Journal of Experimental Psychology: General

Cognitive Control Across the Lifespan: Congruency Effects Reveal Divergent Developmental Trajectories

Christopher D. Erb, Laura Germine, and Joshua K. Hartshorne

Online First Publication, June 8, 2023. https://dx.doi.org/10.1037/xge0001429

AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

https://doi.org/10.1037/xge0001429

BRIEF REPORT

Cognitive Control Across the Lifespan: Congruency Effects Reveal Divergent Developmental Trajectories

Christopher D. Erb[1], Laura Germine[2, 3], and Joshua K. Hartshorne[4]
[1] School of Psychology, University of Auckland
[2] Department of Psychiatry, Harvard Medical School
[3] Institute for Technology in Psychiatry, McLean Hospital, Belmont, Massachusetts, United States
[4] Department of Psychology and Neuroscience, Boston College, Chestnut Hill, Massachusetts, United States

The Simon, Stroop, and Eriksen flanker tasks are commonly used to assess cognitive control across the lifespan. However, it remains unclear whether these three tasks in fact measure the same cognitive abilities and in the same proportion. We take a developmental approach to this question: if the Simon, Stroop, and flanker tasks all roughly measure the same capacity, they should show similar patterns of age-related change. We present data from two massive online cross-sectional studies: Study 1 included 9,585 native English speakers between 10 and 80 years of age who completed the Simon and Stroop tasks, and Study 2 included 13,448 English speakers between 10 and 79 years of age who completed the flanker task. Of the three tasks, only the flanker task revealed an inverted U-shaped developmental trajectory, with performance improving until approximately 23 years of age and declining starting around 40 years of age. Performance on the Simon and Stroop tasks peaked around 34 and 26 years of age, respectively, and did not decline significantly in later life, though it is possible that age-related declines would be observed with more difficult versions of the tasks. Although the Simon and Stroop tasks are commonly interpreted to target similar underlying processes, we observed near zero correlations between the congruency effects observed in each task in terms of both accuracy and response time. We discuss these results in light of recent debates regarding the suitability of these tasks for assessing developmental and individual differences in cognitive control.

Public Significance Statement
Cognitive control—the ability to direct one's thoughts and actions—is a key part of human intelligence. We investigated human cognitive control across the lifespan using three well-known measures of cognitive control. Surprisingly, these three measures provided very different results, challenging current understanding of cognitive control and intelligence over the lifespan.

Keywords: cognitive control, flanker task, lifespan development, Simon task, Stroop task

Supplemental materials: https://doi.org/10.1037/xge0001429.supp

Cognitive control refers to the ability to bring one's ongoing thoughts and actions into alignment with one's current goals and context. Age-related changes in this ability present important implications for a wide range of domains, including education (Blair, 2016), the criminal justice system (Altikriti, 2021), and healthcare (Insel et al., 2006). The development of cognitive control is commonly described as following an inverted U-shaped trajectory, with pronounced gains observed between childhood and early adulthood, and declines in late adulthood (e.g., Dempster, 1992; Zelazo et al., 2004). However, the literature presents a far more complicated picture, with some studies showing evidence of an inverted U-shape (Cohn et al., 1984; Davidson et al., 2006; Erb et al., 2020; Erb & Marcovitch, 2018, 2019; Gathercole et al., 2014; Luna et al., 2004; van der Lubbe & Verleger, 2002; Vu & Proctor,

Christopher D. Erb https://orcid.org/0000-0002-1649-2324
Portions of these results have been presented at the Cognitive Development Society and the Cognitive Science Society.
Data and analysis code are available at https://osf.io/mv8zu/.
Christopher D. Erb served as lead for writing–original draft and served in a supporting role for formal analysis. Laura Germine served as lead for project administration and served in a supporting role for formal analysis and writing–original draft. Joshua K. Hartshorne served as lead for formal analysis and project administration and served in a supporting role for writing–original draft. Christopher D. Erb, Laura Germine, and Joshua K. Hartshorne contributed equally to conceptualization and writing–review and editing.
Correspondence concerning this article should be addressed to Christopher D. Erb, School of Psychology, University of Auckland, 23 Symonds Street, Building 302, Auckland 1010, New Zealand. Email: christopher.erb@auckland.ac.nz

2008; Waszak et al., 2010), others showing no evidence of declines in late adulthood (LaPlume et al., 2022; Verhaeghen & De Meersman, 1998), and some studies reporting *improvements* in late adulthood (Veríssimo et al., 2022).

