category Switch accuracy switch cost (2.752). The fourth community contained all three UCancellation measures: RTV, RT switch cost, and score (concentration performance). The highest centrality estimate of this community, and of the entire IC network model, was UCancellation RTV (8.938).

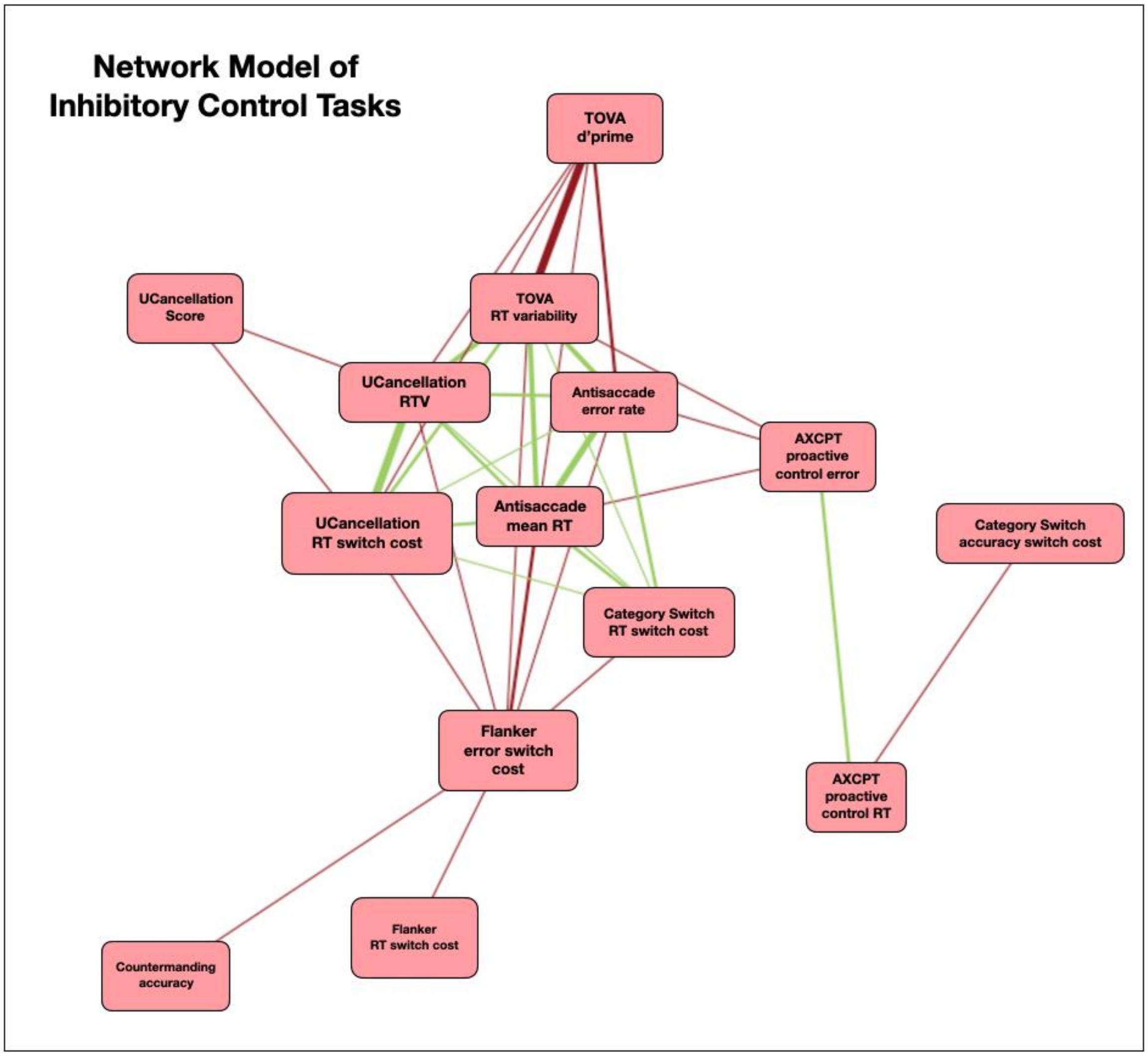


Figure 7. Network model of inhibitory control tasks.

Table 13. Betweenness centrality estimates and community assignment for each IC task. The highest estimate for each community is in bold font.

community	*IC task*	*betweenness centrality estimate*
1	Antisaccade mean RT	**3.938**
	TOVA RTV	3.938
	Category switch RT switch	3.229
	TOVA d'prime	1.757
	Antisaccade error rate	1.071
2	Flanker error switch cost	**8.562**
	Flanker RT switch cost	5.300
	Countermanding accuracy	0.000
3	Category switch accuracy switch	**2.752**
	AX-CPT proactive ctrl RT	2.043
	AX-CPT proactive ctrl errors	1.507
4	UCancellation RTV	**8.938**
	UCancellation RT switch	7.707
	UCancellation score	1.257

Impulsivity Network Model

Next, we calculated an additional network model using only impulsivity questionnaire subscales and observed a moderate network density (0.773), slightly higher than the IC network model. The network model is shown in Figure 8. This model produced three communities, and betweenness centrality estimates ranging from 0.905 to 4.126 (M = 2.500, SD = 0.938). The first community consisted of SUPPS-P positive & negative urgency, BIS/BAS BIS (behavioral inhibition), and Neuroticism. The highest centrality estimate in this community was SUPPS-P negative urgency (2.929). All centrality estimates for this model are listed in Table 14.

The second community consisted of SUPPS-P Sensation Seeking, and all three “BAS” subscales from BIS/BAS: reward responsiveness, drive, and fun-seeking. SUPPS-P

Sensation Seeking had the highest centrality estimate for this community, and for the entire impulsivity network model (4.126). The third community consisted of SWAN Scale Inattention & Hyperactivity, and SUPPS-P lack of premeditation & lack of perseverance. The highest centrality estimate for this community was lack of premeditation (3.560).

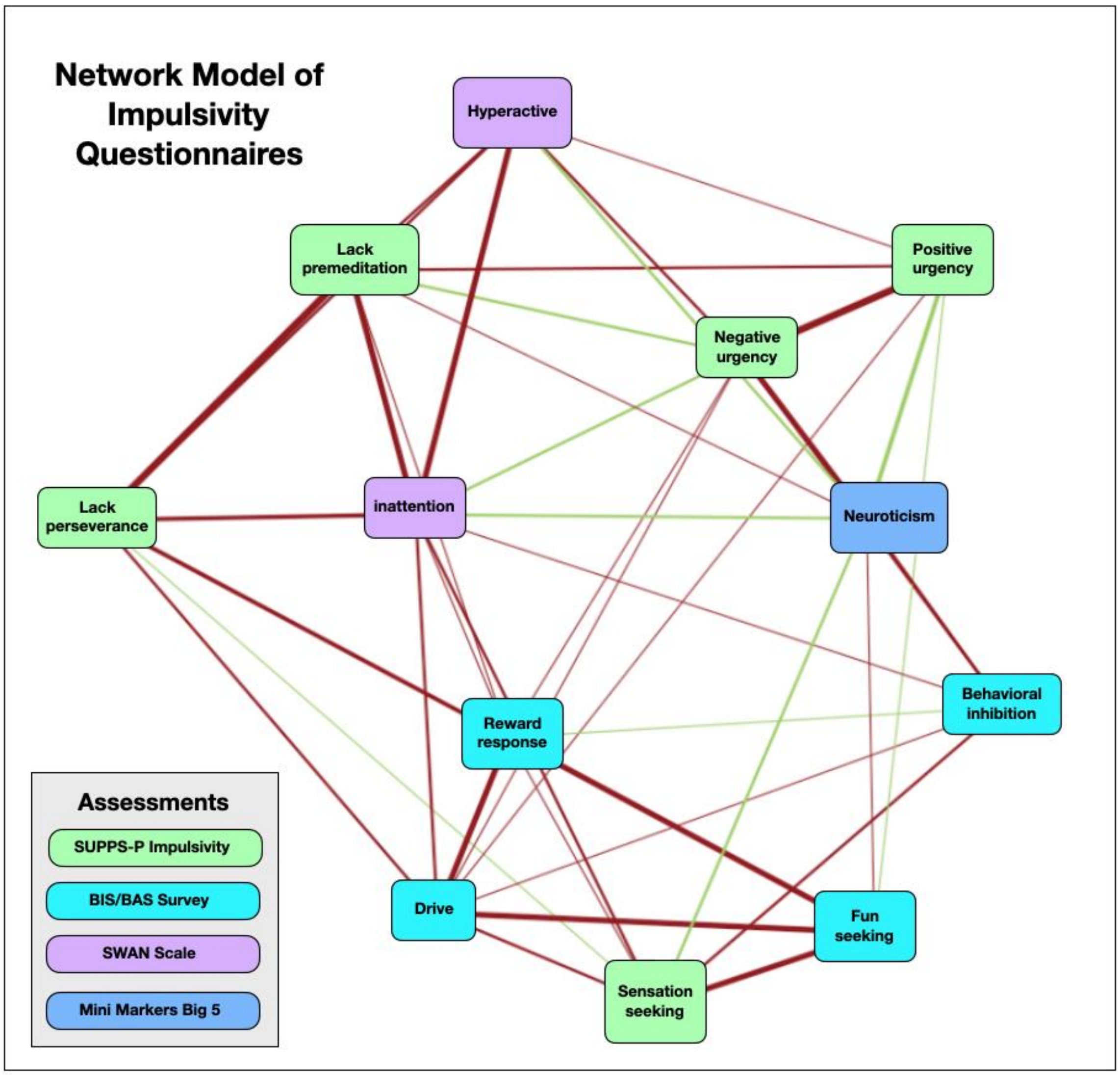


Figure 8. Network model of impulsivity questionnaire subscales.

Table 14. Betweenness centrality estimates and community assignment for each impulsivity subscale. The highest estimate for each community is in bold font.

community	*Impulsivity subscale*	*betweenness centrality estimate*
1	Negative urgency	**2.929**
	Behavioral inhibition (BIS)	2.393
	Positive urgency	2.107
	Neuroticism	1.405
2	Sensation seeking	**4.126**
	Reward responsiveness	3.090
	Drive	2.555
	Fun seeking	1.405
3	Lack of premeditation	**3.560**
	Inattention	3.007
	Hyperactive	2.519
	Lack of perseverance	0.905

Full Network Model

Finally, we calculated a network model containing all variables of IC and impulsivity and observed a moderate, yet relatively smaller density compared to the two previous models (0.658). The network model is shown in Figure 9. Betweenness Centrality estimates in this model ranged from 1.543 to 16.600 (M = 8.538, SD = 4.057). All centrality estimates are listed in Table 15.

The analysis revealed four clusters of nodes, or “communities.” The first community consisted of six variables: SUPPS-P lack of premeditation & lack of perseverance, SWAN Scale Inattention & Hyperactivity, Countermanding accuracy, and UCancellation score. The highest betweenness centrality measure for this community was SUPPS-P lack of premeditation (11.861).

The second community consisted of 11 IC tasks and no impulsivity questionnaires: Flanker error & RT switch costs, TOVA RTV & d’prime, UCancellation RTV & RT switch cost, Antisaccade mean RT & error rate, and AX-CPT proactive control for RT and error rate. The highest centrality estimate in the second community was Flanker error switch cost (16.600).

The third community contained five variables: SUPPS-P negative urgency & positive urgency, Category Switch accuracy switch cost, BIS/BAS BIS (behavioral inhibition), and Neuroticism. The highest centrality estimate of this community was SUPPS-P negative urgency (8.663).

The fourth community contained four variables, all from impulsivity questionnaires: SUPPS-P sensation seeking, and the three “BAS” subscales from BIS/BAS (drive, fun seeking, and reward responsiveness). The highest centrality estimate for this community was SUPPS-P sensation seeking (15.820).

