Psychology in the Schools, Vol. 47(7), 2010
Published online in Wiley InterScience (www.interscience.wiley.com)

DOI: 10.1002/pits.20497

CATTELL–HORN–CARROLL COGNITIVE-ACHIEVEMENT RELATIONS: WHAT WE HAVE LEARNED FROM THE PAST 20 YEARS OF RESEARCH

KEVIN S. McGREW AND BARBARA J. WENDLING

Woodcock-Muñoz Foundation

Contemporary Cattell–Horn–Carroll (CHC) theory of cognitive abilities has evolved over the past 20 years and serves as the theoretical foundation for a number of current cognitive ability assessments. CHC theory provides a means by which we can better understand the relationships between cognitive abilities and academic achievement, an important component of learning disabilities identification and instructional planning. A research synthesis of the extant CHC cognitive-achievement (COG-ACH) research literature is reported. Systematic and operationally defined research synthesis procedures were employed to address limitations present in the only prior attempted synthesis. Nineteen studies met the criteria for inclusion, which yielded 134 analyses. The 134 analyses were organized by three age groups (6–8, 9–13, and 14–19) and by four achievement domains (basic reading skills, reading comprehension, basic math skills, and math reasoning). The results reveal a much more nuanced set of CHC COG-ACH relations than was identified in the only prior review because of (a) breadth of cognitive abilities and measures (broad vs. narrow), (b) breadth of achievement domains (e.g., basic reading skills and reading comprehension vs. broad reading), and (c) developmental (age) status. The findings argue for selective, flexible, and referral-focused intelligence testing, particularly in the context of emerging Response to Intervention (RTI) assessment models. The results suggest that narrow CHC abilities should be the primary focus of instructionally relevant intelligence testing. Furthermore, the finding that more than 90% of the available research is based on the Woodcock–Johnson Battery argues for significant caution in generalizing the findings to other batteries. CHC-based COG-ACH research with other intelligence batteries is recommended. © 2010 Wiley Periodicals, Inc.

The impact of the Response to Intervention (RTI) movement on the assessment activities of school psychologists is in its formative stage. Will RtI supplant traditional cognitive ability testing, supplement it in a complimentary manner, or be a short-term blip on the school psychology radar screen that fades away? We believe that RTI and cognitive ability testing have the potential to form a powerful assessment–intervention monitoring dyad. We also believe, however, that cognitive ability testing practices need to become more purpose driven, flexible, and selective.

The identification of a *psychological process disorder* for specific learning disability (SLD) classification requires some form of cognitive assessment (Newton & McGrew, 2010). We believe that less emphasis should be placed on the overall full-scale IQ and that cognitive assessment should be more selective and focused. For example, there should be selective testing of key markers for screening at-risk children. Researchers advocating early screening as an integral component of some RTI models have identified many abilities (e.g., phonemic awareness, working memory, speed of lexical access or rapid automatic naming speed or vocabulary) that are measured by a number of contemporary intelligence batteries. Given that the major intelligence-testing batteries are among the most psychometrically sound and well-standardized tools available to school psychologists, it makes sense that these tests be used for measurement of at-risk markers in lieu of less technically sound special-purpose instruments. Additionally, treatment resistors will likely need a traditional comprehensive assessment of cognitive strengths and weaknesses as part of the evaluation process for diagnosis and eligibility determination as well as to facilitate intervention planning.

As summarized by Newton and McGrew (2010), the Cattell–Horn–Carroll (CHC) taxonomy of cognitive abilities is the consensus framework from which cognitive abilities are now most often

Correspondence to: Kevin S. McGrew, 1313 Pondview Lane E., St. Joseph, MN 56374. E-mail: iap@earthlink.net

conceptualized and measured. Thus, a critical question is "What CHC broad or narrow cognitive abilities hold promise either as early screening markers or collectively as pattern indicators of a potential SLD process disorder?" Answers to this question reside in the CHC cognitive-achievement (CHC COG-ACH) relations research completed over the past 20 years. It is time to take stock and determine which, if any, CHC cognitive constructs and measures should be included in the school psychologist's assessment arsenal in a new hybrid approach to assessment and intervention—one that combines information from RTI and purpose-driven cognitive ability testing.

Prior CHC COG-ACH Relations Research Summaries

As far as we know, the only attempt to summarize the extant CHC COG-ACH research is a set of regularly updated summary tables published in the CHC cross-battery assessment (CBA) series of books. The first CBA-based synthesis was presented by McGrew and Flanagan (1998). We classify the latest iteration in this series (Flanagan, Ortiz, Alfonso, & Mascolo, 2006) as a *narrative* research synthesis as no attempt was made to systematically quantify the significance of results across studies, an approach typically defined as *meta-analysis* (Cooper, 1998).

Flanagan and colleagues (2006) identified potential CHC COG-ACH research studies through two primary steps. First, research studies and reviews reporting statistically significant relations between cognitive abilities and school achievement were identified via a search of the PsycINFO electronic database. Next, an ancestral search strategy of the references identified in that step revealed other articles for review. Flanagan and colleagues (2006) classified the research reports into three categories: (a) *key CHC studies*–CHC-organized studies that included markers for most broad CHC cognitive abilities; (b) *reviews*–non-CHC organized narrative or meta-analytic research syntheses reporting significant relations between cognitive abilities and school achievement; and (c) *individual studies*–single non-CHC empirical studies that investigated the relations between cognitive abilities and school achievement. For most of the research reports, Flanagan and colleagues (2006) had to translate non-CHC-defined cognitive abilities as per the nomenclature of the CHC taxonomy (e.g., phonemic awareness = narrow ability of phonetic coding (PC) under the broad domain of *Ga*). The Flanagan and colleagues (2006) summary tables included 138 references for reading (8 key CHC studies, 23 reviews, and 107 individual studies) and 37 references for math (3 key CHC studies, 5 reviews, and 29 individual studies).

Brief Summary of Flanagan and Colleagues (2006) Conclusions

Flanagan and colleagues (2006) presented their conclusions in a single summary table (Table 2.14 of Flanagan and colleagues, 2006). The broad *Gc* domain, and the narrow *Gc* abilities of language development (LD), lexical knowledge (VL), and listening ability (LS) were designated as related to both reading and math. The cognitive efficiency broad domains of *Gs* and *Gsm* were also implicated for reading and math, specifically the narrow abilities of perceptual speed (*Gs*-P), memory span (*Gsm*-MS), and working memory (*Gsm*-MW). *Gf,* as well as the narrow abilities of induction (I) and general sequential reasoning (RG), are also listed as related to both reading and math (with a greater relevance for math noted). In reading, Flanagan and colleagues (2006) suggest that *Gf* is primarily related to reading comprehension (RC) and not to basic reading skills (BRS). Finally, reading-specific broad and narrow ability relations listed by Flanagan and colleagues (2006) include *Ga* ("phonological awareness processing" or PC) and *Glr* ("rapid automatic naming" or naming facility [NA]; associative memory [MA]). *Gv* was noted as not displaying any significant relations with reading and math achievement in the studies reviewed, although Flanagan and colleagues (2006) suggest that *Gv* abilities "may be important primarily for higher level or advanced mathematics (e.g., geometry, calculus)" (p. 45).

Although Flanagan and colleagues (2006) mention possible developmental (age)-related findings (e.g., *Gc* "abilities become increasingly more important with age"; PC and NA are "very important during the elementary school years," p. 45), their developmental comments are general and lack specific age-differentiated information. Finally, Flanagan and colleagues (2006) mention possible non-CHC abilities that research suggests are related to school achievement. These include orthographic processing and morphological knowledge in reading.

Limitations of the Flanagan and Colleagues (2006) Synthesis

Although the sheer number of key, review, and individual studies populating the Flanagan and colleagues (2006) research synthesis looks impressive, the syntheses suffer from four primary limitations.

First, it would impossible to replicate the CHC COG-ACH summaries of Flanagan and colleagues (2006) due to the lack of: (a) specified and published database search terms and date ranges, (b) significant CHC-ACH relation significance criteria, (c) study inclusion and exclusion criteria, and (d) procedures to control for possible publication bias (see Cooper, 1998; also see Pillemer & Light, 1980). Second, the use of the broad dependent variable (DV) domains of reading, math, and writing make it impossible to determine if the salient COG-ACH relations generalize across subachievement domains (e.g., BRS vs. RC). Third, the select and limited developmental CHC COG-ACH trends discussed by Flanagan and colleagues (2006) are based only on a handful of *key CHC studies*. The age characteristics of the bulk of the *individual studies* included in the Flanagan and colleagues (2006) research summaries are ignored (i.e., no attempt was made to present CHC COG-ACH results by age-differentiated subsamples).

Finally, *specification error*, sometimes called omitted-variable bias (Keith, 2006), occurs when potentially important variables in predictive or explanatory research are not included in a study's research design. Specification error can lead to biased estimates of the effects (i.e., relative importance) of individual predictor variables on the DV of interest (Pedhazur, 1997). Almost all the *review* and *individual studies* listed by Flanagan and colleagues (2006) likely suffer from an unknown degree of specification error because the intent of their synthesis was to identify the relative importance of the various CHC domains *within the CHC framework*. One example is presented next.

Flanagan and colleagues (2006) list the *individual study* of Wagner and colleagues (1997) as evidence for the importance of *Gc, Ga,* and *Glr* in reading. Specification error is present in this study due to the lack of indicators of the more complete array of important predictor CHC abilities for reading. The CHC cognitive abilities of *Ga, Gc,* and *Glr* may have displayed slightly different degrees of relative importance (and might have demonstrated nonsignificance) had Wagner and colleagues (1997) included measures of *Gsm, Gf, Gs,* and *Gv* in their predictor models.

Purpose of Current CHC COG-ACH Relations Research Synthesis

The purpose of the current review is to take stock of the past 20 years of CHC COG-ACH relations research. More importantly, the primary goals are to address the limitations in the Flanagan and colleagues (2006) synthesis via the use of systematic and operationally defined research synthesis procedures and to add greater specificity to their broad stroke findings as a function of (a) achievement domains (viz., BRS, RC, basic math skills [BMS], math reasoning [MR]); (b) age-related or developmentally moderated findings; and (c) possible "achievement domain × by age × CHC ability" interactions. Our goal is not necessarily to criticize the Flanagan and colleagues (2006) review but, more importantly, to build on their review and extend and refine their work via more explicit and systematic methodological review and synthesis procedures.

Method

Identification of Research Studies

Potential research studies were identified through four methods. First, published research articles were identified via a search of the PsycINFO and Institute for Applied Psychometrics (IAP) reference databases.[1] Second, unpublished doctoral dissertations were located through the ProQuest Digital Dissertations database. Search keywords included "Cattell-Horn-Carroll," "CHC theory," "CHC," "Gf-Gc theory," "Gf-Gc," "Woodcock-Johnson-Revised," "WJ-R," "Woodcock-Johnson III," "WJ III," "Kaufman Assessment Battery for Children–Second Edition," "KABC-II," "Differential Abilities Scales," "DAS," "Differential Abilities Scales–Second Edition (DAS-II)," "Stanford-Binet–Fifth Edition," and "SB-IV." Years included in the searches spanned 1988 to 2009. Third, a request was posted to the CHC and the National Association of School Psychologists (NASP) general purpose listservs[2] asking members for copies of published and unpublished manuscripts or references related to the review. Finally, an *ancestral search* involved examination of the references of the initial research reports (acquired during the first three steps) to determine if they contained references to studies not revealed by the primary search strategy that should be secured for review (Cooper, 1998).

