

Behavioral and Neural Signatures of Working Memory in Childhood

Monica D. Rosenberg,^{1,2} Steven A. Martinez,² Kristina M. Rapuano,² May I. Conley,² Alexandra O. Cohen,³ M. Daniela Cornejo,^{4,5} Donald J. Hagler Jr.,⁴ Wesley J. Meredith,¹ Kevin M. Anderson,² Tor D. Wager,^{6,7} Eric Feczko,^{8,9} Eric Earl,⁸ Damien A. Fair,^{8,10,11} Deanna M. Barch,^{12,13,14} Richard Watts,² and B.J. Casey²

¹Department of Psychology, University of Chicago, Chicago, IL 60637, ²Department of Psychology, Yale University, New Haven, CT 06511, ³Department of Psychology and Neural Science, New York University, New York, NY 10003, ⁴Department of Radiology, University of California, San Diego, San Diego, CA 92122, ⁵Institute of Physics, Pontificia Universidad Católica de Chile, Santiago 8331150, Chile, ⁶Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO 80302, ⁷Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, ⁸Department of Behavioral Neuroscience, Oregon Health & Science University, Portland, OR 97239, ⁹Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, OR 97239, ¹⁰Department of Psychiatry, Oregon Health & Science University, Portland, OR 97239, ¹¹Advanced Imaging Research Center, Oregon Health & Science University, Portland, OR 97239, ¹²Department of Psychological and Brain Sciences, Washington University in St. Louis, St. Louis, MO 63130, ¹³Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, and ¹⁴Department of Radiology, Washington University School of Medicine, St. Louis, MO 63110

Working memory function changes across development and varies across individuals. The patterns of behavior and brain function that track individual differences in working memory during human development, however, are not well understood. Here, we establish associations between working memory, other cognitive abilities, and functional MRI (fMRI) activation in data from over 11,500 9- to 10-year-old children (both sexes) enrolled in the Adolescent Brain Cognitive Development (ABCD) Study, an ongoing longitudinal study in the United States. Behavioral analyses reveal robust relationships between working memory, short-term memory, language skills, and fluid intelligence. Analyses relating out-of-scanner working memory performance to memory-related fMRI activation in an emotional *n*-back task demonstrate that frontoparietal activity during a working memory challenge indexes working memory performance. This relationship is domain specific, such that fMRI activation related to emotion processing during the emotional *n*-back task, inhibitory control during a stop-signal task (SST), and reward processing during a monetary incentive delay (MID) task does not track memory abilities. Together, these results inform our understanding of individual differences in working memory in childhood and lay the groundwork for characterizing the ways in which they change across adolescence.

Key words: development; fMRI; frontoparietal; *n*-back; working memory

Significance Statement

Working memory is a foundational cognitive ability that changes over time and varies across individuals. Here, we analyze data from over 11,500 9- to 10-year-olds to establish relationships between working memory, other cognitive abilities, and frontoparietal brain activity during a working memory challenge, but not during other cognitive challenges. Our results lay the groundwork for assessing longitudinal changes in working memory and predicting later academic and other real-world outcomes.

Introduction

Working memory, a collection of cognitive processes responsible for storing and manipulating information, is a foundational ability that varies widely across individuals. Individual differences in working memory, which appear to be stable over time (Alp,

1994; Ross et al., 2008; Johnson et al., 2013; Tulskey et al., 2014; Xu et al., 2018), have pronounced real-world significance. Although the direction of causality is unclear, measures of verbal

Received Nov. 20, 2019; revised Apr. 17, 2020; accepted Apr. 21, 2020.

Author contributions: M.D.R., R.W., and B.J.C. designed research; S.A.M., M.I.C., M.D.C., D.J.H., and R.W. performed research; M.D.R., K.M.R., A.O.C., M.D.C., D.J.H., K.M.A., T.D.W., E.F., E.E., D.A.F., D.M.B., R.W., and B.J.C. contributed unpublished reagents/analytic tools; M.D.R., S.A.M., W.J.M., and R.W. analyzed data; M.D.R. wrote the paper with contributions from B.J.C. and all authors.

This work was supported in part by National Institutes of Health (NIH) Grants U01 DA041174 (to B.J.C.), U01 DA041120 (to D.M.B.), U01 DA041148 (to D.A.F.), and U24 DA041123 (to B.J.C., M.D.C., D.J.H.) and by the National Science Foundation Grant 1714321 (to A.O.C.). The ABCD Study is supported by the NIH and additional federal partners Grants U01DA041022, U01DA041028, U01DA041048, U01DA041089, U01DA041106, U01DA041117, U01DA041120, U01DA041134, U01DA041148, U01DA041156, U01DA041174, U24DA041123, and U24DA041147. A full list of supporters is available at <https://abcdstudy.org/federal-partners/>. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

and visuospatial working memory explain ~20–30% of the variance in fluid intelligence in children (Engel de Abreu et al., 2010), and visuospatial working memory performance explains >40% of this variance in adults (Fukuda et al., 2010). Furthermore, working memory function, which is related to executive and visuospatial attention (Kane and Engle, 2002; Huang et al., 2012), short-term memory (Alloway et al., 2006), and inhibitory control (Davidson et al., 2006), predicts consequential outcomes in development, including reading and math skills (Bayliss et al., 2003; De Smedt et al., 2009; Alloway and Alloway, 2010; Nouwens et al., 2017). Despite the theoretical and practical importance of characterizing associations between working memory and other mental processes, much remains to be learned about the nature of these relationships during development.

Working memory not only varies across individuals, but also changes across the lifespan. Working memory emerges in infancy and develops rapidly over the first year of life (Diamond and Goldman-Rakic, 1989; Ross-Sheehy et al., 2003; Reynolds and Romano, 2016; Buss et al., 2018). This ability continues to improve during childhood, plateaus in mid-to-late adolescence (Gathercole et al., 2004; Luciana et al., 2005; Conklin et al., 2007; Ullman et al., 2014; Isbell et al., 2015), and declines after age 40–50, albeit less steeply than it changed during early development (Alloway and Alloway, 2013; Nyberg et al., 2014; Eriksson et al., 2015; Swanson, 2017). Developmental gains in working memory follow improvements in attention shifting, attentional maintenance, and distractor suppression (Reynolds and Romano, 2016), whereas changes during later childhood accompany increases in domain-general processing speed and memory capacity (Fry and Hale, 1996, 2000; Pailian et al., 2016) with developmental asymptotic performance by adolescence (Casey et al., 2000, 2005; Klingberg et al., 2002; Steinberg et al., 2009). Decrements in older adulthood relate to declines in processing speed, selective attention, and distractor suppression (Salthouse and Babcock, 1991; Gazzaley et al., 2005; McNab et al., 2015).

Converging neuroimaging evidence suggests that variation in frontoparietal brain systems, which are involved in processes including attention and cognitive control (Woolgar et al., 2011; Scolaro et al., 2015; Assem et al., 2020), accounts for both developmental change in verbal and visuospatial working memory and individual differences in these processes in adulthood. Early work demonstrated that the same aspects of middle and inferior frontal cortex that support working memory performance in adults also support performance in children (Casey et al., 1995). This evidence led to theorizing that the protracted fine-tuning of prefrontal circuitry contributes to working memory improvements during childhood and adolescence (Casey et al., 2000, 2005). Longitudinal studies support this prediction, with evidence that maturation in prefrontal and parietal volume and structural connectivity accompany working memory development (Tamnes et al., 2013; Klingberg and Darki, 2015). Cross-

sectional work suggests that increases in frontoparietal activation during working memory tasks are associated with age-related improvements in performance (Klingberg et al., 2002; Kwon et al., 2002; Crone et al., 2006; Satterthwaite et al., 2013). In the developed brain, individual differences in frontoparietal areas' microstructure, function, and structural and functional connectivity track individual differences in working memory (Osaka et al., 2003; Palva et al., 2010; Burzynska et al., 2011; Takeuchi et al., 2011; Ekman et al., 2016). A subset of developmental studies show similar associations between in-scanner working memory performance (a state-like measure of memory function) and frontoparietal activity during working memory tasks when controlling for age (Crone et al., 2006; Satterthwaite et al., 2013). One study comparing adolescents with higher and lower family incomes found that the higher-income group showed greater frontoparietal activity as a function of load in a working memory (*n*-back) functional MRI (fMRI) task and higher out-of-scanner working memory capacity (Finn et al., 2017a). However, it is not yet known whether frontoparietal network function during working memory challenges specifically, or during cognitive task challenges more generally, predicts individual differences in working memory during development.

Here, we examine behavioral and neural signatures of working memory in childhood. Using data from 11,537 9- to 10-year-olds participating in the Adolescent Brain Cognitive DevelopmentSM (ABCD) Study (Casey, 2018; Volkow et al., 2018), we first establish relationships between working memory and other cognitive and attentional abilities, including short-term memory, language and verbal skills, fluid intelligence, processing speed, attention, inhibitory control, and reward processing. Because the ABCD Study[®] will follow children longitudinally for 10 years, characterizing these associations in childhood not only informs the structure of cognition at a single time point but also facilitates understanding the ways in which this cognitive structure changes across adolescence. We next ask whether performance on an out-of-scanner working memory test is related to frontoparietal brain activity when measured (1) during a working memory challenge and (2) during task challenges unrelated to memory. Together, our results provide insight into individual differences in working memory in childhood, and underscore the importance of task fMRI as a “stress test” for cognition (Finn et al., 2017b) that can reveal task-specific and task-general neural signatures of a mental process or behavior.

Materials and Methods

The ABCD Study

Individual differences in working memory and other cognitive and attentional processes were assessed using data from 11,537 9- to 10-year-olds in the ABCD Study, an ongoing multisite longitudinal study of neurocognitive development (Luciana et al., 2018). Launched in September 2016, the ABCD Study aims to characterize cognitive and neural development with measures of neurocognition, physical and mental health, social and emotional function, and culture and environment. Exclusionary criteria include a diagnosis of schizophrenia, a moderate to severe autism spectrum disorder, an intellectual disability, or a substance use disorder at recruitment. Children with a persistent major neurologic disorder (e.g., cerebral palsy, a brain tumor, stroke, brain aneurysm, brain hemorrhage, subdural hematoma), multiple sclerosis, sickle cell disease, or certain seizure disorders (Lennox–Gastaut syndrome, Dravet syndrome, and Landau Kleffner syndrome) were also excluded.

ABCD Study data collection includes yearly behavioral assessments, interviews, questionnaires, and biosample collection as well as biennial MRI scans (Casey, 2018). Here, we analyze year-one (baseline) demographic and behavioral data collected across 22 sites when children were 9–10 years old and made available as part of curated data release 2.0.1

The ABCD data repository grows and changes over time. The ABCD data used in this report came from National Institute of Mental Health Data Archive DOI 10.15154/1504041. DOIs can be found at <https://nda.nih.gov/study.html?id=721>.

Data used in the preparation of this article were obtained from the ABCD Study (<https://abcdstudy.org>), held in the National Institute of Mental Health Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children age 9–10 and follow them over 10 years into early adulthood. A listing of participating sites and a complete listing of the study investigators can be found at <https://abcdstudy.org/principal-investigators>. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in analysis or writing of this report. This report reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

The authors declare no competing financial interests.

Correspondence should be addressed to Monica D. Rosenberg at mdrosenberg@uchicago.edu or B.J. Casey at bj.casey@yale.edu.