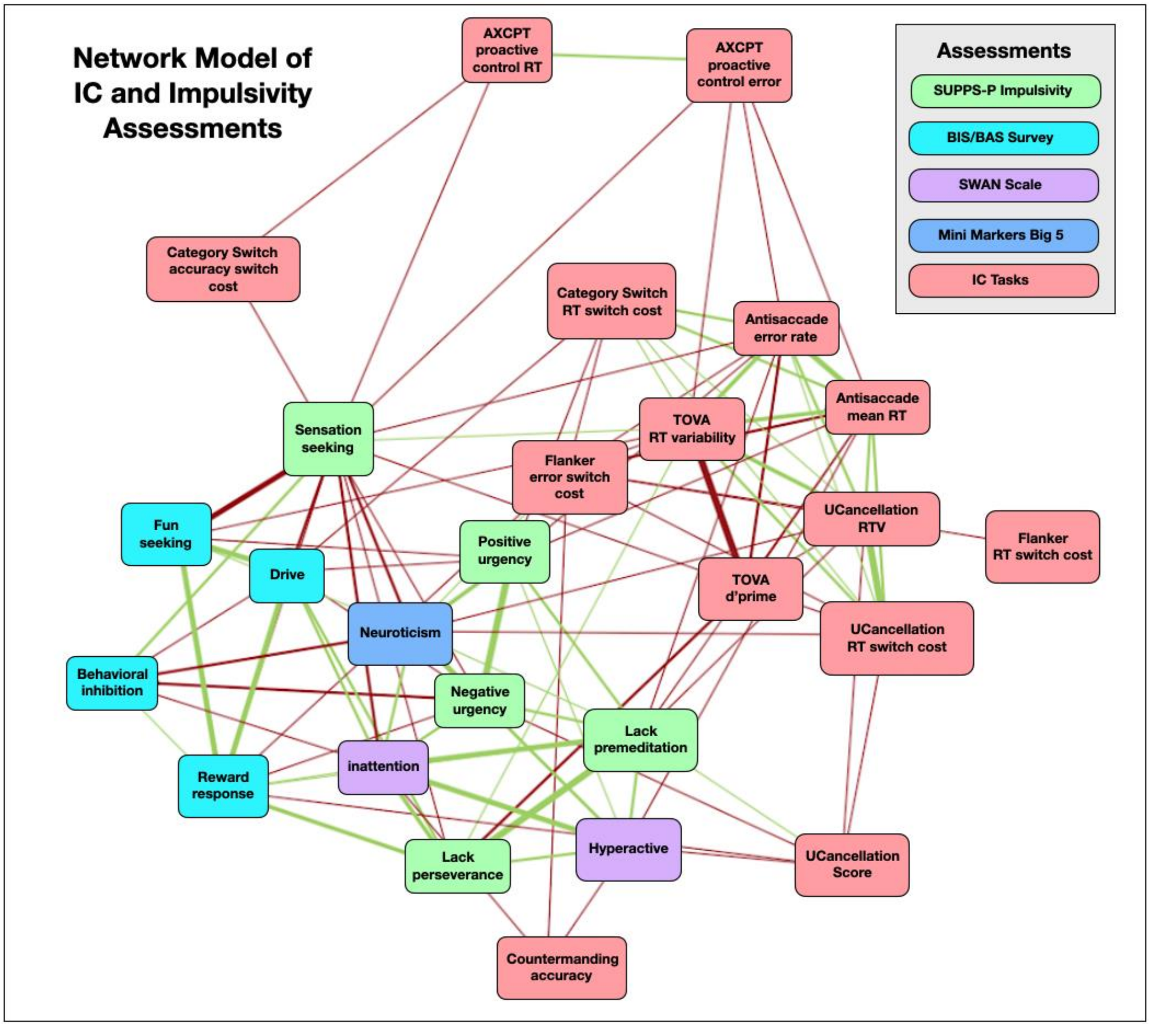


Figure 9. Network model of IC tasks and impulsivity questionnaire subscales.

Table 15. Betweenness centrality estimates and community assignment for all IC and impulsivity variables. The highest estimate for each community is in bold font.

community	*assessment*	*betweenness centrality estimate*
1	Lack of premeditation	**11.862**
	Inattention	10.366
	Hyperactive	8.392
	Cancellation score	4.876
	Countermanding accuracy	3.744
	Lack of perseverance	3.735
2	Flanker error switch cost	**16.600**
	TOVA RT variability	13.293
	UCancellation RT variability	13.123
	Antisaccade mean RT	12.705
	UCancellation RT switch cost	12.538
	Category Switch RT switch cost	9.773
	Flanker RT switch cost	7.675
	Antisaccade error rate	6.600
	TOVA d’prime	5.973
	AX-CPT proactive control (RT)	3.372
	AX-CPT proactive control (error)	1.543
3	Negative urgency	**8.663**
	Category Switch accuracy switch cost	7.928
	Positive urgency	7.113
	Neuroticism	6.214
	Behavioral inhibition (BIS)	6.199
4	Sensation seeking	**15.820**
	Drive	13.079
	Fun seeking	6.081
	Reward responsiveness	4.733

4. Discussion

To better characterize the relationship structure among IC and impulsivity assessments, we calculated three network models and evaluated network density and clusters, or "communities," of assessments in each model. The visualizations of these network models were determined using Spearman correlation coefficients as the edge weights of each model.

The network model of IC measures revealed four communities: the first community contained both TOVA measures, both Antisaccade measures, and Category Switch RT switch cost. This community of highly connected nodes may represent a similar cognitive function, namely selective or sustained attention. The second community contained both Flanker measures, and Countermanding score. This community may represent something unique to the Flanker task (while having some relationship with Countermanding) and could represent a distinct metric related to response inhibition. The third community contained both AX-CPT proactive control measures, and Category Switch accuracy switch cost, and may represent a distinct construct relating to proactive control. The fourth community was dominated by all three UCancellation measures, which suggests it captures a separate dimension of IC, or some quality unique to the task itself.

The impulsivity network model revealed three communities of questionnaire subscales. The first community contained measures from three questionnaires and included the following subscales: positive and negative urgency, neuroticism, and BIS (behavioral inhibition). This community may be tapping into a mechanism related to urgency and avoidance, in line with previous studies that find relationships among pairs of measures from this community (Tianxin et al., 2018; Whiteside & Lynam, 2001). The second community containing sensation seeking, fun seeking, drive, and reward responsiveness may be representing a reward-driven mechanism. The third community contained both SWAN subscales (hyperactivity and inattention) as well as lack of perseverance and lack of premeditation, both from the SUPPS-P impulsivity questionnaire. This community may be indicative of attentional issues.

In the full network model, the first community was identical to the third community in the impulsivity model but also included two accuracy-based IC task measures (Countermanding accuracy and UCancellation score). These scores can be interpreted as relating to attention and may share some variance with questionnaire subscales related to attention.

The community assignments observed in the IC network model were not observed in the full network model, and nearly all the IC tasks were all assigned to community 2, suggesting that IC is largely separate from impulsivity. The grouping of these IC tasks within a single community may be problematic, as it groups together metrics of different types of IC, such as selective attention, switching, and proactive control, and it's important to consider the network architecture of IC tasks from the IC model to better understand their relationship structure before comparing with impulsivity. However, the impulsivity measures were assigned to communities in the full model that mimicked the communities in the impulsivity model. The stronger correlations among impulsivity measures compared to correlations among IC tasks are the likely culprit for how the community detection algorithm assigned communities when all together in the full network model.

The highest centrality estimates within the full network model were SUPPS-P sensation seeking and Flanker error rate switch cost, but these variables are very weakly correlated, suggesting they represent two independent yet important processes. While the difference in format remains an issue, there may be something inherent to a physical response outcome measure that exceeds the awareness of the individual rating themselves on a questionnaire. This should be kept in mind for future studies that conflate measures of IC and impulsivity, while assigning a broad label that blurs the definitions of each. Care should be taken with future studies to make sure definitions of IC and impulsivity are clearly defined, and that composite scores, especially if taken from separate assessments, should be carefully chosen as to avoid redundancy or introduce conflation.

References

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple hypothesis testing. J R Stat Soc B 57:289–300

Bickel, W. K., Jarmolowicz, D. P., Mueller, E. T., Gatchalian, K. M., & McClure, S. M. (2012). Are executive function and impulsivity antipodes? A conceptual reconstruction with special reference to addiction. Psychopharmacology, 221(3), 361–387.

Brickenkamp, R., & Zillmer, E. (1998). The d2 Test of Attention. Seattle: Hogrefe & Huber Publishers. https://doi.org/10.1037/t03299-000

Brites, C., Salgado-Azoni, C. A., Ferreira, T. L., Lima, R. F., & Ciasca, S. M. (2015). Development and applications of the SWAN rating scale for assessment of attention deficit hyperactivity disorder: a literature review. Brazilian Journal of Medical and Biological Research = Revista Brasileira de Pesquisas Medicas E Biologicas / Sociedade Brasileira de Biofisica... [et Al.], 48(11), 965–972.

Burger J, Isvoranu AM, Lunansky G, Haslbeck JMB, Epskamp S, Hoekstra RHA, Fried EI, Borsboom D, Blanken TF. Reporting standards for psychological network analyses in cross-sectional data. Psychol Methods. 2023 Aug; 28(4):806-824. doi: 10.1037/met0000471. Epub 2022 Apr 11. PMID: 35404629.

Christ, S. E., Kester, L. E., Bodner, K. E., & Miles, J. H. (2011). Evidence for selective inhibitory impairment in individuals with autism spectrum disorder. Neuropsychology, 25(6), 690.

Cooper, S. R., Gonthier, C., Barch, D. M., & Braver, T. S. (2017). The Role of Psychometrics in Individual Differences Research in Cognition: A Case Study of the AX-CPT. Frontiers in Psychology,8. https://doi.org/10.3389/fpsyg.2017.01482

Diamond, A. (2013). Executive Functions. Annual Review of Psychology, 64(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. 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(12), 5472–5477.

Epskamp, S., Borsboom, D. & Fried, E.I. Estimating psychological networks and their accuracy: A tutorial paper. Behav Res 50, 195–212 (2018). https://doi.org/10.3758/s13428-017-0862-1

Everling, & Fischer. (1998). The antisaccade: A review of basic research and clinical studies. Neuropsychologia, 36(9), 885-899.

Forbes, G. B. (1998). Clinical utility of the Test of Variables of Attention (TOVA) in the diagnosis of attention-deficit/hyperactivity disorder. Journal of Clinical Psychology, 461–476. https://doi.org/10.1002/(sici)1097-4679(199806)54:4<461::aid-jclp8>3.0.co;2-q

Gray, J. D., Hanna, D., Gillen, A., & Rushe, T. (2016). A closer look at Carver and White's BIS/BAS scales: Factor analysis and age group differences. Personality and Individual Differences, 95, 20–24.

Greenberg, L. (1991). TOVA interpretation manual. Minneapolis (MN): University of Minnesota.

Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and inhibitory control. Psychological Science, 8(1), 60-64. doi:http://dx.doi.org/10.1111/j.1467-9280.1997.tb00545.x

Magnusdottir, B. B., Faiola, E., Harms, C., Sigurdsson, E., Ettinger, U., & Haraldsson, H. M. (2019). Cognitive Measures and Performance on the Antisaccade Eye Movement Task. Journal of Cognition, 2(1), 1–13. https://doi.org/10.5334/joc.52

Mäki-Marttunen, V., Hagen, T., Aminihajibashi, S., Foldal, M., Stavrinou, M., Halvorsen, J. H., Laeng, B., & Espeseth, T. (2018). Ocular signatures of proactive versus reactive cognitive control in young adults. Cognitive, Affective & Behavioral Neuroscience, 18(5), 1049–1063.

