

Within-Individual Variability in Neurocognitive Performance: Age- and Sex-Related Differences in Children and Youths From Ages 8 to 21

David R. Roalf, Raquel E. Gur, Kosha Ruparel,
Monica E. Calkins, Theodore D. Satterthwaite,
and Warren B. Bilker
University of Pennsylvania

Hakon Hakonarson
Children's Hospital of Philadelphia, Philadelphia, Pennsylvania

Lauren Julius Harris
Michigan State University

Ruben C. Gur
University of Pennsylvania and Philadelphia Veterans
Administration Medical Center, Philadelphia, Pennsylvania

Objective: The transition from childhood to adulthood is characterized by improved motor and cognitive performance in many domains. Developmental studies focus on average performance in single domains but ignore consistency of performance across domains. Within-individual variability (WIV) provides an index of that evenness and is a potential marker of development. **Method:** We gave a computerized battery of 14 neurocognitive tests to 9138 youths ages 8–21 from the Philadelphia Neurodevelopmental Cohort. **Results:** As expected, performance improved with age, with both accuracy and speed peaking in adulthood. WIV, however, showed a U-shaped course: highest in childhood, declining yearly into mid-adolescence, and increasing again into adulthood. Young females outperformed and were less variable than males, but by early adulthood male performance matched that of females despite being more variable. **Conclusion:** We conclude that WIV declines from childhood to adolescence as developmental lags are overcome, and then increases into adulthood reflecting the emergence of cognitive specializations related to skill-honing and brain maturation.

Keywords: age differences, cognitive neuroscience, development, sex differences, within-individual variability

Supplemental materials: <http://dx.doi.org/10.1037/neu0000067.supp>

The transition from childhood to adulthood is characterized by improvement in motor (Gallahue, Ozmun, & Goodway, 1998), cognitive (Fischer, 1980), and social domains (Blakemore & Choudhury, 2006). Developmental studies of neurocognition, that is, of performance on tasks linked to brain systems, have focused on single domains, including executive-control (e.g., (Goldberg, Maurer, & Lewis, 2001)), memory (e.g., (Dempster, 1985; Luciana,

Conklin, Hooper, & Yarger, 2005)), language and reasoning (e.g., (Friederici & Wartenburger, 2010)), and more recently, social cognition (e.g., (Burnett, Sebastian, Cohen Kadosh, & Blakemore, 2011; Shaw, Grosbras, Leonard, Pike, & Paus, 2012)).

Studies of performance in single domains are useful but are insensitive to potentially important phenomena related to the relative and uneven pace of development of different abilities (In-

This article was published Online First April 28, 2014.

David R. Roalf, Department of Psychiatry, Neuropsychiatry Section, University of Pennsylvania; Raquel E. Gur, Department of Psychiatry, Neuropsychiatry Section, and Department of Radiology, University of Pennsylvania; Kosha Ruparel, Monica E. Calkins, and Theodore D. Satterthwaite, Department of Psychiatry, Neuropsychiatry Section, University of Pennsylvania; Warren B. Bilker, Department of Biostatistics and Epidemiology, University of Pennsylvania; Hakon Hakonarson, Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; Lauren Julius Harris, Department of Psychology, Michigan State University; Ruben C. Gur, Department of Psychiatry, Neuropsychiatry Section, and Department of Radiology, University of Pennsylvania, and Philadelphia Veterans Administration Medical Center, Philadelphia, Pennsylvania.

We thank Rosetta Chiavacci of Children's Hospital of Philadelphia (CHOP), the recruitment, assessment, and CNB teams, and Larry Macy, Chad T. Jackson, and Allison Mott for systems support.

The Philadelphia Neurodevelopmental Cohort (PNC) is a collaborative project between the Center for Applied Genomics (CAG) at CHOP and the Brain Behavior Laboratory at the University of Pennsylvania. At the time of submission, the sample included 9,138 children and youths, ages 8–21 years, all from the greater Philadelphia area. This study was supported by grants from the National Institute of Mental Health (NIMH) RC2MH089983 (R.E.G.) and RC2MH089924 (H.H.). Dr. Roalf was supported by NIMH T32 MH019112. Dr. Satterthwaite was supported by NIMH K23MH098130 and the Marc Rapport Family Investigator Grant through the Brain & Behavior Research Foundation.

Correspondence concerning this article should be addressed to David R. Roalf, Department of Psychiatry, Neuropsychiatry Section, University of Pennsylvania Perelman School of Medicine, 3400 Spruce St., Gates Building 10th Floor, Philadelphia PA, 19104. E-mail: roalf@upenn.edu

helder & Piaget, 1955). Notably, cognitive systems mature at different rates, and abilities develop unevenly within individuals (e.g., Luna, Garver, Urban, Lazar, & Sweeney, 2004). The study of within-individual variability (WIV) in performance, that is, the variability in performance on different domains within a single individual, can assess individual age-related changes in these phenomena provided that the same measures are available across a wide range of domains and across a developmentally significant span of time. Assessing variability therefore could provide important additional information about typical development (e.g., Van Geert & Van Dijk, 2002). It also could help identify individuals at risk for certain brain-behavior disorders, because variability is higher in focal disorders such as infarction, hemorrhage, trauma, and tumor (Stuss, Murphy, Binns, & Alexander, 2003) and in developmental disorders such as attention deficit–hyperactivity disorder (Leth-Steensen, Elbaz, & Douglas, 2000) and schizophrenia (Roalf, Gur et al., 2013). Finally, in providing additional information about typical development, the study of WIV across domains could offer important new information about individual differences.

Individual differences in performance include sex differences in specific behavioral domains (e.g., R. C. Gur et al., 1995; Halpern et al., 2007). Males outperform females on certain spatial and motor tests (e.g., Harris, 1985; Moreno-Briseño, Diaz, Campos-Romo, & Fernandez-Ruiz, 2010; Thomas & French, 1985), whereas females do better on certain verbal and memory tests (e.g., R. C. Gur et al., 2012; Herlitz, Nilsson, & Backman, 1997). There also is evidence that WIV is higher in males than females (e.g., Deary, Thorpe, Wilson, Starr, & Whalley, 2003; Shea, Lubinski, & Benbow, 2001). If high WIV is an indication of better performance, or “specialization,” in some domains relative to others, then higher WIV in males could be an indication that more males are “cognitive specialists” whereas more females are “cognitive generalists,” performing well across a broad range of functions, or domains. Measuring WIV during development therefore could be another, more sensitive way to document the emergence of individual differences in fundamental patterns of cognitive specialization. Thus, if part of development involves reaching levels of performance that are more nearly equal or even across domains, we would expect WIV to decline with age. However, to our knowledge, WIV has not been examined in large samples with sufficiently broad batteries of tests across a wide age range to permit an evaluation of its development and relevance for overall performance.

It is helpful to distinguish WIV from two other measures of variability that have been examined in earlier studies. One refers to within-individual consistency of performance across time; the other examines sample variance on single measures (e.g., children often have high performance variability even when solving the same problem over short or long intervals; Siegler & McGilly, 1989). Dykiert, Der, Starr, and Deary (2012) found that variability on one test decreased from childhood through adolescence, followed by an increase throughout adulthood and into old age. Sample variance has also been studied in different populations, motivated by Darwin’s observation that evolution is associated with high variability within species (Darwin, 1871). Thorndike (1910) proposed that males have greater variability in performance and that this was why they have reached greater intellectual prominence: “if men differ in intelligence and energy by wider

extremes than do women, eminence in and leadership of the world’s affairs of whatever sort will inevitably belong oftener to men. They will oftener deserve it” (p. 10). Although arguments against this conclusion had already been raised by scholars in Thorndike’s own time (e.g., Hollingworth, 1914), the debate still rages; the empirical support for the higher variability in males seems to differ according to the sample studied and the specific measures used (Halpern et al., 2007).

In two prior studies, we used a Computerized Neurocognitive Battery (CNB), a 1-hr-long battery of neuroscience-based tests that have been validated with neuroimaging studies as recruiting specific brain systems (R. C. Gur et al., 2001; R. C. Gur et al., 2010). The CNB measures performance in 5 domains: executive-control, episodic memory, complex cognitive processing, social cognition, and sensorimotor speed. The battery is designed for large-scale studies and enables efficient data collection and separate scoring for accuracy and speed. The range of domains permits reliable estimates of across-Test WIV. In a prior report (R. C. Gur et al., 2012), we examined global neurocognitive performance (GNP), that is, average performance across domains, for the first 3,500 participants from a sample of more than 9,000 children and youth, ages 8 to 21. We found substantial improvement across the age range in both accuracy and speed with some domain variability in effect size. We also found significant sex differences, although effect sizes were small and moderated by earlier maturation in females. We did not, however, examine WIV, so its developmental course in this cohort is unknown. In the current study, we examined data, including WIV, from the complete sample and tested the following hypotheses: a) Average neurocognitive performance will improve with age; b) Females will reach mature performance (e.g., higher GNP) earlier than males; c) Over the 8- to 21-year span WIV will decrease, showing increasing evenness of performance associated with maturation. Before testing these hypotheses, we examined whether we could confirm in this larger sample the sex difference found in a previously published study: better performance by females than males on tests of attention, memory, and social cognition; better performance by males than females on tests of spatial processing and motor speed. We also tested Thorndike’s hypothesis of greater variance among males, although our single-time measure did not permit evaluation of variability in performance within individuals over time.

Method

Participants

The sample included children and young adults (ages 8–21) recruited through a NIMH-funded Grand Opportunity (GO) study characterizing clinical and neurobehavioral phenotypes in a prospectively acquired community cohort. Participants previously consented for genomic studies when they came for pediatric services at the Children’s Hospital. At that time they provided a blood sample for genetic studies, authorized access to Electronic Medical Records (EMRs), and gave written informed consent/assent to be recontacted for future studies. Of the 50,540 genotyped participants, 18,344 met criteria for inclusion in the current study and, from this group, the participants in the current study were randomly selected, with stratification for age, sex and ethnicity.