<https://doi.org/10.1523/JNEUROSCI.2841-19.2020>

Copyright © 2020 the authors

Table 1. Demographic, neurocognitive, and neuroimaging task performance measures

| | Description | Cognitive process(es) | Performance measure(s) | Data file | Variable name(s) | Percent missing | Percent outliers |
|------------------------|---|--|---|---------------------------------------|--|------------------------------------|------------------|
| Demographics | Age in months | | | abcddemo01 | <i>interview_age</i> | 0.02 | 0 |
| | Sex | | | abcddemo01 | <i>sex</i> | 0.07 | 0 |
| | Autism, epilepsy | | | abcd_screen01 | <i>scrm_asd, scrm_epls</i> | 0 | |
| | Family ID | | | acpspw03 | <i>rel_family_id</i> | 0 | 0 |
| | Site ID | | | abcd_lt01 | <i>site_id_l</i> | 0 | 0 |
| | Scanner ID | | | abcd_mri01 | <i>mri_info_deviceserialnumber</i> | 1.01 | |
| Neurocognitive battery | NIH toolbox cognition battery | Working memory | List Sorting Working Memory Test uss | abcd_tbss01 | <i>nihtbx_list_uncorrected</i> | 1.71 | 1.23 |
| | | | Language | Picture Vocabulary Test uss | abcd_tbss01 | <i>nihtbx_picvocab_uncorrected</i> | 1.32 |
| | | Cognitive control, attention | Flanker Test uss | abcd_tbss01 | <i>nihtbx_flanker_uncorrected</i> | 1.37 | 2.67 |
| | | | Flexible thinking | Dimensional Change Card Sort Test uss | abcd_tbss01 | <i>nihtbx_cardsort_uncorrected</i> | 1.35 |
| | Processing speed | Pattern Comparison Processing Speed Test uss | abcd_tbss01 | <i>nihtbx_pattern_uncorrected</i> | 1.51 | 0.81 | |
| | | Visuospatial sequencing, memory | Picture Sequence Memory Test uss | abcd_tbss01 | <i>nihtbx_picture_uncorrected</i> | 1.40 | 0.59 |
| | Reading | Oral Reading Recognition Test uss | abcd_tbss01 | <i>nihtbx_reading_uncorrected</i> | 1.42 | 2.75 | |
| | Matrix reasoning task | Fluid reasoning | WISC-V matrix reasoning total scaled score | abcd_ps01 | <i>pea_wiscv_tss</i> | 2.18 | 1.40 |
| | Rey auditory verbal learning test (RAVLT) | Learning, memory | Total correct on immediate and delayed recall trials | abcd_ps01 | <i>pea_ravlt_sd_trial_vi_tc, pea_ravlt_id_trial_vii_tc</i> | 1.87 2.31 | 1.95 1.68 |
| | Intertemporal Cash Choice Task | Delay of gratification | Choice of smaller-sooner or larger-later reward | cct01 | <i>cash_choice_task</i> | 1.83 | 0 |
| | Little Man Task | Mental rotation | Efficiency ratio (% accuracy ÷ mean correct-trial RT) | lmtp201 | <i>lmt_scr_efficiency</i> | 2.85 | 1.49 |
| fMRI tasks | Emotional <i>n</i> -back (EN-back) task | Memory, emotion regulation | % correct on 0-back and 2-back blocks | abcd_mri_nback02 | <i>tfmri_nb_all_beh_c0b_rate, tfmri_nb_all_beh_c2b_rate</i> | 20.08 20.08 | 3.19 3.19 |
| | Post-scan EN-back stimuli recognition memory test | Memory | Sensitivity (<i>d'</i>) | mribrec02 | Mean of <i>tfmri_rec_all_beh_posf_dpr, tfmri_rec_all_beh_neutf_dp, tfmri_rec_all_beh_negf_dp, tfmri_rec_all_beh_place_dp</i> | 24.24 | 1.41 |
| | Stop-signal task (SST) | Impulsivity | –1*stop-signal RT | abcd_sst02 | <i>tfmri_sst_all_beh_total_meanrt</i> | 18.97 | 2.51 |
| | Monetary incentive delay (MID) task | Reward processing | Mean earnings | abcd_mid02 | Mean of <i>tfmri_mid_r1_beh_t_earnings, tfmri_mid_r2_beh_t_earnings</i> | 17.26 | 1.95 |

Data were acquired from publicly available ABCD data release 2.0.1 (DOI 10.15154/1504041). Percent missing values represent the percentage of values missing in the full sample of 11,537 children meeting inclusion criteria, although note that recovery of missing data are ongoing. Percent outlier values represent the percentage of data values more than 2.5 SDs from the group mean. uss = uncorrected standard score.

(DOI 10.15154/1504041; $n = 11,537$; 48.4% female). Sample demographics including race, ethnicity, socioeconomic status, and symptoms of internalizing and externalizing disorders are available in Thompson et al. (2019).

Data from children diagnosed with autism spectrum disorder or epilepsy were excluded from the current analysis because moderate to severe forms of autism spectrum disorder and other seizure disorders were exclusionary for the ABCD Study ($n = 338$ of 11,875). Data from children with attention deficit hyperactivity disorder, depression, bipolar disorder, anxiety, and phobias ($n = 1,598$ of 11,537) were not excluded, as these diagnoses were assessed with a single screening question and we aimed to characterize working memory in a heterogeneous developmental population.

Behavioral data

To characterize associations between working memory and other cognitive abilities, we analyzed performance data from all 15 available neurocognitive (Luciana et al., 2018) and neuroimaging (Casey, 2018) tasks, described in detail below (Table 1). Working memory ability outside of the scanner was operationalized as performance on the NIH Toolbox List Sorting Working Memory Task. Performance measures were selected based on previous work including reports of ABCD baseline data (Casey, 2018; Luciana et al., 2018).

NIH Toolbox cognition battery

The NIH Toolbox cognition battery includes seven tasks measuring multiple aspects of cognition (Gershon et al., 2013; Table 1, column 3). Performance is measured using uncorrected standard scores, as age-corrected scores are currently undergoing revision by the NIH Toolbox (Luciana et al., 2018).

The Toolbox List Sorting Working Memory Test measures working memory by asking children to recall stimuli in different orders (Tulsky et al., 2014). Children were first shown pictures of two animals and were asked to repeat them back in order from smallest to largest. They were shown longer lists (with up to seven animals) if they answered correctly. Next, children were shown pictures of both animals and foods and were asked to repeat the animals in order of size and then the foods in order of size. Interleaved lists increased in length from two to seven if children responded accurately, and performance scores reflect the number of accurate responses. Importantly, this list sorting task was developed to assess working memory function in general rather than verbal or visual working memory in particular (Tulsky et al., 2014). Visuospatial working memory was not measured in the ABCD baseline sample and reflects a cognitive ability closely related to but dissociable from verbal working memory (Alloway et al., 2006; Swanson, 2017) that is traditionally tested with tasks such as delayed match-to-sample.

The Toolbox Picture Vocabulary Test measures language and verbal abilities (Gershon et al., 2014). Children hear a series of words and are

asked to choose which of four pictures most closely matches the meaning of the word.

The Toolbox Flanker Task, a flanker task (Eriksen and Eriksen, 1974) used to measure cognitive control and attention, was adapted from the Attention Network Task (Fan et al., 2002; Zelazo et al., 2014). On each trial, children see a row of five arrows. The outer four arrows (distractors, or flankers) all point to the left or right of the screen. The middle arrow (the target) points in the same direction as the flankers on congruent trials, and the opposite direction of the flankers on incongruent trials. Children are asked to indicate whether the center arrow points to the left or to the right. Performance scores are based on speed and accuracy.

The Toolbox Dimensional Change Card Sort Task measures cognitive flexibility (Zelazo, 2006). On each trial, children see two objects on a screen. They are asked to match a third item with one of the initial two based on either color or shape. Children first match items based on one dimension (e.g., color), then match items based on the other dimension (e.g., shape), and finally match based on both shape and color in pseudorandom order. Performance scores are based on speed and accuracy.

The Toolbox Pattern Comparison Processing Speed Test measures visual processing speed (Salthouse et al., 1991; Carlozzi et al., 2015). Children are shown two pictures and are asked to indicate whether they are the same or different. Scores are based on the number of correct responses within a time limit.

The Toolbox Picture Sequence Memory Test measures episodic memory and visuospatial sequencing (Bauer et al., 2013; Dikmen et al., 2014). Children are shown 15 pictures of activities or events and asked to reproduce the presentation order.

The Toolbox Oral Reading Recognition Task measures reading abilities by asking children to pronounce a series of written letters and words (Gershon et al., 2014).

Matrix reasoning

The matrix reasoning subtest of the Wechsler Intelligence Test for Children-V (WISC-V; Wechsler, 2014) measures fluid and spatial reasoning, perceptual organization, visual attention, and sequencing. On each trial, children are shown an array of visual stimuli, and are asked to select one of four stimuli that best completes the pattern. The task continues until a child makes three consecutive errors or completes 32 trials. Performance is measured by converting the number of total correct items to a standard score (Luciana et al., 2018).

Rey Auditory Verbal Learning Test

The Rey Auditory Verbal Learning Test (RAVLT) measures learning and memory. During the test, children hear a list of 15 unrelated words five times. Each time they hear the list, they are asked to recall as many words as possible. After these five learning trials, children hear a distractor list and are again asked to recall as many words as they can. Recall of the initial list is assessed immediately after the distractor list and again 30 min later (Van Den Burg and Kingma, 1999; Luciana et al., 2018). Here, we measure performance as the number of correctly recalled words on these immediate and delayed memory assessments (i.e., RAVLT trials vi and vii).

Intertemporal cash choice task

The Intertemporal Cash Choice Task (Wulfert et al., 2002) assesses children's delay of gratification, motivation, and impulsivity (Luciana et al., 2018). Children are asked, "Let's pretend a kind person wanted to give you some money. Would you rather have \$75 in 3 days or \$115 in 3 months?" Smaller-sooner reward choices were coded with a "1," larger-later reward choices were coded with a "2," and infrequent "don't know" responses were excluded from analysis. Positive correlations between cash choice and performance on another task indicate that children who performed better on the other task were more likely to choose the larger-later option, whereas negative correlations indicate that children who performed worse on the other task were more likely to choose the larger-later option.

Little Man task

The Little Man Task (Acker and Acker, 1982) measures aspects of visuospatial processing including mental rotation. During this task, children see a cartoon of a man holding a briefcase in his left or right hand appear on a computer screen. The man can be right side up or upside down and can appear facing the child or with his back turned. Children are asked to indicate whether the man's briefcase is in his left or right hand via button press. The task includes practice trials and 32 assessment trials. Performance is measured with efficiency (percent accuracy divided by mean correct-trial response time; Luciana et al., 2018).

Emotional n-back task

The in-scanner emotional *n*-back (EN-back) task engages processes related to memory and emotion regulation (Barch et al., 2013; Casey, 2018). During the task, children perform 0-back (low memory load) and 2-back (high memory load) task blocks with four types of stimuli: happy, fearful, and neutral face photographs (Tottenham et al., 2009; Conley et al., 2018) and place photographs. Data are collected during two ~5-min fMRI runs each with four 0-back and 2-back blocks each. Runs included 362 whole-brain volumes after discarded acquisitions. At the start of each 0-back block, children are shown a target stimulus and asked to press a button corresponding to "match" when they see an identical picture and a button corresponding to "no match" when they see a different picture. During 2-back blocks, children are asked to press match when they see a picture identical to the one they saw two trials back. Performance is quantified as percent accuracy on 0-back and 2-back blocks.

Recognition memory task

After scanning, memory for EN-back task stimuli is assessed with a recognition memory test (Barch et al., 2013; Casey, 2018). During this test, children are presented with 48 EN-back stimuli and 48 novel stimuli (i.e., 12 old and new happy, fearful, and neutral face photographs and 12 old and new places), and are asked to rate whether each picture is "old" or "new." Performance is assessed with sensitivity (d') averaged across stimulus types.

Stop-signal task

The in-scanner stop-signal task (SST; Logan, 1994) is designed to measure impulsivity and impulse control (Casey, 2018). SST data are collected during two ~6-min fMRI runs (437 volumes after discarded acquisitions) each with 180 trials each. On each trial, children see an arrow pointing to the left or to the right of the screen (the go signal). They are instructed to indicate the direction of the arrow with a button press as quickly and accurately as possible, except when an upright arrow (the stop signal) appears on the screen (16.67% of trials). The time between go and stop signal onset, the stop-signal delay, is staircased so that each child achieves ~50% accuracy on stop trials. Performance is measured with stop-signal reaction time (SSRT), or the mean stop-signal delay subtracted from the mean reaction time on correct go trials. For consistency with other behavioral measures, SSRTs were reverse scored (multiplied by -1) so that higher scores correspond to better performance.

Monetary incentive delay task

The in-scanner MID task (Knutson et al., 2000; Yau et al., 2012) measures aspects of reward processing, including anticipation and receipt of rewards and losses and motivation to earn rewards and avoid losses (Casey, 2018). Data are collected during two ~5.5-min, 50-trial fMRI runs (403 volumes per run after discarded acquisitions). Trials begin with a cue indicating whether the child can win \$0.20 or \$5, lose \$0.20 or \$5, or earn \$0. After 1500–4000 ms, a target appears for 150–500 ms. Target timing is staircased such that each child achieves ~60% accuracy. Children must respond during the target presentation to achieve the indicated trial outcome. Trials are followed by feedback indicating the outcome.