Marcora, S. M., Staiano, W., & Manning, V. (2009). Mental fatigue impairs physical performance in humans. Journal of Applied Physiology, 106(3), 857–864. https://doi.org/10.1152/japplphysiol.91324.2008

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 1124–1140.

Munakata, Y., Herd, S. A., Chatham, C. H., Depue, B. E., Banich, M. T., & O'Reilly, R. C. (2011). A unified framework for inhibitory control. Trends in Cognitive Sciences, 15(10), 453–459.

Newman, M., Newman, M., Girvan, M., & Girvan, M. (2003). Finding and evaluating community structure in networks. Physical review. E, Statistical, nonlinear, and soft matter physics, 69 2 Pt 2, 026113.

Newman, M. E. (2004). Fast algorithm for detecting community structure in networks. Physical review E, 69(6), 066133. DOI: 10.1103/PhysRevE.69.066133.

Pahor, A., Mester, R. E., Carrillo, A. A., Ghil, E., Reimer, J. F., Jaeggi, S. M., & Seitz, A. R. (2022). UCancellation: A new mobile measure of selective attention and concentration. Behavior Research Methods, 54(5), 2602–2617. https://doi.org/10.3758/s13428-021-01765-5

Roberts, W., Fillmore, M. T., & Milich, R. (2011). Linking impulsivity and inhibitory control using manual and oculomotor response inhibition tasks. Acta Psychologica, 138(3), 419–428.

Schall, J. D., Palmeri, T. J., & Logan, G. D. (2017). Models of inhibitory control. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 372(1718). https://doi.org/10.1098/rstb.2016.0193

Saucier G. Mini-markers: a brief version of Goldberg's unipolar big-five markers. J Pers Assess. 1994 Dec;63(3):506-16. doi: 10.1207/s15327752jpa6303_8. PMID: 7844738.

Sereno, A.B. & Holzman, P.S. (1995). Antisaccades and Smooth Pursuit Eye Movements in Schizophrenia. BIOL PSYCHIATRY, 37, 394-401.

Tianxin Mao, Weigang Pan, Yingying Zhu, Jian Yang, Qiaoling Dong, Guofu Zhou (2018). Self-control mediates the relationship between personality trait and impulsivity, Personality and Individual Differences, Volume 129, (70-75), https://doi.org/10.1016/j.paid.2018.03.013.

Tiego, J., Testa, R., Bellgrove, M. A., Pantelis, C., & Whittle, S. (2018). A Hierarchical Model of Inhibitory Control. Frontiers in Psychology, 9, 1339.

Vervoort, L., De Caluwé, E., Vandeweghe, L., De Decker, A., Wante, L., Van Beveren, M.-L., Goossens, L., Verbeken, S., Sioen, I., Michels, N., & Braet, C. (2019). Parent-Reported BIS/BAS Scales for Children: Factor Structure and Measurement Invariance Across Age and Gender. Assessment, 26(7), 1282–1295.

Whiteside, S. P., & Lynam, D. R. (2001). The Five Factor Model and impulsivity: Using a structural model of personality to understand impulsivity. Personality and Individual Differences, 30(4), 669–689.

Wilbertz, T., Deserno, L., Horstmann, A., Neumann, J., Villringer, A., Heinze, H., Schlagenhauf, F. (2014). Response inhibition and its relation to multidimensional impulsivity. NeuroImage, 103,241-248. doi:http://dx.doi.org/10.1016/j.neuroimage.2014.09.021

Chapter 3: UCancellation: Linking Novel Metrics to Executive Function, Proactive Control, and Sustained Attention

The study described in this chapter focuses on an inhibitory control (IC) task that was developed by our team: UCancellation. This task is based on the traditional D2 Test of Attention and has been redesigned and "gamified" with a smartphone app available for easy data collection. The aim of this study was to validate measures from UCancellation against measures from other previously established IC tasks. A major focus of this chapter was the extraction of a more accurate measure of RT variability. We found that reaction time (RT) measures from UCancellation predicted RT variability from another IC task, known as the Test of Variable Attention (TOVA). This suggests that UCancellation may be a useful alternative in place of the traditional and potentially less-engaging TOVA.

My contributions to this study included overseeing the set-up of the experiment protocol, deciding which assessments to include for analysis, conducting the regression analyses, and helping with writing the manuscript in preparation of submitting to a journal for publication. Special acknowledgements to Elnaz Vafaei for her help in processing the UCancellation data that allowed us to further explore RT variability, and Anja Pahor for her mentorship, guidance, and expertise as the leader of the analysis team for this chapter, and to both of them for their help with writing and revising.

Title: **UCancellation: Linking Novel Metrics to Executive Function, Proactive Control, and Sustained Attention**

Authors: Kamryn Mattingly, Elnaz Vafaei, Radhika Amin, Audrey Carrillo, Jason Reimer, PhD., Susanne Jaeggi, PhD., Aaron Seitz, PhD., Anja Pahor, PhD.

Abstract

Several cognitive functions are required for humans to navigate our day-to-day life, including selective attention, sustained attention, shifting, proactive control, and inhibitory control. Each of these cognitive processes have been studied extensively, and assessments have been developed to study each, however the relationship among them is less understood. Newer assessments have been developed and have begun to show convergent validity as alternatives to older, outdated interfaces, but more research is needed. We collected cognitive task data from healthy adults to identify if reaction-time (RT) or score-based measures from our novel, gamified sustained attention task (UCancellation) could predict similar metrics from other cognitive tasks. We found that UCancellation RT-based measures significantly predicted TOVA RT variability, a well-established measure of sustained attention, suggesting that UCancellation may be a viable alternative to the TOVA in research and clinical settings.

1. Introduction

Effective goal-directed behavior in dynamic environments requires the coordination of various cognitive processes that allow individuals to adapt to changing demands and stay focused on relevant tasks. For example, in tasks such as driving, selective attention ensures focusing on traffic signals, sustained attention enables ongoing environmental monitoring, proactive control facilitates the anticipation of potential hazards (e.g., a changing traffic light), and inhibitory control allows for the suppression of inappropriate or distracting responses (e.g., ignoring a phone notification). Additionally, set shifting plays a crucial role in driving by enabling the driver to efficiently shift attention between different tasks or stimuli, such as switching from focusing on the road to responding to a traffic signal, and then back to monitoring other vehicles or pedestrians. This flexibility in cognitive processes helps drivers adapt to changing situations and manage multiple demands in a dynamic environment. While each of these cognitive processes has been studied extensively in isolation, the interplay between them is less understood, highlighting the need to examine how they relate to each other in both standard and newly developed neuropsychological assessments. Specifically, this study explores whether alternative metrics from a selective attention task predict performance across tasks measuring sustained attention, proactive control, inhibitory control, and set switching.

Selective attention refers to the ability to focus on one or two specific, relevant stimuli while ignoring irrelevant or competing stimuli (also termed focused attention) (Cohen, 2014) . A robust measure of selective attention is the D2 test of attention, which requires participants to quickly identify and mark the target letters "d" with two dashes while ignoring distractors (Brickenkamp & Zillmer, 1998). There is evidence that D2 performance predicts academic skills, even when controlling for age (Arán Filippetti et al., 2022). In our earlier work, we introduced and validated UCancellation, a mobile version of the cancellation task with pictures and letters that also measures selective attention and correlates strongly with D2 performance (Pahor et al., 2022).

In contrast, sustained attention involves maintaining focus and performance over extended periods, ensuring consistent monitoring of tasks or stimuli. It is typically measured using a Continuous Performance Task (CPT), which involves presenting a series of stimuli over an extended period of time, and participants must remain vigilant to detect and respond to targets in a timely manner. One of the most widely used CPT tests is the Test of Variable Attention (TOVA) (Forbes, 1998; Greenberg, 1991). Sustained attention, as measured by CPT, predicts academic skills and is linked to math and reading performance (Gallen et al., 2023), highlighting the importance of attentional control in academic success and the potential for using sustained attention tasks as indicators of cognitive abilities related to learning. Lapses in sustained attention are often associated with diminished performance on tasks requiring response inhibition such as Antisaccade and Flankers (Unsworth et al., 2010). In addition, tests of vigilance and sustained attention, such as the CPT, are among also the most sensitive tools for identifying cognitive deficits in conditions like adult Attention-Deficit/Hyperactivity Disorder (ADHD) (Forbes, 1998).

Selective and sustained attention alone are not enough to ensure optimal performance in complex tasks. Based on the Dual Mechanisms of Control framework, proactive control involves sustained and anticipatory maintenance of goal-relevant information within lateral prefrontal cortex, whereas reactive control involves deploying attention as a “late correction” mechanism, activated only when necessary, typically in response to detecting a high-interference event (Braver, 2012). While the constructs of sustained attention and proactive control might appear similar, they differ in their primary focus and mechanisms: sustained attention focuses on maintaining engagement whereas proactive control involves actively managing and anticipating challenges to optimize performance.

Executive functions, such as inhibitory control and set shifting, are essential for adapting actions to dynamic environmental changes. Inhibitory control, the ability to suppress inappropriate or prepotent responses, can be measured by a variety of tasks such as Stop-signal, Go/NoGo, Countermanding and Antisaccade, which require the participants to suppress the tendency to make a prepotent motor response. These tasks assess how well individuals can control impulsive behaviors in a goal-directed manner. Similarly, set shifting, enables individuals to transition flexibly between mental frameworks, facilitating adaptation to evolving demands and complex situations. Together, these functions highlight the cognitive agility required for effective decision-making and behavior (Cochereau et al., 2021).

Previous research suggested that UCancellation performance, a measure of selective attention, was positively correlated with performance with an executive function composite that included inhibitory control and set shifting tasks. To further examine the role of UCancellation in relation to these cognitive processes, we conducted a study in which college students completed UCancellation Pictures and six widely used tasks: Flanker, Countermanding, and Antisaccade (inhibitory control), Category Switch (set shifting), AX-CPT (proactive control), and TOVA (sustained attention). It is important to note that while the tasks mentioned above are commonly believed to target specific cognitive processes, they actually engage multiple cognitive constructs simultaneously, meaning that results should be interpreted with caution.

The goals of the present study were twofold: (1) to determine the relationship between UCancellation and established measures of executive function, proactive control, and sustained attention; and (2) to explore whether novel metrics derived from UCancellation - such as reaction time variability and reaction time switch cost - predict performance on these established tests. By addressing these objectives, the study aimed to clarify the utility of UCancellation in assessing sustained attention and inhibitory control, potentially advancing its application in both research and clinical practice.

Based on previous findings (Pahor et al., 2022), we anticipated to find a significant positive relationship between UCancellation performance and performance on inhibitory and set shifting tasks. Moreover, we introduced a new metric for UCancellation that estimates reaction time switch cost between two types of targets, and explored how this metric relates to performance on other attention and inhibitory control tasks. We did not anticipate to find a significant relationship between selective attention performance and proactive control, as these are considered to be driven by non-overlapping mechanisms (Schröder et al., 2024). Furthermore, the relationship between UCancellation performance and measures of sustained attention, such as the CPT (e.g., TOVA), remains unknown. Given that intra-individual variability in reaction time is widely recognized as a key indicator of sustained attention (Yamashita et al., 2021), and that this is one of the main outcome measures in TOVA, we calculated reaction time variability (RTV) for UCancellation to explore its potential as a tool for screening sustained attention deficits. This is a relevant area of inquiry, as CPT tasks often require longer durations than cancellation tasks, creating practical challenges for their use in clinical and research settings. By evaluating whether UCancellation RTV can provide insights into sustained attention, we sought to enhance its utility as a faster screening tool for assessing sustained attention.