Table 1
Sample Demographics From the Philadelphia Neurodevelopmental Cohort

Demographic	Overall (<i>n</i> = 9010)	Females (<i>n</i> = 4685)	Males (<i>n</i> = 4325)
Age	13.78 (3.66) [8–21]	14.10 (3.64) [8–21]	13.42 (3.65) [8–21]
Mother education	14.48 (2.46) [0–20]	14.38 (2.46) [0–20]	14.59 (2.44) [0–20]
Father Education	14.27 (2.74) [0–20]	14.16 (2.77) [0–20]	14.38 (2.70) [0–20]
WRAT4 Standard score	102.41 (15.99) [58–145]	101.71 (15.62) [58–145]	103.17 (16.35) [59–145]
Race (<i>n</i>)			
African American	2859	1572	1287
American Indian	14	6	8
Asian	89	53	36
Caucasian	5113	2537	2576
Hawaiian/Pacific Islander	9	6	3
More than one	926	511	415

Note. Means and standard deviations for demographic variables. Minimum and maximum values are included in brackets.

Overall, 9,138 individuals enrolled in the study and are included in the current analysis. All recruitment and study procedures were completed within a 2-year period from November 2009 to November 2011. The sample included ambulatory children in stable health, proficient in English, physically and cognitively capable of participating in an interview and performing the computerized

neurocognitive testing. Youths with disorders that markedly impaired motility or cognition (e.g., significant paresis or palsy, intellectual disability) were excluded. Notably, participants were *not* recruited from psychiatric clinics, and the sample did not include individuals who sought psychiatric help. Participants were not necessarily receiving ongoing medical care at the Children's Hospital but had received care at least once from this hospital and provided a genetic sample. Information about ongoing medical care is available but utilization of health services was not analyzed for this study. Participants and their parents provided informed consent (or assent) and agreed to be recontacted for participation in further studies.

Overall Inclusion Criteria

Several criteria had to be met for inclusion in the study. Participants ages 18 and over had to be able to provide signed informed consent. For participants under age 18, assent and parental consent were required. All participants had to be proficient in English and physically and cognitively able to participate in computerized neurocognitive testing. In addition to the neurocognitive tests (see Procedures and Measures), participants were questioned about general medical conditions, and electronic medical records were reviewed. Proband (age 11–21) and collaterals (of probands age 8–17) also received a structured interview to screen for major domains of psychopathology, which was adapted from the Kiddie-SADS (Kaufman et al., 1997).

Table 2
Standardized neurocognitive performance accuracy by age and sex

Gender	Age	<i>n</i>	ABF_A_z		<i>n</i>	ATT_A_z		<i>n</i>	WM_A_z		<i>n</i>	VMEM_A_z		<i>n</i>	FMEM_A_z		<i>n</i>
			Mean	SD		Mean	SD		Mean	SD		Mean	SD		Mean	SD	
Males	8	444	−0.56	1.11	425	−0.85	1.21	420	−0.83	1.37	440	−0.54	1.39	438	−0.74	0.90	437
	9	361	−0.30	1.05	351	−0.73	1.20	347	−0.62	1.32	359	−0.22	1.08	361	−0.53	0.87	362
	10	395	−0.13	1.02	383	−0.54	1.16	377	−0.26	1.06	394	−0.11	1.10	390	−0.34	0.90	394
	11	333	−0.08	1.01	321	−0.33	1.05	323	−0.16	1.03	333	−0.09	0.96	332	−0.26	0.92	332
	12	335	0.12	0.99	331	−0.15	0.93	328	−0.04	0.90	336	−0.13	1.08	333	−0.09	0.92	331
	13	334	0.06	0.96	328	−0.10	1.02	326	0.08	0.96	334	0.03	0.92	333	0.00	0.94	330
	14	342	0.20	0.89	341	0.07	0.83	337	0.17	0.77	343	−0.02	1.00	342	0.10	0.93	342
	15	368	0.15	0.92	368	0.18	0.88	363	0.15	0.77	366	0.09	0.88	367	0.14	0.94	368
	16	352	0.21	0.96	349	0.27	0.76	348	0.29	0.78	353	0.02	0.93	352	0.23	0.97	353
	17	303	0.28	0.93	303	0.33	0.75	302	0.33	0.71	305	0.10	0.83	304	0.29	0.93	304
	18	332	0.27	0.88	330	0.42	0.73	327	0.33	0.67	333	0.06	0.93	332	0.25	0.90	330
Females	19	223	0.29	0.82	222	0.40	0.76	216	0.33	0.82	223	0.07	1.02	222	0.39	0.96	221
	20	127	0.43	0.87	128	0.55	0.63	124	0.41	0.51	128	0.09	0.88	128	0.20	0.94	128
	21	57	0.23	0.94	57	0.59	0.50	57	0.39	0.78	57	0.05	1.00	56	0.21	0.93	57
	8	330	−0.64	1.08	324	−0.66	1.13	323	−0.95	1.38	334	−0.37	1.33	330	−0.64	0.91	331
	9	339	−0.47	1.09	336	−0.52	1.10	332	−0.60	1.21	340	−0.05	1.03	337	−0.42	0.87	338
	10	335	−0.23	0.93	331	−0.29	0.97	327	−0.36	1.13	336	0.00	1.01	333	−0.22	0.93	335
	11	308	−0.05	0.99	305	−0.18	0.98	296	−0.14	0.95	309	0.04	1.02	308	−0.09	0.94	306
	12	358	−0.08	0.97	355	−0.12	0.93	349	−0.01	0.85	359	0.06	0.88	358	−0.04	1.01	356
	13	360	0.13	0.96	357	0.03	0.96	353	0.10	0.77	361	0.15	0.82	359	0.24	0.99	359
	14	378	0.07	0.89	378	0.15	0.84	373	0.23	0.72	379	0.19	0.76	378	0.13	0.96	379
	15	379	0.06	0.92	377	0.28	0.79	376	0.26	0.69	381	0.06	0.91	379	0.17	0.98	376
16	418	0.05	0.91	416	0.39	0.58	413	0.29	0.73	418	0.09	0.92	418	0.28	0.97	417	
17	393	0.18	0.94	393	0.40	0.63	386	0.30	0.75	391	0.18	0.76	392	0.28	1.02	394	
18	489	0.13	0.92	489	0.49	0.63	482	0.30	0.75	491	0.20	0.89	490	0.33	0.98	488	
19	339	0.25	0.94	339	0.56	0.60	336	0.44	0.63	339	0.18	0.80	337	0.36	0.99	338	
20	159	0.15	0.98	158	0.54	0.57	157	0.37	0.58	160	0.26	0.83	160	0.37	0.95	158	
21	75	0.23	0.97	76	0.57	0.52	75	0.37	0.66	76	0.10	1.00	75	0.34	0.98	75	

Procedures and Measures

The CNB was administered using a Web-based system (see R. C. Gur et al., 2010). The battery consists of 14 tests in 5 domains:

Executive-Control

Abstraction & mental flexibility. The Conditional Exclusion Test presents four objects at a time, and the participant selects the object that does not belong with the rest based on one of three sorting principles. Sorting principles change, and feedback guides their identification (R. C. Gur et al., 2010; Kurtz, Ragland, Moberg, & Gur, 2004).

Attention. The Continuous Performance Test presents 7-segment displays at a rate of 1/sec. The participant must press the space bar whenever the display forms a digit or a letter (R. C. Gur et al., 2001; R. C. Gur et al., 2010; Kurtz, Moberg, Gur, & Gur, 2001).

Working memory. The Letter N-back test displays sequences of uppercase letters. In the 0-back condition, participants respond to the letter X. In the 1-back condition they respond if the letter is identical to that preceding it. In the 2-back condition, they respond if the letter is identical to that presented two trials back (R. C. Gur et al., 2001; R. C. Gur et al., 2010; Ragland et al., 2002).

Episodic Memory

Verbal. The Word Memory Test presents 20 target words that are then mixed with 20 foils equated for frequency, length, concreteness, and imageability. Participants are asked to memorize the target words as they are presented and then to indicate whether a word

presented was included in the target list (Glahn, Gur, Ragland, Censits, & Gur, 1997; R. C. Gur et al., 2001; R. C. Gur et al., 2010).

Facial. The Face Memory Test presents 20 digitized faces subsequently intermixed with 20 foils equated for age, sex, and ethnicity. Participants indicate whether or not they recognize each face (Glahn et al., 1997; R. C. Gur et al., 2001; R. C. Gur et al., 2010).

Spatial. The Visual Object Learning Test uses Euclidean shapes as stimuli with the same paradigm as the word and face memory tests (Glahn et al., 1997; R. C. Gur et al., 2001; R. C. Gur et al., 2010).

Complex Cognition

Language reasoning. The Verbal Reasoning Test consists of verbal analogy problems with simplified instructions and vocabulary (R. C. Gur et al., 1982; R. C. Gur et al., 2001; R. C. Gur et al., 2010).

Nonverbal reasoning. The Matrix Reasoning Test consists of matrices requiring reasoning by geometric analogy and contrast principles (R. C. Gur et al., 2010).

Spatial ability. The Line Orientation Test presents two lines at an angle, and participants click on a button to make one line rotate until it has the same angle as the other. The relative location of the lines and their sizes differ across trials (R. C. Gur et al., 2010).