A variety of factors likely contribute to this muddled picture of cognitive control's development, including differences in task difficulty (Bugg et al., 2007; LaPlume et al., 2022), differences in how conflict effects were computed (Christ et al., 2001), the contributions of contingency learning and feature-integration confounds (Braem et al., 2019), and the extent to which individual differences in processing speed were taken into account (Rey-Mermet & Gade, 2018; Verhaeghen, 2011). There is also increasing concern regarding the extent to which the tasks commonly used to assess cognitive control tap into similar underlying processes. As noted by Draheim et al. (2019), within-participant comparisons of performance on cognitive control tasks often reveal low correlations and poor reliability metrics, likely stemming from the calculation of difference scores in the presence of speed–accuracy tradeoff effects. Finally, most studies are coarse-grained, comparing only a handful of age groups, leaving adolescence undersampled and adulthood nearly entirely unsampled.

The current cross-sectional study therefore aims to compare finely measured age-related changes in performance on three congruency tasks commonly used to assess developmental and individual differences in cognitive control: the Simon task (Simon, 1969), the Stroop task (Stroop, 1935), and the Eriksen flanker task (Eriksen & Eriksen, 1974). We focus on three central questions concerning how cognitive control, as assessed by the size of the congruency effects observed in the tasks, changes across the lifespan: (a) *When is peak performance observed in each task?* (b) *Do the tasks present similar developmental trajectories?* And, (c) *Is performance correlated across congruency tasks?* Addressing these questions is particularly important considering recent work that has used large sample sizes and small age bins to demonstrate considerable heterogeneity in when other fundamental cognitive abilities like intelligence and memory peak (Hartshorne & Germine, 2015).

We therefore collected two massive datasets using online Citizen Science samples (Hartshorne, 2020; Reinecke & Gajos, 2015). Unlike studies that recruit participants with offers of cash or course credit, Citizen Science relies on intrinsically motivating tasks. Critically, data quality matches or exceeds what can be obtained in the lab perhaps precisely because subjects are more likely to be engaged (Germine et al., 2012; Meyerson & Tryon, 2003; Ye et al., 2017).

Method

Participants

Study 1

A total of 9,644 English speakers completed the Stroop and/or Simon tasks: 9,576 participants completed the Stroop task (M_{age} in years = 31, SE = 16; 3,119 male, 6,457 female), 9,585 participants completed the Simon task (M_{age} in years = 31, SE = 16; 3,114 male, 6,471 female), and 9,517 participants completed both. Additional participants were excluded for being repeat participants (N = 377), reporting not-corrected-to-normal vision (N = 469), reporting dyslexia or a neurological disorder (N = 1,073), or software error (Stroop = 163, Simon = 133). Finally, we excluded 81 participants who claimed to be less than 10 or more than 80 years old. This cutoff was determined by finding the oldest age for which we had at least 10 participants and the youngest age over 5 for which we had at least 10 participants.

Study 2

A total of 13,448 English speakers completed the experiment. Additional participants were excluded for being repeat participants (N = 1,915) or software error (N = 1,007; no info was collected on vision or neurological disorders). Finally, we excluded 84 participants who claimed to be less than 10 or more than 79 years old. This cutoff was determined by finding the oldest age for which we had at least 10 participants and then the youngest age over 5 for which we had at least 10 participants.

The research ethics review boards at Massachusetts Institute of Technology (Study 1) and Harvard University (Study 2) approved the protocols. See Figures S1–S4 and the surrounding text in the online supplemental materials for more information regarding the demographic information collected in each study, including distributions of participants by age and gender.

Procedure

Study 1

Participants completed two-alternative forced-choice (2AFC) versions of the Simon task and the Stroop task in a randomized order. To ensure that participants would be willing to complete both tasks, the total number of experimental trials collected in each task was set to 28, which is consistent with the number of trials collected in other validated measures of attention and control such as the NIH Toolbox: Cognition Battery (Zelazo et al., 2013). Each task consisted of 14 congruent trials and 14 incongruent trials, presented in a randomly intermixed order that was held constant across tasks and individuals. On each trial in the Simon task, participants were presented with a left-facing arrow or a right-facing arrow at either the left or right side of their display. Participants were instructed to press the "w" key for arrows pointing to the left and the "o" key for arrows pointing to the right, regardless of which side of the screen the arrow was presented on. On each trial in the Stroop task, participants were presented with the word "orange" or "white" in either orange or white text presented against a gray background. Participants were instructed to press the "w" key if the word appeared in white text and the "o" key if the word appeared in orange text. For both tasks, the intertrial interval (ITI) was 200 ms, and no feedback was given. Participants had as much time to respond as they wished. Each task began with a four-item practice block. If participants made any errors, they were informed of this, and the instructions and practice were repeated.