2. Methods

2.1 Participants

We recruited 511 participants as part of a larger study. All participants were undergraduate students recruited from UC Riverside, UC Irvine, and California State University San Bernardino. Data collection took place online during the onset of Covid-19 lockdown, and many participants did not complete all sessions. Due to missing data, we had full datasets from N = 272 participants, and after removing outliers, our final sample size for this study was N = 243. Outliers for reaction time-based scores were defined as responses made too quickly (< 200 ms) and for score-based measures, outliers were considered scores that exceeded four standard deviations. Participants included in our final sample were between the ages of 18 and 57 years old (M= 21.23 years; SD = 4.85 years). All participants provided informed consent and received course credit for their participation. They reported having normal or corrected-to-normal vision and hearing. In addition to a demographics questionnaire, participants completed several impulsivity-related questionnaires, though the latter were not analyzed in the present study.

2.2 Procedure

The study consisted of three sessions in which participants completed UCancellation, AX-CPT, Antisaccade, Flanker, TOVA, Category Switch Task, and Countermanding tasks on their personal tablets and computers. Participants also completed additional questionnaires as part of a larger study; these questionnaires are not included in this analysis. The assessments were administered via a custom-built app, Recollect the Study, and Inquisit Web via Millisecond Software (version 6.1.0.0), while being monitored by a researcher via Zoom. The interval between sessions was at most two weeks and each session lasted one hour or less. Participants took a 2-minute break between tasks, with the option to skip it. To account for fatigue effects, all assessments were counterbalanced across sessions. The order of assessments was also counterbalanced across data collection sites, and each site implemented two different orders of assessments, alternating with each participant.

2.3 Materials

2.3.1 UCancellation

This task is based on the D2 Test of Attention and is delivered via tablet or smartphone (Brickenkamp & Zillmer, 1998). Instead of letters as the stimuli, our version introduces a gamified makeover, using pictures of cartoon dogs and monkeys (Pahor et al., 2022). Each row displays 8 items, containing 3-5 targets. Every 10 rows include 40 targets, and each row is shown for 6 seconds. The objective is to select as many targets as possible and progress through as many rows as possible, without responding to non-targets. The participant must select the targets (upright dog and upside-down monkey) separately in single blocks and together in a mixed block. The primary outcome measure for UCancellation is Concentration Performance in the mixed block, calculated as ∑Hits - ∑False Alarms.

We aimed to extract reaction time data from this task but first addressed key considerations. Since the UCancellation task records reaction times only for targets and false alarms, we accounted for each target's position and the time elapsed between consecutive targets. This adjustment for calculating RTV was analyzed using correct mixed block trials, each of which featured both dogs and monkeys and consisted of three distinct target types: "First Target," "Switched Target," and "Non-switched Target". Figure 10 provides a visual representation of the sequence. The "First Target" represents the initial stimulus or object that participants are instructed to identify and respond to at the beginning of the sequence. A "Switched Target" denotes a scenario where the task requires a shift in attention from one category of target to another, reflecting a change in cognitive focus or demand. In contrast, a "Non-switched Target" corresponds to a consecutive presentation of the same target type, enabling sustained attention on a single category without necessitating a shift.

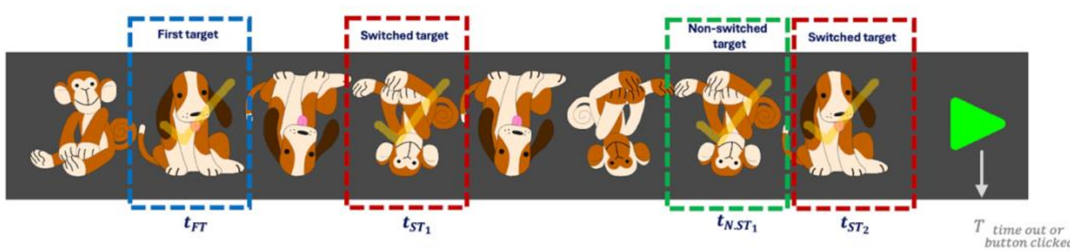


Figure 10. An example of a mixed block sequence including a "First Target" at t1, "Switched Targets" at t2 and t4, and a "Non-switched Target" at t3.

The RT for each target was categorized into RT for the First Target(RT(FT)), RT for the Switched Target (RT(ST)), and RT for the Non-switched Target (RT(N.ST)), calculated as follows (Figure 11):

$$\begin{cases} RT_{FT} = t_{FT} - T_{\substack{\textit{time out or}\\ \textit{button clicked}\\ \textit{of previous line}}} - T_{\text{Screen Pause}} \\ RT_{ST} = t_{ST_n} - t_{ST_{n-1}} \\ RT_{N.ST} = t_{N.ST_n} - t_{N.ST_{n-1}} \end{cases}$$

Figure 11. Equations for calculating for RT of first and switched targets.

If a participant completes a row before the time limit, they can press a button to proceed to the next row. T (representing either the timeout or the button click) denotes the moment when the participant either clicks the button to move to the next row or when the current row's time limit is reached. Due to a 1-second blank screen interval between rows, this time point must be considered when calculating the RT for the "First Target" to ensure accurate measurement.

To compare RTs for Switched targets and Non-switched targets accurately we corrected for the number of items processed since the last target. Without this adjustment, the RTs may reflect differences in processing demands rather than the true effects of target switching. By normalizing RTs based on the number of intervening items, we aim to isolate the cognitive mechanisms involved in responding to switched versus non-switched targets, providing a more precise analysis of task performance. Therefore, switched targets and non-switched targets are treated as part of the same distribution. This approach assumes that both types of targets, despite their differences in switching context, are drawn from a common underlying process or set of factors influencing response time. By merging the distributions, the aim is to examine the overall pattern of responses without biasing the analysis toward any one type of target. This correction allowed us to proceed with a more accurate calculation of RT-based measures, namely RT variability (RTV) and RT switch cost.

Our second outcome measure for UCancellation was RTV, commonly used in other tasks to evaluate how stable a participant's reaction times are when performing tasks that require rapid responses to stimuli (Antonini et al., 2013). Specifically, RTV was calculated based on the standard deviation of the response times across rows, providing an indicator of how much an individual’s reaction times fluctuate during the task. High RTV suggests more inconsistency in the participant’s responses, while low RTV indicates a more stable and predictable pattern of performance (Anguera et al., 2022; Yamashita et al., 2021; Ziegler et al., 2019).

Our third outcome measure from UCancellation was RT Switch Cost, which refers to the increase in RTs observed when participants switch from one task or response set to another, compared to when they perform a task continuously without switching. This phenomenon is often attributed to cognitive processes such as attentional shifts, reorganization of cognitive resources, or conflict resolution (Dykstra et al., 2022). Tasks that are more complex, demand higher cognitive effort, or involve multi-step processes typically result in greater switch costs due to the additional cognitive effort required to adapt to the new task demands. In the context of the Cancellation Task, which requires participants to identify and mark specific targets while ignoring distractors, switch cost is calculated as the difference in RTs between Switched Targets and Non-Switched Targets. Larger values of RT switch cost suggests that participants are slowing down for Switched Targets, while smaller RT switch cost values suggest participants are responding within a similar duration regardless of target type (Draheim et al., 2019).

2.3.2 AX-CPT.

This task was delivered via the online platform Inquisit version 6.1.0.0 and is a continuous performance task (CPT) with a modified duration of 9 minutes (Cooper et al., 2017; Marcora et al., 2009). Participants are instructed to select “E” on the keyboard when a target sequence is presented and “I” for all other sequences (non-targets). Target sequences were presented by the letter A, which in this context is the “cue” and followed by the “probe,” which was the letter X. The other possible (non-target) sequences were AY, BX, and BY. A fixation cross was shown between each letter. This task contains a total of 180 trials, with 126 target trials, and 18 non-target trials. Our measure of interest from this task was proactive control, sometimes referred to as the Proactive Behavioral Index (PBI), which we calculated twice: once using the average RT (ms), and again using errors from AY and BX trials: (AY-BX)/(AY+BX) (Braver et al., 2009; Mäki-Marttunen et al., 2018). The range of proactive control spans -1 and 1, where negative values represent “reactive” responding, and positive values represent more “proactive” control.

2.3.3 Antisaccade.

This task was delivered using the online platform Inquisit and was based on the original antisaccade task, which is traditionally used as a measure of inhibition. The antisaccade task is intended to probe how well one can prevent the automatic reflex of looking towards a target stimuli (Everling & Fischer, 1998; Sereno & Holzman, 1995). A fixation cross was presented on the center of screen, and a yellow square appeared to the left or right of the fixation cross for 150 ms. At the same time the yellow square is shown, there is an “O” or “Q” displayed on the side of the screen opposite of the square for 175 ms. The objective of this task is to resist the reflex of looking towards the yellow square so participants can have enough time to look towards the other side of the screen to identify if an “O” or “Q” was shown. Once the target letter has been shown, it’s immediately masked with “##” and the participant is instructed to press the letter that was displayed for that trial. The task lasted 7 minutes and the duration of target letter presentation was decreased to have a higher level of difficulty, and to reduce the chance of participants seeing the target letter in the event of failing to suppress looking at the yellow square. Our measure of interest was error rate, with larger values indicating more errors, and smaller values indicating less errors (Friedman et al., 2008).

2.3.4 Flanker.

This task was administered using the online software Inquisit and requires participants to correctly indicate the orientation of a target stimuli amidst distracting stimuli presented on either side of the target. For every trial, there are five arrows presented in a horizontal line, spatially jittered as to avoid fixation of gaze, and each arrow points to the left or right. Participants were instructed to indicate the direction of the arrow in the center (target) by pressing the button on the keyboard that matches the center arrow. Congruent trials include only trials where the center arrow matches the direction of the side arrows, and incongruent trials consist of trials with the center arrow pointing the opposite direction of the side arrows. Incongruent trials create an increased cognitive demand, which typically results in slower response times (Christ et al., 2011). By calculating the difference between congruent and incongruent trials, we can estimate the average RT switch cost. This metric is similar to RT switch cost described in 2.3.1 for the UCancellation task.

2.3.5 Category Switch Task.

This is a rule-switching task, administered online using Inquisit, and lasts approximately 7 minutes (Mayr & Kliegl, 2000). For each trial, a word appears on the screen, and participants are prompted to respond according to two rules: choose if the displayed word is “living” or “non-living,” or if the word is “smaller” or “bigger” than a basketball. The practice block consists of 32 trials (16 each rule type), and the test block consists of 64 trials (32 each rule type), and all trials are randomized. Congruent trials are defined as two subsequent trials of the same rule, while incongruent trials switch from one rule to the other. The outcome measure of interest for this task is accuracy switch cost, calculated as the difference in mean number of correct responses for congruent versus incongruent trials. As incongruent trials present a higher cognitive demand, it’s expected that participants will have a greater number of errors for incongruent trials.

2.3.6 Countermanding.

This task was administered online through the app Recollect the Study. The design of this task is a combination of the Simon task and the spatial Stroop task (Diamond, 2013). Participants are presented with two different conditions requiring two different responses. In the first condition (congruent) an image of a cartoon dog is presented on the right side of the screen, and the participant must select the green button below on the same side.

For incongruent trials, participants are presented with an image of a cartoon monkey on the right side of the screen but must select the green button on the opposite (left) side of the screen. Participants advance upon responding, or after a 15 second timeout. There are three test blocks: congruent (12 trials), incongruent (12 trials), and a randomized mixed block (48 trials). The main variable of interest for this task is average accuracy across mixed trials. Due to a coding error, reaction time data was not available for the full sample of participants.