Selection and Description of Studies

A research investigation had to meet *all* of the following criteria to be included in the current review. First, the study had to be explicitly designed as per the CHC (or *Gf-Gc*) cognitive abilities framework. Second, the study must have empirically investigated the relations between the primary CHC cognitive variables (independent variables [IVs]) and achievement variables in reading and math (DVs). Third, the research report had to report quantitative information (e.g., reporting of statistical difference tests, regression weights, effect sizes) regarding the relevant strength of each CHC IV and the respective achievement DV domain. Finally, the study had to include markers from five or more of the seven primary CHC cognitive domains.[3]

Nineteen investigations met all criteria. Each investigation was described (see Table 1) according to type of sample, age or grade range of samples,[4] size of sample(s), number of samples analyzed,

---

[1] IAP Reference database described at: http://tinyurl.com/dcvrdm

[2] CHC listserv address: http://groups.yahoo.com/group/IAPCHC/; NASP listerv address: http://groups.yahoo.com/group/NASP-Listserv/

[3] In our introduction we argued that a limitation of prior CHC COG-ACH relations synthesis was the inclusion of studies that included a narrow range of CHC IVs (specification error). We could similarly be asked why we included investigations (viz., Ganci, 2004; Keith, 1999; Vanderwood, McGrew, Flanagan, & Keith, 2002) that did not include measures of the major seven broad (*Gf, Gc, Glr, Gsm, Ga, Gv, Gs*) domains. We included these studies as, in each instance, the omission of one or more domains was an a priori decision articulated by the researchers based on the failure to find a significant relation for the excluded broad domain (e.g., *Gv*) in the researchers' review of the then extant research literature.

[4] Only school-age samples were included in the review. A number of research investigations reported results for young adult to older adult samples, but collectively the number of adulthood samples was too sparse to allow for meaningful synthesis across studies.

Table 1
Studies Included in CHC COG-ACH Relations Research Synthesis: School-Age Studies

Study Category, Number, Authors, Date	Type of Sample[a]	Age (A) or Grade (G) Range	Sample *n* (# Samples Analyzed)	Cog. Batteries and CHC IVs	CHC IVs Strata Included	Ach DV Domains	Type of Ach. DV[b]	Analysis Method[c]
Manifest variables-no *g* (full scale IQ)								
1. McGrew (1993)	National norm	A 5–19	59–325 (15)	WJ-R	Broad narrow	BRS RC	WJ-R C	MR
2. McGrew & Hessler (1995)	National norm	A 5–19	113–325 (15)	WJ-R	Broad narrow	MCS, MR	WJ-R C	MR
3. Evans, Floyd, & McGrew (2002)	National norm	A 6–19	96–431 (14)	WJ III	Broad narrow	BRS RC	WJ III C	MR
4. Floyd, Evans, & McGrew (2003)	National norm	A 6–19	131–468 (14)	WJ III	Broad narrow	MCS MR	WJ III C	MR[f]
5. McGrew (2007)	National norm	A 6–19	892–1,742 (3)	WJ III	Narrow	BRS RC BMS MR	WJ III T	MR
6. Miller (2000)[d]	Middle school RD	A 11–14	30 (1)	WJ-R/DAS/ SB-IV/WISC-III	Broad narrow	BRS RC	WJ III T	MR
7. Ganci (2004)[d]	Elem. RD/NRD	A 6–12	25/25 (2)	WISC-III/NEPSY	Broad	BRS RC	RD/NRD G	MANCOVA/DFA
8. Proctor, Floyd, & Shaver (2005)	Low/Avg BMS/MR	A 6–18	52–68 (4)	WJ III	Broad	BMS MR	WJ III C	ANOVA
9. Floyd, Bergeron, & Alfonso (2006)	Low RC/Low Ach/Avg Ach	G 2–12	28–50 (3)	WJ III	Broad narrow	RC	WJ III C	ANOVA
10. Hale, Fiorello, Dumont, Willis, Rackley, & Elliott (2008)[d]	Linking MLD & normals	A 6–17	371/42 (2)	DAS-II	Broad narrow	BMS MR	WIAT-II T	RCA
Latent variables – *g* included								
11. McGrew, Keith, Flanagan, & Vanderwood (1997)[e]	National norm	G 1–12	222–255 (10)	WJ-R	*g* broad narrow	BRS RC MR	WJ-R T	SEM(1)
12. Keith (1999)[d]	National norm	G 1–12	100–846 (9)	WJ-R	*g* broad narrow	BRS RC BMS MR	WJ-R T	SEM(1)
13. Flanagan (2000)	Elem. normals	G 3–4	166 (1)	WISC-R/WJ-R	*g* broad narrow	BRS RC	WJ-R T	SEM(1)
14. Vanderwood, McGrew, Flanagan, & Keith (2002)[e]	National norm	G 1–12	222–255 (10)	WJ-R	*g* broad narrow	BRS RC	WJ-R T	SEM(1)

(Continued)

Table 1
Continued

Study Category, Number, Authors, Date	Type of Sample[a]	Age (A) or Grade (G) Range	Sample *n* (# Samples Analyzed)	Cog. Batteries and CHC IVs	CHC IVs Strata Included	Ach DV Domains	Type of Ach. DV[b]	Analysis Method[c]
15. Floyd, Keith, Taub, & McGrew (2007)	National norm	A 5–19	110–1,007 (8)	WJ III	*g* broad narrow	BRS	WJ III C	SEM(2)
16. Taub, Floyd, Keith, & McGrew (2008)	National norm	A 5–19	110–1,007 (8)	WJ III	*g* broad	BM	WJ III C	SEM(2)
17. McGrew (2008)	Elem normals	G 3–5	148 (1)	WISC-III/WJ III	*g* broad	BRS RC BMS MR	WJ III C	SEM(1,2)
18. Benson (2008)	National norm	G K–12	269–275 (6)	WJ III	*g* broad narrow	BRS RC RF	WJ III C	SEM(3)
19. Bensen & Moseley (2009)	National norm	A 5–19	245–1,105 (8)	WJ III	*g* broad	BMS MR	WJ III C	SEM(3)

Notes. Studies are organized into two broad categories as a function of the type of variables analyzed: studies that analyzed MVs and did not include a full-scale (*g*-type) IQ score and studies that analyzed LVs and included a *g* factor (see Keith, 2006). Studies 1, 2, 11, 12, and 14 all used the WJ-R standardization data and are thus not independent samples. They either focused on different Ach DVs (Rdg vs. Math) or employed different data analytic methods or IV→DV models. A similar situation exists for WJ III studies 3–5, 15, 16, 18, and 19. A similar situation exists for WJ III studies 3–5, 15, 16, 18, and 19.

[a] National norm are samples composed of subjects from test battery standardization samples. The two samples from Hale et al. (2008; normal and math LD-MLD) were drawn from DAS-II/WIAT-II special linking sample.

[b] C = composite or cluster; T = individual test; G = group membership.

[c] MANCOVA = multivariate analysis of covariance; DFA = discriminant function analysis; ANOVA = analysis of variance; RCA = regression commonality analysis; SEM = structural equation modeling. A more detailed discussion of these methods, including visuographic representations of each approach, are available at http://tinyurl.com/nhjr23

[d] All studies, except four, included indicators from seven CHC cognitive domains (*Gf, Gc, Glr, Ga, Gv, Gsm, Gs*). Miller (2000) did not include *Gv*; Ganci (2004) did not include *Gv, Gf*; Hale et al. (2008) did not include *Ga*; Keith (1999) did not include *Gv, Glr*.

[e] McGrew et al. (1997; study #11) presented summary results of separate reading and math analyses. The detailed reading results were later published in Vanderwood et al. (2001; study #14) and are coded for study #14 and not study #11. The detailed math results were never published—these results are coded for study #11.

[f] Stepwise (backward stepping) multiple regression.

cognitive assessment batteries providing the IVs, CHC strata of IVs,[5] achievement DVs, type of achievement DV, and type of analyses (see notes to Table 1 for additional information). Inspection of the 19 investigations reflected a division between studies that analyzed observable or *manifest* IVs (and no full scale or *g* score) and those that analyzed *latent* IVs (with or without a latent *g* factor).[6]

The majority of investigations reported results for more than one sample. For example, in Study 1 (McGrew, 1993), 15 subsamples (sample *n* range of 59–325) of the WJ-R norm data were analyzed separately with multiple regression methods across the ages of 5–19 years. From this point forward, however, we refer to the *number of analyses* and not the *number of samples*. This distinction is important given that many analyses used the same norm subjects (e.g., WJ-R or WJ III standardization data) but divided the same subject pool into different age groups, analyzed the same set of subject data but employed different analysis methods, or analyzed test versus cluster scores with the same subject pool (see second note to Table 1). Thus, many of the *samples* were not independent—but the analyses were. There were a total of 134 analyses—54.5% were manifest variable (MV; $n = 73$) and 45.5% were latent variable (LV; $n = 61$).

Of the 134 analyses, 126 (94%) were exclusive to the WJ-R or WJ III norm data. Given the preponderance of WJ-R and WJ III studies, the coded age ranges were driven, in large part, by the age ranges reported in these investigations. The result was the division of the analyses into the age groups of 6–8, 9–13, and 14–19 years.[7] For investigations that analyzed a finer gradation of development (e.g., McGrew, 1993, reported regression results for each year from ages 5–19), the appropriate subsamples (e.g., results for ages 5–8) were treated as a single age group (ages 6–8).[8] Certain studies (Studies 7–10) were based on a single wide-range sample (e.g., Ganci, 2004; samples of elementary reading disabled (RD) and non–reading disabled (NRD) students across ages 6–12), thus these single sets of results were included in each of the broad three age groups for which the single sample included subjects. This resulted in some single wide-age samples being reported in two or three of the three major age-grouped analyses.

Coding and Summarization of Analysis Results

For each of the three broad age categories, analyses were coded by the four achievement domains (BRS, RC, BMS, and MR). This coding produced 12 separate CHC-ACH domain summary tables (3 age groups × 4 achievement domains). One sample coding summary analysis (BRS for ages 6–8) is presented in Table 2. The total number of analyses included by achievement domains, ordered by

[5] Classification of IVs as g (general; stratum III), broad (stratum II), or narrow (stratum I) was based on the current authors' analysis and not necessarily that reported by the original researchers. For example, it was a common practice in early CHC research literature to refer to certain measures or factors as reflecting broad abilities (e.g., WJ-R *Ga* cluster) when, in retrospect, later CHC literature recognized that these measures or factors represented narrow CHC abilities (e.g., WJ-R *Ga* cluster was subsequently determined to be measuring the narrow *Ga* ability of PC; see McGrew & Flanagan, 1998 and McGrew &Woodcock, 2001). Another example of IV stratum reclassification was our reclassification of the Hale, Kaufman, Naglieri, & Kavale (2006) *Glr* IV as measuring the narrow *Glr* ability of free recall memory (M6) because this was a single test measure from the DAS (viz., Recall of Objects) and did not measure two or more qualitatively different narrow CHC abilities within the respective broad domain.

[6] See Keith (2006) for the definitions and differences between manifest (MV) and latent variables (LV).

[7] Samples described by grade level (vs. age) were reclassified by age using the following general scheme (6–8 years = Grades K-3/4; 9–13 years = Grades 4/5–8; 14–19 years = Grades 9–12).