Response time and accuracy are not used as MID performance measures because target timing was individualized. Overall task performance is instead summarized as the average amount of money earned during both runs. This metric is correlated with mean accuracy ($r_s = 0.68$, $p < 0.001$) and correct-trial RT ($r = -0.29$, $p < 0.001$), but may capture additional variance in reward-related behavior. For example, two children who achieve 60% accuracy could earn different amounts of money

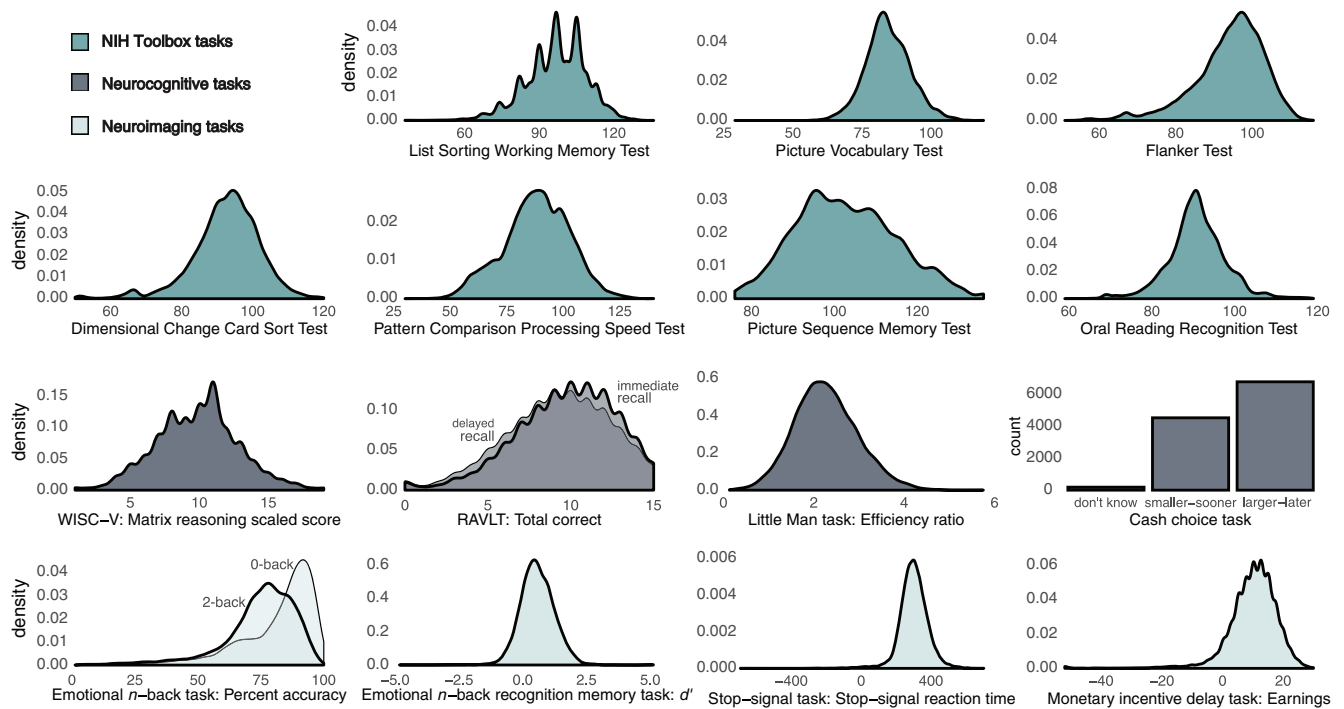


Figure 1. Kernel density estimates, or smoothed histograms, show performance in the full sample of 11,537 9- to 10-year-olds, including statistical outliers. NIH Toolbox performance is measured with uncorrected standard scores. Responses on the cash choice task, whether a child preferred to receive a smaller-sooner reward, a larger-later reward, or could not choose, are visualized with a histogram. Although “don’t know” responses on this task are included here, they were excluded from formal analysis.

if one preferentially responds to win large rewards and avoid large losses, perhaps due to greater reward motivation.

Relationships between behavioral measures

To characterize associations between working memory and other cognitive abilities, we first cross-correlated all 17 behavioral measures using data from all children meeting inclusion criteria ($n = 11,537$). Although normality was not evaluated with formal tests, which reject the null hypothesis for near-normal distributions in large samples (Ghasemi and Zahediasl, 2012), rank correlation was applied because visual inspection indicated that behavioral measures were not normally distributed (Fig. 1). To establish whether relationships were robust to potential confounds such as age, sex, missing data, outliers, and statistical dependence introduced by family structure, data collection method, and site, we subsequently replicated this analysis using:

(1) Data from one child per family, based on self-report ($n = 9,750$). For families with multiple children in the 11,537-participant cohort, the child whose randomly assigned NDAR Global Unique Identifier (GUID) came first in alphabetical order was included in this sample. (2) Data from all children meeting inclusion criteria using Spearman partial correlation to control for age and sex. (3) Data values within 2.5 SDs of the group mean (Table 1). (4) Data from children with no missing values (i.e., complete cases; $n = 7,504$). (For this analysis, “don’t know” responses on the cash choice task were counted as missing values.) (5) Data from children who completed the emotional n -back task, SST, and MID task during MRI data collection rather than on a laptop outside the scanner ($n = 9,452$). (Of the children whose data were analyzed here, 1,102 did not complete any task in the scanner or did not have this information available, 399 completed one task in the scanner, 584 completed two tasks in the scanner, and 9,452 completed all three tasks in the scanner.) (6) Data from children without performance flags on the emotional n -back task, SST, and MID task ($n = 6,441$). Performance flags, available in the curated ABCD data, were assigned based on the following criteria: <60% 0-back or 2-back accuracy on the emotional n -back task; <150 go trials, <60% go trial accuracy, >30% incorrect go trial percentage, >30% late go trial percentage, >30% “no response” go trials, <30 stop trials, or <20% or >80% stop trial accuracy on the SST; less than three positive and negative feedback events for large reward, small reward,

large loss, small loss, or no stakes trials on the MID task. (7) Data from a conservative subsample excluding outlier values, incomplete cases, children who completed neuroimaging tasks outside of the scanner, children with neuroimaging task performance flags, and related children. Age and sex were controlled with partial correlation ($n = 4,393$). (8) Data from each of the 21 data collection sites with more than 100 participants separately ($n = 328$ – 988 ; mean $n = 524.4$; SD = 192.1; 22nd site with 36 children excluded from this analysis).

Because of the presence of missing data, relationships between behavioral measures were evaluated with pairwise correlations rather than with data reduction techniques such as principal component analysis (PCA), which do not typically allow for missing data. 34.96% of children were missing at least one performance measure, and neuroimaging task performance data were missing in 20.14% of the sample on average (Table 1; note that recovery of missing data are ongoing). Although Bayesian probabilistic PCA can account for missing data as well as the nesting of participants in families and data collection sites (Thompson et al., 2019), this approach assumes that missing data occur randomly, independent of other sample features (Oba et al., 2003). This assumption is violated in the current sample, as, for example, children with better working memory function are less likely to be missing other behavioral measures (Spearman correlation between list sorting performance and number of missing performance measures = -0.09 , $p < 2.2 \times 10^{-16}$). [Of note, Thompson et al. (2019) performed Bayesian probabilistic PCA on a subset of the behavioral measures analyzed here, specifically those collected outside the scanner. A smaller percentage of these data are missing from curated release 2.0.1 (<3% per measure; Table 1), reducing issues associated with missingness.]

Neuroimaging data collection

ABCD scan sessions included a localizer and acquisition of a high-resolution anatomic scan, two runs of resting state fMRI, diffusion weighted images, 3D T2-weighted spin echo images, two more runs of resting state fMRI, and task-based fMRI. [Sites with Siemens scanners used framewise integrated real-time MRI monitoring (FIRMM; Dosenbach et al., 2017) to monitor children’s head motion during data collection; scan operators at these sites may have stopped resting-state data collection

after three runs if 12.5 min of low-motion resting-state data had been collected.] Image acquisition order was fixed, but fMRI task order was randomized across participants (Casey, 2018). Data were collected on Siemens Prisma, Phillips and GE 750 3T scanners, with detailed acquisition parameters reported in previous work (Casey, 2018; Hagler et al., 2019). Functional images were collected using a multiband gradient EPI sequence with the following parameters: TR = 800 ms, TE = 30 ms, flip angle = 52°, 60 slices acquired in the axial plane, voxel size = 2.4 mm³, multiband slice acceleration factor = 6.

Image preprocessing

Task-based data were processed by the ABCD Study Data Analysis and Informatics Center (DAIC) using approaches described in detail in Hagler et al. (2019). Preprocessing steps included motion correction with 3dvolreg in AFNI, B₀-distortion (i.e., field-map) correction, gradient nonlinearity distortion correction, and resampling scans into alignment with cubic interpolation using a mid-session scan as the reference. Registration between T2-weighted spin echo scans, field maps, and T1-weighted structural images was performed using mutual information. Functional images were aligned to T1-weighted images using rigid-body transformation (Hagler et al., 2019).

The equivalent of 16 volumes was removed from the start of each run. For Siemens and Philips scanners, eight volumes were removed because the first eight volumes were used as the multiband reference scans. For GE scanners running DV25 software, five volumes were removed because the first 12 volumes were used as the multiband reference. The images were then combined into a single volume and saved as the initial TR (leaving a total of five frames to be discarded). For GE scanners running DV26 software, 16 volumes were removed.

Relationships between fMRI activity and working memory function

After preprocessing, voxel-wise time series data were normalized within run. Task-related activity estimates were generated for each child using general linear models (GLMs) with 3dDeconvolve in AFNI (Hagler et al., 2019). GLMs included nuisance regressors accounting for baseline and quadratic trends as well as motion estimates and their derivatives, temporally filtered to attenuate 0.31- to 0.43-Hz signals related to respiration (Fair et al., 2020). Volumes with framewise displacement values >0.9 mm were censored (Siegel et al., 2014).

In addition to fixation, the emotional *n*-back task GLM included predictors for happy, fearful, and neutral face as well as place stimuli in the 0-back and 2-back conditions. Task blocks (~24 s) were modeled as square waves convolved with a two-parameter gamma basis function (Hagler et al., 2019). The SST model included predictors for correct and incorrect stop and go trials, modeled as instantaneous. The MID model included small and large reward and loss cues and feedback and no stakes cues, modeled as instantaneous (Hagler et al., 2019). GLM beta coefficients for cortical gray matter voxels were sampled into surface space. [This step differs from the processing pipeline described in Hagler et al. (2019), in which preprocessed data were sampled onto the cortical surface, but does not affect the beta values.]

To characterize relationships between out-of-scanner List Sorting Working Memory performance and fMRI activation in response to a working memory challenge, we first estimated memory-related activity with a linear contrast of 2-back versus 0-back emotional *n*-back task blocks (6,965 datasets available; *n* = 3,750 after exclusion). Subject-specific beta weights were entered into a multiple regression model, including list sorting performance as a predictor, with FSL's PALM software (Winkler et al., 2014). Covariates were also included in the model: age and sex (to account for effects present in the uncorrected list sorting standard scores; Luciana et al., 2018), scanner [to account for magnet-related differences between the 26 scanners as well as effects of participant population (e.g., family income, education, race and ethnicity)], fluid intelligence (to account for non-specific effects of cognitive function), and mean frame-to-frame head motion during the emotional *n*-back task. Head motion was measured with *tfmri_nback_all_beta_mm* from curated data release 2.0.1 sheet *nback_bwroi02.txt*. Degrees of freedom (*tfmri_nback_all_beta_dof*) and number of frames with a framewise displacement of <0.2 mm (*tfmri_nback_ab_subthn vols*) were not

included as covariates because they were highly correlated with frame-to-frame motion ($r = -0.96$ and -0.95 , respectively). Nonparametric significance was assessed with permutation testing using PALM's tail approximation acceleration method (Winkler et al., 2016). For each contrast, 1,000 permutations were run, a generalized Pareto distribution (GPD) was fit to the tail of the resulting null distribution, and *p* values below 0.10 were computed from the GPD. Regression coefficients surviving a family-wise error-corrected *p* value threshold of 0.05 were considered significant.