2.3.7 Test of Variable Attention (TOVA).

This task was administered online using Inquisit and lasts approximately 14 minutes (Forbes, 1998; Greenberg, 1991). Participants are presented with a central fixation point and must wait for a stimulus (small square) to appear on the screen. If the stimulus is above the fixation point, this is a target stimulus, and participants must press the spacebar. If the stimulus appears below the fixation point, this is a non-target, and participants do not respond. This task contains a practice session consisting of 50 trials, randomized, with half the trials including a target stimulus. There are two types of test blocks: low frequency and high frequency. Low frequency blocks have a lower proportion of target stimuli (22.2% of stimuli are targets), compared to high frequency blocks, which have a target frequency of 77.8%. Each test block lasts about 3 minutes, and participants complete each type of test block twice. This task produces several outcome variables, but we chose reaction time variability (RTV) as our main measure of interest in this task, for a more even comparison with RTV in our UCancellation task.

2.4 Statistical Analysis

To identify shared variance across all measures of interest, we calculated the Spearman correlation coefficients for each pair, correcting for multiple comparisons according to the false discovery rate (FDR) (Benjamini & Hochberg, 1995). To characterize how much UCancellation can predict other cognitive tasks, we conducted multiple regression analyses using UCancellation measures and age as predictors, and the outcome variables as measures from other cognitive tasks. We kept the origin of measures congruent within each analysis (i.e., if the outcome variable was accuracy-based, our predictor(s) were also accuracy-based; same for RT-based measures).

3. Results

When comparing UCancellation measures to other cognitive tasks, we found that UCancellation scores significantly correlated with some measures from other cognitive tasks. Furthermore, UCancellation RTV and UCancellation RT switch cost significantly predict the variance of TOVA RTV. Additionally, age was a significant predictor for Flanker RT switch cost and Countermanding accuracy.

3.1 Descriptive statistics

Descriptive statistics for all variables are shown in Table 16 and violin plots of Z-scores of all measures are shown in Figure 12. Histograms of UCancellation variables are shown in Figure 13.

Table 16. Descriptive statistics of all variables. N = 243.

Descriptive Statistics	*Mean*	*S.D.*	*Median*	*Min*	*Max*	*Range*	*Skew*	*S.E.*
Age (years)	20.864	3.557	20.00	18.00	39.000	21.000	2.517	0.228
AX-CPT proactive control (errors)	0.200	0.565	0.000	-0.947	0.957	1.904	-0.309	0.036
AX-CPT proactive control (RT)	0.149	0.100	0.160	-0.200	0.420	0.620	-0.177	0.006
Antisaccade error rate	0.226	0.126	0.189	0.011	0.589	0.578	0.895	0.008
Flanker RT switch (ms)	83.003	52.398	78.750	-109.96	250.19	360.15	0.161	3.361
Category Switch accuracy switch cost	-0.025	0.053	-0.021	-0.208	0.104	0.313	-0.589	0.003
TOVA RTV (ms)	99.242	37.750	89.859	0.000	210.18	210.18	0.668	2.422
Counter -manding accuracy	96.759	3.388	97.917	85.417	100.00	14.583	-1.130	0.217
Cancellation score	279.765	40.882	277.000	137.00	391.00	254.00	-0.037	2.623
Cancellation RTV (ms)	262.780	60.206	257.489	114.16	406.04	291.85	0.107	3.862
Cancellation RT switch cost (ms)	292.024	114.207	275.194	-10.133	622.70	632.84	0.534	7.326

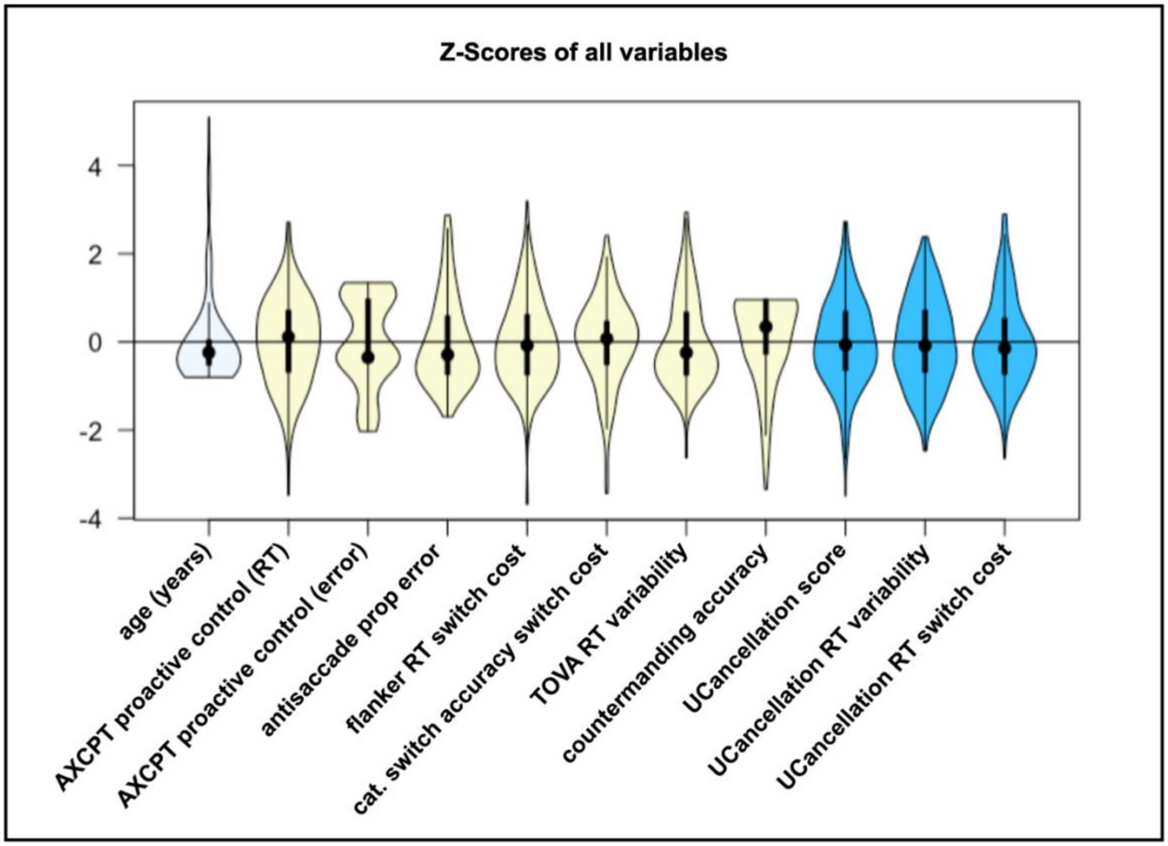
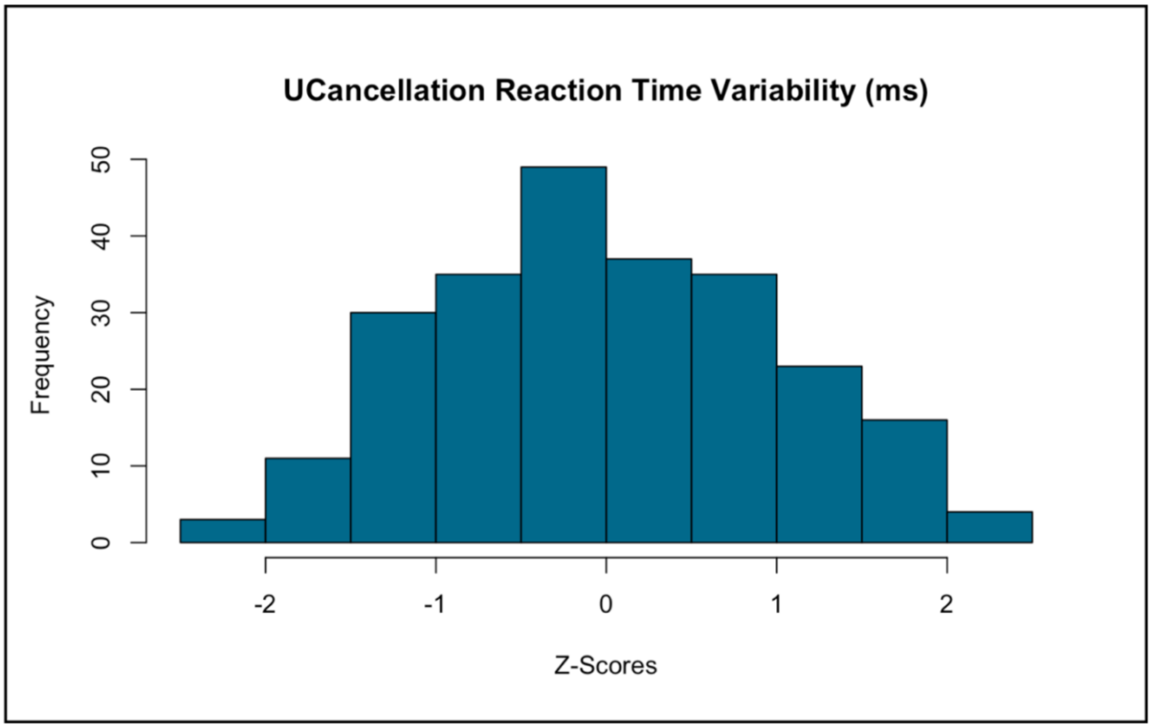
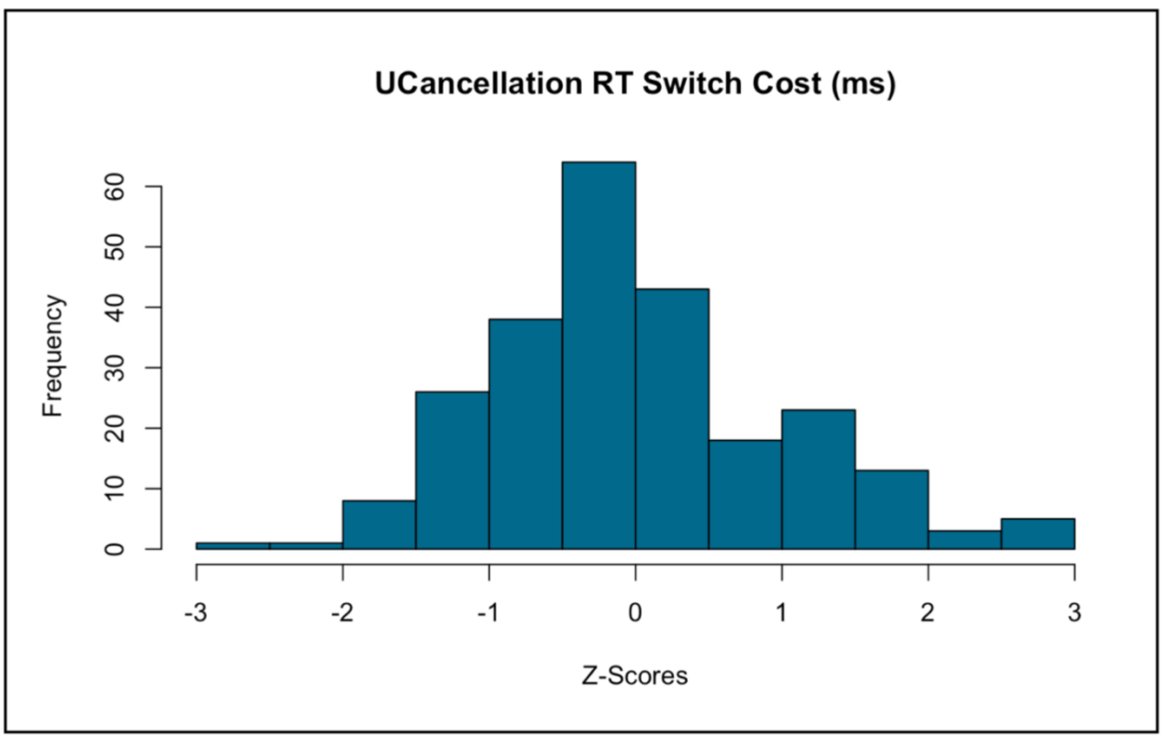


Figure 12. Distributions of all measures.