Study 2

Participants completed a 2AFC version of the Eriksen flanker task in which a stimulus array consisting of five arrows appeared on each trial. On congruent trials, each of the arrows cued the same response (e.g., ←←←←←). On incongruent trials, the centrally presented target arrow cued a different response than the surrounding distractors (e.g., →→←→→). The distractors were presented before the target for 100 ms. The distractors and target were then presented together for 50 ms, followed by a blank screen for 70 ms. A fixation cross then remained on the screen until a response was given or 3,000 ms elapsed. The ITI was set at 1,000 ms. Participants responded by pressing the "x" key (on keyboards) or a left-side software button (on touchscreens) for left-facing target arrows and the "m" key (on keyboards) or a right-facing software button (on touchscreens) for right-facing arrows. The task consisted of 96 trials, with congruent/incongruent and left/right responses fully crossed. Response types were in the same order for all participants. Trials in which no response was provided within the 3,000 ms limit were marked as inaccurate (0.28% of all trials; 2.2% of incorrect trials). The task began with two four-item blocks of practice: the first at a slower speed and the second at the true speed. Feedback was given after each response in the practice, and incorrect responses prompted that item to repeat.

Data Processing

The first trial of each block was excluded from the analysis (one trial each for Stroop and Simon; three for flanker). To control for post-error performance adjustments (e.g., Danielmeier & Ullsperger, 2011), all inaccurate trials and trials following an inaccurate trial were excluded from the analysis of response times (RTs). This resulted in the exclusion of 13.58% of Stroop trials, 8.10% of Simon trials, and 25.82% of flanker trials. Additionally, all responses that were faster than 100 ms or slower than 99.5% of the responses provided within each task were excluded from analysis, resulting in the further exclusion of 0.5% of Stroop trials, 0.4% of Simon trials, and 0.4% of flanker trials (the latter number includes trials that timed out at 3,000 ms). Investigation revealed this was sufficient to eliminate very extreme values while minimizing differential effects on subjects of different ages. To improve linearity and minimize the effects of processing speed differences (Veríssimo et al., 2022), RTs were transformed with the natural logarithm. To minimize the contribution of feature-integration effects (Braem et al., 2019; Erb & Marcovitch, 2018; Hommel, 2004; Nieuwenhuis et al., 2006), data analysis was restricted to response alternation trials. This resulted in the inclusion of 18 trials from the Simon task, 18 trials from the Stroop task, and 43 trials from the flanker task after the first trial of each block was excluded. After exclusions, there were on average 16.26 RTs per subject for Simon (95% between 14 and 17), 15.38 for Stroop (95% between 11 and 17), and 33.71 in flanker (95% between 16 and 42). Results from analyses including response repetition trials are included in the online supplemental materials, as are analyses evaluating how the congruency effects changed over the course of each task.

Transparency and Openness

Data and analysis code are available at https://osf.io/mv8zu/ (Hartshorne et al., 2023). This study was not preregistered.

Results

Internal Reliability

To evaluate the internal reliability of our congruency effect measures, we estimated Cronbach's α for the average split-half correlation, averaged over 100 random splits, using the Spearman–Brown correction. Reliability was strong for accuracy in the Simon task (0.90, CI: [0.89, 0.92]), marginally adequate in the flanker task (0.71, [0.66, 0.74]), and comparatively low in the Stroop task (0.63, [0.58, 0.68]). Log-transformed RTs revealed low reliability for the flanker task (0.36, [0.20, 0.44]) and essentially no reliability for the Simon task (−0.10, [−0.35, 0.10]) or the Stroop task (0.05, [−0.18, 0.20]). These results echo concerns about the reliability of standard congruency tasks (e.g., Draheim et al., 2019).

Peak Performance and Developmental Trajectories

To evaluate age-related changes in the congruency effect for each task, we performed separate Bayesian thin plate spline regressions for each task using the brms package (Bürkner, 2017; Carpenter et al., 2017). The resulting curves are shown in Figure 1 (accuracy) and Figure 2 (RTs). Note that the RT results should be interpreted with caution given the low reliability metrics presented above.

To quantify how the curves change with age, we calculated the local slope at every 1/100th of a year. This was repeated for each sampled curve, allowing us to estimate uncertainty. Intervals where the slope's 95% credible interval excluded 0 are intervals of significant developmental change. Note that this method inherently corrects for multiple comparisons (Simpson, 2016).