To assess the specificity of relationships between list sorting scores and memory-related fMRI activity, we repeated this analysis using contrast maps related to three other processes: emotion processing, inhibitory control, and reward processing. Subject-specific beta coefficient maps reflecting emotion processing-related activity were computed by contrasting emotional versus neutral face stimulus blocks on the emotional *n*-back task (6,965 datasets available; *n* = 3,750 after exclusion). Beta coefficient maps reflecting inhibitory control-related activity were computed by contrasting successful versus unsuccessful stop trials on the SST (6,906 datasets available; *n* = 4,316 after exclusion). Beta coefficient maps reflecting activity related to reward sensitivity were computed by contrasting successful versus unsuccessful reward trials (i.e., volumes corresponding to positive vs negative feedback) on the MID task (6,984 datasets available; *n* = 4,277 after exclusion). Multiple regression models including working memory performance, age, sex, scanner, fluid intelligence, and mean frame-to-frame head motion during the relevant task runs were applied to predict subject-specific beta weights.

Neuroimaging data exclusion

Neuroimaging data from all Philips scanners (~13% of the sample) were excluded from analysis due to an error in curated ABCD data preprocessing. Neuroimaging data from children with poor structural scan quality, determined with curated data release 2.0.1 sheet *freesqc01.txt*, were also excluded from analysis. Participants with a score of zero for *fsqc_qc* and/or a score greater than one for *fsqc_qu_motion*, *fsqc_qu_pia-lover*, *fsqc_qu_wmunder*, or *fsqc_qu_inhomogeneity* were excluded. For each task contrast, participants with fewer than 550 degrees of freedom in preprocessed, concatenated fMRI time series, missing grayordinate (i.e., gray-matter vertex or voxel) values, and/or extreme values (>3 SDs from the group mean) for the mean or SDs of beta weights over all grayordinates were also excluded. All fMRI analyses were performed using data from only one child per family to avoid confounds introduced by family structure. We elected to take this conservative approach rather than attempt to control for family structure with multilevel block permutation (Winkler et al., 2015) given the complexity of possible familial relationships (e.g., triplets, monozygotic and dizygotic twins, full siblings, half siblings, cousins, etc.).

Results

Working memory performance in childhood

Working memory function, operationalized with NIH Toolbox List Sorting Working Memory Test performance, approximated the normative population mean (uncorrected standard score mean = 96.7, SD = 12.0, range = 36–136; normative mean = 100, SD = 15). Working memory was positively correlated with age ($r_s = 0.148$, $p < 2.2 \times 10^{-16}$) and differed by sex, albeit with a negligible effect size (female mean = 96.3; male mean = 97.2; Welch $t_{(11317)} = -3.99$; $p = 6.8 \times 10^{-5}$; Cohen's $d = -0.075$). Performance on all other neurocognitive measures in the ABCD task battery, which assess short term memory, fluid intelligence, visuospatial attention, reading and language skills, cognitive control, processing speed, flexible thinking, learning, delay of gratification, emotion regulation, impulsivity, and reward processing, is visualized in Figure 1.

Behavioral signatures of working memory

Although a rich literature in cognitive psychology describes relationships between working memory and cognitive and

attentional processes in adulthood, how these associations emerge in development is less well understood. Thus, a primary goal of the current work is to characterize these associations in childhood to understand how they change over time.

To relate working memory to cognitive and attentional abilities, we computed pairwise Spearman correlations between performance scores on all tasks included in the dataset (i.e., all behavioral measures visualized in Fig. 1). Correlation coefficients are reported without corresponding *p* values because effect sizes as small as $r^2 = 3.34 \times 10^{-4}$ are significant at $p < 0.05$ in a sample of 11,537, and statistical dependence introduced by family relatedness, site effects, and the inclusion of multiple performance measures per test precludes parametric *r*-to-*p* conversion. Furthermore, the goal of this analysis is to establish a pattern of behavioral relationships rather than to evaluate the statistical significance of particular associations.

Across individuals, working memory was most strongly related to language skills measured with the NIH Toolbox Picture Vocabulary and Oral Reading Recognition tests (r_s values = 0.40); memory-related performance on the emotional *n*-back task, Picture Sequence Memory Test, and RAVLT; and fluid intelligence measured with matrix reasoning ($r_s = 0.35$; Figs. 2, 3).

Correlations between list sorting performance and performance on other memory tests also revealed relationships between different aspects of memory. The emotional *n*-back task, collected during fMRI, measured performance during high memory load (2-back) and low memory load (0-back) task blocks. During 2-back blocks, children were asked to indicate when they saw a picture identical to the one they saw two trials back. During 0-back blocks, children were shown a target picture and instructed to indicate when they saw a matching image. Working memory was more strongly related to 2-back than to 0-back accuracy ($r_s = 0.36$ vs 0.32; Steiger's $z = 5.36$, $p < 0.0001$), indicating that, as predicted, working memory ability is reflected to a greater degree by performance on high-load versus low-load *n*-back blocks. [Interestingly, 2-back and 0-back accuracy scores were highly correlated ($r_s = 0.62$), suggesting that common processes such as working memory, attention, and motivation contribute to performance on both tasks.] Recognition memory for

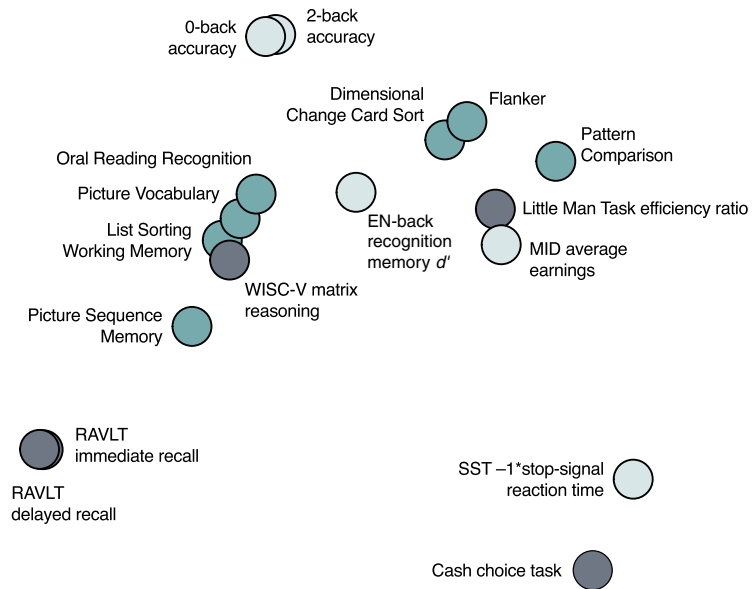


Figure 2. Multidimensional scaling plot illustrating 2D distance between behavioral metrics in children with no missing data ($n = 7,504$). Classical multidimensional scaling was applied to the complete-case sample to avoid assumptions associated with imputing missing values. Distance was calculated as the Euclidean distance between each pair of behavioral measures after mean-centering and scaling each measure across participants. NIH Toolbox measures are shown in dark green, other neurocognitive measures in dark gray, and neuroimaging task measures in light green.

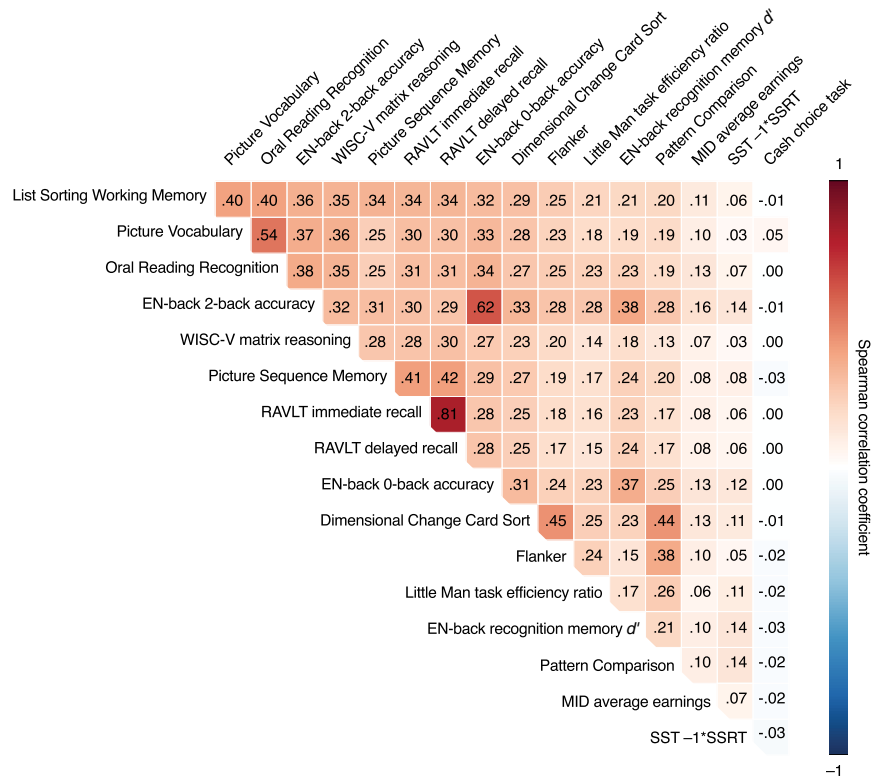


Figure 3. Spearman correlations between performance measures in the full 11,537-child sample. Measures are ordered according to the strength of their relationship with working memory, operationalized as NIH Toolbox List Sorting Working Memory Test. Because the outcome of the cash choice task is binary, relationships with performance on this measure are equivalent to point-biserial Spearman correlation coefficients.

emotional *n*-back stimuli (happy, fearful, and neutral face photographs and place photographs) was tested after fMRI data collection. Although working memory was less highly correlated with recognition memory than with performance on

visual attention tasks, including the Flanker task ($r_s = 0.21$ vs 0.25 ; Steiger's $z = -4.01$, $p = 0.0001$), this may reflect low overall memory for specific stimuli, especially face photographs, at this age (Casey, 2018). Finally, the RAVLT assessed immediate recall of a word list as well as recall after a 30-min delay, and the NIH Toolbox Picture Sequence Memory Test measured episodic memory and visuospatial sequencing. Working memory was similarly related to immediate and delayed recall on the RAVLT and performance on the Picture Sequence Memory Test (r_s values = 0.3389, 0.3372, and 0.3428, respectively).

Of note, correlations between list sorting performance and other behavioral measures are influenced by both the similarity of the constructs they measure and the reliability of the measures themselves. Previous work reports good to excellent reliability for all NIH Toolbox tasks [test-retest intraclass correlation coefficient (ICC) = 0.61–0.90; Dikmen et al., 2014; Gershon et al., 2014; Tulsy et al., 2014; Zelazo et al., 2014; Carlozzi et al., 2015], the WISC-V matrix reasoning test (test-retest $r = 0.78$; Luciana et al., 2018), and the RAVLT (test-retest $r = \sim 0.60$ – 0.70 ; Van Den Burg and Kingma, 1999; Luciana et al., 2018). In the current sample, ICC, calculated by comparing performance during participants' first and second fMRI runs, is 0.64 for the 0-back task and 0.68 for the 2-back task. However, the reliability of MID earnings is poor (ICC = 0.27). Thus, the low correlation between list sorting performance and measures including MID earnings may in part reflect low measure reliability. (Test-retest reliability estimates for the Intertemporal Cash Choice Task, Little Man Task, and post-scan n -back stimuli recognition memory test are not available because these tests were only completed once and reliability estimates have not been published elsewhere; test-retest reliability was not calculated for the SST because run-specific stop-signal delay values were not available.)

Behavioral relationships replicate across independent data releases

To assess the reproducibility of behavioral relationships, we cross-correlated all performance measures using data from curated ABCD releases 1.1 ($n = 4,397$) and 2.0.1 ($n = 7,140$ new individuals) separately. The pattern of relationships was stable across releases (spatial $r_s = 0.982$), suggesting that the observed pattern is generalizable rather than idiosyncratic to a particular sample.

Behavioral relationships are not influenced by family structure

The full 11,537-child cohort includes 3,532 related children from 9,750 unique families (based on self-report). Because relatedness affects the independence of behavioral measures and could have affected relationships between them, we replicated correlations between cognitive and attentional abilities in a subset of data from only one child per family. The pattern of behavioral relationships in this unrelated subsample was nearly identical to that observed in the full sample: the Spearman spatial correlation between the two samples' vectorized behavioral cross-correlation matrices was 0.999.