Social Cognition

Emotion identification. The Emotion Identification Test displays 5 faces, one of which shows an emotionally neutral expres-

SMEM_A_z			LAN_A_z			NVR_A_z			SPA_A_z			EID_A_z			EDI_A_z			ADI_A_z		
Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
-0.17	1.06	424	-1.16	0.94	444	-0.78	0.71	413	-0.73	0.77	445	-0.57	1.22	442	-1.12	1.10	432	-1.15	1.01	
-0.03	1.03	362	-0.78	0.97	364	-0.55	0.79	334	-0.49	0.87	365	-0.43	1.02	360	-0.75	0.98	359	-0.87	0.92	
-0.07	0.99	395	-0.47	0.94	393	-0.37	0.74	372	-0.27	0.90	395	-0.30	0.93	393	-0.43	0.99	392	-0.52	0.99	
-0.02	0.96	332	-0.24	0.87	333	-0.31	0.77	324	-0.21	0.86	333	-0.18	1.05	334	-0.30	1.00	334	-0.39	0.95	
0.13	0.87	335	-0.02	0.94	335	-0.10	0.90	322	0.01	0.87	336	-0.09	0.95	336	-0.16	0.94	332	-0.22	0.94	
0.02	0.91	334	0.10	0.94	333	-0.03	0.86	318	0.16	0.90	334	-0.01	0.90	333	0.03	0.98	330	0.03	0.92	
0.08	0.97	344	0.29	0.83	343	0.12	0.96	335	0.27	0.91	344	0.15	0.88	343	0.19	0.80	341	0.10	0.86	
0.02	1.06	368	0.33	0.88	368	0.24	1.01	359	0.38	0.96	368	0.11	0.91	366	0.19	0.86	368	0.16	0.84	
0.17	0.96	353	0.45	0.80	352	0.41	1.04	343	0.51	0.95	353	0.16	0.89	353	0.32	0.81	352	0.35	0.81	
0.17	0.96	304	0.52	0.81	304	0.44	1.05	295	0.55	0.90	305	0.24	0.89	304	0.37	0.85	303	0.33	0.78	
0.01	1.09	332	0.45	0.82	331	0.52	1.03	330	0.58	0.99	332	0.18	0.75	332	0.33	0.80	330	0.45	0.84	
0.15	1.02	222	0.59	0.71	222	0.80	1.05	218	0.69	1.05	223	0.21	0.85	223	0.37	0.77	222	0.50	0.83	
0.02	1.05	128	0.59	0.71	128	0.77	1.00	126	0.70	1.06	128	0.19	0.82	128	0.47	0.62	128	0.64	0.68	
0.12	0.99	57	0.69	0.77	56	0.62	1.08	57	0.73	0.96	57	0.11	1.15	57	0.35	0.97	57	0.43	0.97	
-0.43	1.00	331	-1.10	0.89	334	-0.81	0.70	293	-0.91	0.69	332	-0.45	1.03	331	-0.80	1.03	331	-0.86	0.96	
-0.20	0.94	340	-0.80	0.91	339	-0.55	0.79	292	-0.77	0.76	342	-0.41	1.20	340	-0.59	1.07	338	-0.66	1.00	
-0.05	1.01	337	-0.46	0.87	336	-0.45	0.78	312	-0.57	0.82	337	-0.26	0.97	336	-0.28	0.93	336	-0.37	0.92	
0.10	0.97	308	-0.25	0.93	309	-0.23	0.82	287	-0.30	0.89	308	-0.18	0.89	308	-0.16	0.94	307	-0.16	0.86	
0.05	0.97	360	-0.05	0.88	359	-0.16	0.86	339	-0.31	0.83	360	0.01	0.91	359	0.01	0.91	358	0.00	0.88	
0.00	1.01	360	0.17	0.80	360	-0.05	0.80	347	-0.14	0.89	360	0.17	0.98	361	0.21	0.89	358	0.20	0.84	
0.05	0.99	379	0.23	0.81	379	0.02	0.90	365	0.00	0.93	379	0.18	0.86	379	0.32	0.80	377	0.28	0.81	
0.08	0.89	380	0.31	0.86	380	0.04	0.90	367	0.05	0.93	382	0.34	0.88	381	0.29	0.83	378	0.30	0.83	
0.08	1.01	418	0.33	0.80	417	0.14	0.92	407	0.13	0.85	418	0.29	0.85	418	0.38	0.76	415	0.41	0.75	
0.04	1.07	394	0.42	0.77	393	0.27	1.01	384	0.16	0.88	394	0.32	0.89	393	0.40	0.79	386	0.43	0.76	
-0.07	1.06	490	0.35	0.84	490	0.44	1.01	476	0.30	0.99	491	0.29	0.81	491	0.45	0.77	489	0.53	0.77	
0.08	0.93	341	0.41	0.79	339	0.61	1.00	336	0.32	1.03	341	0.24	0.88	341	0.42	0.75	338	0.60	0.69	
-0.16	0.97	160	0.34	0.79	160	0.48	1.00	157	0.30	0.98	159	0.34	0.78	160	0.36	0.73	160	0.53	0.78	
-0.17	0.95	76	0.43	0.82	75	0.42	1.02	75	0.37	1.08	76	0.41	0.86	76	0.29	0.84	75	0.51	0.74	

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

sion, whereas 4 express the emotions happy, sad, anger, and fear. The faces are presented one at a time, and the participant is asked to identify the emotion displayed. The faces are balanced for sex, age, and ethnicity (R. C. Gur et al., 2002; R. E. Gur et al., 2006; Mathersul, Palmer, Gur, Gur, & Cooper, 2009).

Emotion differentiation. The Emotion Differentiation Test presents pairs of faces showing emotional expressions, each pair obtained from the same individual expressing the same emotion, one more intense than the other or of equal intensity.

Age differentiation. The Age Differentiation Test asks the participant whether two faces of two different persons are of the same different age, and if different, to select the face that appears older (R. C. Gur et al., 2010).

Sensorimotor Speed

Sensorimotor processing speed. The Motor Praxis task requires moving the mouse and clicking on a green square that disappears after the click. With each click, the square gets smaller and changes from one of a number of locations to another location in a random fashion (R. C. Gur et al., 2001; R. C. Gur et al., 2010).

Simple motor speed. The Finger Tapping Test measures how quickly the participant can press the spacebar using only the index finger (R. C. Gur et al., 2001; R. C. Gur et al., 2010).

Each test provides measures of accuracy and speed, except for sensorimotor tests, which measured only speed. A brief standardized reading test from the Wide Range Achievement Test

(WRAT4; (Wilkinson & Robertson, 2006) was administered to determine participants’ ability to complete the battery and to provide an estimate of IQ.

Participants were given the option of being tested at home or in the laboratory. The option of home testing was given to increase representativeness of the sample by minimizing recruitment obstacles (e.g., transportation, distance, time, scheduling) and to accommodate the large number of participants: as many as 165 participants enrolled in the study per week. Approximately 70% preferred home assessment. To ensure uniformity in procedure, a trained tester sat at the table next to the participant and read aloud the instructions that appeared on the screen. At the time of appointment scheduling, recruiters gauged the appropriateness of a home appointment by inquiring about the availability of a quiet, private space in the home that could be set up for testing. Specifically, they requested a quiet space without distractions, where others (including parents) could not overhear, and a flat surface, like a table or a desk, to set up the laptop, with two chairs that could be pulled up to work at the laptop. If such a space was not available, an in-office appointment was scheduled. Moreover, if assessors arrived at the home and the testing situation could not accommodate the above conditions, the appointment was rescheduled for an in-office appointment. Tests were administered in a fixed order that, based on prior experience, assured novelty to preserve participant interest (R. C. Gur et al., 2012). Brief breaks were offered between tests. Overall, 47 individuals had invalid

Table 3
Standardized neurocognitive performance speed by age and sex

	Age	ABF_S_z		ATT_S_z		WM_S_z		VMEM_S_z		FMEM_S_z		SMEM_S_z		LAN_S_z		n
		n	Mean SD	n	Mean SD	n	Mean SD	n	Mean SD	n	Mean SD	n	Mean SD	n	Mean SD	
Males	8	444	-0.27 1.02	425	-1.13 1.27	419	-0.71 1.25	440	-1.19 1.29	438	-0.67 1.17	437	-0.46 1.09	423	-1.00 1.55	442
	9	361	-0.11 0.89	351	-0.68 1.16	345	-0.52 1.21	359	-0.74 1.07	361	-0.38 0.93	362	-0.35 0.94	362	-0.68 1.26	364
	10	395	0.03 0.84	383	-0.42 0.92	377	-0.25 1.05	394	-0.38 0.95	390	-0.11 0.87	394	-0.10 0.89	395	-0.25 1.06	392
	11	333	0.06 0.79	321	-0.09 0.82	323	-0.03 0.94	333	-0.14 0.79	332	-0.05 0.99	332	-0.02 0.92	332	-0.08 0.94	332
	12	335	0.03 1.14	331	0.05 0.67	328	0.12 0.85	336	-0.05 0.87	333	-0.01 0.93	331	0.03 0.87	333	-0.01 0.83	334
	13	334	0.12 1.14	328	0.06 0.79	326	0.12 0.84	334	0.09 0.88	333	0.11 1.00	330	0.11 0.96	334	0.18 0.73	332
	14	342	0.14 0.90	341	0.29 0.68	337	0.20 0.88	343	0.22 0.74	342	0.19 0.89	342	0.14 0.95	344	0.15 0.90	343
	15	368	0.05 0.98	368	0.39 0.70	363	0.29 0.86	366	0.20 0.95	367	0.07 1.10	368	0.12 1.06	368	0.14 0.92	368
	16	352	0.18 1.06	349	0.54 0.62	348	0.45 0.71	353	0.37 0.76	352	0.25 0.88	353	0.28 0.89	353	0.30 0.71	352
	17	303	-0.01 1.14	303	0.59 0.61	302	0.52 0.65	305	0.36 0.78	304	0.18 0.95	304	0.17 1.04	304	0.27 0.67	303
	18	332	0.02 1.09	330	0.65 0.61	327	0.37 0.77	333	0.38 0.74	332	0.18 0.89	330	0.20 1.01	332	0.11 0.92	331
19	223	-0.19 1.32	221	0.50 0.65	216	0.29 0.84	223	0.39 0.77	222	0.14 1.07	221	0.10 1.06	222	0.12 0.80	222	
20	127	-0.16 1.19	128	0.54 0.60	124	0.35 0.71	128	0.34 0.81	128	0.07 0.98	128	0.00 1.09	128	0.07 0.80	128	
21	57	-0.16 1.09	57	0.65 0.67	57	0.37 0.69	57	0.43 0.76	56	0.11 1.00	57	0.02 1.00	57	-0.02 0.96	56	
Females	8	330	-0.35 0.98	324	-1.47 1.05	323	-1.10 1.18	334	-1.08 1.28	330	-0.66 1.03	331	-0.60 1.13	331	-0.88 1.33	330
	9	339	-0.24 0.88	336	-1.04 1.03	331	-0.70 1.14	340	-0.62 0.87	337	-0.41 0.92	338	-0.42 1.02	340	-0.57 1.15	339
	10	335	-0.11 0.80	331	-0.77 0.82	327	-0.50 1.02	336	-0.31 0.85	333	-0.23 1.02	335	-0.14 0.89	337	-0.24 0.89	335
	11	308	-0.03 0.90	305	-0.36 0.91	296	-0.32 1.04	309	-0.14 0.98	308	-0.08 0.84	306	-0.12 0.99	308	-0.09 1.01	308
	12	358	0.04 1.01	355	-0.10 0.74	349	-0.04 0.88	359	0.09 0.68	358	0.05 0.86	356	0.05 0.88	360	0.19 0.70	359
	13	360	0.09 0.91	357	-0.01 0.76	353	0.06 0.84	361	0.17 0.84	359	0.15 0.86	359	0.14 0.93	360	0.26 0.78	359
	14	378	0.09 0.96	377	0.21 0.67	373	0.21 0.85	379	0.26 0.71	378	0.12 0.98	379	0.11 0.92	379	0.30 0.71	379
	15	379	0.15 0.84	377	0.33 0.71	376	0.30 0.71	381	0.32 0.70	379	0.21 0.89	376	0.24 0.84	380	0.39 0.61	380
	16	418	0.13 0.90	416	0.45 0.58	413	0.25 0.82	418	0.39 0.73	418	0.27 0.92	417	0.19 0.89	418	0.36 0.66	417
	17	393	0.07 1.01	393	0.40 0.65	386	0.30 0.81	391	0.36 0.76	392	0.17 1.09	394	0.14 1.06	394	0.36 0.70	393
	18	489	0.05 1.10	489	0.52 0.62	482	0.27 0.77	491	0.40 0.84	490	0.21 0.87	488	0.10 1.01	490	0.29 0.75	490
19	339	0.06 0.99	339	0.53 0.58	336	0.15 0.80	339	0.42 0.73	337	0.20 0.96	338	0.06 1.01	341	0.19 0.87	339	
20	159	-0.10 1.10	158	0.46 0.75	157	0.10 0.89	160	0.49 0.61	160	0.17 1.06	158	0.14 0.89	160	0.12 0.85	160	
21	75	0.04 1.00	76	0.65 0.62	75	0.23 0.75	76	0.36 0.90	75	0.25 0.93	75	0.08 1.06	76	0.27 0.66	75	
Total N		8966		8869		8769		8978		8944		8939		8961		8962