[8] Studies 1–4 used a similar multiple regression approach in which year-by-year standardized regression weights were plotted (as a function of age), and a smooth polynomial curve was fit to the complete set of weights across all ages. For the current review we divided the smoothed curves into the three broad age groups and classified the IVs as significant if the smoothed population parameter curve estimates were significant 50% or more of the time across the broad age group.

Table 2
Summary of CHC Cognitive-Basic Reading Skills Studies: 6–8 Years

		General, Broad and Narrow CHC Abilities Included in Studies[b]																						
	Sample or	Brd																			*Gc*			
Study[a]	Subsample	Rdg	*Gs*	P	RE/R4	AC/EF	*Gsm*	MW	MS	*Ga*	PC	US/UR	*Glr*	MA	NA	MM	*Gf*	I	RG	RQ	LD/VL	K0	LS	VL
Manifest variables																								
1. McGrew (1993)	a. 6–8 yrs[c]			X					X		X			O			O				X			
3. Evans et al. (2002)	a. 6–8 yrs[c]		X				X	X		X	X		X				O				X			
5. McGrew (2007)	a. 6–8 yrs[d]			X	O	O		X	X		X	O		X	O	O		O	O	X	X	X	X	O
	b. 6–8 yrs[e]			X	O	O		X	O		X	O		X	O	O		O	O	X	X	O	O	O
7. Ganci (2004)	6–12 yrs[h]		O						O	X			O								X			
Latent variables	**#s / #t**		**1/2**	**3/3**	0/2	0/2	**1/1**	**3/3**	**2/4**	**2/2**	**4/4**	0/2	**1/2**	**2/3**	0/2	0/2	0/2	0/2	0/2	**2/2**	**5/5**	**1/2**	**1/2**	0/2
12. Keith (1999)	a. 1–4th gr[f]	X		O					O		X						O				O			
13. Flanagan (2000)	3–4th gr[f]	X		O					O		X			O			O				O			
14. Vanderwood et al.	a. 1–2nd gr[f]	X		O					O		X			O			O				O			
(2002)	b. 3–4th gr[f]	X		O					O		X			O			O				O			
15. Floyd et al.	a. 5–6 yrs[c]		X				O			O			X				O				O			
(2007)[g]	b. 7–8 yrs[c]		X				X			O			O				O				X			
	c. 5–6 yrs[c]		X				O	O	O	O	O	O	X	X	O		O				O	O	O	O
	d. 7–8 yrs[c]		X				O	O	O	O	O	O	O	O	O		O				X	O	X	O
18. Benson (2008)	a. K–3rd gr[c]		O				O			O	O		O		O		O				O			
	#s/#t	**4/4**	**4/5**	0/4			1/5	0/2	0/6	0/5	**4/7**	0/2	2/5	1/5	0/3		0/9				2/9	0/2	**1/2**	0/2
	Grand #s/#t	**4/4**	**5/7**	3/7	0/2	0/2	2/6	**3/5**	2/10	2/7	**8/11**	0/4	3/7	3/8	0/5	0/2	0/11	0/2	0/2	2/2	**7/14**	1/3	**2/4**	0/4

Notes. *Gv* domain was eliminated from table to save space. All *Gv* results were nonsignificant. Complete table available on the Internet (see Footnote 11). X = significant effect/relation reported; O = no significant effect/relation reported for cognitive ability that was included as an IV. Blank space indicates that cognitive ability was not included as an IV. #s/#t = # times cognitive ability was significant/total # of times cognitive ability was included in analysis. 50+% in bold font.
[a] See Table 2 for summary of study characteristics.
[b] See Newton & McGrew (current issue) for definitions of broad and narrow CHC abilities.
[c] DV was WJ-R or WJ III BRS cluster or LV defined by the WJ-R/WJ III tests (Letter-Word Identification; Word Attack) that compose the BRS cluster.
[d] DV was WJ-III Letter-Word Identification test.
[e] DV was WJ III Word Attack test.
[f] WJ-R/WJ III Letter-Word Identification and Word Attack tests represented separate DV (LVs) in a single SEM model. Significance (X) was recorded for a cognitive ability if it was significantly associated with either test (or both).
[g] Floyd et al. (2007) samples a/c and b/d based on the same samples of subjects, but SEM models included different number of LVs at different strata. Samples a/b evaluated g + broad models. Samples c/d evaluated g + broad + narrow models.
[h] DV as classification of subjects as RD or NRD in basic reading skills, reading comprehension, or both. Thus, Ganci (2004) is included in both the BRS and RC summary tables.

the three age groups (6–8, 9–13, and 14–19 years) were: BRS (14, 14, 9); RC (11, 13, 8); BMS (12, 12, 10); and MR (13, 14, 12).[9]

As reported in Table 2, for each analysis a "vote tally" method of CHC IV significance was recorded. The reported general, broad, or narrow IVs were (a) recorded as significantly related to the achievement domain (X), (b) included in the study but were not significantly related to the achievement domain (O), or (c) not included as an IV in the analyses (blank space). A significant IV relation was designated if it was reported as significant (t test; F test; regression weight; standard error of the mean [*SEM*] effect size) by the studies' researchers.[10] The significance findings were tabulated separately for the MV and LV analyses, as well as across MV and LV studies (MV+LV). A ratio of number of significant findings (*#s*) to total number of analyses which included the IV (*#t*) was recorded. The "*Grand #s/#t*" (see Table 2) was the primary metric analyzed in the current synthesis. For example, across the seven BRS analyses (at ages 6–8) that included a broad *Gs* IV, five were significant (5/7; see Table 2). Each *Grand #s/#t* ratio was converted to a *percent of significant analyses* value, reflecting the percentage of times that a CHC IV was significantly related to the achievement DV in studies in which the IV was included.

The *percent of significant analyses* for all CH IVs were organized in graphs (see Figure 1).[11] To facilitate interpretation of the summary figures, results were plotted only for CHC IVs that were based on at least two analyses and that had percentages greater than or equal to 20%. Results were further classified as per consistency of significance: high (≥80%), medium (50%–79%), low (30%–49%), or tentative/speculative.[12] Our operational definition of "consistency of significance" was constructed logically based on the post hoc examination of group "breaks" or trends in the summary figures. We recognize that these operational criteria may be too arbitrary or liberal for some, whereas others may argue that they are too conservative. Other scholars are free to invoke different criteria and reinterpret the summary coding tables (see Footnote[9]).

Results and Discussion

The most consistent CHC COG-ACH relation findings for each reading and math subdomain as a function of breadth of the cognitive predictor variables (broad vs. narrow) are discussed in this section. The tentative/speculative findings (particularly at the narrow ability level) and the lack of significance for some abilities (e.g., the "*Gv* mystery") are discussed in a later section.

[9] All 12 summary coding sheets are available at http://tinyurl.com/nord9r

[10] The initial goal was to conduct a meta-analysis with a common effect size metric. Due to the limited number of research investigations, the lack of sufficient reported statistics for the calculation of effect sizes in some manuscripts, lack of independence of samples, and variety of statistical procedures used and results reported, the use of a common meta-analytic effect size metric was deemed impossible. Instead, a relationship was defined as significant (t test; F test; regression weight; SEM effect size) if so reported by the studies' researchers. No attempt was made to separate *direct* and *indirect* (see Keith, 2006) effects for IVs—an IV was considered significant if the combined *total* (direct + indirect) effect was deemed significant in each respective analysis. In analyses where only direct IV effects were reported, we calculated the indirect effects and then used the combined total effect value to determine if a significant relation was present. In analysis that reported split-sample model calibration and validation samples for the same age (e.g., Floyd, Keith, Taub, & McGrew, 2007), only the results from the cross-validated validation sample were used to determine significance. Keith (1997) reported separate IV effect sizes for White, Black, and Hispanic samples. Via multiple group confirmatory factor analysis (MGCFA), Keith (1997) reported the models to be invariant across groups. The median effect sizes across the three subsamples were used for significance classification in this review.

[11] To save space, only one summary figure is included for illustrative purposes. All four summary figures are available at http://tinyurl.com/nord9r

[12] Tentative or speculative results were those that were (a) between 20% and 29% in consistency, (b) based on a small number of analyses (e.g., $n = 2$), and/or (c) based only on McGrew's (2007) exploratory multiple regression analysis of WJ III MVs at the individual IV test level.

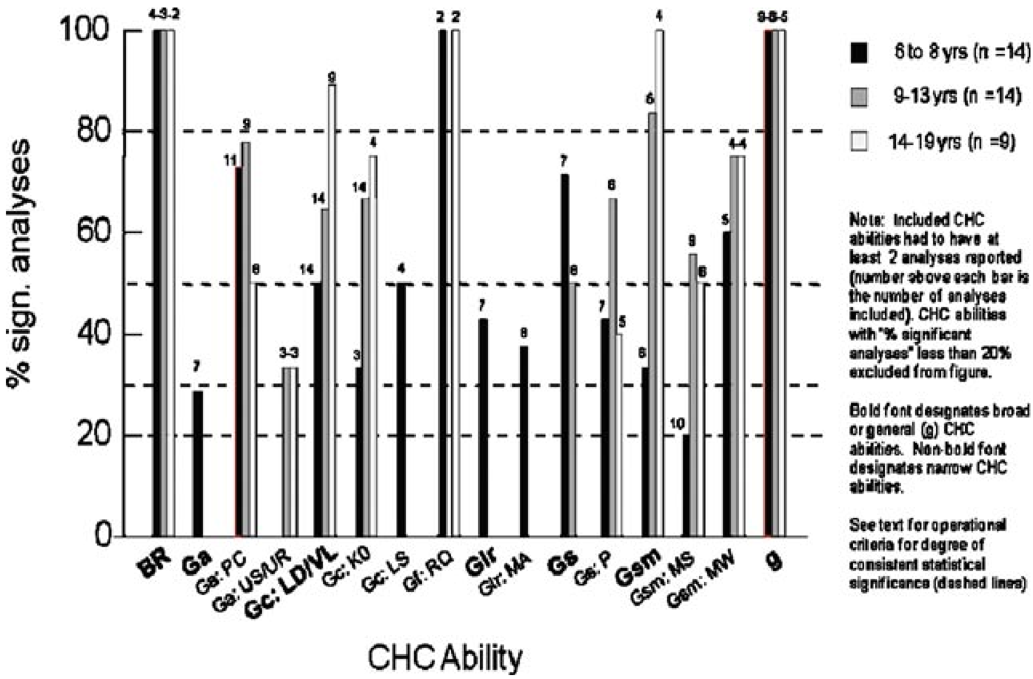

FIGURE 1. % of Significant Analyses in Prediction of Basic Reading Skills by CHC Abilities by Three Age Groups.

BRS

BRS are primarily decoding and word recognition skills. Tests measuring BRS typically include tasks such as reading real or nonsense words, an application of the alphabetic principle—knowing how sounds and symbols go together.

Broad CHC Abilities and BRS. The following broad CHC cognitive abilities were consistently significant (low, medium, or high) in the prediction of BRS (see Figure 1) at one or more age group: comprehension-knowledge (*Gc*)[13] (medium at ages 6–13, high at ages 14–19); long-term retrieval (*Glr*) (low at ages 6–8); processing speed (*Gs*) (medium at ages 6–13); and short-term memory (*Gsm*) (low at ages 6–8, high at ages 9–19).