Given the importance of accounting for family structure in big data samples, we next characterized effects of relatedness on behavioral relationships with an additional analysis. First, we computed the absolute difference between all 136 pairwise behavioral correlations in the full sample ($n = 11,537$) and the unrelated subsample ($n = 9,750$ after excluding 1,787 related children). Next, we randomly excluded 1,787 children from the full sample, re-calculated pairwise behavioral relationships, and recorded the difference between the full-sample correlations and these random subsample correlations. We repeated this process 1,000 times to generate a null distribution of correlation

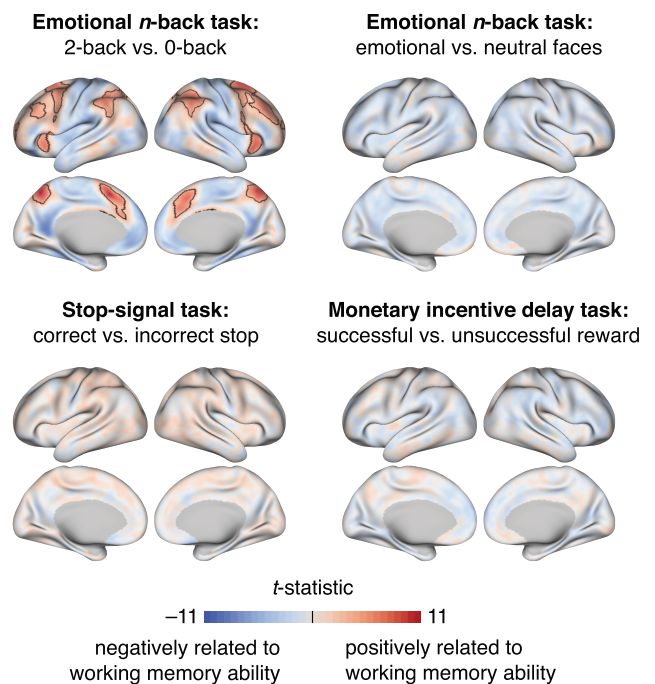


Figure 4. Relationships between fMRI activation and working memory function, measured with an out of scanner list sorting task, across individuals. Analyses control for age, sex, scanner, fluid intelligence, and mean frame-to-frame head motion during the relevant fMRI runs. Unthresholded t statistics (regression coefficients divided by their SE) are visualized on the inflated cortical surface. Black outlines indicate vertices significant at family-wise error-corrected, two-tailed $p < 0.05$.

coefficient differences for each pair of behavioral measures. Non-parametric p values were generated by comparing each true correlation difference, $|r(i,j)_{\text{full sample}} - r(i,j)_{\text{unrelated subsample}}|$, to its corresponding null distribution. We used this conservative sub-sampling approach rather controlling for family relatedness with linear mixed models because of the range of possible familial relationships and the fact that relatedness may be inaccurately captured with self-report measures.

Using the subsampling approach, we found that none of the 136 pairwise behavioral relationships differed between the full sample and the unrelated subsample more than they differed between the full sample and the random subsamples (all $>$ Bonferroni-corrected $p = 0.05/136 = 3.68 \times 10^{-4}$). Therefore, excluding family members from the sample did not disproportionately affect pairwise behavioral relationships, and including related participants does not bias the current results.

Behavioral relationships are robust to age, sex, outliers, and missing data

Control analyses confirmed that behavioral relationships were robust to other potential confounds. Specifically, the overall pattern of relationships was consistent after controlling for age and sex with partial correlation ($n = 11,537$; $r_s = 0.995$), excluding outlier values (>2.5 SDs from the group mean; $n = 11,537$; $r_s = 0.995$), excluding children with any missing behavioral scores ($n = 7,504$; $r_s = 0.997$), excluding children who completed any neuroimaging task (i.e., the emotional n -back task, SST, or MID task) on a laptop outside the scanner ($n = 9,452$; $r_s = 0.997$), and excluding children with neuroimaging task performance flags provided by the ABCD Study ($n = 6,441$; $r_s = 0.982$). The overall pattern of relationships was replicated to a lesser degree in a conservative subsample of children excluding relatives, outlier

values, incomplete cases, children who completed neuroimaging tasks outside the scanner, and children with neuroimaging task performance flags, and controlling for age and sex ($n = 4,393$; $r_s = 0.973$). Finally, associations between behavioral measures were similar across data collection sites despite differences in target sociodemographics (Garavan et al., 2018). Similarity between site-specific behavioral cross-correlation patterns ranged from $r_s = 0.80$ to $r_s = 0.94$ (mean $r_s = 0.86$, $SD = 0.03$).

Neural signatures of working memory

To identify a vertex-wise map of associations with working memory, we related out-of-scanner List Sorting Working Memory performance to fMRI activation in response to a working memory challenge. Working memory was significantly related to 2-back versus 0-back (i.e., high vs low memory load) activation in regions of frontal and parietal cortex including bilateral intraparietal sulci, dorsal premotor cortex/frontal eye fields, dorsolateral prefrontal cortex, anterior insula, dorsal anterior cingulate cortex extending into the presupplementary motor area, and precuneus (Fig. 4). In line with previous work highlighting the importance of frontoparietal regions for working memory in development (Klingberg et al., 2002; Satterthwaite et al., 2013; Klingberg and Darki, 2015), children with better out-of-scanner working memory performance showed increased activity during high relative to low memory load task blocks in this distributed set of regions that overlap with frontoparietal and dorsal attention networks (Power et al., 2011; Yeo et al., 2011; Fig. 5).

Relationships between working memory and fMRI activity are domain specific

Are patterns of fMRI activation that track individual differences in working memory driven by general task demands, or are they driven by working memory engagement per se? We performed two analyses to disentangle these alternatives. First, we examined the association between individual differences in working memory performance and activation revealed by a contrast emotional of versus neutral face blocks in the n -back task. Although these emotion-related activation patterns were measured during a working memory task, the contrast of these two types of face blocks does not reflect a working memory challenge. Therefore, significant relationships between these patterns and working memory would suggest that neural signatures of working memory are domain general rather than domain specific. Second, we examined the relationships between individual differences in working memory and activation patterns reflecting distinct cognitive processes in distinct task contexts: inhibitory control during an SST and reward processing during a MID task.

Results revealed that working memory was not significantly associated with emotion-related activation during the emotional n -back task, inhibitory control-related activation during the SST, or reward-related activation during the MID task (Fig. 4). Although we did not compare regression coefficients across conditions because participant samples were overlapping but not identical, more participants and time points were available for the SST and

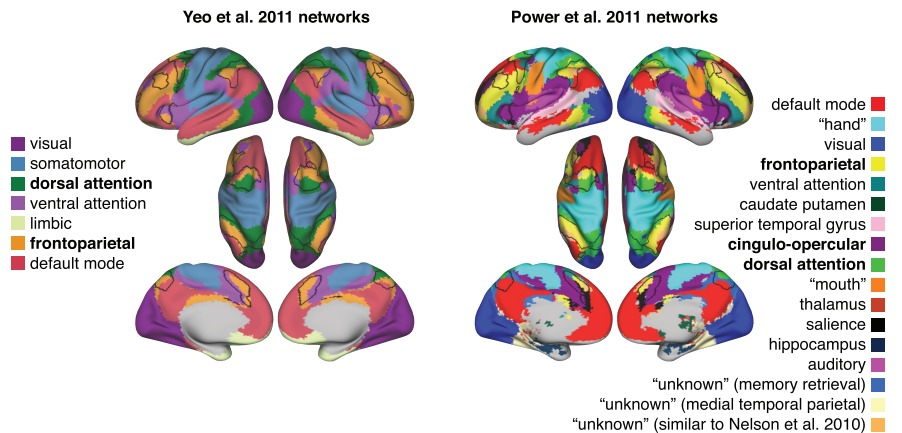


Figure 5. Overlap between neural signatures of working memory in childhood and canonical resting-state functional networks from Yeo et al. (2011) and Power et al. (2011). Black outlines indicate significant relationships between 2-back versus 0-back activation and working memory function across individuals (family-wise error-corrected, two-tailed $p < 0.05$).

MID than for the emotional n -back task. Furthermore, within-subject reliability of fMRI task activations (characterized by comparing beta weights in regions of the Desikan cortical brain atlas from participants' first and second task runs, as included in curated data release 2.0.1) was numerically highest for the MID task contrast. Thus, the presence of significant effects for the working memory contrast, but not the inhibitory control or reward processing contrasts, is not attributable to sample size, amount of data per individual, or increased within-subject fMRI activation reliability for the working memory contrast, and results suggest that frontoparietal activity is a domain-specific rather than a domain-general signature of working memory.

Memory-related frontoparietal activity reflects in-scanner and out-of-scanner working memory performance

One potential explanation of the current results is that in-scanner emotional n -back performance, a state-like measure of working memory and task engagement rather than a measure of individual differences in working memory per se, drives the selective relationship between working memory and 2-back versus 0-back frontoparietal activation. To evaluate this possibility, we replicated the analysis relating out-of-scanner working memory performance to 2-back versus 0-back activation with age, sex, scanner, fluid intelligence, mean frame-to-frame head motion during the n -back task, and 0-back and 2-back accuracy included in the model as covariates. Results revealed significant clusters in frontoparietal regions (Fig. 6), demonstrating that memory-related activation reflects both in-scanner and out-of-scanner working memory performance.

As an exploratory analysis, we next related 2-back versus 0-back activity to behavior with a multiple regression model that included all 17 behavioral measures reported in Figure 1 as well as age, sex, scanner, and mean frame-to-frame head motion during the n -back task. As expected, n -back activity predicted in-scanner 2-back and 0-back accuracy, even when controlling for performance on all out-of-scanner cognitive tasks. In addition, significant (albeit qualitatively weaker) relationships were observed between 2-back versus 0-back activity and performance on the list sorting task and the out-of-scanner measures with which it was most highly correlated, the NIH Toolbox Picture Vocabulary and Oral Reading Recognition tasks and the WISC-V matrix reasoning test. Activity was also significantly related to

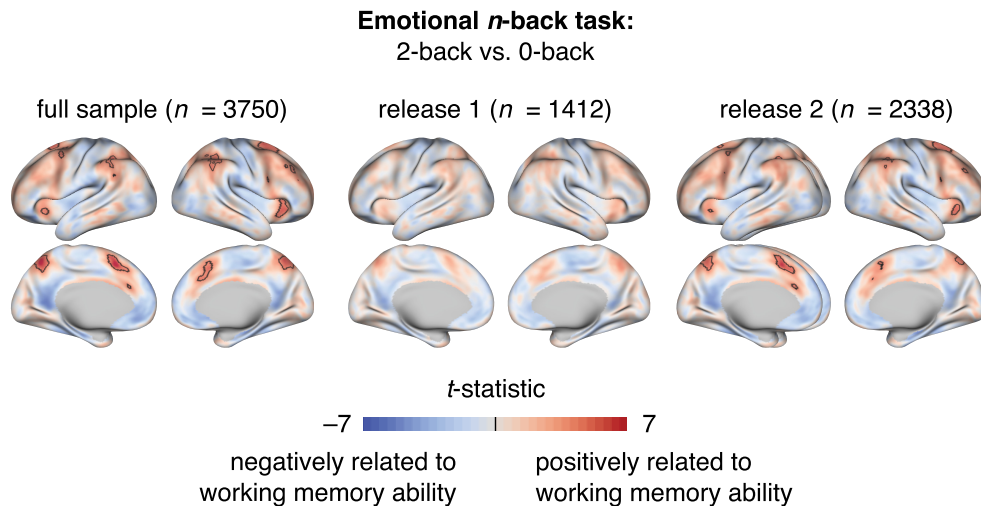


Figure 6. Relationships between 2-back versus 0-back activation and out-of-scanner working memory performance across individuals, controlling for age, sex, scanner, fluid intelligence, mean frame-to-frame head motion, and in-scanner 0-back and 2-back accuracy, in the full sample and in ABCD data releases 1.1 (“release 1”) and 2.0.1 (“release 2”) separately. Black outlines indicate vertices significant at family-wise error-corrected, two-tailed $p < 0.05$. Results demonstrate that frontoparietal activation in 2-back versus 0-back contrasts reflects trait-like in addition to state-like working memory abilities.

performance on the Little Man Task. Together, these results suggest that frontoparietal activity reflects both in-scanner and out-of-scanner measures of working memory as well as behavioral performance on statistically related measures. Given that a limitation of this exploratory analysis is strong correlation among predictor variables, future work relating independent measures of cognitive and attentional processes to frontoparietal activity observed in multiple task contexts can inform the degree to which working memory-related activity specifically reflects working memory behavioral performance.