Simon Task

The accuracy congruency effect significantly decreased from the youngest age group until approximately 34 years of age, with a hiatus during most of the 20 s, with no further significant age-related changes observed. Response times revealed no significant age-related changes in the size of the congruency effect.

Stroop Task

The accuracy congruency effect significantly decreased from the youngest age group until approximately 26 years of age. Although the accuracy congruency effect began to increase again in later life, this numerical trend did not reach statistical significance. Response times showed the opposite effect early in the lifespan, with the congruency effect increasing significantly from 14 to 26 years of age. Although numerically there appears to be a decrease in the RT congruency effect starting in middle age, this trend did not reach significance.

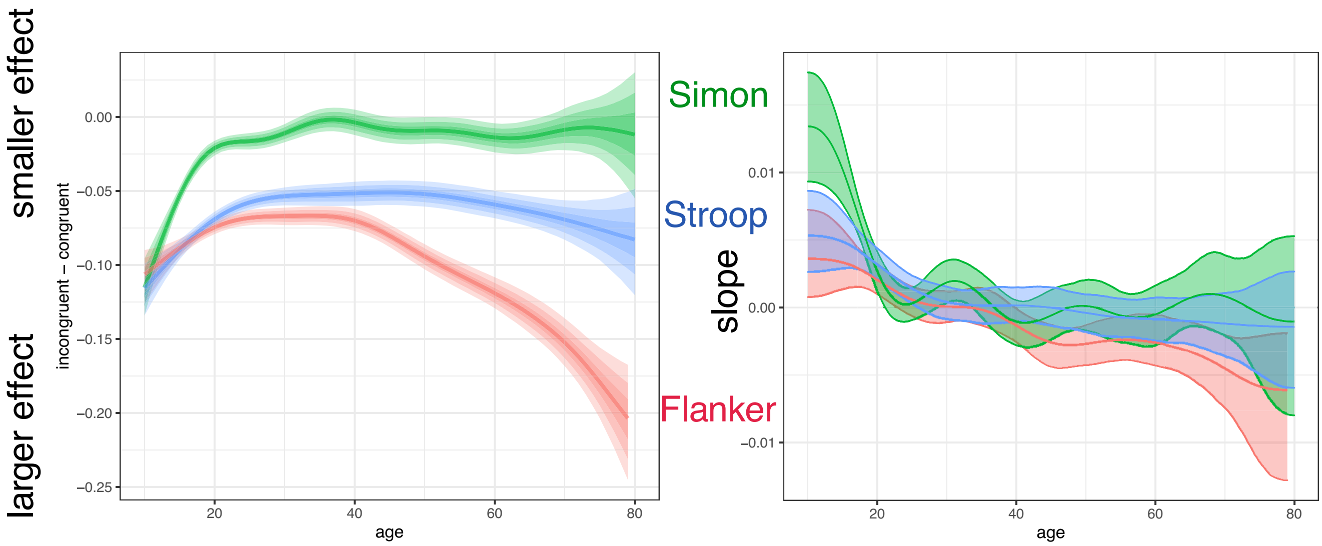
Flanker Task

The accuracy congruency effect significantly decreased from the youngest age group until approximately 23 years of age and then started to increase at approximately 40 years of age through the oldest age group tested. The RT congruency effect was initially stable and then increased from approximately 37–72 years of age.

Correlations Between Simon and Stroop Effects

The correlation between congruency effects for the Simon and Stroop tasks was near zero for both dependent measures (accuracy: $r = .03$ [0.01, 0.05], $p = .005$; RT: $r = .01$ [−0.01, 0.03], $p = .24$), despite the two tasks having identical trial structure and closely matched procedures. As illustrated in Figure 3, the correlations were largely unchanged across the age range investigated.