It is not surprising that *Gc* has strong BRS relations as ample evidence exists that general language and vocabulary development, aspects of *Gc,* are necessary for acquiring reading skills (Cooper, 2006; Shaywitz, Morris, & Shaywitz, 2008; Torgesen, 2002; Vellutino, Tunmer, Jaccard, & Chen, 2007). Also, the ability to form, store, and retrieve sound–symbol relations and efficiently retrieve lexical and general knowledge (*Glr*) has been linked to early reading development (Cooper, 2006; Perfetti, 2007; Shaywitz et al., 2008; Vellutino et al., 2007). The importance of *Gsm* for reading

[13] As indicated in Table 1 and in all figures, broad *Gc* was coded as consisting of the combination of the narrow abilities of LD and VL. This reflects the fact that the WJ-R *Gc* cluster and the WJ III Verbal Comprehension test, which were the primary *Gc* variables included in this review, have been noted to be combined measures of LD and VL (see Flanagan, Ortiz, Alfonso, & Mascolo, 2006; also see McGrew & Woodcock, 2001).

is consistent with research that has implicated auditory/verbal/phonological working or short-term memory (Cooper, 2006; Hammill, 2004; Kintsch, 2005; Kintsch & Rawson, 2005). Finally, different researchers (Berninger, Abbott, Thomson, & Raskind, 2001; Kintsch, 2005; Kintsch & Rawson, 2005; Shaywitz et al., 2008; Wolf, Bowers, & Biddle, 2000) have emphasized a variety of speed or fluency constructs in early reading skill acquisition (e.g., rapid automatic naming; naming speed; speed of semantic or lexical access; verbal efficiency; automaticity), findings consistent with the significant *Gs* finding in this review. The consistency of significance of *Glr* and *Gs* tend to decline with age whereas the influence of other abilities, such as *Gc* and *Gsm,* increase with age (see Figure 1).[14] It is likely that *Glr* and *Gs* are more important during the beginning stages of reading when basic skills are first being acquired and, after BRS are in place, *Gc* and *Gsm* become more important in the development of reading.

The broad abilities of auditory processing (*Ga*), fluid reasoning (*Gf*), and visual processing (*Gv*) were not consistently significantly related to BRS. Most surprising may be that *Ga* did not meet the criteria for low, medium, or high significance at any of the ages. The reason for the lack of broad *Ga* significance is apparent when one examines the results of the research at the narrow ability level.

Narrow CHC Abilities and BRS. *Ga*-PC was classified as medium at all three age levels (see Figure 1), a finding supporting the importance of phonemic awareness in BRS, despite the lack of significance for a broad *Ga*/BRS relationship. This finding is consistent with research (Berninger et al., 2006; Cooper, 2006; Shaywitz et al., 2008; Torgesen, 2002) indicating that awareness of sounds is a prerequisite skill for mastering the alphabetic principle in reading (e.g., Adams, 1990; Ehri, 1998) and that a phonological core deficit exists in many individuals with dyslexia (e.g., Morris et al., 1998; Stanovich & Siegel, 1994).

The broad *Gc*/BRS significant findings appear related to three different narrow *Gc* abilities. General information (*Gc*-K0) was consistently related to BRS at all ages, with a trend toward increased importance with increasing age (low at ages 6–8; medium at ages 9–19). This finding is consistent with the importance of prior background knowledge, knowledge integration, and general fund of cultural knowledge in reading (Cooper, 2006; Kintsch & Rawson, 2005). *Gc*-LS was classified as medium at the youngest age group (6–8 years), a finding consistent with research that has implicated the ability to comprehend spoken language (i.e., listening comprehension) in reading development (Hoover & Gough, 1990; Joshi & Aaron, 2000).

Three different narrow memory abilities (*Gsm* and *Glr*), particularly those of a short-term nature (*Gsm*), appear important for BRS. Memory span (*Gsm*-MS) was not significant at ages 6–8 but was medium for ages 9–19. In contrast, working memory (*Gsm*-MW) was consistently classified medium at all age levels (see broad *Gsm* discussion for references regarding the importance of memory span and working memory in reading). *Glr*-MA was only related (low) to BRS at the youngest ages (6–8 years), a finding consistent with research that has demonstrated that paired-associate learning, such as learning phoneme–grapheme relationships (Cooper, 2006; Hammill, 2004; Perfetti, 2007; Shaywitz et al., 2008), makes a unique contribution to predicting early reading (Windfuhr & Snowling, 2001).

The driving force behind the broad *Gs*/BRS finding appears to be the influence of the narrow perceptual speed (*Gs*-P) ability, which was consistently significant at all ages (low at ages 6–8 and 14–19; medium at ages 9–13). The importance of *Gs*-P is not surprising given the confirmed

[14] When discussing possible age-related or developmental trends, it is important to note that it is the consistency of significance that is being discussed.

relationship between perceptual speed, speed of processing, and the need for automaticity in integrating phonological and orthographic codes in word reading (e.g., Barker, Torgesen, & Wagner, 1992; Berninger, 1990; Hale & Fiorello, 2004; Joshi & Aaron, 2000; Urso, 2008).

RC

RC is constructing meaning from text through a complex process that integrates multiple linguistic factors. Tests measuring RC include tasks that require word knowledge and understanding passages usually demonstrated through multiple-choice, open-ended, or cloze procedures.

Broad CHC Abilities and RC. Several broad CHC abilities were consistently significant (low, medium, or high) in predicting RC at one or more of the age groups: auditory processing (*Ga*) (medium at ages 6–8), comprehension-knowledge (*Gc*) (high at all ages), long-term retrieval (*Glr*) (low ages 9–13), and short-term memory (*Gsm*) (low at ages 6–8 and 14–19).

Broad *Gc* displayed the most consistent relation to RC across all age levels, a finding reinforced by studies that have demonstrated the important role of general language development and vocabulary (e.g., Baker, Simmons, & Kame'enui, 1995; Coyne, Simmons, Kame'enui, & Stoolmiller, 2004; Jenkins, Fuchs, Van den Broek, Espin, & Deno, 2003; Perfetti, 2007) and prior knowledge (e.g., Anderson & Pearson, 1984; Kintsch & Rawson, 2005; Nation, 2005) in RC. In fact, Floyd, Bergeron, and Alfonso (2006) reported broad *Gc* levels as the primary difference between individuals with good and poor comprehension. Research supporting the significant relations for *Ga* and *Glr* abilities with reading, in this case RC, was discussed in the previous BRS section and will not be repeated here.

Broad abilities not consistently significant at any of the three age groups include processing speed (*Gs*), fluid reasoning (*Gf*), and visual processing (*Gv*). *Gs* was classified, however, as tentative/speculative at the younger ages (6–13), which is consistent with Keith's (1999) research. The *Gf* tentative/speculative classification, only in the oldest age group (14–19 years), is suggestive of *Gf* involvement at higher levels of RC, a finding consistent with other research linking *Gf* and RC (e.g., Floyd et al., 2006; McGrew, 1993; Nation, Clarke, & Snowling, 2002). The tentative/speculative RC findings for *Gs* and *Gf* may be partially explained by the narrow abilities in the current research synthesis.

Narrow CHC Abilities and RC. Given the complex cognitive demanding nature of RC (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Perfetti, 2007), it is not surprising that working memory (*Gsm*-MW) displayed a high classification at all ages. Memory span (*Gsm*-MS) was consistently significant (medium) only at ages 14–19. We hypothesize that this finding is related to the inclusion of longer items (lengthy oral sentences) at the top end of most MS tasks, items which increase the demand for listening comprehension or *Gc*-LS versus simpler rote MS. Listening comprehension is frequently cited as a good predictor of RC (Aaron & Joshi, 1992; Cooper, 2006). The importance of broad *Gc* for RC appears due to the narrow abilities of general information (*Gc*-K0) and *Gc*-LS, which were consistently significant (high for all age groups).

Although broad *Ga* was consistently significant (medium) only at ages 6–8, *Ga*-PC was implicated at all ages (low at ages 6–8 and 14–19; tentative/speculative at 9–13). A possible explanation of these age-differentiated findings is the observation (from inspection of the online summary coding table for RC) that six of the nine significant PC-RC analyses came from McGrew's (2007) multiple regression MV study that used the WJ III tests (vs. clusters) as IVs. The WJ III Sound Awareness test has been reported to be a possible mixed measure of PC (*Ga*-PC) and working memory (*Gsm*-MS; see Schrank, 2006). When Sound Awareness was significantly related to a WJ III achievement DV, McGrew (2007) double-coded it as reflecting significance for both PC and MW.

Thus, it is possible that the *Ga*-PC/RC significance may be related to influence of *Ga*-PC abilities at the youngest ages (6–8) and the more complex *Gsm*-MW component at the older ages (14–19).

Again demonstrating the importance of narrow (vs. broad) cognitive abilities in understanding achievement are the *Glr* and *Gs* broad and narrow ability findings. Meaningful memory (*Glr*-MM) had a high classification at ages 9–19; *Glr*-NA was medium at ages 9–13 and low at ages 14–19; and broad *Glr* was only salient (low) at ages 9–13. Individuals with RC difficulties often display problems with verbal fluency, word retrieval, NA, or speed and quality of lexical access (e.g., Kintsch & Rawson, 2005; Nation, Marshall, & Snowling, 2001; Shaywitz et al., 2008), which is consistent with the *Glr*-NA/RC findings reported here. Similarly, despite broad *Gs* not being identified as being consistently significant at any age, narrow perceptual speed (*Gs*-P) was significant at all ages (medium, medium, low). A strong relationship between RC and quick and automatic processing of letters and words (fluency) and word reading speed (e.g., Fuchs, Fuchs, Hosp, & Jenkins, 2001; Jenkins et al., 2003) is consistent with the strong perceptual speed (*Gs*-P) and RC findings.

BMS

BMS include arithmetic and computational skills. Tests measuring BMS usually include tasks ranging in difficulty from basic math facts to solving more complex algorithmic computations.

Broad CHC Abilities and BMS. The following broad CHC cognitive abilities were consistently significant (low, medium, or high) in the prediction of BMS at one or more age groups: comprehension-knowledge (*Gc*) (medium at ages 9–19); fluid reasoning (*Gf*) (medium at all ages); and processing speed (*Gs*) (medium at all ages). Language skills (aspects of broad *Gc*) have been linked to math (e.g., Fiorello & Primerano, 2005; Flanagan et al., 2006; Floyd, Evans, & McGrew, 2003; Fuchs et al., 2006; McGrew, 2008; Swanson & Jerman, 2006) especially as *Gc* relates to the development of number concepts (e.g., Carey, 2004; Gelman & Butterworth, 2005) and the retrieval of math facts (Chong & Siegel, 2008). The importance of *Gf* in the prediction of math achievement is consistent with considerable research (e.g., Fiorello & Primerano, 2005; Flanagan et al., 2006; Floyd et al., 2003; Fuchs et al., 2006; Geary, 1993, 2007; Rourke & Conway, 1997). The *Gs*-to-arithmetic skill link (Bull & Johnston, 1997; Fiorello & Primerano, 2005; Floyd et al., 2003; Fuchs et al., 2006) has been explained in terms of counting speed (Geary, 1993, 2007), numerical processing fluency (Swanson & Jerman, 2006), and efficiency and consistency of execution of simple cognitive tasks during math (Fuchs et al., 2006; Geary, 1993, 2007).