Neuroimaging findings replicate across independent data releases

To assess the reproducibility of observed brain–behavior relationships, we repeated fMRI analyses using data from participants first included in ABCD releases 1.0 and 2.0.1 separately. Consistent with the full-sample findings, List Sorting Working Memory Task performance was significantly related to 2-back versus 0-back activity in frontoparietal regions in both subsamples. List sorting performance was not significantly related to emotional versus neutral face, successful versus unsuccessful stop, or successful versus unsuccessful reward activity in either subsample (Fig. 7). When controlling for in-scanner *n*-back accuracy, the pattern of results was qualitatively similar, but significant relationships were observed in release-2 data only (release 1 $n = 1,412$; release 2 $n = 2,338$; Fig. 6).

Discussion

Working memory is a foundational cognitive function that changes over development and varies across individuals. Here, we characterize relationships between working memory, cognitive and attentional processes, and task-related brain activity in childhood using behavioral and fMRI data from the largest developmental neuroimaging sample to date. Behavioral analyses demonstrate that children with stronger working memory abilities perform better on a range of cognitive tasks, and revealed close relationships between working memory, performance on other memory tasks, language abilities, and fluid intelligence. fMRI analyses of emotional *n*-back task, SST, and MID task data provide evidence

that frontoparietal activation in response to an explicit working memory challenge, but not in response to general task demands, is a reliable marker of working memory ability. Finally, a control analysis suggests that memory-related frontoparietal activity reflects individual differences in working memory above and beyond ongoing task performance.

Positive associations between working memory, language abilities, and fluid intelligence replicate previous work on the structure of cognition in children and adults (Engle et al., 1999; Gathercole, 1999; Thompson et al., 2019). As expected, children with stronger working memory performance (measured with the List Sorting Working Memory Test) also showed better performance on tests of episodic memory (Picture Sequence Memory), short-term memory (Rey Auditory Verbal Learning), and low-load and high-load working memory (emotional *n*-back 0- and 2-back conditions, respectively). Correlations between these measures in the full sample of 11,537 children ranged from 0.32 to 0.36, suggesting that they reflect both distinct and overlapping aspects of memory function. Somewhat surprisingly given established links between working memory and processing speed (Conway et al., 2002), working memory was less closely related to performance on the Pattern Comparison Processing Speed Test than to performance on every cognitive task except the SST, MID task, and Intertemporal Cash Choice task. Although the strength of the relationship between working memory and processing speed ($r_s = 0.20$) is numerically similar to previous findings with the same tasks in 8- to 12-year-olds ($r = 0.26$; Carullozzi et al., 2015), individual differences in working memory were more strongly related to processes including executive attention and cognitive flexibility than to processing speed in the current cohort. Together, these results reveal relationships between working memory and cognitive and attentional processes in childhood.

This behavioral cross-correlation pattern was consistent after controlling for age and sex and excluding statistical outliers, incomplete cases, and neuroimaging task data collected outside the scanner. These behavioral patterns remained unchanged when measured in a subsample of the data that did not include relatives (i.e., only including one child per identified family). Thus, although it is important to account for these factors in

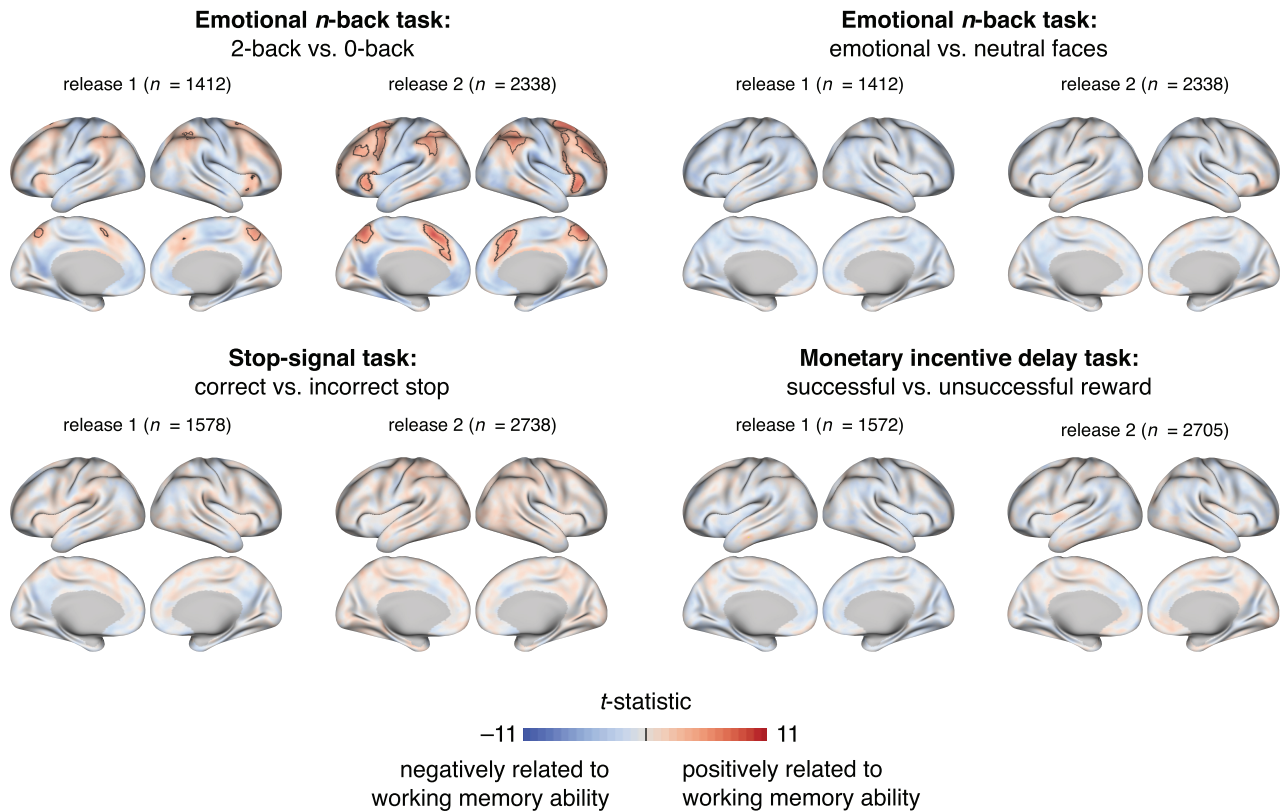


Figure 7. Relationships between task activation and out-of-scanner working memory performance across individuals, controlling for age, sex, scanner, fluid intelligence, and mean frame-to-frame head motion. Black outlines indicate vertices significant at family-wise error-corrected, two-tailed $p < 0.05$. Analyses were run using data from each release separately. Participants included in the full sample were included in the “release 1” analysis if they appeared in curated ABCD data release 1.1, and were included in “release 2” otherwise.

large datasets such as the ABCD sample, the current results appear robust to effects of statistical dependence and outliers. Furthermore, the results are not biased by the inclusion of related children.

Neuroimaging results likewise align with previous work, providing evidence that frontoparietal activity reflects differences in working memory function during development (Klingberg et al., 2002; Satterthwaite et al., 2013). The narrow age range of the current sample, however, allowed us to disentangle individual differences from developmental changes, providing novel evidence that frontoparietal brain function underlies variability in working memory both within and across individuals. Furthermore, assessing relationships between working memory and fMRI activity related to memory, emotion processing, inhibitory control, and reward processing demonstrated that frontoparietal activation is a domain-specific rather than a task-general neural signature of working memory. Accounting for in-scanner emotional *n*-back performance, which reflects individual differences in working memory and attentional processes as well as transient attentional state, revealed relationships between out-of-scanner working memory performance and memory-related fMRI activation in regions of superior parietal and presupplementary motor cortex. Children with stronger working memory abilities, therefore, show increased frontoparietal activation during high relative to low memory load task blocks in part because they simply perform better on these tasks, but also because of individual differences in their ability to hold and manipulate information in mind.

The current results suggest that frontoparietal activation is a domain-specific neural signature of working memory in that

individual differences in working memory are selectively reflected in 2-back versus 0-back frontoparietal activity. Results of an exploratory analysis also suggest some degree of specificity in the reverse direction: 2-back versus 0-back frontoparietal activity is uniquely related to both in-scanner and out-of-scanner working memory measures, as well as a subset of measures correlated with working memory task performance. Importantly however, frontoparietal activity does not only support working memory function but is also related to processes including attention and cognitive control (Corbetta and Shulman, 2002; Vincent et al., 2008; Spreng et al., 2010, 2013; Ptak, 2012). Recent work has emphasized the multifunctional nature of the frontoparietal network, proposing that it represents a domain-general “cognitive core” of the brain (Assem et al., 2020). Our results are not inconsistent with this conceptualization. Rather, they demonstrate that a high versus low memory load contrast reveals a frontoparietal activity signature of working memory, and leave open the possibility that an attention or cognitive control contrast could reveal a frontoparietal activity signature of attention or cognitive control. Future work that expands the collection of attention and control tasks and varies their cognitive demands will provide additional insights into the functional significance of overlapping and distinct patterns of frontoparietal activity across psychological tasks with development.

A neural signature of working memory based on task activation data complements a growing body of work identifying neuromarkers of individual differences from functional brain connectivity. In particular, patterns of task-based and resting-state functional connectivity, or statistical dependence between two brain regions’ activity time courses, have been used to

predict individual differences in abilities including attention, fluid intelligence, and aspects of memory (Finn et al., 2015; Galeano Weber et al., 2017; Rosenberg et al., 2017; Lin et al., 2018; Rudolph et al., 2018; Yamashita et al., 2018; Avery et al., 2020). Recent work suggests that models based on task connectivity generally outperform those based on resting-state connectivity for predicting behavior, potentially because tasks engage circuits related to a process of interest to magnify individual differences in behaviorally relevant neural phenotypes, thereby improving predictions (Finn et al., 2017b; Greene et al., 2018; Rosenberg et al., 2018; Yoo et al., 2018). It is an open question, however, whether tasks selectively enhance the prediction of task-relevant behaviors. Here, motivated by previous work relating frontoparietal activation to developmental change in working memory (Klingberg et al., 2002; Kwon et al., 2002; Crone et al., 2006; Satterthwaite et al., 2013), we address this question with task activation rather than functional connectivity analyses. The current result, that frontoparietal activity indexes working memory only when working memory is explicitly taxed, suggests that task challenges may reveal neural signatures of task-relevant behaviors, and underscores the importance of multitask or multicondition data for elucidating state-specific and state-general biomarkers of behavior.

The goal of the current work was to characterize a brain-based biomarker and behavioral signature of working memory in childhood not just for the sake of understanding these relationships at a single point in time, but also for ultimately understanding their trajectories across development. Because the ABCD Study will follow children from age 9–10 to age 19–20, longitudinal work can provide new insights into associations between working memory, cognitive and attentional processes, and real-world outcomes across adolescence and young adulthood. Biennial MRI sessions, during which participants will complete the same emotional *n*-back task, SST, and MID task that they completed at ages 9 through 10, will also facilitate the discovery of changing neural signatures of abilities and behavior. For example, will there be changes in the distinct and overlapping brain activity patterns associated with working memory, inhibitory control, and reward processing with age? Will the domain specificity and domain generality of these signatures vary over time? Are there different developmental trajectories for frontoparietal organization of function across these processes? A fruitful way to frame the current findings is as a single point along a nonlinear trajectory rather than as a summary of working memory function in development as a whole.