Figure 1
Accuracy Congruency Effects Across Tasks



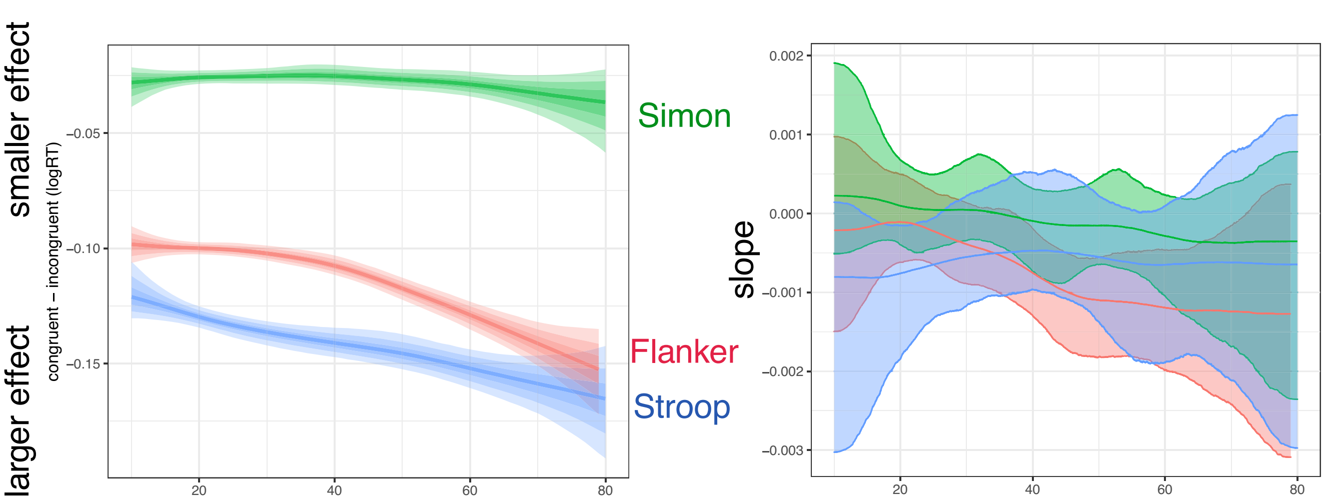
Note. Left: accuracy congruency effects observed in the Simon task (green), Stroop task (blue), and Eriksen flanker task (red) as a function of age in years. Shading denotes 50%, 80%, and 95% confidence intervals. Right: slopes, with 95% confidence intervals. See the online article for the color version of this figure.

Discussion

Cognitive control is commonly proposed to follow an inverted U-shaped developmental trajectory. However, a detailed understanding of cognitive control's development has proven elusive due to a range of factors, including issues with reliability (Draheim et al., 2019), small correlations among congruency effects (Rey-Mermet et al., 2018), and the prevalence of confounds (Braem et al., 2019). The current study focused on yet another factor that is likely to contribute to inconsistencies in the literature: the use of small sample sizes, with large age bins, and categorical designs that compare a limited number of age groups. To address this limitation, we conducted two massive online cross-sectional studies to track the development of cognitive control in the Simon, Stroop, and Eriksen flanker tasks.

In contrast to the notion that these tasks exhibit similar developmental trajectories, we observed pronounced differences across the

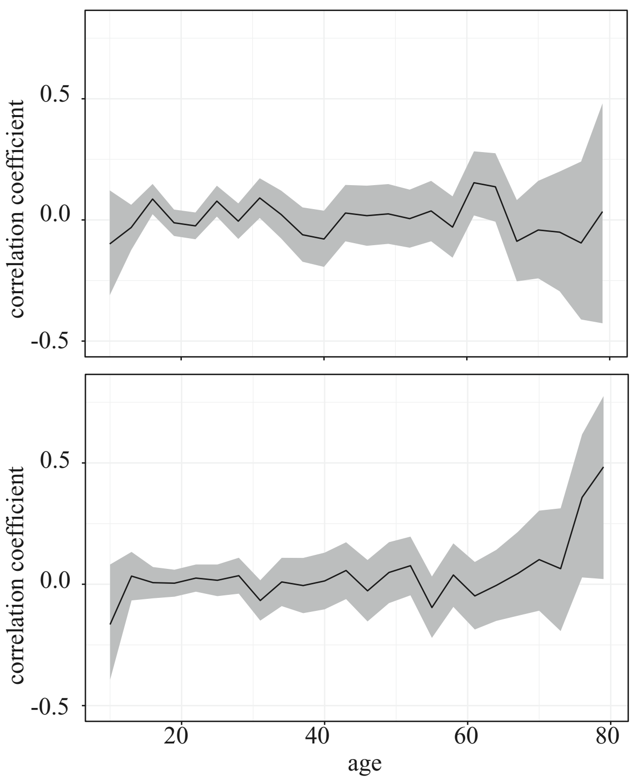
Figure 2
Response Time Congruency Effects Across Tasks



Note. Left: response time congruency effects observed in Simon task (green), Stroop task (blue), and Eriksen flanker task (red) as a function of age in years. Shading denotes 50%, 80%, and 95% confidence intervals. Right: slopes, with 95% confidence intervals. See the online article for the color version of this figure.