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

data for the CNB. The main reasons for an invalid session were as follows: a) participant disability (e.g., unable to understand directions, vision problems, illiterate ($n = 19$); b) a combination of participant behavior and disability ($n = 15$); c) participant behavior (e.g., lack of effort, preferred to not continue, sleepy; $n = 10$); or d) loss of environmental control ($n = 3$). Individuals ($n = 128$) who did not complete at least half of the CNB (7 tests) and at least one test from each domain were excluded from analysis. The sample demographics for completers ($n = 9010$) are presented in Table 1 and online Supplemental Table S1.

Data Analysis

Accuracy and speed values were transformed to their standard equivalents (z scores) based on means and SDs for the entire sample. A global index of neurocognitive performance (GNP) was calculated by averaging the standardized scores across all CNB tests. An index of within-individual variability (WIV) across all tests was calculated for each participant in SD units (see Holtzer, Verghese, Wang, Hall, & Lipton, 2008; Roalf, Ruparel, et al., 2013). Both GNP and WIV were calculated for accuracy and speed separately. These values were entered as dependent measures in ANCOVAs with age, age-squared, and sex as factors. The age-square factor was used because initial data evaluation using Generalized Additive Models indicated significant nonlinear effects (all $ps < .001$) with quadratic functions for both GNP and WIV. Parental education was measured as the average of mother and father education. Equality of variance between sexes was compared using Levene’s F tests. Post hoc contrasts were used to

interpret interactions; Satterthwaite corrections were used when unequal variances could not be assumed; corrected degrees of freedom are reported where appropriate. Cohen’s d values are presented for age-specific WIV comparisons. All ANCOVAs were corrected for race and average parental education. SAS 9.3 was used for all statistical analyses.

Results

Replicating Sex Differences Across Age

To determine whether the current complete sample replicates sex differences, we followed Gur et al.’s (2012) procedure of calculating z scores for each measure. The z scores were calculated in SAS using the PROC STANDARD routine. Age-adjusted scores were used by employing the “by age” (in whole years) option. The sample size was sufficient to generate reliable z scores for each year. When examining age effects, this option was removed so that the z scores would be based on the entire sample. Performance data by age and sex, and the number of participants completing each task are presented in Tables 2 and 3 and online Supplemental Table S2. Overall, performance was better in older (15+) than younger (under 15) participants. The results replicated previous findings for females of better accuracy in attention and better accuracy and speed for word and face memory and all social cognition tests, and for males better accuracy in spatial cognition and greater simple motor speed (see Supplemental materials for additional notes for Results; see Figure 1).

NVR_S_z			SPA_S_z			EID_S_z			EDI_S_z			ADI_S_z			MOT_S_z			SM_S_z		
Mean	SD	<i>n</i>	Mean	SD	<i>n</i>	Mean	SD	<i>n</i>	Mean	SD	<i>n</i>	Mean	SD	<i>n</i>	Mean	SD	<i>n</i>	Mean	SD	<i>n</i>
0.07	0.60	413	-0.12	1.08	445	-0.98	1.08	442	-0.42	1.04	432	-0.24	0.87	434	-1.18	0.74	445	-0.68	1.00	
0.09	0.73	334	0.07	0.84	365	-0.67	1.05	360	-0.26	1.11	359	-0.17	0.92	354	-0.78	0.81	366	-0.36	0.92	
0.11	0.77	372	0.18	0.73	395	-0.35	0.91	393	-0.07	0.97	392	-0.03	0.97	388	-0.53	0.71	397	-0.16	0.86	
0.19	0.67	324	0.18	1.04	333	-0.18	0.83	334	0.00	0.86	334	-0.03	0.96	327	-0.17	0.72	334	0.05	0.80	
0.02	1.09	322	0.24	0.82	336	-0.08	0.95	336	-0.02	1.05	332	-0.01	0.97	331	-0.04	0.80	336	0.16	0.84	
0.07	0.91	318	0.18	0.98	334	0.04	0.90	333	0.10	0.92	330	0.10	0.96	329	0.14	0.82	334	0.14	0.88	
-0.04	1.11	335	0.15	0.94	344	0.19	0.90	343	0.05	1.08	341	0.02	1.08	340	0.30	0.74	344	0.27	1.03	
-0.14	1.13	359	0.12	0.91	368	0.13	0.95	366	0.04	1.07	368	-0.06	1.07	364	0.45	0.79	368	0.23	0.94	
-0.20	1.23	343	0.19	0.93	353	0.23	0.90	353	0.13	0.94	352	0.02	1.11	350	0.71	0.81	353	0.37	0.75	
-0.32	1.23	295	0.10	1.01	305	0.20	0.86	304	-0.04	1.05	303	-0.03	1.01	303	0.77	0.81	304	0.31	0.85	
-0.26	1.29	330	-0.10	1.02	332	0.09	0.89	332	0.15	0.93	330	0.11	0.98	326	0.86	0.83	333	0.21	0.96	
-0.51	1.23	218	-0.25	0.95	223	0.02	1.14	223	0.00	0.99	222	-0.11	1.28	220	0.89	0.92	223	0.30	0.77	
-0.44	1.20	126	-0.37	1.30	128	-0.10	1.08	128	-0.12	1.05	128	-0.13	1.15	126	0.81	0.89	128	0.22	0.98	
-0.55	1.76	57	-0.60	1.14	57	-0.24	1.18	57	-0.16	1.25	57	-0.19	1.18	54	0.87	0.94	57	0.28	1.03	
0.17	0.61	293	-0.35	1.22	332	-0.68	0.85	331	-0.45	0.98	331	-0.23	0.81	332	-1.25	0.79	334	-0.74	1.05	
0.09	0.88	292	-0.18	1.11	342	-0.43	0.83	340	-0.22	0.90	338	-0.15	0.98	333	-0.91	0.79	341	-0.45	1.01	
0.20	0.73	312	-0.13	1.06	337	-0.22	0.85	336	-0.24	1.06	336	-0.17	0.98	333	-0.70	0.72	336	-0.19	0.92	
0.19	0.75	287	-0.11	1.14	308	0.00	0.84	308	-0.02	0.88	307	-0.07	1.02	301	-0.42	0.78	308	-0.07	0.98	
0.24	0.84	339	0.13	0.84	360	0.18	0.82	359	0.11	0.90	358	0.06	0.90	357	-0.21	0.77	360	0.07	0.80	
0.22	0.81	347	0.12	1.03	360	0.28	0.85	361	0.17	0.91	358	0.13	0.91	357	-0.11	0.87	361	0.09	0.94	
0.14	0.86	365	0.07	0.92	379	0.31	0.86	379	0.05	1.03	377	0.10	1.00	371	0.15	0.77	380	0.11	0.91	
0.09	1.03	367	0.14	0.81	382	0.37	0.82	381	0.16	0.99	378	0.11	1.04	381	0.15	0.81	381	0.14	0.82	
0.10	0.95	407	0.09	0.93	418	0.48	0.71	418	0.20	1.01	415	0.08	1.01	415	0.35	0.72	417	0.20	0.86	
-0.02	1.08	384	0.03	0.91	394	0.42	0.91	393	0.15	0.93	386	0.09	1.03	393	0.36	0.81	394	0.17	0.90	
-0.13	1.13	476	-0.14	1.01	491	0.36	0.91	491	0.15	0.95	489	0.17	0.95	486	0.41	0.85	490	0.15	0.85	
-0.16	1.09	336	-0.26	1.06	341	0.38	0.90	341	0.21	0.93	338	0.22	0.91	334	0.45	0.79	341	0.21	0.84	
-0.29	1.33	157	-0.40	1.21	159	0.32	0.95	160	0.11	1.02	160	0.08	1.11	157	0.42	0.89	160	0.11	0.89	
-0.19	1.08	75	-0.31	1.03	76	0.30	0.89	76	0.24	0.91	75	0.10	1.09	74	0.48	0.72	76	0.15	0.97	
		8583			8997			8978			8926			8870			9001			