UCancellation Reaction Time Variability (ms)
Frequency
0
10
20
30
40
50
-2
-1
0
1
2
Z-Scores


UCancellation RT Switch Cost (ms)
Frequency
0
10
20
30
40
50
60
-3
-2
-1
0
1
2
3
Z-Scores

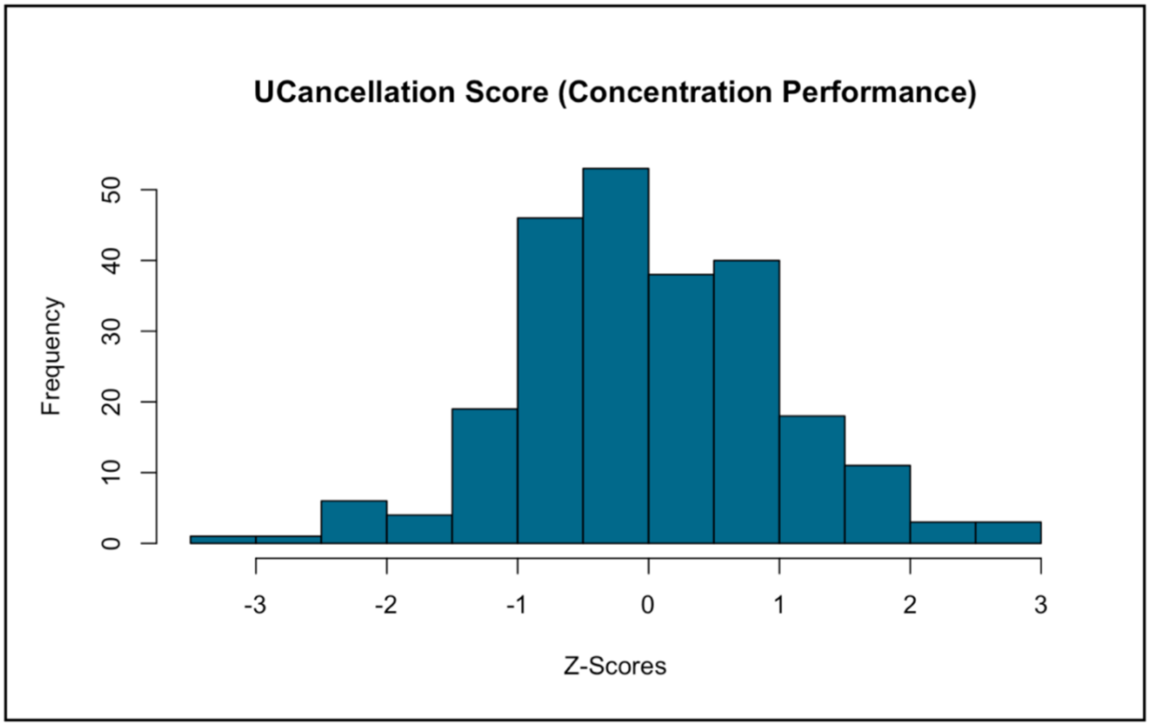


Figure 13. Histograms of UCancellation RT Variability, RT Switch Cost, and Score.

3.2 Correlation Analysis

Spearman correlation coefficients are shown in Table 17, with bold values indicating significance of p < 0.05, using a False Discovery Rate (FDR) correction for multiple comparisons (Benjamini & Hochberg, 1995). Age showed a significant positive correlation with Flanker RT switch cost (*r* = 0.174) and Countermanding accuracy (*r* = 0.170). Antisaccade error rate significantly correlated with TOVA RTV (*r* = 0.348), UCancellation RTV (*r* = 0.297) and UCancellation RT switch cost *(r* = 0.188). TOVA RTV was also significantly correlated with UCancellation RTV (*r* = 0.325) and UCancellation RT switch cost (*r* = 0.231). UCancellation RTV and UCancellation RT switch cost were significantly correlated with each other (*r* = 0.552). UCancellation score was significantly correlated with UCancellation RTV (*r* = -0.189). AX-CPT proactive control (RT) was significantly correlated with AX-CPT proactive control (error) (r = 0.296). Category Switch Task accuracy switch cost did not significantly correlate with any other measure in this analysis.

Table 17. Spearman correlation matrix of all cognitive task variables. Bold values indicate significance of $p < 0.05$, after correction for multiple comparisons using False Discovery Rate (FDR).

Spearman Correlations	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***
1 *Age*	1										
2 *AXCPT proactive control (RT)*	0.133	1									
3 *AXCPT proactive control (error)*	0.091	**0.296**	1								
4 *Anti-saccade error rate*	0.044	0.098	-0.033	1							
5 *Flanker RT switch cost (ms)*	**0.174**	0.142	0.017	0.117	1						
6 *Category Switch accuracy switch cost*	-0.083	-0.070	-0.072	-0.040	-0.048	1					
7 *TOVA RTV (ms)*	-0.142	0.005	-0.160	**0.348**	0.097	0.096	1				
8 *Counter-manding accuracy*	**0.170**	0.065	0.000	0.071	0.093	0.068	0.029	1			
9 *Cancel score*	-0.100	0.077	0.074	-0.006	-0.021	-0.031	-0.135	-0.015	1		
10 *Cancel RTV (ms)*	-0.084	0.019	-0.021	**0.297**	0.026	0.051	**0.325**	-0.203	**-0.189**	1	
11 *Cancel RT switch cost (ms)*	-0.152	0.090	0.092	**0.188**	0.026	0.009	**0.231**	-0.071	-0.085	**0.552**	1

3.3 Regression Analysis

To determine how much variance of commonly used inhibition tasks is predicted by UCancellation measures, we conducted six multiple linear regressions analyses. Due to some scores originating from reaction measures and others calculated as an accuracy measure, we used congruent types of UCancellation measures as predictors/regressors in each regression equation. Age was included as a covariate for every regression analysis.

To test for normality of residuals of regression analyses, we conducted a Shapiro-Wilk test on each regression model. Residuals were normally distributed for the regression analysis of AX-CPT proactive control (RT) ($p = 0.519$), but not for AX-CPT proactive control (errors), TOVA, Category Switch, Flanker, Antisaccade, or Countermanding ($p < 0.05$), suggesting residuals from these models are less likely to follow a normal distribution. To correct non-normal residuals in our regression analyses, we applied transformations to the outcome variables as detailed in subsequent sections.

3.3.1 AX-CPT Proactive Control.

Since AX-CPT proactive control (error) is a score-based measure, (as opposed to RT-based) we calculated a multiple regression analysis using UCancellation score and age as regressors to predict the variance of AX-CPT proactive control (error). This model was not statistically significant $F(2,240) = 0.1191$, $R^2 = 0.001$, $p = 0.888$. Neither age nor UCancellation score predicted any significant amount of variance of AX-CPT proactive control ($\beta = 0.010$, $t = -0.183$, $p = 0.855$ and $\beta = 0.001$, $t = 0.419$, $p = 0.676$, respectively).

For predicting AX-CPT proactive control (RT) we calculated a separate multiple regression using UCancellation RTV, UCancellation RT switch cost, and age as regressors. This model was not statistically significant $F(3,239) = 1.512$, $R^2 = 0.006$, $p = 0.212$. As predicted, none of the regressors in this model predicted any significant amount of variance of AX-CPT proactive control (RT): age ($\beta = 0.002$, $t = 1.827$, $p = 0.069$), UCancellation RTV (β = 0.001, t = -0.918, p = 0.360), and UCancellation RT switch cost (β = 0.001, t = 1.353, p = 0.177).

3.3.2 Antisaccade Error Rate.

A multiple regression analysis was conducted to determine whether UCancellation score predicts Antisaccade error rate. The residuals for this model were non-normal based on the results of a Shapiro-Wilk test ($p < 0.001$) so we applied a cube-root transformation and achieved normality of residuals ($p = 0.08$). The model was not significant ($F(2,240)$ = 1.044, R^2 = 0.0086, p = 0.354) and neither UCancellation score nor age predicted Antisaccade error rate (β = 0.001, t = 0.403 , p = 0.687 and β = 0.0021., t = 1.433 , p = 0.153, respectively).

3.3.3 Flanker Reaction Time (RT) Switch Cost.

Since the primary outcome variable for Flanker is reaction time-based, a multiple regression analysis was conducted with Flanker RT switch cost as the dependent variable, using UCancellation RT variability and UCancellation RT switch cost as regressors and age as a covariate. The residuals for this model were non-normal based on the results of a Shapiro-Wilk test ($p < 0.001$) so we applied a square-root transformation and achieved normality of residuals ($p = 0.8201$). The model was significant ($F(3,239)$ = 3.363, R^2 = 0.0417, p = 0.0195), however age was the significant predictor (β = 1.256, t = 4.077 , $p < 0.001$). Neither UCancellation RT variability nor UCancellation RT switch cost predicted any significant amount of variance of Flanker RT switch cost (β = 0.0034, t = 0.393, p = 0.6946 and β = 0.002, t = 0.993 , p = 0.217, respectively).

3.3.4 Category Switch Task Accuracy Switch Cost.

A multiple regression analysis was conducted to examine whether UCancellation score predicts accuracy switch cost in the Category Switch Task. The residuals for this model were non-normal based on the results of a Shapiro-Wilk test ($p < 0.001$) so we applied a series of transformations to achieve normality of the residuals. None of the attempted transformations (i.e., log transformation, square root, and cube root) achieved normality as calculated in a Shapiro-Wilk test, but the closest approximation was achieved using a BoxCox transformation ($p = 0.01$) (Box & Cox, 1964). This model was not significant ($F(2,240) = 0.1862$, $R^2 = 0.0015$, $p = 0.8302$) and neither UCancellation score nor age predicted any significant amount of variance of Category Switch Task accuracy switch cost ($\beta = 0.001$, $t = -0.489$, $p = 0.625$ and $\beta = 0.001$, $t = -0.436$, $p = 0.663$, respectively).

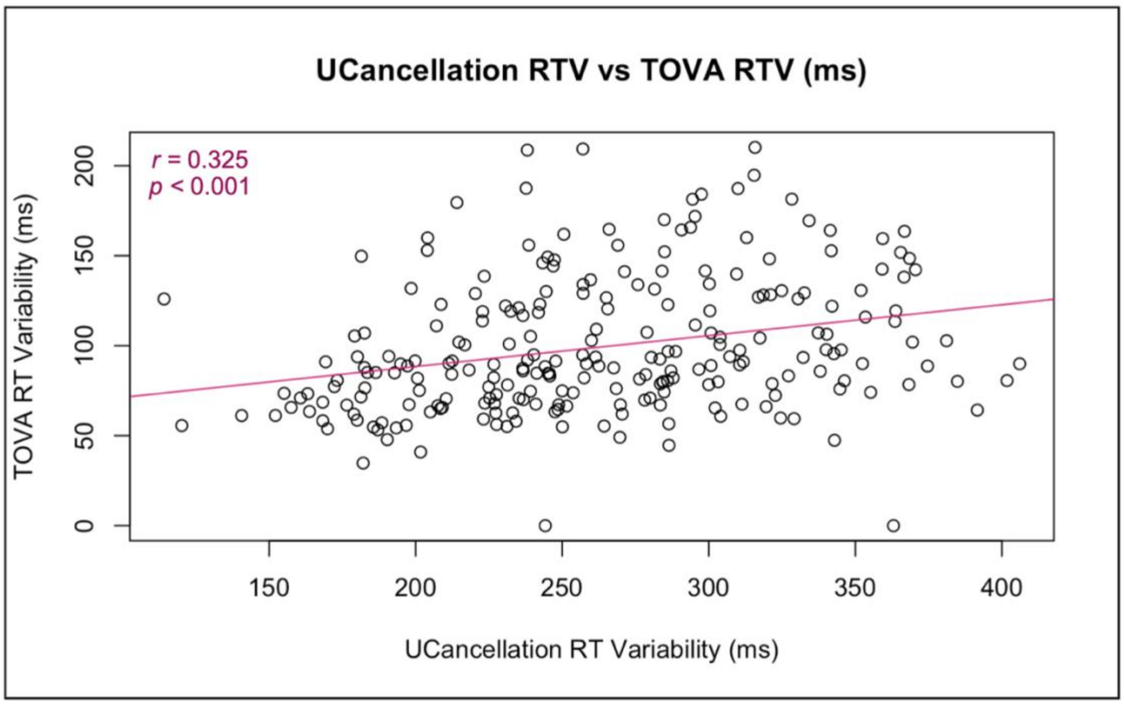
3.3.5 Countermanding Accuracy.