Figure 3

Correlations Between the Simon and Stroop Congruency Effects Observed in Accuracy (Top) and Response Time (Bottom) as a Function of Age



Note. Age groups were presented in 3-year bins. Shading denotes 95% confidence intervals.

tasks. For instance, peak performance occurred at approximately 23 years of age in the flanker task and at approximately 34 years of age in the Simon task. Although the Stroop task revealed age-related improvements in accuracy until approximately 26 years of age, these improvements coincided with age-related decreases in RT performance (i.e., larger congruency effects) between 14 and 26. Thus, the tasks exhibited notably different trajectories between childhood and early adulthood.

The tasks also revealed different effects of aging across late adulthood. Neither the Simon task nor the Stroop task revealed significant age-related declines in performance between early and late adulthood, with RTs in the Stroop task showing a nonsignificant trend of age-related improvements in older adulthood. The lack of an inverted U-shaped developmental trajectory in Stroop performance is consistent with a recent massive online study by LaPlume et al. (2022) that failed to observe significant age-related declines in interference processing in a number-naming version of the Stroop task. Only the flanker task showed strong evidence of an inverted U-shaped trajectory in the current study, with accuracy revealing age-related declines in performance starting at approximately 40 years of age and continuing until the oldest age group tested. This finding is generally consistent with the trajectory observed by Waszak et al. (2010) but inconsistent with the results of Veríssimo et al. (2022) from the Attentional Network Test (Fan et al., 2002).

Two important limitations of the current study should be acknowledged. First, both experiments were cross-sectional and, consequently, were unable to provide information about how performance changed over time within individuals. As such, it is possible that the results of the current study reflect cohort effects, particularly in relation to technology use (Fozard & Wahl, 2012). Second, the three congruency tasks used in the current study were not equated for difficulty. The flanker task was particularly difficult because the distractor arrows were presented before the target arrow and the target arrow was presented for a limited time. Given that the flanker task was the only task to reveal a robust inverted U-shaped developmental trajectory, it is possible that significant age-related declines in Simon and Stroop performance would be observed in late adulthood if the tasks were modified to be more difficult. Such a demonstration would further underscore the importance of efforts to develop adaptive, threshold-based congruency tasks (Draheim et al., 2021). Regardless, the present results call into question the notion that standard versions of the task reveal inverted U-shaped developmental trajectories, particularly in light of the large sample size and small age bins collected in the current study. Similarly, the tasks showed very different within-task learning effects, indicating that results will depend differentially on task length (Figure S5 and surrounding text in the online supplemental materials).

Our correlational analyses revealed a small but significant link between the accuracy congruency effects observed in the Simon and Stroop tasks. Recent diffusion modeling by Hedge et al. (2022) indicates that correlations between congruency effects are only weakly informative regarding the presence or absence of shared control processes, especially when other factors such as response caution and processing speed are not considered. Consequently, the correlations observed in the current study do not necessarily indicate that the tasks are of limited utility for targeting common cognitive control processes. However, taken together with the divergent developmental trajectories observed across the tasks, our correlational results further underscore the need for caution when attempting to form general conclusions about the development of cognitive control by synthesizing results across different congruency tasks.

Conclusion

Recent research has called into question the extent to which commonly used versions of the Simon, Stroop, and Eriksen flanker task can be used to measure the same cognitive abilities in the same proportion (Draheim et al., 2019; Hedge et al., 2022; Rey-Mermet et al., 2018). We explored this question from a developmental perspective, reasoning that the tasks should show similar patterns of age-related change if they tap into shared cognitive abilities. Our results revealed markedly different developmental trajectories, with only the flanker task conforming to an inverted U-shape. These findings caution against using standard congruency tasks to draw general conclusions about the development of cognitive control and underscore the importance of developing more psychometrically rigorous measures (Draheim et al., 2019) and approaches that allow for the time-course of response conflict to be better characterized across tasks (Hardwick et al., 2019). Our findings also highlight the value of using large sample sizes and small age bins to test for heterogeneity in the development of cognitive abilities (Hartshorne, 2020; Hartshorne & Germine, 2015).

Constraints on Generality

The participant population was restricted to native English speakers in Experiment 1 and any English speakers in Experiment 2 (roughly ½ billion and 2 billion, respectively; Crystal, 2008), but was otherwise demographically diverse.