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

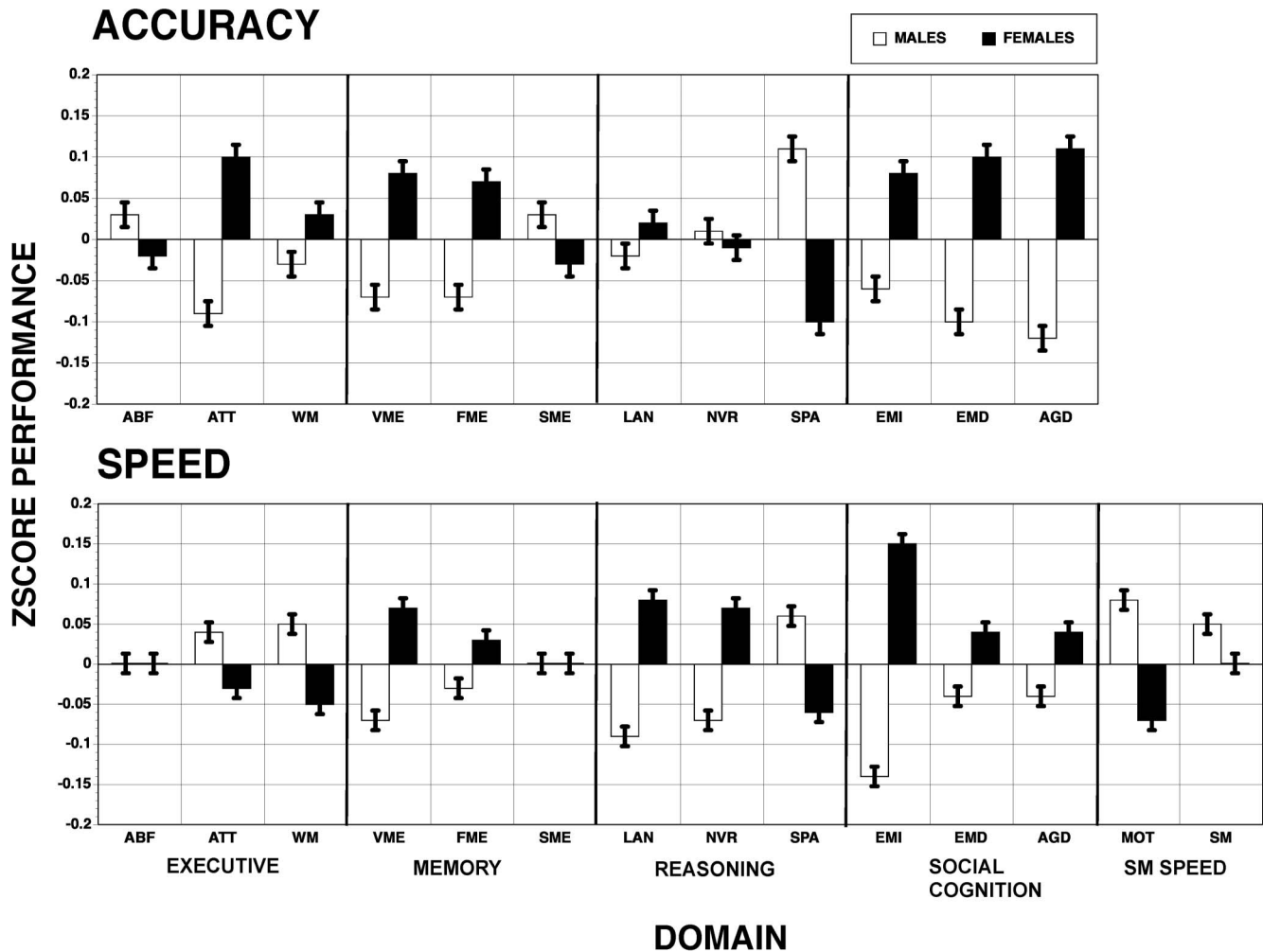


Figure 1. Means (+SEM) of age-adjusted z scores for accuracy (A) and speed (B) in females (red bars) and males (blue bars) across the entire sample ($n = 9010$) on Executive-Control (ABF = Abstraction & Mental Flexibility; ATT = Attention; WM = Working Memory), Episodic Memory (VMEM = Verbal; FMEM = Facial, SMEM = Spatial), Complex Cognition (LAN = Language Reasoning; NVR = Nonverbal [matrix] Reasoning; SPA = Spatial Ability), Social Cognition (EMI = Emotion Identification; EMD = Emotion Differentiation; AGD = Age Differentiation), and Sensorimotor Speed (SM = Sensorimotor speed; MOT = Simple Motor Speed). Sample sizes vary by task: Females (4437–4679); Males (4146–4322).

Global Neurocognitive Performance: Effects of Age and Sex

For GNP accuracy (Figure 2A), the main effects of age, $F(1, 8921) = 4919.55$, $p < .0001$, age-square, $F(1, 8921) = 484.99$, $p < .0001$, and sex, $F(1, 8921) = 14.29$, $p = .0002$, were all significant. The interaction of age and sex, $F(1, 8921) = 9.65$, $p = .0019$, but not age-square and sex, was significant. As expected, GNP accuracy increased with age, peaking for females at age 19 and for males at age 20. Variance of GNP accuracy was higher in males, Levene $F(1, 9008) = 21.60$, $p < .0001$.

For GNP speed (Figure 2B), the main effects of age, $F(1, 8921) = 2189.58$, $p < .0001$, age-square, $F(1, 8921) = 783.35$, $p < .0001$, and sex, $F(1, 8921) = 9.55$, $p = .0020$, were likewise significant. The interaction of age and sex, $F(1, 8921) = 9.47$, $p = .0021$, but not age-square and sex, was

significant. As expected, GNP speed increased with age, males being faster than females at early ages, females drawing even after age 11. Both sexes peaked at age 16 and declined thereafter, more sharply in males. Variance of GNP speed did not differ between the sexes.

Within-Individual Variability: Effects of Age and Sex

WIV for accuracy showed a nonlinear (U-shaped) relationship with age (Figure 3A). ANCOVAs showed main effects of age, $F(1, 8921) = 643.20$, $p < .0001$, and age-square $F(1) = 257.73$, $p < .0001$, but not sex. The interaction of age-square and sex, $F(1, 8921) = 9.67$, $p = .0019$, but not age and sex, was significant. As hypothesized, WIV decreased with age but then, unexpectedly, increased significantly after age 17 and into early adulthood, especially in males. Eight-year-old males, also had

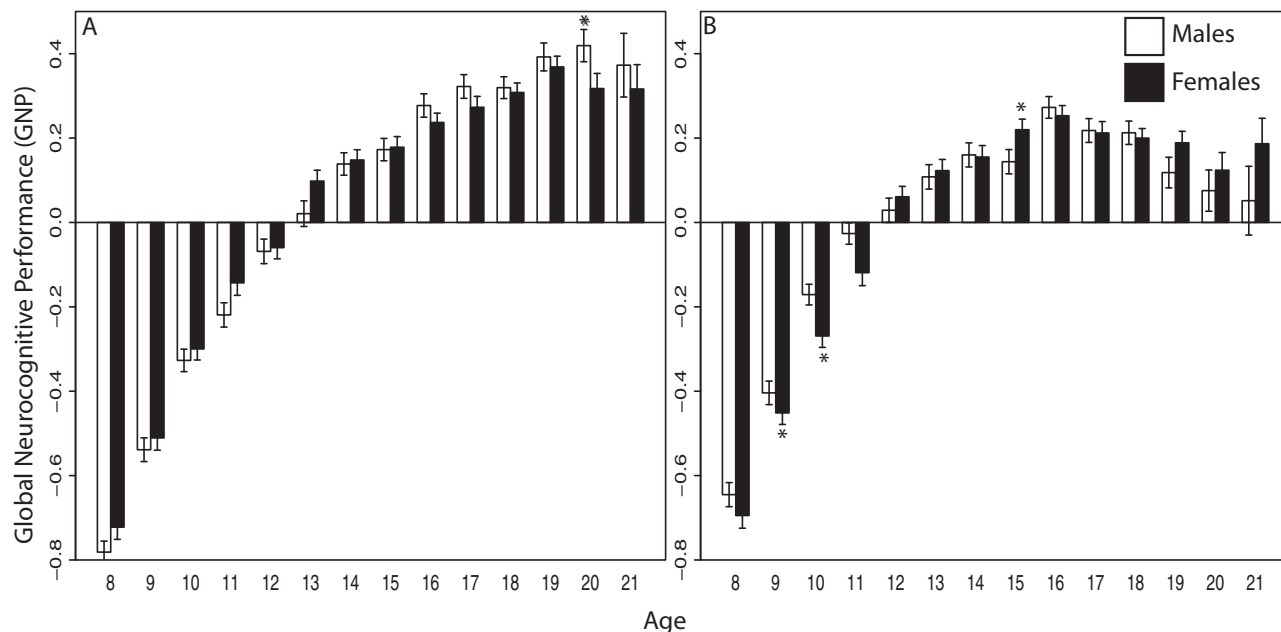


Figure 2. Means (+SEM) global neurocognitive score (GNP) for accuracy (A) and speed (B) in females (red bars) and males (blue bars) across the entire sample ($n = 9010$). As expected, GNP accuracy and speed improved with age. Overall, females had higher GNP for accuracy and speed scores than males. Females reach mature performance earlier; however, young adult males outperform females in accuracy but not speed. Asterisks (*) denote age-specific sex differences.

higher WIV for accuracy than their female counterparts, $t(768.50) = 3.57$, $p = .0004$; Cohen's $d = 0.26$. However, during most of adolescence, sex differences were absent until, starting in early adulthood, males again showed higher values, in particular at age 19, $t(430.58) = 2.61$, $p = .0094$; Cohen's $d = 0.24$. Variance of WIV for accuracy was higher in males, $F(1, 9008) = 29.51$, $p < .0001$.