The broad abilities of *Ga, Gv, Glr,* and *Gsm* were not consistently significant at any age, findings inconsistent with prior research. *Glr* and *Gsm* were found to be predictive of math skills throughout childhood and adolescence by Floyd et al. (2003). Semantic retrieval, an aspect of *Glr,* has also been reported to contribute to math disabilities in college students (Cirino, Morris, & Morris, 2002). Furthermore, the ability to fluently retrieve math facts from memory, a *Glr* function, is the most consistent BMS deficit associated with MD (e.g., Garnett, Frank, & Fleischner, 1983; Geary, 1990, 1993; Geary, Hamson, & Hoard, 2000; Goldman, Pellegrino, & Mertz, 1988). Most of the above inconsistencies can be further explained and, to some extent, resolved by inspecting the narrow ability findings for BMS.

Narrow CHC Abilities and BMS. As was the case for BRS and RC, the current findings suggest that the narrow cognitive abilities may play an important role in the prediction of basic math achievement even when the corresponding broad ability does not. For example, broad *Ga* was not significant at any age group, but *Ga*-PC displayed a medium level of consistent significance at ages 6–13 and was tentative/speculative at ages 14–19. Phonological processing (*Ga*-PC in CHC taxonomy) has been reported to predict arithmetic achievement (e.g., Leather & Henry, 1994;

Rasmussen & Bisanz, 2005) and is associated with MD and low math achieving children with fact fluency deficits (Chong & Siegel, 2008). Phonological processing has been hypothesized to influence computation skills because speech sound processes are used when solving problems (Bull & Johnston, 1997; Geary, 1993; Rourke & Conway, 1997), and counting requires retrieval of the phonological codes for number words (Geary, 1993; Logie & Baddeley, 1987).

Perceptual speed (*Gs*-P) was consistently significant at a high level for all age groups. The *Gs*-P/BMS link may be a function of "subitizing," which is the ability to instantly see "how many" without counting. Subitizing underlies development of fluency with math facts, which is thought to be a core deficit in MD (e.g., Geary, 1993; Jordan, Hanich, & Kaplan, 2003). Alternatively, because all the significant *Gs*-P findings involved the WJ III Visual Matching test (either alone or when combined with another perceptual speed test; Keith, 1999; McGrew, 2007; McGrew & Hessler, 1995), a test requiring rapid recognition of similar digit pairs in a line of numbers, an alternative explanation may be that the *Gs*-P/BMS may be due to number facility (*Gs*-N) ability. Working memory (*Gsm*-MW), which was consistently significant (high for all ages) is frequently cited as important for BMS (e.g., Geary, 1993, 2007; Geary, Brown, & Samaranayake, 1991; Hitch & McAuley, 1991; Passolunghi, Mammarella, & Altoè, 2008; Swanson & Jerman, 2006) and is also suggested as a core deficit for individuals with MD (e.g., Bull, Espy, & Wiebe, 2008; Chong & Siegel, 2008; Geary, 2003; Koontz & Berch, 1996; Passolunghi & Siegel, 2001).

MR

MR is problem-solving skills in math. Tests measuring MR typically include word problems, number series, concepts, and application of mathematical operations and concepts. Achievement in MR will depend, to some extent, on proficiency with BMS.

Broad CHC Abilities and MR. Several broad CHC abilities were consistently significant (low, medium, or high) in predicting MR at one or more of the age groups: comprehension-knowledge (*Gc*) (low at ages 6–8, medium at ages 9–13, high at ages 14–19); fluid reasoning (*Gf*) (high at ages 6–13 and medium at ages 14–19); processing speed (*Gs*) (medium at ages 6–13); and short-term memory (*Gsm*) (low at ages 14–19). Much research supports the strong predictive nature of *Gc* for mathematics (e.g., Hale, Fiorello, Kavanagh, Hoeppner, & Gaither, 2001; Keith, 1999; McGrew, Flanagan, Keith, & Vanderwood, 1997) and the association between MD and limited oral language abilities (Fuchs et al., 2008; Proctor, Floyd, & Shaver, 2005). Furthermore, the importance of *Gc* increased with age, a finding consistent with the increased linguistic demands of more complex MR tasks (e.g., Fuchs et al., 2006, 2008; Geary, 1994).

The *Gf* finding is consistent with a body of research that identifies *Gf* as important to MR (e.g., Fiorello & Primerano, 2005; Floyd et al., 2003; Fuchs et al., 2006; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; McGrew, 2008; Rourke & Conway, 1997). Fluid reasoning is a strong predictor of math achievement (Hale, Fiorello, Kavanagh, Holdnack, & Aloe, 2007) and is often significantly impaired in individuals with MD (Geary, 2007; Proctor et al., 2005). Finally, *Gs* has also been identified as an important predictor of MR (e.g., Floyd et al., 2003; Geary, 2007; Swanson & Jerman, 2006), possibly facilitating reasoning by freeing up cognitive resources in working memory for higher-level complex processing and thinking (e.g., Bull & Johnston, 1997; Fuchs et al., 2006).

The current review did not identify broad *Ga, Glr,* or *Gv* as consistently significant for MR at any age group. Whereas these nonsignificant findings are consistent with the review by Flanagan and colleagues (2006), other research is contradictory. Floyd and colleagues (2003) reported a moderate relationship between math achievement and *Glr* and *Ga*. As will be discussed next, the current review found several narrow *Glr* or *Ga* abilities to be significant predictors of MR even though the broad cognitive abilities were not.

Narrow CHC Abilities and MR. *Ga*-PC was classified as medium in consistency of significance at ages 6–8 and low for ages 9–19. A number of studies have implicated the phonological system as underlying individual differences in math problem-solving (e.g., Furst & Hitch, 2000; Gathercole & Pickering, 2000; Geary & Brown, 1991; Swanson & Sachse-Lee, 2001). As noted for BMS, research has implicated the importance of perceptual speed (*Gs*-P), such as rapid processing of numbers, and counting speed, for math performance (e.g., Fuchs et al., 2006; Geary et al., 2007; McGrew, 2008). This implication was supported in the current review when perceptual speed (*Gs*-P) displayed a medium level of consistent significance with MR across all age groups.

Two narrow *Gsm* abilities were important in the prediction of MR at one or more age groups: memory span (*Gsm*-MS) was low for ages 6–8 and working memory (*Gsm*-MW) was high at all ages. Other research supports these findings as memory span (*Gsm*-MS) has been related to performing mental arithmetic and how quickly numbers can be counted (Geary, 1993; Holmes & Adams, 2006). Considerable research confirms the importance of working memory (*Gsm*-MW) for MR (e.g., Fuchs et al., 2008; Geary, 2007; Passolunghi, 2006). Both the phonological loop (e.g., Furst & Hitch, 2000; Gathercole & Pickering, 2000; Geary, 2007) and visuospatial sketchpad (e.g., Geary, 2007; Holmes & Adams, 2006), components of working memory (*Gsm*-MW), contribute to MR performance.

The "Gv Mystery" and Tentative/Speculative CHC COG-ACH Findings

In this review, as in most other prior CHC and non-CHC COG-ACH reviews, broad visuospatial processing (*Gv*) was the "odd man out" as it was not found to predict reading or math achievement. How can this be when other research has implicated certain *Gv* abilities (e.g., visuospatial; length estimation) as a core deficit in MD and as important for success in mathematics (Geary, 1993, 2007; Hale et al., 2008; Osmon, Smertz, Braun, & Plambeck, 2006; Pinel, Piazza, Le Bihan, & Dehaene, 2004; Rourke, 1993; Zorzi, Priftis, & Umiltá, 2002)? Below is a brief discussion of possible explanations, together with two tentative/speculative *Gv* narrow ability findings that did emerge, which is followed by a brief discussion of the other tentative/speculative findings found in this review, findings that must be viewed as suggestive and in need of additional research, particularly because most of these findings are based on single WJ III tests from a handful of analyses.

Specification error, as described earlier in this article, may be present in the current review. It is possible that those *Gv* abilities related to academic learning simply are missing from the current collection of intelligence batteries used in school achievement research. The types of *Gv* tests in current intelligence batteries (e.g., block design, spatial relations; memory for designs or pictures) may not measure the *Gv* abilities important for reading and math. For example, the visual aspects of orthographic processing or awareness (the ability to rapidly map graphemes to phonemes; rapid processing of visual symbols; etc.) have been reported as important for reading (e.g., Barker, Torgesen, & Wagner, 1992; Berninger, 1990; Berninger et al., 2006; Flanagan et al., 2006; Hale & Fiorello, 2004; Urso, 2008) and are absent from intelligence batteries. Additionally, more complex visuospatial processing (not measured by current intelligence tests) may be important for school learning, such as *Gv* tasks that measure complex visuospatial working memory (e.g., see Holmes, Adams, & Hamilton, 2008; Maehara & Saito, 2007; Mammarella, Pazzaglia, & Cornoldi, 2008). It is also possible that the *Gv* mystery may be a DV or criterion variable problem. The math achievement DV measures used in the extant CHC COG-ACH research may not tap the higher-level mathematics (e.g., geometry, trigonometry, calculus) that draw heavily on *Gv* abilities.

Also, lack of significance does not mean that *Gv* abilities are not involved in reading and math. Obviously, individuals use their eyes when reading and when processing diagrams and figures during reading and math. *Gv* measures, as currently designed in intelligence batteries, simply may have no achievement variance to account for because the more powerful predictors (e.g., *Gc, Gsm, Ga*) account for the lion's share of reliable variance in the achievement variables. To illustrate this concept we will use breathing. One must breathe to read and do math, yet breathing oxygen is not included as a predictor in COG-ACH research as it would not emerge as statistically significant. Like breathing, basic *Gv* processes may function as a threshold ability—you need a minimal amount to read and perform math, but beyond the minimal threshold level "more *Gv*" does not improve performance.

Two narrow *Gv* abilities were identified as tentative/speculative in the current review. Visual memory (*Gv*-MV) was so classified at ages 14–19 for RC, possibly related to the positive effect of visual imagery on RC (e.g., Gambrell & Jawitz, 1993). Spatial scanning (*Gv*-SS) was similarly classified at ages 6–8 for BMS. Both of these isolated and tentative findings, which were based on single test indicators in a handful of studies, should be viewed with caution and warrant additional investigation.

Gf: Quantitative Reasoning (RQ) and RG and Reading and Math. It is not surprising that quantitative reasoning (*Gf*-RQ) and general sequential (deductive) reasoning (*Gf*-RG) displayed significant relations with both math domains (BMS, MR). Other research supports the relationship between mathematics and quantitative and deductive reasoning (e.g., Fuchs et al., 2005, 2006; Geary et al., 2007; Rourke & Conway, 1997). These two narrow *Gf* ability findings also suggest that the broad *Gf* relation with math achievement may be driven largely by measures of deductive reasoning (e.g., WJ III Analysis-Synthesis) and/or tests that require the use of both inductive (I) and deductive (RQ) reasoning, but only when involving reasoning with numbers and numerical relations (e.g., WJ III Number Series and Number Matrices tests). *Gf*-RQ was also related to both domains of reading (BRS and RC), adding tentative support to the role of some aspects of *Gf* in reading, especially RC.

Ga: Speech Sound Discrimination (US) and Resistance to Auditory Stimulus Distortion (UR) and Reading and Math. *Ga*-US/UR was related to BRS and MR at certain age levels. These findings were due solely to the presence of the WJ III Sound Patterns-Voice test in a handful of analyses. Research has supported, however, the importance of sound discrimination abilities in learning to read (Berninger et al., 2006; McBride-Chang, Chang, & Wagner, 1997). Additionally, research has implicated speech sound processing, temporal processing, auditory perception, and encoding and maintaining phonological representations in working memory as related to math achievement (e.g, Fuchs et al., 2005; McGrew, 2008; Rourke & Conway, 1997; Swanson & Jerman, 2006).