References

Altikriti, S. (2021). Toward integrated processual theories of crime: Assessing the developmental effects of executive function, self-control, and decision-making on offending. *Criminal Justice and Behavior*, *48*(2), 215–233. https://doi.org/10.1177/0093854820942280

Blair, C. (2016). Developmental science and executive function. *Current Directions in Psychological Science*, *25*(1), 3–7. https://doi.org/10.1177/0963721415622634

Braem, S., Bugg, J. M., Schmidt, J. R., Crump, M. J. C., Weissman, D. H., Notebaert, W., & Egner, T. (2019). Measuring adaptive control in conflict tasks. *Trends in Cognitive Sciences*, *23*(9), 769–783. https://doi.org/10.1016/j.tics.2019.07.002

Bugg, J. M., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2007). Age differences in Stroop interference: Contributions of general slowing and task-specific deficits. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *14*(2), 155–167. https://doi.org/10.1080/138255891007065

Bürkner, P. C. (2017). *Advanced Bayesian multilevel modeling with the R package brms*. arXiv preprint arXiv:1705.11123. https://doi.org/10.48550/arXiv.1705.11123

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Christ, S. E., White, D. A., Mandernach, T., & Keys, B. A. (2001). Inhibitory control across the life span. *Developmental Neuropsychology*, *20*(3), 653–669. https://doi.org/10.1207/S15326942DN2003_7

Cohn, N. B., Dustman, R. E., & Bradford, D. C. (1984). Age-related decrements in Stroop color test performance. *Journal of Clinical Psychology*, *40*(5), 1244–1250. https://doi.org/10.1002/1097-4679(198409)40:5<1244::AID-JCLP2270400521>3.0.CO;2-D

Crystal, D. (2008). Two thousand million? *English Today*, *24*(1), 3–6. https://doi.org/10.1017/S0266078408000023

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, *2*, Article 233. https://doi.org/10.3389/fpsyg.2011.00233

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*(1), 45–75. https://doi.org/10.1016/0273-2297(92)90003-K

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), 508–535. https://doi.org/10.1037/bul0000192

Draheim, C., Tsukahara, J. S., Martin, J. D., Mashburn, C. A., & Engle, R. W. (2021). A toolbox approach to improving the measurement of attention control. *Journal of Experimental Psychology: General*, *150*(2), 242–275. https://doi.org/10.1037/xge0000783

Erb, C. D., & Marcovitch, S. (2018). Deconstructing the Gratton effect: Targeting dissociable trial sequence effects in children, pre-adolescents, and adults. *Cognition*, *179*, 150–162. https://doi.org/10.1016/j.cognition.2018.06.007

Erb, C. D., & Marcovitch, S. (2019). Tracking the within-trial, cross-trial, and developmental dynamics of cognitive control: Evidence from the Simon task. *Child Development*, *90*(6), e831–e848. https://doi.org/10.1111/cdev.13111

Erb, C. D., Touron, D. R., & Marcovitch, S. (2020). Tracking the dynamics of global and competitive inhibition in early and late adulthood: Evidence from the flanker task. *Psychology and Aging*, *35*(5), 729–743. https://doi.org/10.1037/pag0000435

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, *16*(1), 143–149. https://doi.org/10.3758/BF03203267

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. https://doi.org/10.1162/089892902317361886

Fozard, J. L., & Wahl, H.-W. (2012). Age and cohort effects in gerontechnology: A reconsideration. *Gerontechnology (Valkenswaard)*, *11*(1), 10–21. https://doi.org/10.4017/gt.2012.11.01.003.00

Gathercole, V. C. M., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Viñas Guasch, N., Roberts, E. J., Hughes, E. K., & Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology*, *5*, Article 11. https://doi.org/10.3389/fpsyg.2014.00011

Germine, L., Nakayama, K., Duchaine, B. C., Chabris, C. F., Chatterjee, G., & Wilmer, J. B. (2012). Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. *Psychonomic Bulletin and Review*, *19*(5), 847–857. https://doi.org/10.3758/s13423-012-0296-9

Hardwick, R. M., Forrence, A. D., Krakauer, J. W., & Haith, A. M. (2019). Time-dependent competition between goal-directed and habitual response preparation. *Nature Human Behaviour*, *3*(12), 1252–1262. https://doi.org/10.1038/s41562-019-0725-0

Hartshorne, J. K. (2020). How massive online experiments (MOEs) can illuminate critical and sensitive periods in development. *Current Opinion in Behavioral Sciences*, *36*, 135–143. https://doi.org/10.1016/j.cobeha.2020.09.005

Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, *26*(4), 433–443. https://doi.org/10.1177/0956797614567339

Hartshorne, J. K., Germine, L. T., & Erb, C. (2023). Erb, Germine, & Hartshorne. *Data*. Advance online publication. https://doi.org/10.17605/OSF.IO/MV8ZU