For WIV speed (Figure 3B), the main effects of age, $F(1, 8921) = 86.24$, $p < .0001$, age-square, $F(1, 8921) = 416.90$, $p < .0001$ and sex, $F(1, 8921) = 8.92$, $p = .0042$, were significant. The interaction of age and sex, $F(1, 8921) = 18.21$, $p < .0001$, but not age-square and sex, was significant. As hypothesized, WIV for speed decreased in childhood but here again for males more than for females, it increased significantly in adolescence and into early adulthood. Overall, males had higher WIV for speed than females, which became more evident with increased age. Speed WIV was comparable in males and females during childhood (8–10) and decreased yearly in both sexes until about age 15, when WIV began to increase and became higher in males than females, in particular at age 15 [$t(687.12) = 3.89$, $p < .0001$; Cohen's $d = 0.27$], 16 [$t(671.75) = 2.70$, $p = .0071$; Cohen's $d = 0.20$], 17 [$t(603.46) = 2.26$, $p = .0239$; Cohen's $d = 0.20$], 19 [$t(407.97) = 4.13$, $p < .0001$; Cohen's $d = 0.35$], and 21 [$t(98.38) = 2.33$, $p = .0218$; Cohen's $d = 0.42$]. Variance of WIV speed was higher in males, $F(1, 9008) = 17.39$, $p < .0001$. Overall, WIVs for accuracy and speed were highest in childhood, lowest around age 15, and then higher into adulthood. Males had generally higher values than females, most clearly in late adolescence and early adulthood.

Potential Contribution of Medical and Psychiatric Conditions

Although the participants were physically and cognitively able to perform the tests using the computerized format, because the sample was drawn from a pediatric hospital, it is possible that, for some participants, certain medical or psychiatric conditions affected their performance. In a follow-up analysis, we excluded all participants ($n = 2167$; 51% female) with significant medical conditions, such as respiratory disease (e.g., cystic fibrosis), gastrointestinal disorders (e.g., inflammatory bowel disease), immune dysfunction (e.g., common variable immunodeficiency), and endocrine/metabolic disorders (e.g., diabetes). To compare differences in health status between younger and older participants, the sample was split into two groups, younger and older than 15. The overall rate of significant medical conditions within this sample is 24.04%. Of participants under the age of 15, 20.30% had a significant medical condition compared to 28.61% of participants age 15 and older. We also generated a composite score from responses on the general mental health screen and entered it as a covariate for all participants. Neither procedure altered the pattern of results (see online Supplemental Figure S1).

Performance Associations With WRAT4 and Parental Education

WRAT4 score differed by age, $F(913,8982) = 39.77$, $p < .0001$ (see online Supplemental Table S1). There was a significant difference between ages 8 and 9, 12 and 13, and 15–18. As a group, individuals under 15 [$104.20(15.40)$] had higher WRAT4 standard

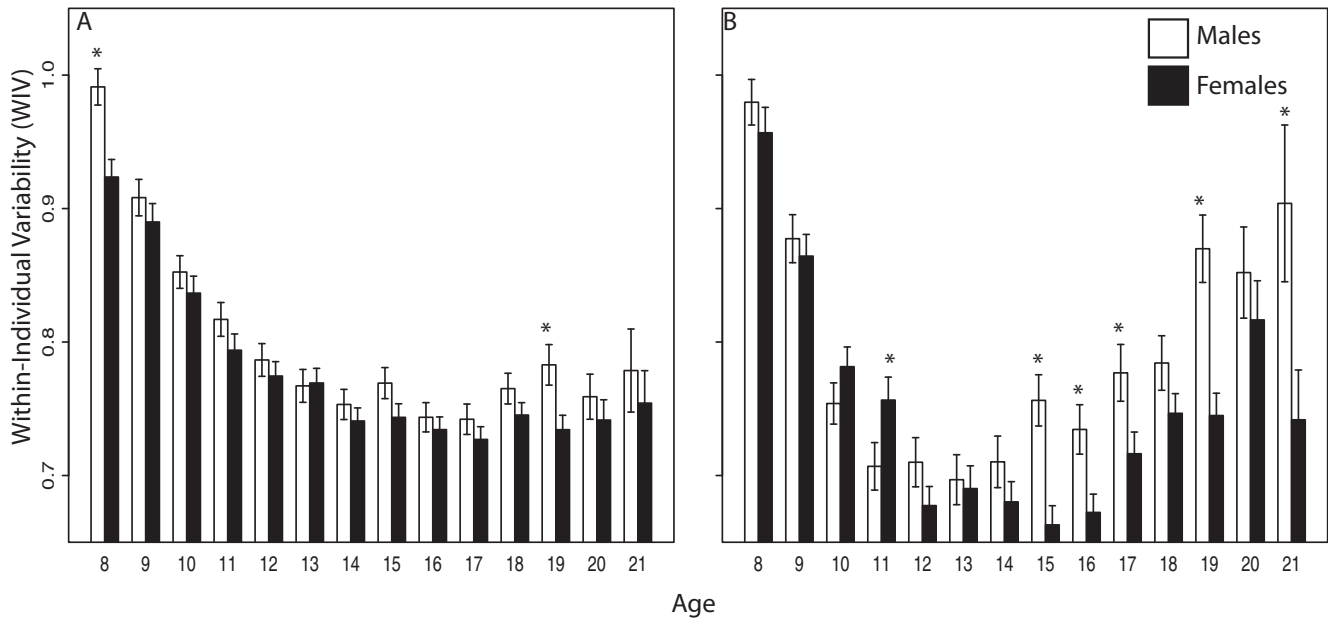


Figure 3. Means (\pm SEM) for across-test within-individual variability (WIV) for accuracy (A) and speed (B) in females (red bars) and males (blue bars) across the entire sample ($n = 9010$). As expected, males have higher accuracy and speed WIV as compared with females. In general, accuracy WIV decreases with age and younger males (age 8) show the highest variability. For speed WIV, early adolescent females (11–12) are more variable (speed) than males; however, this pattern reverses by age 15. WIV is dramatically higher in late adolescent/early adult males as compared to females. Higher values represent higher variability. Asterisks (*) denote age-specific sex differences. Note: scale begins at 0.7.

scores than individuals 15 years and older [100.20(16.42)]. Although statistically significant, the groups differed by only 4 points on the standard scale.

Overall, parental education differed by age, $F(2, 8925) = 32.08$, $p < .0001$ (see online Supplemental Table S1). This was most evident between the parents of the 8-year-olds and 21-year-olds. Individuals under 15 [14.43 (2.34)] had higher parental education than those 15 and older [14.19 (2.31)]. The difference, however, was less than 1 year.

Correlations for each sex for CNB test scores, parental education, and WRAT4 are listed in Table 4. In general, they were similar for males and females. GNP was negatively correlated with WIV for both accuracy and speed. Parental education was positively correlated with GNP accuracy and speed and negatively

correlated with WIV accuracy and speed. Estimated IQs based on the WRAT4 were average for both sexes, with an age-adjusted mean scale score of ~ 103 and SD of ~ 16 for males and ~ 102 and SD of ~ 16 for females, indicating that the IQs for our population-based sample were similar to those for the standardization sample. Overall, the correlations between performance and age-adjusted scale scores on the WRAT4 were in the expected directions: positively correlated with GNP accuracy and speed, negatively correlated with WIV accuracy and speed.

Discussion

By examining within-individual variability, we found a heretofore unreported age-related change in within-individual variability

Table 4
Correlations Among Performance Parameters, Parental Education, and WRAT in Males and Females

Correlations	GNP accuracy	GNP speed	WIV accuracy	WIV speed	Parental education	WRAT STD
GNP accuracy	—	.36	-.54	-.24	.25	.38
GNP speed	.36	—	-.23	.68	.08	.13
WIV accuracy	-.55	-.25	—	.21	-.14	-.22
WIV speed	-.19	-.67	.21	—	-.07	-.13
Parental education	.21	.04	-.11	-.02	—	.33
WRAT STD	.37	.11	-.19	-.10	.39	—

Note. Females are displayed above the diagonal, and males below the diagonal. Significant ($p < .001$) correlations are in bold type.

of neurocognitive performance. At 8 years, the youngest age group sampled, WIV was at its highest level and then decreased every year for both accuracy and speed to its lowest level in early adolescence. This decrease with age was predicted by the hypothesis of uneven development, whereby some domains mature before others. But we also observed, unexpectedly, that after about age 14, WIV increased every year into adulthood, although never reaching the levels shown by young children. This pattern was clearer for speed than accuracy and was stronger for males than for females. We propose that this post-pubescence increase in WIV reflects the unfolding of greater specialization related to skill honing and brain maturation. Such increased specialization in speed, as indexed by increased WIV for speed that continues into adulthood, is consistent with evidence that myelination of brain white matter in some cortical regions is still incomplete at age 21, especially for males (Giedd & Rapoport, 2010; Raznahan et al., 2011). This disproportionate development of white matter with age and sex may underlie superior speed performance in some domains but poorer performance in other domains. For example, young adult males are faster than adolescent males in motor performance, but slower during emotion identification. To the extent that certain tasks, such as motor and visual tasks, also require rapid interhemispheric transfer of information, increased speed of performance in some tasks but not others is consistent with evidence that the myelination of the corpus callosum is still incomplete at age 21 (Giedd et al., 1999; Pujol, Vendrell, Junque, Marti-Vilalta, & Capdevila, 1993).

Our results for across-task WIV are similar to those from a previous study of variability in a single reaction time (RT) task, which, like our results, showed decreased variability from childhood through adolescence, followed by an increase throughout adulthood and into old age (Dykiert et al., 2012). In our study, however, this increase began in mid-adolescence. No multitask WIV has been reported previously. Our measure of variability, which has been used before in aging research (MacDonald, Li, & Backman, 2009), is within a single person but occurs across tasks in several domains and therefore provides an index of evenness, or consistency, of performance across domains. Our finding that our youngest participants (8 years) were more uneven than our mid-adolescents (12–14 years) fits with prior reports that children show higher variation in performance (Siegler, 1994). Indeed, as previously noted, children often have high variability when solving the same problem over short or long intervals (Siegler & McGilly, 1989). What has not been reported before and was unexpected was the increase in variability in late adolescence (after age 14) and into early adulthood. We do not believe this pattern in older adolescents and young adults should be seen as a reversion to a childlike pattern of variability reflecting uneven development. Instead, we propose that it reflects increased cognitive specialization, meaning that some individuals are improving more on tasks in certain domains relative to tasks in other domains. If, indeed, this pattern reflects increased cognitive specialization, it is noteworthy that it is more evident in speed than in accuracy, at least within the time parameters of this study. Assuming we can do follow-up studies with the same participants when they are older, it is conceivable that the pattern also will be seen in accuracy.