Attention/Concentration (AC) and Executive Functions (EF) and Math. A potentially intriguing tentative finding was the significant relation between the WJ III Pair Cancellation test (classified as AC/EF by McGrew, 2007) and BMS and MR (but not reading) at the youngest age group (ages 6–8). It has been suggested that the WJ III Pair Cancellation test taps response inhibition, interference control, and sustained attention (Cooper, 2006). Additionally, Poock (2005) reported that Pair Cancellation was one of three WJ III tests that reliably differentiated Attention-Deficit/Hyperactivity Disorder (ADHD) and non-ADHD subjects. Compromised executive functioning, including poor attention and inhibitory control, has been associated with the problems in development of math computation skills and with individuals with MD (Fuchs et al., 2006; Geary, 2007; Geary et al., 2007; McLean & Hitch, 1999; Swanson, 1993; Swanson & Jerman, 2006; Swanson & Sachse-Lee, 2001).

Gc: LS and General Information (K0) and Math. Gc-LS was consistently significant in its relation with BMS at two ages and at all ages for MR. General information (K0) was also consistently related to MR at all ages. These findings are consistent with the prior discussions of the importance of broad *Gc* and multiple narrow *Gc* abilities and mathematics.

Glr: MA, MM, and NA and Math. Although broad *Glr* was not significantly related to BMS or MR in the current research synthesis, a number of narrow *Glr* abilities were identified as tentative/speculative. *Glr*-NA was predictive of BMS at all ages, and *Glr*-MA and *Glr*-MM were predictive of BMS and MR at one or more age levels. The importance for all three narrow *Glr* abilities is consistent with the finding that MD students often have difficulty forming and later retrieving or accessing long-term memory representations of math facts (Geary, 1993, 2007). Learning math facts is a paired-associate learning task requiring MA (Geary, 2007; Osman et al., 2006). Additionally, verbal counting, an aspect of *Glr*-NA, has been mentioned as a precursor to early math achievement (Mazzocco & Thompson, 2005; Passolunghi, Vercelloni, & Schadee, 2007).

Limitations and Suggestions for Additional Research

A research synthesis of this scope, with numerous operational and procedural decisions, will have aspects that will be questioned by other researchers and will leave many important questions and issues untouched. We, too, felt constrained in not being able to discuss the results and methodological factors in greater detail. Thus, we limit our discussion of limitations and suggest additional research to four issues.

First, the dominance of the WJ-R or WJ III studies (94% of analysis based on WJ battery) suggests that the results and implications are best characterized as a referendum on WJ CHC COG-ACH relations. Generalization of the review results to other CHC-based intelligence batteries (DAS-II, KABC-II, SB-V) or assessment approaches (CBAs) should be approached with caution. Similarly classified broad and narrow CHC measures from other batteries, particularly the narrow CHC test classifications (which are primarily based on logical expert consensus methods), cannot be assumed to display the same CHC COG-ACH relation patterns reported here. For example, although empirically classified (as per a CHC-designed Wechsler Intelligence Scale for Children, Third Edition [WISC-III] + WJ III cross-battery CFA study; Phelps, McGrew, Knopik, & Ford, 2005) as narrow *Gsm* measures of working memory (Gsm-MW), the reported MW factor loadings for WJ III Numbers Reversed (.65), WJ III Auditory Working Memory (.59), and WISC-III Digit Span (.70) tests suggest that they are not interchangeable MW measures. More importantly, other intelligence battery tests or composites that may be classified the same (either empirically or logically) as a WJ III measure may not necessarily display the same strength of relation with achievement domains. For example, in the Phelps and colleagues (2005) dataset, the WJ III Visual Matching (*Gs*-P) test correlated .42 with WJ III Letter-Word Identification and .40 with WJ III Passage Comprehension, whereas the similarly *Gs*-P classified tests (see Flanagan et al., 2006) of WJ III Cross Out (.35, .27) and WISC-III Symbol Search (.32, 27) correlated at lower levels. Empirical support for CHC COG-ACH interpretations beyond the WJ batteries is limited to nonexistent.

Second, as previously discussed, our operational criterion for the significance consistency classifications was admittedly post hoc and arbitrary. Given the lack of a prior systematic CHC COG-ACH research synthesis, we erred on the side of leniency as we viewed the current review as exploratory and suggestive in nature. This was our intent—to identify possible significant COG-ACH relations warranting further study and discussion.

Third, a number of interesting CHC COG-ACH relations were classified as tentative/speculative. Additional research with the same or similar measures of these tentatively identified abilities is needed. There has been 20 years of CHC COG-ACH research, but most of it has consisted of analysis of the WJ-R and WJ III battery measures and norm data. Additional research is needed with other measures and in different norm samples.

Fourth, space did not allow for analyses by methodological factors. Most important was the possibility of different conclusions when comparing MV versus LV research (see Table 1) and, more importantly, what the MV/LV $\rightarrow$ ACH differential findings suggest for future research and current assessment practice. Inspection of the complete set of online summary coding tables reveals an obvious trend for MV studies to report more significant COG-ACH relations than LV studies. For example, across all achievement domains, CHC IVs, and ages, MV analyses were significant approximately 1.5 times more frequently than were IV analyses (MV = 40.1%; LV = 26.2%). To disentangle the possible MV/LV × ACH domain × age group interactions requires a separate analysis and article. We provide the online summary tables in hopes that others will explore these methodological nuances and their implications for practice. Although we did not undertake such detailed exploration, we believe that there is a strong probability that the MV > LV COG-ACH significance finding is most likely due to the absence (MV) or presence (LV) of a general intelligence (g) factor in the research designs. This topic deserves greater deliberation, analysis, discussion, and debate than we can offer here.

The combined limitations make one overarching conclusion clear: The extant CHC COG-ACH literature of the past 20 years has been restricted to a mosaic of methodological approaches that have been primarily applied to samples that frequently have not been independent (i.e., WJ-R and WJ III standardization sample subjects). We believe that the salient COG-ACH relations reported are those that are the most robust—they managed to "bubble to the surface" despite the methodological twists and turns across research studies.

Summary and Implications

What have we learned from 20 years of CHC COG-ACH relations research? What are the implications for the changing landscape of school psychology assessment practice? We present a half dozen primary conclusions and implications (in no order of implied importance):

1. *The current review extends and expands on the earlier Flanagan et al. (2006) narrative research synthesis.* The Flanagan and colleagues (2006) CHC COG-ACH research synthesis served a valuable function during the formative stages of bridging the gap between CHC intelligence theory and the practice of intellectual assessment. We choose not to devote pages to detailed comparisons of the similarities and differences between the conclusions of the current review and that of Flanagan and colleagues (2006). The results and conclusions of the current review are intended to improve upon the specificity of the CHC COG-ACH relations literature to better inform assessment practice. The current review reveals a much more nuanced set of CHC COG-ACH relations as a function of (a) breadth of cognitive abilities and measures (broad vs. narrow), (b) breadth of achievement domains (e.g., BRS and RC vs. broad reading), and (c) developmental (age) status.
2. *Many analyses have been completed.* A large number of CHC COG-ACH analyses have been completed since the CHC model of human cognitive abilities was first operationalized in an applied intelligence battery in 1989. Assessment professionals now have a more solid, empirically-based foundation upon which to make CHC-based COG-ACH–related assessment decisions and interpretations.
3. *Almost all of the available research is limited exclusively to one cognitive battery.* Almost all (94%) analyses have been completed with the WJ battery (WJ-R; WJ III). This is good news for the WJ III as it is the only individually administered intelligence battery with an empirical knowledge base from which to inform CHC-based assessment, diagnosis, and intervention planning. Until additional CHC COG-ACH research is completed with other (non-WJ) intelligence batteries, users of these other batteries must proceed with caution when forming COG-ACH relations-based diagnostic, interpretative, and intervention hypotheses. Given the brief evidence we presented for the lack of direct 1:1 broad or narrow CHC construct measure equivalence and predictive validity (of achievement) across measures, which is supported by Floyd, Bergeron, McCormack, Anderson, and Hargrove-Owens (2005), additional CHC composite exchangeability analyses across major intelligence batteries are needed. It is recommended that independent researchers, test authors, or publishers of other CHC-based or interpreted batteries begin completing similar studies, preferably in age-differentiated

subgroups of test battery standardization sample data. Studies with students with clinically diagnosed learning disabilities are also needed for all intelligence batteries.

4. *The primary action is at the narrow ability level.* Our discussions of broad and narrow abilities were lengthy and difficult to write—due to the need to explain why some broad CHC abilities did not display strong relations. A resolution of these findings typically occurred when we examined the narrow ability results. It is our conclusion that the most important focus for CHC COG-ACH relations is at the narrow ability level. Broad CHC composites may demonstrate the best average predictive validity across a broad array of academic and non-academic criterion measures, but when attempting to understand and develop potential interventions for subareas of reading (e.g., word attack, sight vocabulary, RC) and math (e.g., learning math facts, solving applied math problems), narrow is better. Space does not allow for further explanation, but this finding is consistent with the classic "bandwidth-fidelity tradeoff" (Cronbach & Gleser, 1957) and the "attenuation paradox" (Boyle, 1991; Loevinger, 1954) issues in the validity and reliability measurement literature, respectively. Broad best predicts and explains broad. Narrow best predicts and explains narrow. We believe that validated narrow cognitive ability indicators need to be the focus of assessment personnel working in the schools and should be featured in future cognitive battery test development. The Flanagan and colleagues CBA approach (see Flanagan et al., 2006), which primarily focuses on narrow CHC abilities, is an example consistent with this recommended focus. The addition of more narrow CHC ability clusters to the WJ III Battery system,as a result of the publication of the WJ III Diagnostic Supplement (Woodcock, McGrew, Mather, & Schrank, 2003), is another example.
5. *"Intelligent" intelligence testing.* The extant CHC COG-ACH literature, even the more conservative LV studies that include a general intelligence factor (g; full-scale score proxy), confirms the conclusion that a number of broad and narrow CHC abilities are important above and beyond the influence of *g* when predicting school achievement. We believe that this argues for more judicious, flexible, selective, "intelligent" (Kaufman, 1979) intelligence testing where practitioners select sets of tests most relevant to each academic referral. Unless there is a need for a full-scale IQ *g* score for diagnosis (e.g., MR, gifted), professionals need to break the habit of "one complete battery fits all" testing. The CHC COG-ACH research summarized here should assist assessment professionals in making better decisions regarding which measures from an intelligence battery may provide the most diagnostic and instructionally relevant information for different academic domains at different ages or grades. Selective referral-focused assessments, with branching tree decision-rules for follow-up testing, need to be encouraged in practice and in pre- and postprofessional training (see McGrew, 2009, for examples[15]). Before conducting assessments for reading and math problems, practitioners need to ask the following questions when designing their initial assessment: What is (are) the subdomain(s) of concern? What is the age of the student? What CHC abilities does research suggest are most related to this (these) domain(s) at this age level? Our findings suggest that the design of assessments requires "intelligent" decisions that recognize CHC cognitive domain $\times$ achievement subdomain $\times$ age level interactions. Keith (1994) stated that "intelligence is important, intelligence is complex." Kaufman (1979) argued for "intelligent" intelligence testing. We agree and add that the intelligent design of individual assessments is critically important and complex (but not difficult or impossible), and must recognize the complexity of the domains of human cognitive abilities and achievements and the nuanced differential interactions between different CHC abilities and achievement domains. The intelligent design of assessments does not come from a higher power—it comes from integrating the research synthesis presented here with professional and clinical experience.