Hedge, C., Powell, G., Bompas, A., & Sumner, P. (2022). Strategy and processing speed eclipse individual differences in control ability in conflict tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(10), 1448–1469. https://doi.org/10.1037/xlm0001028

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*(11), 494–500. https://doi.org/10.1016/j.tics.2004.08.007

Insel, K., Morrow, D., Brewer, B., & Figueredo, A. (2006). Executive function, working memory, and medication adherence among older adults. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *61*(2), 102–107. https://doi.org/10.1093/geronb/61.2.P102

LaPlume, A. A., Anderson, N. D., McKetton, L., Levine, B., & Troyer, A. K. (2022). When I'm 64: Age-related variability in over 40,000 online cognitive test takers. *The Journals of Gerontology: Series B*, *77*(1), 104–117. https://doi.org/10.1093/geronb/gbab143

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*(5), 1357–1372. https://doi.org/10.1111/j.1467-8624.2004.00745.x

Meyerson, P., & Tryon, W. W. (2003). Validating internet research: A test of the psychometric equivalence of internet and in-person samples. *Behavior*

Research Methods, Instruments, and Computers, *35*(4), 614–620. https://doi.org/10.3758/BF03195541

Nieuwenhuis, S., Stins, J. F., Posthuma, D., Polderman, T. J., Boomsma, D. I., & de Geus, E. J. (2006). Accounting for sequential trial effects in the flanker task: Conflict adaptation or associative priming? *Memory and Cognition*, *34*(6), 1260–1272. https://doi.org/10.3758/BF03193270

Reinecke, K., & Gajos, K. Z. (2015, March 14–18). *LabintheWild: Conducting large-scale online experiments with uncompensated samples*. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing, Vancouver, BC, Canada (pp. 1364–1378). ACM Press.

Rey-Mermet, A., & Gade, M. (2018). Inhibition in aging: What is preserved? What declines? A meta-analysis. *Psychonomic Bulletin and Review*, *25*(5), 1695–1716. https://doi.org/10.3758/s13423-017-1384-7

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(4), 501–526. https://doi.org/10.1037/xlm0000450

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, *81*(1), 174–176. https://doi.org/10.1037/h0027448

Simpson, G. (2016, December 15). Simultaneous intervals for smooths revisited: Correcting a silly mistake. *From the Bottom of the Heap*. Retrieved from https://fromthebottomoftheheap.net/2016/12/15/simultaneous-interval-revisited/

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

van der Lubbe, R. H., & Verleger, R. (2002). Aging and the Simon task. *Psychophysiology*, *39*(1), 100–110. https://doi.org/10.1111/1469-8986.3910100

Verhaeghen, P. (2011). Aging and executive control: Reports of a demise greatly exaggerated. *Current Directions in Psychological Science*, *20*(3), 174–180. https://doi.org/10.1177/0963721411408772

Verhaeghen, P., & De Meersman, L. (1998). Aging and the Stroop effect: A meta-analysis. *Psychology and Aging*, *13*(1), 120–126. https://doi.org/10.1037/0882-7974.13.1.120

Veríssimo, J., Verhaeghen, P., Goldman, N., Weinstein, M., & Ullman, M. T. (2022). Evidence that ageing yields improvements as well as declines across attention and executive functions. *Nature Human Behaviour*, *6*(1), 97–110. https://doi.org/10.1038/s41562-021-01169-7

Vu, K. P. L., & Proctor, R. W. (2008). Age differences in response selection for pure and mixed stimulus-response mappings and tasks. *Acta Psychologica*, *129*(1), 49–60. https://doi.org/10.1016/j.actpsy.2008.04.006

Waszak, F., Li, S. C., & Hommel, B. (2010). The development of attentional networks: Cross-sectional findings from a life span sample. *Developmental Psychology*, *46*(2), 337–349. https://doi.org/10.1037/a0018541

Ye, T., Reinecke, K., & Robert, L. P., Jr. (2017, February 25–March 1). *Personalized feedback versus money: The effect on reliability of subjective data in online experimental platforms*. In Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, Portland, OR, United States (pp. 343–346). ACM Press.

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., & Weintraub, S. (2013). II. NIH Toolbox Cognition Battery (CB): Measuring executive function and attention. *Monographs of the Society for Research in Child Development*, *78*(4), 16–33. https://doi.org/10.1111/mono.12032

Zelazo, P. D., Craik, F. I., & Booth, L. (2004). Executive function across the life span. *Acta Psychologica*, *115*(2–3), 167–183. https://doi.org/10.1016/j.actpsy.2003.12.005

Received June 13, 2022
Revision received March 16, 2023
Accepted March 23, 2023 ■