Our hypothesis of greater across-domain variability in males was strongly supported. Males were more uneven, most notably for speed after age 15. Our results suggest that with maturation

there is greater specialization in males for some domains, reflected in differential cognitive processing, whereas in females, specialization more often develops more nearly equally across a range of domains. These results are also consistent with studies indicating, as previously noted, that males are more likely than females to be “cognitive specialists,” performing some functions better than others, whereas females are more likely to become “cognitive generalists,” performing at more nearly equal levels across domains (Halpern et al., 2007). In addition to finding higher WIV in males, our results support Thorndike’s (1910) hypothesis of higher variance in males. Higher variance in males also is consistent with studies of sex differences in IQ and spatial ability (e.g., Deary et al., 2003; Shea et al., 2001), findings that we now have extended over a broader age range. Our findings also align with research showing males to be overrepresented in both extremes of the ability distribution (Thorndike, 1910, but see Hollingworth, 1914), similar to sex differences in IQ in young children (Deary et al., 2003). The sex difference in variability is small, but insofar as the differences are maximized in the tails of the distribution, it may explain why more men both over- and underachieve in comparison with women in the real world. The results support the potential utility of WIV as a dimension or metric related to cognitive development. Initial evaluation of the construct validity of this new dimension suggests that WIV correlates negatively with average performance and with parental education, which is itself positively correlated with average performance. It appears that higher parental education is associated with an environment where both better performance and greater evenness of performance are fostered. These effects also could be genetically mediated, and incorporation of the genomic data available on this sample could help shed light on this issue.

Our GNP data add to the substantial literature showing increases in neurocognitive ability throughout development (e.g., R. C. Gur et al., 2012; Steinberg, 2005). Adolescence and early adulthood are marked by significant increases in brain maturation and subsequent increases in behavioral and cognitive functioning, including improvement in information processing and reasoning, especially in abstract, organized, and logical thinking (Steinberg, 2005). The domains captured in the computerized neurocognitive battery show significant improvement with age, particularly in children and young adults (R. C. Gur et al., 2012) and can be linked to age changes in functional or neuroanatomical systems (e.g., Ragland et al., 2002). In general, males and females showed similar age-related differences in performance. The main difference was that females were more accurate earlier (age 13), whereas males were faster earlier (ages 10–11). The accuracy results fit reports of earlier physical (Hills & Byrne, 2010) and behavioral maturation (e.g., Tiemeier et al., 2010) in females. Perhaps males are more inclined to trade accuracy for speed or they are more impulsive and less able to inhibit their initial responses (Cross, Copping, & Campbell, 2011; Feingold, 1994). So far as children are concerned, any differences in performance also could be related to developmental differences in the pace of brain maturation, in particular prefrontal myelination (Gogtay et al., 2004). Notably, these sex differences do not appear to be modulated by sex hormones or puberty during early adolescence (Herlitz, Reuterskiold, Loven, Thilers, & Rehnman, 2013), although these factors may play a role in late adolescence.

The cross-sectional nature of our study limits our ability to separate developmental effects from cohort effects. However, many of the participants have agreed to be followed longitudinally and these studies are underway. Participant enrollment through pediatric clinics at the Children's Hospital may affect generalizability as a result of comorbid medical conditions, but we have taken steps to address this potential limitation in our statistical analyses. Other cohort factors may also unintentionally influence our data. For example, our sample may be influenced by the recruitment methods used by the Children's Hospital. However, children and young adults with physical or mental conditions likely to affect performance were excluded. Moreover, because the study was completed in a short period of time (24 months), within-cohort bias should be reduced. This is especially true in comparison with studies that required a decade or more to complete or that attempt to combine disparate data sets to achieve a large sample.

Although our sample was large and demographically diverse, we included individuals with medical conditions (e.g., neurological or neurosurgical) that could have affected their performance and skewed some of the data. The prevalence of chronic health conditions in our sample (24.04%) is similar to a recent study of approximately 5,000 children (from three cohorts) over the course of 6 years (20.8%; Van Cleave, Gortmaker & Perrin, 2010). Indeed, our overall prevalence is *lower* than the prevalence reported in one of the cohorts (26.6%) in this study. Furthermore, the presence of *significant* medical conditions does not imply *permanence* because medical conditions change over time, more effective treatments and access to health care may become more available, and the natural course of development may affect the condition. Van Cleave et al. also note that the persistence of medical conditions within a given individual in their study was relatively low, approximately 7%. Thus, simply excluding individuals based on significant medical conditions at one time point may ignore natural developmental milestones that help resolve these medical issues and their impact on cognition. Finally, our follow-up analysis that excluded these individuals and controlled for general mental health showed similar results. Notably, the average WRAT4 score for our sample was identical to that obtained in the normative sample, indicating that our sample can provide a valid approximation of typical development. Another limitation is that our final sample included more females than males. Because standardization of CNB scores was based on the entire sample, estimations for some effects in certain age groups could have been imprecise because of disparities in the numbers of males and females, although standardization across the sample should have minimized any systematic distortion of parameter estimates. Given the modest number of tests in each neurobehavioral domain, global scores used for performance and variability may have been biased by the test instruments used. However, because previous studies (R. C. Gur et al., 2012; R. C. Gur et al., 2010) have consistently shown that the CNB is balanced for sex differences, we believe the instrument provides a robust set of tests for measuring global performance and variability. Differences in the parental education and WRAT4 scores of younger and older participants were minimal but could potentially explain some of the differences in performance. However, including parental education and WRAT4 in the statistical model did not change the results for any measure, indicating minimal effects of these mod-

erating variables. Although age-related differences in WIV look systematic, many factors are likely to contribute to variability. Differences in WIV therefore may depend to a large extent on the specific constellation of abilities being measured; if so, its generalizability needs further scrutiny. Longitudinal follow-up studies could expand upon the variability findings by examining variability across time within specific tests and assess additional underlying factors that may affect variability such as comorbid health conditions, substance use, or other social or environmental factors.

Notwithstanding these limitations, the current study of more than 9,000 children and youths demonstrates age- and sex-related differences in global neurocognitive performance along with significant within-individual variability. These data provide a framework for the longitudinal study of individual neurocognitive developmental trajectories and potential deviations from those trajectories similar to those found in longitudinal studies of physical growth, intelligence, and personality (e.g., Bayley & Schaefer, 1964; Bradley & Caldwell, 1976; Crano & Mendoza, 1987; Kagan & Moss, 1959). The difference is the new window that WIV offers into behavioral architecture and its development. It could signify a normal variant of degree of cognitive specialization and, when abnormally high, it could index vulnerability or giftedness. Measuring variability of performance also could aid in the identification of critical developmental periods within individuals and provide a benchmark for evaluation of developmental disorders. By screening children for existing neurocognitive problems (e.g., higher than normal WIV for a given sex/age), cognitive weaknesses can be assessed and monitored in a way that can lead to targeted educational or clinical interventions (Sege & De Vos, 2008). In combination with follow-up clinical evaluation and measurement of neurobiological markers (e.g., neuroimaging data, genomics), such data have the potential to unveil specific mechanisms that may underlie developmental traits as well as genetic variants possibly underlying key individual differences, including those related to age and sex. The next step will be to make the neuro-cognitive connection. For example, do individual differences in prefrontal development predict individual differences in variability of performance on cognitive tests that presumably tap into prefrontal functions as well as on tests in other cognitive domains? We have extensive neuroimaging data on a subsample of about 1500 children that could be used to test such hypotheses.

References

- Bayley, N., & Schaefer, E. S. (1964). Correlations of maternal and child behaviors with the development of mental abilities: Data from the Berkeley Growth Study. *Monographs of the Society for Research in Child Development*, 29, 1–80. doi:10.2307/1165805
- Blakemore, S. J., & Choudhury, S. (2006). Development of the adolescent brain: Implications for executive function and social cognition. *Journal of Child Psychology and Psychiatry*, 47, 296–312.
- Bradley, R. H., & Caldwell, B. M. (1976). The relation of infants' home environments to mental test performance at fifty-four months: A follow-up study. *Child Development*, 47, 1172–1174. doi:10.2307/1128457
- Burnett, S., Sebastian, C., Cohen Kadosh, K., & Blakemore, S. J. (2011). The social brain in adolescence: Evidence from functional magnetic resonance imaging and behavioural studies. *Neuroscience and Biobehavioral Reviews*, 35, 1654–1664. doi:10.1016/j.neubiorev.2010.10.011
- Crano, W. D., & Mendoza, J. L. (1987). Maternal factors that influence children's positive behavior: Demonstration of a structural equation