Finally, as sample sizes in psychology and human neuroscience rapidly increase, it is important to note limitations of big data cohort-based approaches. First, behavioral and neuroimaging task batteries for these studies are determined by committee to address specific scientific goals. Although the resulting task sets often assess a range of mental processes, they may not be optimal for answering all questions. In the ABCD Study neuroimaging battery, for example, cognitive control demands and task difficulty are not equated across the emotional *n*-back task, SST, and MID task. Thus, the 2-back versus 0-back contrast may reflect processes such as cognitive control and attention that are not reflected in the three control contrasts. In addition, the individualized nature of the SST and MID task (which results in different timing parameters for different children) may obscure activation patterns related to individual differences in behavior. Furthermore, recent work highlights potential issues in the ABCD stop-signal task that may impact interpretation of the SST

data (Bissett et al., 2020). Future work relating individual differences in working memory to fMRI activity reflecting cognitive control, attentional engagement, and other processes in contexts matched for task difficulty will further inform the domain specificity and generality of neural signatures of working memory. Second, large samples are not necessarily representative samples, and the ABCD cohort, while geographically, demographically, and socioeconomically diverse, should not be considered representative of the country or the world as a whole (Garavan et al., 2018). Looking ahead, future work relating cognitive and neural measures in weighted samples (LeWinn et al., 2017) can complement existing studies of single-site and multisite datasets. Third, just as the ABCD participant population may not represent youth as a whole, the structure of neurocognition in nine- and 10-year-olds likely does not reflect that of children at other ages. Longitudinal analyses of the ABCD cohort can inform changes in brain–behavior relationships across adolescence, and data collection efforts such as the Human Connectome Project (HCP) Development Study (Somerville et al., 2018) and HCP Aging Study (Bookheimer et al., 2019) can inform these associations in younger and older individuals. Finally, because even small effects can reach significance when samples are large, it is helpful to distinguish statistical from practical significance. Here, we focused on statistical significance as a proof-of-principle demonstration that memory-related frontoparietal activity tracks individual differences in working memory in childhood. Future work can evaluate practical or applied significance by testing whether models based on task activation patterns generalize to predict real-world outcomes including academic performance or changes in these outcomes over time.

Despite these limitations, the current results provide the most well powered characterization of relationships between working memory, cognitive and attentional processes, and task-based fMRI activation in development to date. Overall, they replicate established findings that children with stronger working memory function perform better on a variety of cognitive tasks, particularly those assessing other aspects of memory, language skills, and fluid intelligence. Furthermore, they provide evidence that frontoparietal network activation in response to an explicit working memory challenge is a robust and domain-specific marker of individual differences in working memory ability at age nine and ten. Together, these results inform understanding of the structure of neurocognition in childhood, and highlight the importance of evaluating brain–behavior relationships in multiple task contexts to demarcate the specificity and generality of neural signatures of abilities and behavior.

References

- Acker W, Acker C (1982) Bexley Maudsley automated processing screening and Bexley Maudsley category sorting test manual. Windsor: NFER-Nelson Publishing.
- Alloway TP, Alloway RG (2010) Investigating the predictive roles of working memory and IQ in academic attainment. *J Exp Child Psychol* 106:20–29.
- Alloway TP, Alloway RG (2013) Working memory across the lifespan: a cross-sectional approach. *J Cogn Psychol* 25:84–93.
- Alloway TP, Gathercole SE, Pickering SJ (2006) Verbal and visuospatial short-term and working memory in children: are they separable? *Child Dev* 77:1698–1716.
- Alp IE (1994) Measuring the size of working memory in very young children: the imitation sorting task. *Int J Behav Dev* 17:125–141.
- Assem M, Glasser MF, Van Essen DC, Duncan J (2020) A domain-general cognitive core defined in multimodally parcellated human cortex. *Cereb Cortex*, bhaa023.
- Avery EW, Yoo K, Rosenberg MD, Greene AS, Gao S, Na DL, Scheinost D, Constable TR, Chun MM (2020) Distributed patterns of functional

- connectivity predict working memory performance in novel healthy and memory-impaired individuals. *J Cogn Neurosci* 32:241–255.
- Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, Glasser MF, Curtiss S, Dixit S, Feldt C, Nolan D, Bryant E, Hartley T, Footer O, Bjork JM, Poldrack R, Smith S, Johansen-Berg H, Snyder AZ, Van Essen DC, et al. (2013) Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage* 80:169–189.
- Bauer PJ, Dikmen SS, Heaton RK, Mungas D, Slotkin J, Beaumont JL (2013) III. NIH Toolbox Cognition Battery (CB): measuring episodic memory. *Monogr Soc Res Child Dev* 78:34–48.
- Bayliss DM, Jarrold C, Gunn DM, Baddeley AD (2003) The complexities of complex span: explaining individual differences in working memory in children and adults. *J Exp Psychol Gen* 132:71–92.
- Bissett PG, Hagen MP, Poldrack RA (2020) A cautionary note on stop-signal data from the Adolescent Brain Cognitive Development [ABCD] study. *bioRxiv*. Advance online publication. Retrieved May 15, 2020. doi:10.1101/2020.05.08.084707.
- Bookheimer SY, Salat DH, Terpstra M, Ances BM, Barch DM, Buckner RL, Burgess GC, Curtiss SW, Diaz-Santos M, Elam JS, Fischl B, Greve DN, Hagy HA, Harms MP, Hatch OM, Hedden T, Hodge C, Japardi KC, Kuhn TP, Ly TK, et al. (2019) The Lifespan Human Connectome Project in Aging: an overview. *Neuroimage* 185:335–348.
- Burzynska AZ, Li S-C, Lindenberger U, Nagel IE, Heekeren HR, Bäckman L, Preuschhof C (2011) Microstructure of frontoparietal connections predicts cortical responsivity and working memory performance. *Cereb Cortex* 21:2261–2271.
- Buss AT, Ross-Sheehy S, Reynolds GD (2018) Visual working memory in early development: a developmental cognitive neuroscience perspective. *J Neurophysiol* 120:1472–1483.
- Carlozzi NE, Beaumont JL, Tulskey DS, Gershon RC (2015) The NIH Toolbox Pattern Comparison Processing Speed Test: normative data. *Arch Clin Neuropsychol* 30:359–368.
- Casey BJ (2018) The Adolescent Brain Cognitive Development (ABCD) Study: imaging acquisition across 21 sites. *Dev Cogn Neurosci* 32:42–54.
- Casey BJ, Cohen JD, Jezzard P, Turner R, Noll DC, Trainor RJ, Giedd J, Kaysen D, Hertz-Pannier L, Rapoport JL (1995) Activation of prefrontal cortex in children during a nonspatial working memory task with functional MRI. *Neuroimage* 2:221–229.
- Casey BJ, Giedd JN, Thomas KM (2000) Structural and functional brain development and its relation to cognitive development. *Biol Psychol* 54:241–257.
- Casey BJ, Tottenham N, Liston C, Durston S (2005) Imaging the developing brain: what have we learned about cognitive development? *Trends Cogn Sci* 9:104–110.
- Conklin HM, Luciana M, Hooper CJ, Yarger RS (2007) Working memory performance in typically developing children and adolescents: behavioral evidence of protracted frontal lobe development. *Dev Neuropsychol* 31:103–128.
- Conley MI, Dellarco DV, Rubien-Thomas E, Cohen AO, Cervera A, Tottenham N, Casey BJ (2018) The racially diverse affective expression (RADIATE) face stimulus set. *Psychiatry Res* 270:1059–1067.
- Conway ARA, Cowan N, Bunting MF, Theriault DJ, Minkoff SRB (2002) A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence* 30:163–183.
- Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–215.
- Crone EA, Wendelken C, van Leijenhorst L, Donohue S, Bunge SA (2006) Neurocognitive development of the ability to manipulate information in working memory. *Proc Natl Acad Sci USA* 103:9315–9320.
- Davidson MC, Amso D, Anderson LC, 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* 44:2037–2078.
- De Smedt B, Janssen R, Bouwens K, Verschaffel L, Boets B, Ghesquière P (2009) Working memory and individual differences in mathematics achievement: a longitudinal study from first grade to second grade. *J Exp Child Psychol* 103:186–201.
- Diamond A, Goldman-Rakic PS (1989) Comparison of human infants and rhesus monkeys on Piaget's AB task: evidence for dependence on dorso-lateral prefrontal cortex. *Exp Brain Res* 74:24–40.
- Dikmen SS, Bauer PJ, Weintraub S, Mungas D, Slotkin J, Beaumont JL, Gershon R, Temkin NR, Heaton RK (2014) Measuring episodic memory across the lifespan: NIH Toolbox Picture Sequence Memory Test. *J Int Neuropsychol Soc* 20:611–619.
- Dosenbach NUF, Koller JM, Earl EA, Miranda-Dominguez O, Klein RL, Van AN, Snyder AZ, Nagel BJ, Nigg JT, Nguyen AL, Wesevich V, Greene DJ, Fair DA (2017) Real-time motion analytics during brain MRI improve data quality and reduce costs. *Neuroimage* 161:80–93.
- Ekman M, Fiebach CJ, Melzer C, Tittgemeyer M, Derrfuss J (2016) Different roles of direct and indirect frontoparietal pathways for individual working memory capacity. *J Neurosci* 36:2894–2903.
- Engel de Abreu PMJ, Conway ARA, Gathercole SE (2010) Working memory and fluid intelligence in young children. *Intelligence* 38:552–561.
- Engle RW, Tuholski SW, Laughlin JE, Conway ARA (1999) Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J Exp Psychol Gen* 128:309–331.
- Eriksen BA, Eriksen CW (1974) Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept Psychophys* 16:143–149.
- Eriksson J, Vogel EK, Lansner A, Bergström F, Nyberg L (2015) Neurocognitive architecture of working memory. *Neuron* 88:33–46.
- Fair DA, Miranda-Dominguez O, Snyder AZ, Perrone A, Earl EA, Van AN, Koller JM, Feczko E, Tisdall MD, van der Kouwe A, Klein RL, Mirro AE, Hampton JM, Adeyemo B, Laumann TO, Gratton C, Greene DJ, Schlaggar BL, Hagler DJ, Watts R, et al. (2020) Correction of respiratory artifacts in MRI head motion estimates. *Neuroimage* 208:116400.
- Fan J, McCandliss BD, Sommer T, Raz A, Posner MI (2002) Testing the efficiency and independence of attentional networks. *J Cogn Neurosci* 14:340–347.
- Finn AS, Minas JE, Leonard JA, Mackey AP, Salvatore J, Goetz C, West MR, Gabrieli CFO, Gabrieli JDE (2017a) Functional brain organization of working memory in adolescents varies in relation to family income and academic achievement. *Dev Sci* 20:e12450.
- Finn ES, Scheinost D, Finn DM, Shen X, Papademetris X, Constable RT (2017b) Can brain state be manipulated to emphasize individual differences in functional connectivity? *Neuroimage* 160:140–151.
- Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM, Papademetris X, Constable RT (2015) Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat Neurosci* 18:1664–1671.
- Fry AF, Hale S (1996) Processing speed, working memory, and fluid intelligence: evidence for a developmental cascade. *Psychol Sci* 7:237–241.
- Fry AF, Hale S (2000) Relationships among processing speed, working memory, and fluid intelligence in children. *Biol Psychol* 54:1–34.
- Fukuda K, Vogel E, Mayr U, Awh E (2010) Quantity not quality: the relationship between fluid intelligence and working memory capacity. *Psychon Bull Rev* 17:673–679.
- Galeano Weber EM, Hahn T, Hilger K, Fiebach CJ (2017) Distributed patterns of occipito-parietal functional connectivity predict the precision of visual working memory. *Neuroimage* 146:404–418.
- Garavan H, Bartsch H, Conway K, Decastro A, Goldstein RZ, Heeringa S, Jernigan T, Potter A, Thompson W, Zals D (2018) Recruiting the ABCD sample: design considerations and procedures. *Dev Cogn Neurosci* 32:16–22.
- Gathercole SE (1999) Cognitive approaches to the development of short-term memory. *Trends Cogn Sci* 3:410–419.
- Gathercole SE, Pickering SJ, Ambridge B, Wearing H (2004) The structure of working memory from 4 to 15 years of age. *Dev Psychol* 40:177–190.
- Gazzaley A, Cooney JW, Rissman J, D'Esposito M (2005) Top-down suppression deficit underlies working memory impairment in normal aging. *Nat Neurosci* 8:1298–1300.
- Gershon RC, Wagster MV, Hendrie HC, Fox NA, Cook KF, Nowinski CJ (2013) NIH toolbox for assessment of neurological and behavioral function. *Neurology* 80:S2–S6.
- Gershon RC, Cook KF, Mungas D, Manly JJ, Slotkin J, Beaumont JL, Weintraub S (2014) Language measures of the NIH toolbox cognition battery. *J Int Neuropsychol Soc* 20:642–651.
- Ghasemi A, Zahediasl S (2012) Normality tests for statistical analysis: a guide for non-statisticians. *Int J Endocrinol Metab* 10:486–489.
- Greene AS, Gao S, Scheinost D, Constable RT (2018) Task-induced brain state manipulation improves prediction of individual traits. *Nat Commun* 9:2807.