Since our Countermanding measure is accuracy-based, we chose UCancellation score and age as regressors in a multiple regression analysis with Countermanding accuracy as the dependent variable. The residuals for this model were non-normal based on the results of a Shapiro-Wilk test ($p < 0.001$) so we applied a series of transformations to achieve normality of the residuals. None of the attempted transformations (i.e., log transformation, square root, and cube root) achieved normality as calculated in a Shapiro-Wilk test, but the closest approximation was achieved using a BoxCox transformation (p = 3.02 x 10^-11). The model was significant ($F(2,240) = 5.222$, $R^2 = 0.042$, $p = 0.006$), however age was the only regressor that predicted any significant amount of variance ($\beta = 5.78$, $t = 3.203$, $p = 0.002$). UCancellation score did not predict Countermanding accuracy ($\beta = 0.503$, $t = 0.072$, $p = 0.943$).

3.3.6 TOVA Reaction Time (RT) Variability.

Since the primary outcome variable for TOVA is RT-based, we chose UCancellation RT variability and UCancellation RT switch cost as regressors (in separate analyses). First, we calculated a multiple regression analysis with TOVA RT variability as the dependent variable and UCancellation RTV as the regressor. The residuals for this model were non-normal based on the results of a Shapiro-Wilk test (p < 0.001) so we applied a series of transformations to achieve normality of the residuals. None of the attempted transformations (i.e., log transformation, square root, and cube root) achieved normality as calculated in a Shapiro-Wilk test, but the closest approximation was achieved using a BoxCox transformation (p = 7.095 x 10^-6). The model was significant (F(2,240) = 12.83, R^2 = 0.097, p < 0.0001). UCancellation RTV and age were both significant predictors of TOVA RTV (β = 0.010, t = 4.419 , p < 0.0001 and β = 0.169, t = -2.539 , p = 0.0118, respectively). Next, we calculated a multiple regression analysis with TOVA RT variability as the dependent variable and UCancellation RT switch cost as the regressor. We calculated another BoxCox transformation and attempted to achieve normality of residuals as calculated in a Shapiro-Wilk test (p = 1.251 x 10^-5). The model was significant (F(2,240) = 8.287, R^2 = 0.0646, p < 0.001). UCancellation RT switch cost and age were both significant predictors of TOVA RTV (β = 0.0045, t = 3.278, p = 0.0012 and β = 0.1443, t = -1.982, p = 0.0487, respectively). The scatter plot and regression line between UCancellation RT switch cost and TOVA RTV is shown in Figure 14.

To examine which UCancellation measure better predicts TOVA RT variability, both regressors were included in the model simultaneously. We again calculated a BoxCox transformation for this model to attempt normality of the residuals (p = 7.93 x 10^-06). This model was significant ($F(3,239) = 8.985$, $R^2 = 0.1014$, $p < 0.001$). UCancellation RT variability and age were both significant predictors ($\beta = 0.012$, $t = 3.092$, $p = 0.0022$ and $\beta = 0.171$, $t = -2.330$, $p = 0.021$, respectively). On the other hand, UCancellation RT switch cost did not predict TOVA RT variability ($\beta = 0.0063$, $t = 1.123$, $p = 0.2625$). The scatter plot and regression line between UCancellation RTV and TOVA RTV is shown in Figure 14.



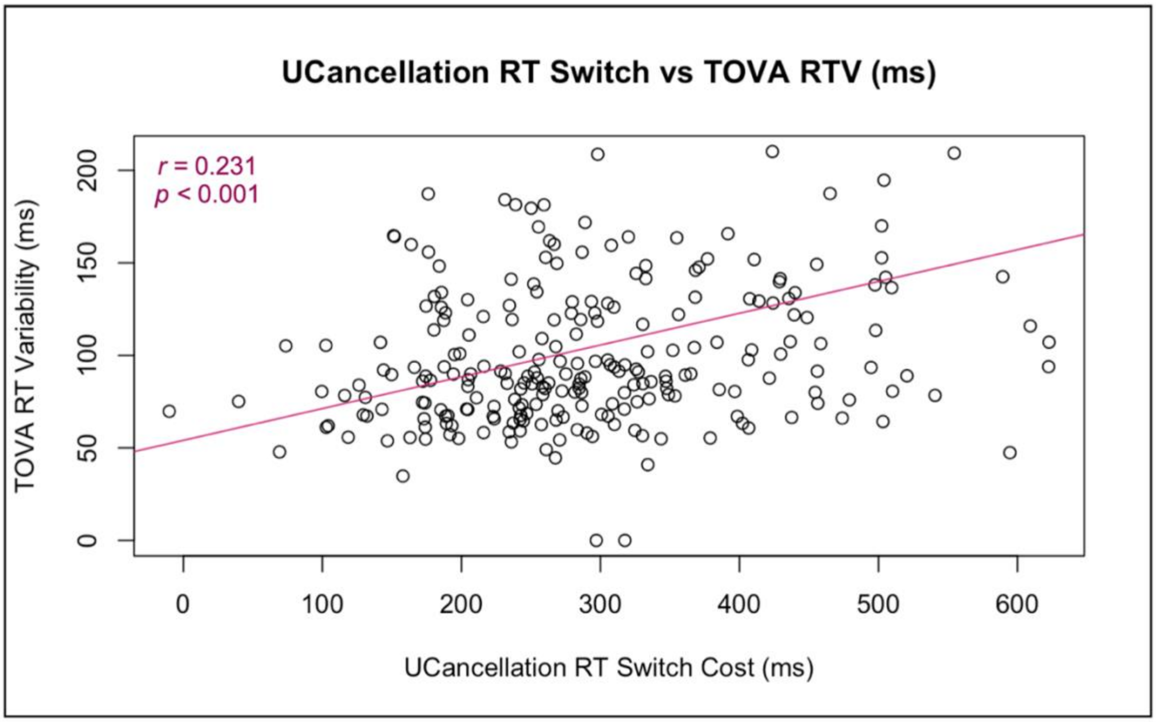


Figure 14. Scatter plots showing the significant relationships between TOVA RTV with UCancellation RTV (top) and UCancellation Switch Cost (bottom).

These results show that UCancellation score does not significantly predict score-based measures of AX-CPT, Antisaccade, or Category Switch. However, UCancellation RT variability does significantly predict RT variability in a measure of sustained attention (TOVA). Despite also being an RT-based measure, UCancellation switch cost did not significantly predict the variance of RT-based measures from other cognitive tasks, especially when controlling for age. Age was consistently a significant regressor in most RT-based regression analyses, but an insignificant regressor in all score-based regression analyses, except for Countermanding.

4. Discussion

In this study, we tested whether measures from our novel, gamified UCancellation task predicted other commonly used cognitive tasks in a sample of healthy young adults. We found that UCancellation RT-based measures significantly predicted TOVA RTV and may serve as a more user-friendly alternative to the TOVA.

We hypothesized that measures from our UCancellation task would significantly predict outcome measures from other commonly used inhibition tasks, specifically that RT-based measures of UCancellation would predict RT-based measures on other tasks, and accuracy/error-based scores would predict accuracy/error-based scores on other cognitive tasks. The relationships between traditional cognitive tasks and measures extracted from our UCancellation revealed some significant correlations, however many of these relationships did not survive significance when assessed in a regression analysis.

Other significant correlations were observed that did not fall within the scope of our analysis. Namely, Antisaccade error rate was significantly correlated with UCancellation RTV and UCancellation RT switch cost, but we did not analyze RT-based measures as predictors of accuracy/error-based scores in our regression analyses. The relationship between Antisaccade error rate and RT-based measures may reflect the task design for Antisaccade, where the timing of the distractor stimulus presentation is tightly bound with the opportunity for the participant to successfully view the target stimulus. Failing to avoid the distractor stimulus in Antisaccade results in insufficient time to see the target and answer the trial prompt correctly. Future studies may find it useful to explore the relationship between error rates on fast-paced cognitive tasks and RT measures from other types of slower-paced tasks.

There was no evidence that UCancellation score is significantly related to either measure of AX-CPT proactive control (RT or error-based), which is consistent with previous findings exploring selective attention and continuous performance tasks (Schröder et al., 2024).

Age was significantly correlated with, and significantly predicted Flanker RT switch cost and Countermanding accuracy. While the majority of our sample was younger adults (age 18-25 years), we had participants up to 39 years old. The positive correlation between age and Flanker RT switch cost suggests some relationship could exist between aging and RT-based measures of task-switching. In this case, a greater value of Flanker RT switch cost indicates that participants are slowing down for incongruent trials due to the increased cognitive demand, and that this impact on performance becomes worse with age. However, age was also positively correlated with Countermanding accuracy. This accuracy measure is that of “concentration performance” and this would indicate that task performance improves with increased age. Given the smaller range of age in our sample (18-39 years old), age-related findings should be taken with caution, as brain development isn’t complete in humans until approximately mid-20s, and cognitive function decline in healthy adults isn’t typically observed until much later in life (age 50+ years) (Kray & Lindenberger, 2020; Knox et al., 2020; Pujol et al., 1993). A larger age range would be more informative to probe possible age-related effects as they relate to inhibition and other cognitive functions.

In our regression analyses, UCancellation RTV and RT switch cost both significantly predicted TOVA RTV. This suggests that UCancellation RT-based measures capture a similar cognitive process associated with response time. Implementing UCancellation may be a viable alternative for testing RTV in future studies and has the added benefit of a more engaging interface with cartoon stimuli that are more appealing and motivating than the traditional stimulus used in the TOVA (a white square over a black background). Despite this finding, there are several other available outcome measures from the TOVA that we did not extract from UCancellation or focus on in our analysis, such as average accuracy, average reaction time, and measures specific to low and high-frequency conditions. Future studies may find it beneficial to include both TOVA and UCancellation, but in certain cases, UCancellation may be sufficient for capturing comparable RTV measures depending on the experimental design.

Given that UCancellation takes less than ten minutes to complete, it provides a practical alternative for assessing sustained attention and cognitive control without the added risk of fatigue-related interference. This brevity allows it to be feasibly incorporated alongside other cognitive tasks within a single assessment session, facilitating more comprehensive evaluations of attentional processes while minimizing participant burden.

References

Anguera, J. A., Volponi, J. J., Simon, A. J., Gallen, C. L., Rolle, C. E., Anguera-Singla, R., Pitsch, E. A., Thompson, C. J., & Gazzaley, A. (2022). Integrated cognitive and physical fitness training enhances attention abilities in older adults. Npj Aging, 8(1), 12. https://doi.org/10.1038/s41514-022-00093-y

Antonini, T. N., Narad, M. E., Langberg, J. M., & Epstein, J. N. (2013). Behavioral correlates of reaction time variability in children with and without ADHD. Neuropsychology, 27(2), 201.