[15] Page limitations did not allow for the presentation of selective referral-focused intelligence testing based on the findings of the current review. McGrew (2009) has presented examples of how to use the results of the current review to complete such assessments. An online PowerPoint slide show can be viewed at http://www.slideshare.net/iapsych/chc-selective-referralfocused-assessment-scenarios

6. *There is a future for intelligence testing.* We believe that the current results are consistent with the call for an integration of RTI and intelligent intelligence testing (see Hale, Kaufman, Naglieri, & Kavale, 2006). The cognitive markers mentioned by RTI early screening advocates correspond nicely with many CHC broad and narrow abilities identified in the current review. A variety of researchers have argued for early screening for at-risk students based on cognitive markers (Berninger, 2006; Fuchs, Compton, Fuchs, Paulsen, Bryant, & Hamlett, 2005; Torgesen, 2002) that may be "*precursors* to manifest disabilities" (Fletcher et al., 2002). In addition to the more academic variables (e.g., letter identification or knowledge of concepts of print), cognitive markers often mentioned as relevant to reading by proponents of some RTI models (Fletcher et al., 2002; Torgesen, 2002) have included (with our corresponding CHC broad or narrow ability classification) picture naming or receptive and expressive vocabulary (*Gc*-VL; lexical knowledge); sentence recall or verbal short-term memory (*Gsm*-MS; memory span); phonological awareness skills or processing (*Ga*-PC), rapid automatic naming of objects, numbers, or letters (*Glr*-NA, *Gs*-P); working memory (*Gsm*-MW; working memory); general oral language comprehension and development (*Gc*-LD: language development); and verbal knowledge (*Gc*-K0; general information). Although less researched, some (but not all) math cognitive markers mentioned by those supporting some RTI models (e.g., Fuchs et al., 2005, 2006, 2008) have included efficiency of execution of cognitive tasks (*Gs*; processing speed), short-term memory (*Gsm*-MS; memory span), working memory (*Gsm*-MW; working memory), fluid intelligence (*Gf*; fluid intelligence), language ability (*Gc*; comprehension-knowledge), and vocabulary knowledge (*Gc*-VL; lexical knowledge). Interestingly, most of these reading and math markers were identified as significant COG-ACH relations in the current review as well as in the Flanagan and colleagues (2006) narrative review. A number of today's CHC-based intelligence batteries include reliable and valid tests that can serve as psychometrically sound markers as articulated by RTI advocates. We believe that those who argue against the use of any cognitive ability tests in the new RTI environment (a) have failed to examine the abilities measured by many contemporary CHC intelligence batteries, (b) have not taken the time to do the RTI marker-to-CHC ability terminology "crosswalk" (as demonstrated earlier in text), or (c) may have an agenda that is more sociopolitical than empirical.

Times and tests have changed during the past 20 years. Progress has been made in constructing more comprehensive (broader array of broad and narrow abilities sampled), CHC-based cognitive assessment batteries. Contemporary intelligence tests should be viewed as valuable toolboxes, with each tool carefully selected by intelligent craftsman to match the presenting problem. The current research synthesis provides empirical evidence to help guide assessment practices. We do not believe the current review is the end of the journey, but rather an important step toward a more complete understanding of the relationships between cognitive abilities and school achievement.

References

*Studies included in the research synthesis (see Table 1).

Aaron, P. G., & Joshi, M. R. (1992). Reading problems: Consultation and remediation. New York: Guilford.

Adams, M. J. (1990). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson, R. Barr, M. L. Kamil, & P. B. Mosenthal (Eds.), Handbook of reading research (Vol. 1, pp. 255–291). New York: Longman.

Baddeley, A. D., Logie, R. H., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. Journal of Memory and Language, 24, 119–131.

Baker, S. K., Simmons, D. C., & Kame'enui, E. J. (1995). Vocabulary acquisition: Synthesis of the research. Technical Report 13. Eugene, OR: University of Oregon.

Barker, T. A., Torgesen, J. K., & Wagner, R. K. (1992). The role of orthographic processing skills on five different reading tasks. Reading Research Quarterly, 27, 335–345.

*Benson, N. (2008). Cattell–Horn–Carroll cognitive abilities and reading achievement. Journal of Psychoeducational Assessment, 26(1), 27–41.

*Benson, N., & Moseley, B. (under revision). Integrating psychometric and information processing perspectives to clarify the process of mathematical reasoning. Journal of School Psychology.

Berninger, V. W. (1990). Multiple orthographic codes: Key to alternative instructional methodologies for developing the orthographic phonological connections underlying word identification. School Psychology Review, 19, 518–533.

Berninger, V. W. (2006). Research-supported ideas for implementing reauthorized IDEA with intelligent professional psychological services. Psychology in the Schools, 43(7), 781–796.

Berninger, V. W., Abbott, R., Thomson, J., & Raskind, W. (2001). Language phenotype for reading and writing disability: A family approach. Scientific Studies in Reading, 5, 59–105.

Berninger, V. W., Abbott, R., Thomson, J., Wagner, R., Swanson, H. L., Wijsman, E., et al., (2006). Modeling developmental phonological core deficits within a working memory architecture in children and adults with developmental dyslexia. Scientific Studies in Reading, 10, 165–198.

Boyle, G. J. (1991). Does item homogeneity indicate internal consistency of item redundancy in psychometric scales? Personality and Individual Differences, 12(3), 291–294.

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. Developmental Neuropsychology, 33(3), 205–228.

Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. Journal of Experimental Child Psychology, 65, 1–24.

Carey, S. (2004). Bootstrapping and the origin of concepts. Daedulus, 133, 59–68.

Chong, S. L., & Siegel, L. S. (2008). Stability of computational deficits in math learning disability from second through fifth grades. Developmental Neuropsychology, 33(3), 300–317.

Cirino, P. T., Morris, M. K., & Morris, R. D. (2002). Neuropsychological concomitants of calculation skills in college students referred for learning difficulties. Developmental Neuropsychology, 21, 201–218.

Cooper, H. (1998). Synthesizing research: A guide for literature reviews. Thousand Oaks, CA: SAGE Publications.

Cooper, K. L. (2006). A componential reading comprehension task for children. Unpublished doctoral thesis, University of New England, Portland, Maine.

Coyne, M. D., Simmons, D. C., Kame'enui, E. J., & Stoolmiller, M. (2004). Teaching vocabulary during shared storybook readings: An examination of differential effects. Exceptionality, 12, 145–162.

Cronbach, L. J., & Gleser, G. C. (1957). Psychological tests and personnel decisions. Urbana, IL: University of Illinois Press.

Ehri, L. C. (1998). Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), Word recognition in beginning literacy (pp. 3–40). Mahwah, NJ: Lawrence Erlbaum.

*Evans, J. J., Floyd, R. G., McGrew, K. S., & Leforgee, M. H. (2002). The relations between measures of Cattell–Horn–Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. School Psychology Review, 31, 246–262.

Fiorello, C. A., & Primerano, D. P. (2005). Research into practice: Cattell–Horn–Carroll cognitive assessment in practice: Eligibility and program development issues. Psychology in the Schools, 42(5), 525–536.

*Flanagan, D. P. (2000). Wechsler-based CHC cross-battery assessment and reading achievement: Strengthening the validity of interpretations drawn from Wechsler test scores. School Psychology Quarterly, 15, 295–329.

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Mascolo, J. T. (2006). The achievement test desk reference, 2nd ed. (ATDR-2). New York: Wiley.

Fletcher, J. M., Foorman, B. R., Boudousquie, A., Barnes, M. A., Schatschneider, C., & Francis, D. J. (2002). Assessment of reading and learning disabilities: A research-based intervention-oriented approach. Journal of School Psychology, 40(1), 27–63.

*Floyd, R. G., Bergeron, R., & Alfonso, V. C. (2006). Cattell–Horn–Carroll cognitive ability profiles of poor comprehenders. Reading and Writing, 19(5), 427–456.

Floyd, R. G., Bergeron, R., McCormack, A. C., Anderson, J. L., & Hargrove-Owens, G. L. (2005). Are Cattell–Horn–Carroll broad ability composite scores exchangeable across batteries? School Psychology Review, 34(3), 329–357.

*Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between measures of Cattell–Horn-Carroll (CHC) cognitive abilities and mathematics achievement across the school-age years. Psychology in the Schools, 40(2), 155–171.

*Floyd, R. G., Keith, T. Z., Taub, G. E., & McGrew, K. S. (2007). Cattell–Horn–Carroll cognitive abilities and their effects on reading decoding skills: g has indirect effects, more specific abilities have direct effects. School Psychology Quarterly, 22(2), 200–233.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). Responsiveness to intervention: Preventing and identifying mathematics disability. Teaching Exceptional Children, March/April, 60–63.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. Journal of Educational Psychology, 97, 493–513.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., et al. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. Journal of Educational Psychology, 98(1), 29–43.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. Scientific Studies of Reading, 5, 239–256.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? Journal of Educational Psychology, 100, 30–47.

Furst, A., & Hitch, G. J. (2000). Separate roles for executive and phonological components in mental arithmetic. Memory and Cognition, 28, 774–782.

Gambrell, L. B., & Jawitz, P. B. (1993). Mental imagery, text illustrations, and children's story comprehension and recall. Reading Research Quarterly, 28, 264–276.

*Ganci, M. (2004). The diagnostic validity of a developmental neuropsychological assessment (NEPSY)– Wechsler Intelligence Scale for Children–third edition (WISC-III) based cross battery assessment. Retrieved from ProQuest UMI Dissertation Publishing (UMI Microform 3150999).

Gamett, K., Frank, B., & Fleischner, J. E. (1983). Automatization and basic fact performance in learning disabled children. Learning Disabilities Quarterly, 6, 223–230.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. British Journal of Educational Psychology, 70, 177–194.

Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. Journal of Experimental Child Psychology, 49, 363–383.

Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. Psychological Bulletin, 114, 345–362.

Geary, D. C. (1994). Children's mathematical development: Research and practical applications. Washington, DC: American Psychological Association.

Geary, D. C. (2003). Learning disabilities in arithmetic: Problem solving differences and cognitive deficits. In H. L. Swanson, K. Harris, & S. Graham (Eds.), Handbook of learning disabilities (pp. 199–212). New York: Guilford.

Geary, D. C. (2007). An evolutionary perspective on learning disabilities in mathematics. Developmental Neuropsychology, 32(1), 471–519.

Geary, D. C., & Brown, S. C. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled children. Developmental Psychology, 27, 398–406.

Geary, D. C., Brown, S. C., & Samaranayake, V. A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed of processing differences in normal and mathematically disabled children. Developmental Psychology, 27, 787–797.

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. Journal of Experimental Child Psychology, 77, 236–263.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. Child Development, 78, 1343–1359.

Gelman, R., & Butterworth, B. (2005). Number and language: How are they related? Trends in Cognitive Sciences, 9, 6–10.

Goldman, S. R., Pellegrino, J. W., & Mertz, D. L. (1988). Extended practice of basic addition facts: Strategy changes in learning-disabled students. Cognition & Instruction, 5, 223–266.