- analysis of selected data from the Berkeley Growth Study. *Child Development*, 58, 38–48. doi:10.2307/1130290
- Cross, C. P., Copping, L. T., & Campbell, A. (2011). Sex differences in impulsivity: A meta-analysis. *Psychological Bulletin*, 137, 97. doi:10.1037/a0021591
- Darwin, C. (1871). *The descent of man, and selection in relation to sex* (facsimile of original published by J Murray in 1981, London). Princeton, NJ: Princeton University Press.
- Deary, I. J., Thorpe, G., Wilson, V., Starr, J. M., & Whalley, L. J. (2003). Population sex differences in IQ at age 11: The Scottish mental survey 1932. *Intelligence*, 31, 533–542. doi:10.1016/S0160-2896(03)00053-9
- Dempster, F. N. (1985). Short-term memory development in childhood and adolescence. In C. J. B. S. M. Pressley (Ed.), *Basic processes in memory development* (pp. 209–248). New York, NY: Springer-Verlag. doi:10.1007/978-1-4613-9541-6_5
- Dykiert, D., Der, G., Starr, J. M., & Deary, I. J. (2012). Sex differences in reaction time mean and intraindividual variability across the life span. *Developmental Psychology*, 48, 1262–1276. doi:10.1037/a0027550
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116, 429. doi:10.1037/0033-2909.116.3.429
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87, 477–531. doi:10.1037/0033-295X.87.6.477
- Friederici, A. D., & Wartenburger, I. (2010). Language and brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1, 150–159. doi:10.1002/wcs.9
- Gallahue, D. L., Ozmun, J. C., & Goodway, J. (1998). *Understanding motor development: Infants, children, adolescents, adults*. New York, NY: McGraw-Hill.
- Gbedd, J. N., Blumenthal, J., Jeffries, N. O., Rajapakse, J. C., Vaituzis, A. C., Liu, H., . . . Castellanos, F. X. (1999). Development of the human corpus callosum during childhood and adolescence: A longitudinal MRI study. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 23, 571–588. doi:10.1016/S0278-5846(99)00017-2
- Giedd, J. N., & Rapoport, J. L. (2010). Structural MRI of pediatric brain development: What have we learned and where are we going? *Neuron*, 67, 728–734. doi:10.1016/j.neuron.2010.08.040
- Glahn, D. C., Gur, R. C., Ragland, J. D., Censits, D. M., & Gur, R. E. (1997). Reliability, performance characteristics, construct validity, and an initial clinical application of a visual object learning test (VOLT). *Neuropsychology*, 11, 602. doi:10.1037/0894-4105.11.4.602
- Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., . . . Thompson, P. M. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America*, 101(21), 8174–8179. doi:10.1073/pnas.0402680101
- Goldberg, M. C., Maurer, D., & Lewis, T. L. (2001). Developmental changes in attention: The effects of endogenous cueing and of distractors. *Developmental Science*, 4, 209–219. doi:10.1111/1467-7687.00166
- Gur, R. C., Gur, R. E., Obrist, W. D., Hungerbuhler, J. P., Younkun, D., Rosen, A. D., . . . Reivich, M. (1982). Sex and handedness differences in cerebral blood flow during rest and cognitive activity. *Science*, 217, 659–661. doi:10.1126/science.7089587
- Gur, R. E., Kohler, C. G., Ragland, J. D., Siegel, S. J., Lesko, K., Bilker, W. B., & Gur, R. C. (2006). Flat affect in schizophrenia: Relation to emotion processing and neurocognitive measures. *Schizophrenia Bulletin*, 32, 279–287. doi:10.1093/schbul/sbj041
- Gur, R. C., Mozley, L. H., Mozley, P. D., Resnick, S. M., Karp, J. S., Alavi, A., . . . Gur, R. E. (1995). Sex differences in regional cerebral glucose metabolism during a resting state. *Science*, 267, 528–531. doi:10.1126/science.7824953
- Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., . . . Gur, R. E. (2001). Computerized neurocognitive scanning: I. Methodology and validation in healthy people. *Neuropsychopharmacology*, 25, 766–776. doi:10.1016/S0893-133X(01)00278-0
- Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., . . . Gur, R. E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology*, 26, 251–265. doi:10.1037/a0026712
- Gur, R. C., Richard, J., Huggett, P., Calkins, M. E., Macy, L., Bilker, W. B., . . . Gur, R. E. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: Standardization and initial construct validation. *Journal of Neuroscience Methods*, 187, 254–262. doi:10.1016/j.jneumeth.2009.11.017
- Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Huggett, P., Macy, L., . . . Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *Journal of Neuroscience Methods*, 115, 137–143. doi:10.1016/S0165-0270(02)00006-7
- Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, 8, 1–51. doi:10.1111/j.1529-1006.2007.00032.x
- Harris, L. J. (1985). Delicacy of fibres in the brain. In J. Ghesquire, R. D. Martin, & F. Newcombe (Eds.), *Sexual dimorphism: Symposia of the Society for the Study of Human Biology* (Vol. XXIV, pp. 283–321). Philadelphia, PA: Taylor & Francis.
- Herlitz, A., Nilsson, L.-G., & Backman, L. (1997). Gender differences in episodic memory. *Memory & Cognition*, 25, 801–811. doi:10.3758/BF03211324
- Herlitz, A., Reuterskiold, L., Loven, J., Thilers, P. P., & Rehnman, J. (2013). Cognitive sex differences are not magnified as a function of age, sex hormones, or puberty development during early adolescence. *Developmental Neuropsychology*, 38, 167–179. doi:10.1080/87565641.2012.759580
- Hills, A. P., & Byrne, N. M. (2010). An overview of physical growth and maturation. *Medicine and Sport Science*, 55, 1–13. doi:10.1159/000321968
- Hollingworth, L. S. (1914). Variability as related to sex differences in achievement: A critique. *American Journal of Sociology*, 19, 510–530. doi:10.1086/212287
- Holtzer, R., Verghese, J., Wang, C., Hall, C. B., & Lipton, R. B. (2008). Within-person across-neuropsychological test variability and incident dementia. *JAMA: Journal of the American Medical Association*, 300, 823–830. doi:10.1001/jama.300.7.823
- Inhelder, B., & Piaget, J. (1955). *De la logique de l'enfant a la logique de l'adolescent* (The growth of logical thinking from childhood to adolescence). Paris, France: Presses Universitaires de France.
- Kagan, J., & Moss, H. A. (1959). Parental correlates of child's IQ and height: A cross-validation of the Berkeley growth study results. *Child Development*, 30, 325–332.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., . . . Ryan, N. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36, 980–988. doi:10.1097/00004583-199707000-00021
- Kurtz, M. M., Moberg, P. J., Gur, R. C., & Gur, R. E. (2001). Approaches to cognitive remediation of neuropsychological deficits in schizophrenia: A review and meta-analysis. *Neuropsychology Review*, 11, 197–210. doi:10.1023/A:1012953108158
- Kurtz, M. M., Ragland, J. D., Moberg, P. J., & Gur, R. C. (2004). The Penn Conditional Exclusion Test: A new measure of executive-function with alternate forms of repeat administration. *Archives of Clinical Neuropsychology*, 19, 191–201. doi:10.1016/S0887-6177(03)00003-9
- Leth-Steensen, C., Elbaz, Z. K., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: A

- response time distributional approach. *Acta Psychologica*, 104, 167–190. doi:10.1016/S0001-6918(00)00019-6
- Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, 76, 697–712. doi:10.1111/j.1467-8624.2005.00872.x
- Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, 75, 1357–1372. doi:10.1111/j.1467-8624.2004.00745.x
- MacDonald, S. W., Li, S. C., & Backman, L. (2009). Neural underpinnings of within-person variability in cognitive functioning. *Psychology and Aging*, 24, 792–808. doi:10.1037/a0017798
- Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., & Cooper, N. G. (2009). Explicit identification and implicit recognition of facial emotions: II. Core domains and relationships with general cognition. *Journal of Clinical and Experimental Neuropsychology*, 31, 278–291. doi:10.1080/13803390802043619
- Moreno-Briseño, P., Diaz, R., Campos-Romo, A., & Fernandez-Ruiz, J. (2010). Sex-related differences in motor learning and performance. *Behavioral and Brain Functions*, 6, 74. doi:10.1186/1744-9081-6-74
- Pujol, J., Vendrell, P., Junque, C., Martí-Vilalta, J. L., & Capdevila, A. (1993). When does human brain development end? Evidence of corpus callosum growth up to adulthood. *Annals of Neurology*, 34, 71–75. doi:10.1002/ana.410340113
- Ragland, J. D., Turetsky, B. I., Gur, R. C., Gunning-Dixon, F., Turner, T., Schroeder, L., . . . Gur, R. E. (2002). Working memory for complex figures: An fMRI comparison of letter and fractal n-back tasks. *Neuropsychology*, 16, 370–379. doi:10.1037/0894-4105.16.3.370
- Raznahan, A., Shaw, P., Lalonde, F., Stockman, M., Wallace, G. L., Greenstein, D., . . . Giedd, J. N. (2011). How does your cortex grow? *The Journal of Neuroscience*, 31, 7174–7177.
- Roalf, D. R., Gur, R. C., Almasy, L., Richard, J., Gallagher, R. S., Prasad, K., . . . Gur, R. E. (2013). Neurocognitive performance stability in a multiplex multigenerational study of schizophrenia. *Schizophrenia Bulletin*, 39, 1008–1017. doi:10.1093/schbul/sbs078
- Roalf, D. R., Ruparel, K., Verma, R., Elliott, M. A., Gur, R. E., & Gur, R. C. (2013). White matter organization and neurocognitive performance variability in schizophrenia. *Schizophrenia Research*, 143, 172–178. doi:10.1016/j.schres.2012.10.014
- Sege, R. D., & De Vos, E. (2008). Care for children and evidence-based medicine. *Pediatric annals*, 37, 168–172. doi:10.3928/00904481-20080301-13
- Shaw, D. J., Grosbras, M.-H., Leonard, G., Pike, G. B., & Paus, T. (2012). Development of the action observation network during early adolescence: A longitudinal study. *Social Cognitive and Affective Neuroscience*, 7, 64–80. doi:10.1093/scan/nsq105
- Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, 93, 604–614. doi:10.1037/0022-0663.93.3.604
- Siegler, R. S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science*, 3, 1–5. doi:10.1111/1467-8721.ep10769817
- Siegler, R. S., & McGilly, K. (1989). Strategy choices in children's time-telling. In I. Levin & D. Zakay (Eds.), *Time and human cognition: A life span perspective*, Vol. 59 (pp. 185–218). Amsterdam, The Netherlands: Elsevier Science.
- Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, 9, 69–74. doi:10.1016/j.tics.2004.12.005
- Stuss, D. T., Murphy, K. J., Binns, M. A., & Alexander, M. P. (2003). Staying on the job: The frontal lobes control individual performance variability. *Brain: A Journal of Neurology*, 126, 2363–2380. doi:10.1093/brain/awg237
- Thomas, J. R., & French, K. E. (1985). Gender differences across age in motor performance: A meta-analysis. *Psychological Bulletin*, 98, 260–282. doi:10.1037/0033-2909.98.2.260
- Thorndike, E. L. (1910). The original nature of man. In *Educational psychology*, Vol 1. New York, NY: Columbia University Press.
- Tiemeier, H., Lenroot, R. K., Greenstein, D. K., Tran, L., Pierson, R., & Giedd, J. N. (2010). Cerebellum development during childhood and adolescence: A longitudinal morphometric MRI study. *NeuroImage*, 49, 63–70. doi:10.1016/j.neuroimage.2009.08.016
- Van Cleave, J., Gortmaker, S. L., & Perrin, J. M. (2010). Dynamics of obesity and chronic health conditions among children and youth. *Journal of the American Medical Association*, 303(7), 623–630. doi:10.1001/jama.2010.104
- Van Geert, P., & Van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior & Development*, 25, 340–374. doi:10.1016/S0163-6383(02)00140-6
- Wilkinson, G. S., & Robertson, G. J. (2006). *Wide Range Achievement Test—Fourth edition*. Lutz, FL: P.A. Resources.

Received August 19, 2013

Revision received January 2, 2014

Accepted January 6, 2014 ■