Arán Filippetti, V., Gutierrez, M., Krumm, G., & Mateos, D. (2022). Convergent validity, academic correlates and age- and SES-based normative data for the d2 Test of attention in children. Applied Neuropsychology: Child, 11(4), 629–639. https://doi.org/10.1080/21622965.2021.1923494

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple hypothesis testing. J R Stat Soc B 57:289–300

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. Trends in Cognitive Sciences, 16(2), 106–113. https://doi.org/10.1016/j.tics.2011.12.010

Braver, T. S., Paxton, J. L., Locke, H. S., & Barch, D. M. (2009). Flexible neural mechanisms of cognitive control within human prefrontal cortex. Proceedings of the National Academy of Sciences, 106(18), 7351–7356. https://doi.org/10.1073/pnas.0808187106

Brickenkamp, R., & Zillmer, E. (1998). The d2 Test of Attention. Seattle: Hogrefe & Huber Publishers. https://doi.org/10.1037/t03299-000

Christ, S. E., Kester, L. E., Bodner, K. E., & Miles, J. H. (2011). Evidence for selective inhibitory impairment in individuals with autism spectrum disorder. Neuropsychology, 25(6), 690.

Cochereau, J., Zandvoort, M., Santarius, T., & Mandonnet, E. (2021). Set Shifting. In E. Mandonnet & G. Herbet (Eds.), Intraoperative Mapping of Cognitive Networks: Which Tasks for Which Locations (pp. 273–285). Springer International Publishing. https://doi.org/10.1007/978-3-030-75071-8_17

Cohen, R. A. (2014). Focused and Sustained Attention. In R. A. Cohen (Ed.), The Neuropsychology of Attention (pp. 89–112). Springer US. https://doi.org/10.1007/978-0-387-72639-7_6

Cooper, S. R., Gonthier, C., Barch, D. M., & Braver, T. S. (2017). The Role of Psychometrics in Individual Differences Research in Cognition: A Case Study of the AX-CPT. Frontiers in Psychology, 8. https://doi.org/10.3389/fpsyg.2017.01482

Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. Psychological Bulletin, 145(5), 508.

Dykstra, T., Smith, D. M., Schumacher, E. H., & Hazeltine, E. (2022). Measuring task structure with transitional response times: Task representations are more than task sets. Psychonomic Bulletin & Review, 29(5), 1812–1820. https://doi.org/10.3758/s13423-021-02035-3

Forbes, G. B. (1998). Clinical utility of the Test of Variables of Attention (TOVA) in the diagnosis of attention-deficit/hyperactivity disorder. Journal of Clinical Psychology, 461–476. https://doi.org/10.1002/(sici)1097-4679(199806)54:4<461::aid-jclp8>3.0.co;2-q

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. Journal of Experimental Psychology: General, 137(2), 201.

Gallen, C. L., Schaerlaeken, S., Younger, J. W., Younger, J. W., O'Laughlin, K. D., Anguera, J. A., Bunge, S. A., Ferrer, E. E., Hoeft, F., McCandliss, B. D., Mishra, J., Rosenberg-Lee, M., Gazzaley, A., Uncapher, M. R., Anguera, J. A., Gazzaley, A., & Project iLEAD Consortium. (2023). Contribution of sustained attention abilities to real-world academic skills in children. Scientific Reports, 13(1), 2673. https://doi.org/10.1038/s41598-023-29427-w

Greenberg, L. (1991). TOVA interpretation manual. Minneapolis (MN): University of Minnesota.

Knox PC, Pasunuru N. Age-related alterations in inhibitory control investigated using the minimally delayed oculomotor response task. PeerJ. 2020 Jan 9;8:e8401. doi: 10.7717/peerj.8401. PMID: 31942260; PMCID: PMC6955103.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. Psychology and Aging, 15(1), 126–147.

Mäki-Marttunen, V., Hagen, T., Aminihajibashi, S., Foldal, M., Stavrinou, M., Halvorsen, J. H., Laeng, B., & Espeseth, T. (2018). Ocular signatures of proactive versus reactive cognitive control in young adults. Cognitive, Affective, & Behavioral Neuroscience, 18(5), 1049–1063. https://doi.org/10.3758/s13415-018-0621-5

Marcora, S. M., Staiano, W., & Manning, V. (2009). Mental fatigue impairs physical performance in humans. Journal of Applied Physiology, 106(3), 857–864. https://doi.org/10.1152/japplphysiol.91324.2008

Oeseburg, B., Jansen, D. E. M. C., Dijkstra, G. J., Groothoff, J. W., & Reijneveld, S. A. (2010). Prevalence of chronic diseases in adolescents with intellectual disability. Research in Developmental Disabilities, 31(3), 698–704. https://doi.org/10.1016/j.ridd.2010.01.011

Pahor, A., Mester, R. E., Carrillo, A. A., Ghil, E., Reimer, J. F., Jaeggi, S. M., & Seitz, A. R. (2022). UCancellation: A new mobile measure of selective attention and concentration. Behavior Research Methods, 54(5), 2602–2617. https://doi.org/10.3758/s13428-021-01765-5

Pujol, J., Vendrell, P., Junqué, C., Martí-Vilalta, J.L. and Capdevila, A. (1993), When does human brain development end? Evidence of corpus callosum growth up to adulthood. Ann Neurol., 34: 71-75. https://doi.org/10.1002/ana.410340113

Schröder, E., Ingels, A., Dumitrescu, A., Kornreich, C., & Campanella, S. (2024). Proactive and Reactive Inhibitory Control Strategies: Exploring the Impact of Interindividual Variables on an ERP Continuous Performance Task (AX-CPT). Clinical EEG and Neuroscience, 55(3), 317–328. https://doi.org/10.1177/15500594221145905

Unsworth, N., Redick, T. S., Lakey, C. E., & Young, D. L. (2010). Lapses in sustained attention and their relation to executive control and fluid abilities: An individual differences investigation. Intelligence, 38(1), 111–122. https://doi.org/10.1016/j.intell.2009.08.002

Yamashita, A., Rothlein, D., Kucyi, A., Valera, E. M., Germine, L., Wilmer, J., DeGutis, J., & Esterman, M. (2021). Variable rather than extreme slow reaction times distinguish brain states during sustained attention. Scientific Reports, 11(1), 14883. https://doi.org/10.1038/s41598-021-94161-0

Ziegler, D. A., Simon, A. J., Gallen, C. L., Skinner, S., Janowich, J. R., Volponi, J. J., Rolle, C. E., Mishra, J., Kornfield, J., Anguera, J. A., & Gazzaley, A. (2019). Closed-loop digital meditation improves sustained attention in young adults. Nature Human Behaviour, 3(7), 746–757. https://doi.org/10.1038/s41562-019-0611-9

General Discussion

This dissertation sought to evaluate and characterize the relationships among assessments used for measuring inhibitory control and impulsivity in a healthy adult sample. The findings from these analyses support the hypothesis that IC and impulsivity are separate constructs possibly driven by independent brain mechanisms. However, some overlap may exist that could be detected across accuracy-based scores and self-report measures related to selective attention. Despite these findings, assessments of IC and impulsivity were largely unrelated, and the difference in format remains a confounding factor. Future studies must be cautious with how IC is defined and operationalized as to avoid redundancy and conflation across measures chosen.

Additionally, the UCancellation task from Chapters 1 and 2 was brought to greater focus in Chapter 3 for further validation as a measure of sustained attention, for its novel interface and accessibility via a smartphone app. The findings from this study support that UCancellation RT variability may be a useful alternative for the traditional IC task TOVA. Ultimately, the results of this dissertation provide a critique of how IC and impulsivity measures are operationalized and compared across a large scale, with a focus on including specific subscales and multiple metrics from IC tasks and avoiding composite scores when possible.

Chapter 1 revealed a factor structure that separated IC from impulsivity, with three factors containing only impulsivity subscales and one factor with all IC measures. Chapter 2 revealed a network structure that mimicked the factor analysis in Chapter 1. Conceptually, they are both calculated with a correlation matrix as the input, with different statistical decisions made for each. The grouping of all IC measures into a single factor poses an issue as there were multiple dimensions of IC captured across the measures chosen for this analysis. While the factor analysis in Chapter 1 may not have delineated between IC measures, the IC network model in Chapter 2 provided a dissected perspective of how the IC measures in our sample compared to one another. Despite the assignment of IC measures into different categories, impulsivity subscales reliably fell into multiple factors in Chapter 1 and into multiple network communities in Chapter 2, for both the impulsivity model and the full network model.

This may partially be due to the development of many self-report questionnaires. Traditionally, surveys and questionnaires follow a psychometric standard that is often validated by using different types of factor analyses, which are based on correlations among survey questions to produce sub-scales. These sub-scales are labeled according to theoretical constructs, and the score for each sub-scale is believed to measure certain trait or quality about a person.

One limitation for Chapters 1 and 2 is that impulsivity questionnaires, due to the nature of their development and methods for validity, are inherently more well-behaved in subsequent factor analyses, and correlation-based analyses such as the network model in Chapter 2. IC tasks have been shown to less reliable than questionnaires in previous studies and this may contribute to the mismatch of IC and impulsivity assessments. Further, the correlations among impulsivity questionnaires were higher than among IC tasks. The measures selected from IC tasks spanned a wider range of behaviors (sustained attention, switching, proactive control, etc.) compared to our selected impulsivity measures, and this is consistent with the correlational structure observed across all measures in our sample.

Chapter 3 demonstrated that UCancellation, a novel, user-friendly alternative to the traditional cognitive task TOVA may be a valid option for measuring RT-based measures of sustained attention. While multiple metrics can be extracted from each IC task in this dataset, we opted to focus on accuracy and RT-based measures in traditional cognitive tasks that measure attention, control, and switching. Future studies may want to include other metrics from additional cognitive tasks to explore newer and more accessible alternatives to outdated traditional cognitive assessments.

Limitations during this study were rampant as the onset of Covid-19 prohibited in-person data collection. Because every participant was administered assessments over Zoom video calls, we encountered a very high number of technical issues. Given the unpredictability of this period, we had a high proportion of participants who did not complete all sessions of data collection. For data we did collect, it was difficult to enforce compliance on task performance, and we received a significant portion of unusable data. For this reason, we had to exclude the Stroop task, typically known as a standard for measuring inhibitory control. Similar studies in the future should highly consider including the Stroop in their list of assessments to better characterize how other measures of IC and impulsivity may relate to its scores.

Future studies should also consider an adolescent population for a similar analysis, as frontal brain regions related to higher-level cognitive control are still developing. Many of the assessments used in this dissertation are commonly used for screening clinical disorders such as ADHD, in both adults and children. Many of the assessments used here have child-specific versions that capture similar dimensions of IC and impulsive behavior.

Finally, the main confounding limitation was format difference—future studies may wish to include “tasks” of impulsivity, such as the BART (Balloon Analogue Risk Test), or questionnaires of “inhibition,” such as the MSCS (Multi-Dimensional Self-Control Survey) (Lejuez, 2002; Nilsen, 2020). By including formats that measure both IC and impulsivity, one can potentially eliminate this confound, and determine if IC and impulsivity are more related than they revealed to be using standard assessments in our analyses.

In conclusion, this dissertation had the objective of characterizing inhibitory control and impulsivity in a healthy adult sample, and we found that IC and impulsivity are not related when using traditional cognitive tasks and impulsivity questionnaires. We highlighted the importance of operationalization and definition of what metrics are used from certain assessments and how these measures relate to each other. The use of assessments used for screening IC and impulsivity in a clinical setting is costly and time-consuming, and these findings can inform which of these measures may be redundant or most informative depending on the goals of the clinician. By eliminating redundancy of assessments used, one can determine the most ideal battery of assessments to administer to a patient or research participant for robust and informative characterization of inhibitory control and impulsive behavior.