Hale, J. B., & Fiorello, C. A. (2004). School neuropsychology: A practitioner's handbook. New York: Guilford.

*Hale, J. B., Fiorello, C. A., Dumont, R., Willis, J. O., Rackley, C., & Elliott, C. (2008). Differential Ability Scales–Second Edition (Neuro) psychological predictor of math performance for typical children and children with math disabilities. Psychology in the Schools, 45(9), 838–858.

Hale, J. B., Fiorello, C. A., Kavanagh, J. A., Hoeppner, J. B., & Gaither, R. A. (2001). WISC-II predictors of academic achievement for children with learning disabilities: Are global and factor scores comparable? School Psychology Quarterly, 16(1), 31–55.

Hale, J. B., Fiorello, C. A., Kavanagh, J. A., Holdnack, J. A., & Aloe, A. M. (2007). Is the demise of IQ interpretation justified? A response to special issue authors. Applied Neuropsychology, 14, 37–51.

Hale, J. B., Kaufman, A., Naglieri, J. A., & Kavale, K. A. (2006). Implementation of IDEA. Integrating response to intervention and cognitive assessment methods. Psychology in the Schools, 43(7), 753–770.

Hammill, D. (2004). What we know about the correlates of reading. Exceptional Children, 70(4), 453–468.

Hitch, G. J., & McAuley, E. (1991). Working memory in children with specific arithmetical learning difficulties. British Journal of Psychology, 82, 375–386.

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. Educational Psychology, 26, 339–366.

Holmes, J., Adams, J. W., & Hamilton, C. J. (2008). The relationship between visuospatial sketchpad capacity and children's mathematical skills. European Journal of Cognitive Psychology, 20(2), 272–289.

Hoover, W., & Gough, P. B. (1990). The simple view of reading. Reading and Writing: An Interdisciplinary Journal, 2, 127–160.

Jenkins, J. R., Fuchs, L. S., Van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. Journal of Educational Psychology, 95, 719–729.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A longitudinal investigation. Journal of Experimental Child Psychology, 85, 103–119.

Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. Reading Psychology, 21, 85–97.

Kaufman, A. (1979). Intelligent testing with the WISC-R. New York: Wiley.

Keith, T. Z. (1994). Intelligence is important, intelligence is complex. School Psychology Quarterly, 9(3), 209–221.

Keith, T. Z. (1997). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (pp. 373–402). New York: Guilford.

*Keith, T. Z. (1999). Effects of general and specific abilities on student achievement: Similarities and differences across ethnic groups. School Psychology Quarterly, 14, 239–262.

Keith, T. Z. (2006). Multiple regression and beyond. Boston: Allyn & Bacon.

Kintsch, E. (2005). Comprehension theory as a guide for the design of thoughtful questions. Topics in Language Disorders, 25, 51–64.

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling & C. Hulme (Eds.), The science of reading: A handbook (pp. 209–226). Oxford, England: Blackwell Publishing.

Koontz, K. L., & Berch, D. B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithmetic learning disabled children. Mathematical Cognition, 2, 1–23.

Leather, C. V., & Henry, L. A. (1994). Working memory span and phonological awareness tasks as predictors of early reading ability. Journal of Experimental Child Psychology, 58, 88–111.

Loevinger, J. (1954). The attenuation paradox in test theory. Psychological Bulletin, 51(5), 493–504.

Logie, R. H., & Baddeley, A. D. (1987). Cognitive processes in counting. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 310–326.

Maehara, Y., & Saito, S. (2007). The relationship between processing and storage in working memory span: Not two sides of the same coin. Journal of Memory and Language, 56, 212–228.

Mammarella, I. C., Pazzaglia, F., & Comoldi, C. (2008). Evidence for different components in children's visuospatial working memory. British Journal of Developmental Psychology, 26, 337–355.

Mazzocco, M. M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. Learning Disabilities Research & Practice, 20, 142–155.

McBride-Chang, C., Chang, L., & Wagner, R. K. (1997). Growth modeling of phonological awareness. Journal of Educational Psychology, 89(4), 621–630.

*McGrew, K. (2007, January). Prediction of WJ III reading and math achievement by WJ III cognitive and language tests. Retrieved January, 11, 2008, from http://intelligencetesting.blogspot.com/2008/01/prediction-of-wj-iii-readingmath-ach-by.html

*McGrew, K. (2008). WISC-III/WJ III cross-battery g+specific abilities findings. Retrieved October 21, 2008, from http://intelligencetesting.blogspot.com/2008/10/wisc-iiiwj-iii-cross-battery-gspecific.html

McGrew, K. (2009, Feb). CHC cognitive and achievement relations research synthesis: What we've learned from 20 years of research. Mini-skills workshop presented at the 40th annual convention of the National Association of School Psychologists, Boston, MA.

*McGrew, K. S. (1993). The relationship between the WJ-R Gf-Gc cognitive clusters and reading achievement across the lifespan. Journal of Psychoeducational Assessment, Monograph Series: WJ R Monograph, 39–53.

McGrew, K. S., & Flanagan, D. P. (1998). The intelligence test desk reference (ITDR): Gf-Gc Cross-Battery assessment. Boston: Allyn & Bacon.

*McGrew, K. S., Flanagan, D. P., Keith, T. Z., & Vanderwood, M. (1997). Beyond g: The impact of Gf-Gc specific cognitive abilities research on the future use and interpretation of intelligence tests in the schools. School Psychology Review, 26, 189–210.

*McGrew, K. S., & Hessler, G. L. (1995). The relationship between the WJ-R Gf-Gc cognitive clusters and mathematics achievement across the life-span. Journal of Psychoeducational Assessment, 13, 21–38.

McGrew, K. S., & Woodcock, R. W. (2001). Technical manual. Woodcock Johnson III. Rolling Meadows, IL: Riverside.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. Journal of Experimental Child Psychology, 74, 240–260.

*Miller, B. D. (2001). Using Cattell–Horn–Carroll cross-battery assessment to predict reading achievement in learning disabled middle school students. Retrieved from ProQuest UMI Dissertation Publishing (UMI Microform 9997281).

Morris, R. D., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Stuebing, K. K., Fletcher, J. M., et al. (1998). Subtypes of reading disability: Variability around a phonological core. Journal of Educational Psychology, 90, 347–373.

Nation, K. (2005). Children's reading comprehension difficulties. In M. J. Snowling & C. Hulme, (Eds.), The science of reading: A handbook (pp. 248–265). Oxford, England: Blackwell Publishing.

Nation, K., Clarke, P., & Snowling, M. J. (2002). General cognitive ability in children with reading comprehension difficulties. British Journal of Educational Psychology, 72, 549–560.

Nation, K., Marshall, C. M., & Snowling, M. J. (2001). Phonological and semantic contributions to children's picture naming skill: Evidence from children with developmental reading disorders. Language and Cognitive Processes, 16, 241–259.

Newton, J. H., & McGrew, K. S. (2010). Introduction to the special issue: Current research in Cattell-Horn-Carroll-based assessment. Psychology in the Schools, 47(7), 621–634.

Osmon, D. C., Smertz, J. M., Braun, M. M., & Plambeck, E. (2006). Processing abilities associated with math skills in adult learning disability. Journal of Clinical and Experimental Neuropsychology, 28, 84–95.

Passolunghi, M. C. (2006).Working memory and mathematical disability. In T. Packiam Alloway & S. Gathercole (Eds.), Working memory and neurodevelopmental condition (pp. 113–138). Hove, UK: Psychology Press.

Passolunghi, M. C., Mammarella, I. C., & Altoè, G. (2008). Cognitive abilities as precursors of the early acquisition of mathematical skills during first through second grades. Developmental Neuropsychology, 33, 229–250.

Passolunghi, M. C., & Siegel, L. S. (2001). Short term memory, working memory, and inhibitory control in children with specific arithmetic learning disabilities. Journal of Experimental Child Psychology, 80, 44–57.

Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). The precursors of mathematics learning: Working memory, phonological ability and numerical competence. Cognitive Development, 22, 165–184.

Pedhazur, E. J. (1997). Multiple regression in behavioral research. New York: Wadsworth.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. Scientific Studies of Reading, 11(4), 357–383.

Phelps, L., McGrew, K. S., Knopik, S. N., & Ford, L. A. (2005). The general (g), broad, and narrow CHC stratum characteristics of the WJ III and WISC-III tests: A confirmatory cross-battery investigation. School Psychology Quarterly, 20, 66–88.

Pillemer, D. B., & Light, R. J. (1980). Synthesizing outcomes: How to use research evidence from many studies. Harvard Educational Review, 50(2), 176–194.

Pinel, P., Piazza, D., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. Neuron, 41, 1–20.

Poock, M. A. J. (2005). Validation of the attention battery of the Woodcock Johnson–Third Edition Tests of Cognitive Abilities (WJ III, COG) for children with attention deficit hyperactivity disorder. Retrieved from ProQuest UMI Dissertation Publishing (UMI Microform 3199438).

*Proctor, B. E., Floyd, R. G., & Shaver, R. B. (2005). Cattell–Horn–Carroll broad cognitive ability profiles of low math achievers. Psychology in the Schools, 42(1), 1–12.

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. Journal of Experimental Child Psychology, 91, 137–157.

Rourke, B. P. (1993). Arithmetic disabilities, specific and otherwise: A neuropsychological perspective. Journal of Learning Disabilities, 26, 214–226.

Rourke, B. P., & Conway, J. A. (1997). Disabilities of arithmetic and mathematical reasoning: Perspectives from neurology and neuropsychology. Journal of Learning Disabilities, 30, 34–46.

Schrank, F. A. (2006). Specification of the cognitive processes involved in the performance on the Woodcock–Johnson III (Assessment Service Bulletin No. 7). Rolling Meadows, IL: Riverside Publishing.

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. Annual Review of Psychology, 59, 451–475.

Stanovich, K., & Siegel, L. (1994). Phenotypic profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. Journal of Educational Psychology, 86(1), 24–53.

Swanson, H. L. (1993). Executive processing in learning disabled readers. Intelligence, 17, 117–149.

Swanson, H. L., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. Review of Educational Research, 76, 249–274.

Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. Journal of Experimental Child Psychology, 79, 294–321.

*Taub, G. E., Floyd, R. G., Keith, T. Z., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. School Psychology Quarterly, 23(2), 187–198.

Torgesen, J. K. (2002). The prevention of reading difficulties. Journal of School Psychology, 40(1), 7–26.

Urso, A. (2008). Processing speed as a predictor of poor reading. Unpublished doctoral dissertation, University of Arizona, Tucson.

*Vanderwood, M., McGrew, K. S., Flanagan, D. P., & Keith, T. Z. (2002). The contribution of general and specific cognitive abilities to reading achievement. Learning and Individual Differences, 13, 159–188.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, S. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. Scientific Studies of Reading, 11(1), 3–32.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. Developmental Psychology, 33, 468–479.

*Windfuhr, K. L., & Snowling, M. J. (2001). The relationship between paired associate learning and phonological skills in normally developing readers. Journal of Experimental Child Psychology, 80, 160–173.

Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. Journal of Learning Disabilities, 33, 387–407.

Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. A. (2003). Woodcock–Johnson III Diagnostic Supplement to the Tests of Cognitive Abilities. Itasca, IL: Riverside Publishing.

Zorzi, M., Priftis, K., & Umiltá, C. (2002). Neglect disrupts the mental number line. Nature, 417, 138–139.