- Hagler DJ, Hatton SN, Cornejo MD, Makowski C, Fair DA, Dick AS, Sutherland MT, Casey BJ, Barch DM, Harms MP, Watts R, Bjork JM, Garavan HP, Hilmer L, Pung CJ, Sicut CS, Kuperman J, Bartsch H, Xue F, Heitzeg MM, et al. (2019) Image processing and analysis methods for the adolescent brain cognitive development study. *Neuroimage* 202:116091.
- Huang L, Mo L, Li Y (2012) Measuring the interrelations among multiple paradigms of visual attention: an individual differences approach. *J Exp Psychol Hum Percept Perform* 38:414–428.
- Isbell E, Fukuda K, Neville HJ, Vogel EK (2015) Visual working memory continues to develop through adolescence. *Front Psychol* 6:696.
- Johnson MK, McMahon RP, Robinson BM, Harvey AN, Hahn B, Leonard CJ, Luck SJ, Gold JM (2013) The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology* 27:220–229.
- Kane MJ, Engle RW (2002) The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. *Psychon Bull Rev* 9:637–671.
- Klingberg T, Darki F (2015) The role of fronto-parietal and fronto-striatal networks in the development of working memory: a longitudinal study. *Cereb Cortex* 25:1587–1595.
- Klingberg T, Forssberg H, Westerberg H (2002) Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *J Cogn Neurosci* 14:1–10.
- Knutson B, Westdorp A, Kaiser E, Hommer D (2000) fMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage* 12:20–27.
- Kwon H, Reiss AL, Menon V (2002) Neural basis of protracted developmental changes in visuo-spatial working memory. *Proc Natl Acad Sci USA* 99:13336–13341.
- LeWinn KZ, Sheridan MA, Keyes KM, Hamilton A, McLaughlin KA (2017) Sample composition alters associations between age and brain structure. *Nat Commun* 8:874.
- Lin Q, Rosenberg MD, Yoo K, Hsu TW, O'Connell TP, Chun MM (2018) Resting-state functional connectivity predicts cognitive impairment related to Alzheimer's disease. *Front Aging Neurosci* 10:94.
- Logan GD (1994) On the ability to inhibit thought and action: a users' guide to the stop signal paradigm. In: *Inhibitory processes in attention, memory, and language* (Dagenbach D, Carr TH, eds), pp 189–239. San Diego: Academic Press.
- Luciana M, Conklin HM, Hooper CJ, Yarger RS (2005) The development of nonverbal working memory and executive control processes in adolescents. *Child Dev* 76:697–712.
- Luciana M, Bjork JM, Nagel BJ, Barch DM, Gonzalez R, Nixon SJ, Banich MT (2018) Adolescent neurocognitive development and impacts of substance use: overview of the Adolescent Brain Cognitive Development (ABCD) baseline neurocognition battery. *Dev Cogn Neurosci* 32:67–79.
- McNab F, Zeidman P, Rutledge RB, Smittenaar P, Brown HR, Adams RA, Dolan RJ (2015) Age-related changes in working memory and the ability to ignore distraction. *Proc Natl Acad Sci USA* 112:6515–6518.
- Nouwens S, Groen MA, Verhoeven L (2017) How working memory relates to children's reading comprehension: the importance of domain-specificity in storage and processing. *Read Writ* 30:105–120.
- Nyberg L, Andersson M, Kauppi K, Lundquist A, Persson J, Pudas S, Nilsson LG (2014) Age-related and genetic modulation of frontal cortex efficiency. *J Cogn Neurosci* 26:746–754.
- Oba S, Sato MA, Takemasa I, Monden M, Matsubara KI, Ishii S (2003) A Bayesian missing value estimation method for gene expression profile data. *Bioinformatics* 19:2088–2096.
- Osaka M, Osaka N, Kondo H, Morishita M, Fukuyama H, Aso T, Shibasaki H (2003) The neural basis of individual differences in working memory capacity: an fMRI study. *Neuroimage* 18:789–797.
- Pailian H, Libertus ME, Feigenson L, Halberda J (2016) Visual working memory capacity increases between ages 3 and 8 years, controlling for gains in attention, perception, and executive control. *Atten Percept Psychophys* 78:1556–1573.
- Palva JM, Monto S, Kulashakar S, Palva S (2010) Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proc Natl Acad Sci USA* 107:7580–7585.
- Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE (2011) Functional network organization of the human brain. *Neuron* 72:665–678.
- Ptak R (2012) The frontoparietal attention network of the human brain: action, saliency, and a priority map of the environment. *Neuroscientist* 18:502–515.
- Reynolds GD, Romano AC (2016) The development of attention systems and working memory in infancy. *Front Syst Neurosci* 10:15.
- Rosenberg MD, Casey BJ, Holmes AJ (2018) Prediction complements explanation in understanding the developing brain. *Nat Commun* 9:589.
- Rosenberg MD, Finn ES, Scheinost D, Constable RT, Chun MM (2017) Characterizing attention with predictive network models. *Trends Cogn Sci* 21:290–302.
- Ross-Sheehy S, Oakes LM, Luck SJ (2003) The development of visual short-term memory capacity in infants. *Child Dev* 74:1807–1822.
- Ross RG, Wagner B, Heinlein S, Zerbe GO (2008) The stability of inhibitory and working memory deficits in children and adolescents who are children of parents with schizophrenia. *Schizophr Bull* 34:47–51.
- Rudolph MD, Graham AM, Feczko E, Miranda-Dominguez O, Rasmussen J, Nardos R, Entringer S, Wadhwa PD, Buss C, Fair DA (2018) Maternal IL-6 during pregnancy can be estimated from newborn brain connectivity and predicts future working memory in offspring. *Nat Neurosci* 21:765–772.
- Salthouse TA, Babcock RL (1991) Decomposing adult age differences in working memory. *Dev Psychol* 27:763–776.
- Salthouse TA, Babcock RL, Shaw RJ (1991) Effects of adult age on structural and operational capacities in working memory. *Psychol Aging* 6:118–127.
- Satterthwaite TD, Wolf DH, Erus G, Ruparel K, Elliott MA, Gennatas ED, Hopson R, Jackson C, Prabhakaran K, Bilker WB, Calkins ME, Loughhead J, Smith A, Roalf DR, Hakonarson H, Verma R, Davatzikos C, Gur RC, Gur RE (2013) Functional maturation of the executive system during adolescence. *J Neurosci* 33:16249–16261.
- Scolari M, Seidl-Rathkopf KN, Kastner S (2015) Functions of the human frontoparietal attention network: evidence from neuroimaging. *Curr Opin Behav Sci* 1:32–39.
- Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, Petersen SE (2014) Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Hum Brain Mapp* 35:1981–1996.
- Somerville LH, Bookheimer SY, Buckner RL, Burgess GC, Curtiss SW, Dapretto M, Elam JS, Gaffrey MS, Harms MP, Hodge C, Kandala S, Kastman EK, Nichols TE, Schlaggar BL, Smith SM, Thomas KM, Yacoub E, Van Essen DC, Barch DM (2018) The Lifespan Human Connectome Project in Development: a large-scale study of brain connectivity development in 5–21 year olds. *Neuroimage* 183:456–468.
- Spreng RN, Stevens WD, Chamberlain JP, Gilmore AW, Schacter DL (2010) Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage* 53:303–317.
- Spreng RN, Sepulcre J, Turner GR, Stevens WD, Schacter DL (2013) Intrinsic architecture underlying the relations among the default, dorsal attention, and fronto-parietal control networks of the human brain. *J Cogn Neurosci* 25:74–86.
- Steinberg L, Cauffman E, Woolard J, Graham S, Banich M (2009) Are adolescents less mature than adults?: minors' access to abortion, the juvenile death penalty, and the alleged APA “flip-flop.” *Am Psychol* 64:583–594.
- Swanson HL (2017) Verbal and visual-spatial working memory: what develops over a life span? *Dev Psychol* 53:971–995.
- Takeuchi H, Taki Y, Sassa Y, Hashizume H, Sekiguchi A, Fukushima A, Kawashima R (2011) Verbal working memory performance correlates with regional white matter structures in the frontoparietal regions. *Neuropsychologia* 49:3466–3473.
- Tammes CK, Walhovd KB, Grydeland H, Holland D, Østby Y, Dale AM, Fjell AM (2013) Longitudinal working memory development is related to structural maturation of frontal and parietal cortices. *J Cogn Neurosci* 25:1611–1623.
- Thompson WK, Barch DM, Bjork JM, Gonzalez R, Nagel BJ, Nixon SJ, Luciana M (2019) The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: findings from the ABCD Study's baseline neurocognitive battery. *Dev Cogn Neurosci* 36:100606.
- Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Marcus DJ, Westerlund A, Casey BJ, Nelson C (2009) The NimStim set of facial

- expressions: judgments from untrained research participants. *Psychiatry Res* 168:242–249.
- Tulsky DS, Carlozzi N, Chiaravalloti ND, Beaumont JL, Kisala PA, Mungas D, Conway K, Gershon R (2014) NIH Toolbox Cognition Battery (NIHTB-CB): list sorting test to measure working memory. *J Int Neuropsychol Soc* 20:599–610.
- Ullman H, Almeida R, Klingberg T (2014) Structural maturation and brain activity predict future working memory capacity during childhood development. *J Neurosci* 34:1592–1598.
- Van Den Burg W, Kingma A (1999) Performance of 225 Dutch school children on Rey's Auditory Verbal Learning Test (AVLT): parallel test-retest reliabilities with an interval of 3 months and normative data. *Arch Clin Neuropsychol* 14:545–559.
- Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL (2008) Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J Neurophysiol* 100:3328–3342.
- Volkow ND, Koob GF, Croyle RT, Bianchi DW, Gordon JA, Koroshetz WJ, Pérez-Stable EJ, Riley WT, Bloch MH, Conway K, Deeds BG, Dowling GJ, Grant S, Howlett KD, Matochik JA, Morgan GD, Murray MM, Noronha A, Spong CY, Wargo EM, et al. (2018) The conception of the ABCD Study: from substance use to a broad NIH collaboration. *Dev Cogn Neurosci* 32:4–7.
- Wechsler D (2014) Wechsler intelligence scale for children: technical and interpretive manual, Ed 5. Bloomington: Pearson's Clinical Assessment.
- Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE (2014) Permutation inference for the general linear model. *Neuroimage* 92:381–397.
- Winkler AM, Webster MA, Vidaurre D, Nichols TE, Smith SM (2015) Multi-level block permutation. *Neuroimage* 123:253–268.
- Winkler AM, Ridgway GR, Douaud G, Nichols TE, Smith SM (2016) Faster permutation inference in brain imaging. *Neuroimage* 141:502–516.
- Woolgar A, Hampshire A, Thompson R, Duncan J (2011) Adaptive coding of task-relevant information in human frontoparietal cortex. *J Neurosci* 31:14592–14599.
- Wulfert E, Block JA, Santa Ana E, Rodriguez ML, Colsman M (2002) Delay of gratification: impulsive choices and problem behaviors in early and late adolescence. *J Pers* 70:533–552.
- Xu Z, Adam KCS, Fang X, Vogel EK (2018) The reliability and stability of visual working memory capacity. *Behav Res Methods* 50:576–588.
- Yamashita M, Yoshihara Y, Hashimoto R, Yahata N, Ichikawa N, Sakai Y, Yamada T, Matsukawa N, Okada G, Tanaka SC, Kasai K, Kato N, Okamoto Y, Seymour B, Takahashi H, Kawato M, Imamizu H (2018) A prediction model of working memory across health and psychiatric disease using whole-brain functional connectivity. *Elife* 7.
- Yau WY, Zubieta JK, Weiland BJ, Samudra PG, Zucker RA, Heitzeg MM (2012) Nucleus accumbens response to incentive stimuli anticipation in children of alcoholics: relationships with precursive behavioral risk and lifetime alcohol use. *J Neurosci* 32:2544–2551.
- Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125–1165.
- Yoo K, Rosenberg MD, Hsu W-T, Zhang S, Li C-S, Scheinost D, Constable RT, Chun MM (2018) Connectome-based predictive modeling of attention: comparing different functional connectivity features and prediction methods across datasets. *Neuroimage* 167:11–22.
- Zelazo PD (2006) The dimensional change card sort (DCCS): a method of assessing executive function in children. *Nat Protoc* 1:297–301.
- Zelazo PD, Anderson JE, Richler J, Wallner-Allen K, Beaumont JL, Conway KP, Gershon R, Weintraub S (2014) NIH toolbox cognition battery (CB): validation of executive function measures in adults. *J Int Neuropsychol Soc* 20:620–629.