Understanding the Cognitive Deficits Related to Mathematics Difficulties: A Meta-Analysis

Peng Peng, Ph.D. (kevpp2004@hotmail.com)
Department of Special Education and Communication Disorders
University of Nebraska, Lincoln

Introduction

While mathematical competency is critical for competing successfully in today’s high technology world, learning mathematics is a big challenge for many children. Converging evidence shows that approximately 6% of the school-aged population show some form of mathematics difficulties (MD) even with average or higher IQ and adequate instruction (Berch & Mazzocco, 2007; Geary, 2011). Despite the prevalence of MD, MD is much less studied compared to reading difficulties (RD) (Gersten, Clarke, & Mazzocco, 2007). In recent years, an increasing number of studies have examined the factors associated with MD and suggested two major approaches to understanding the deficit profiles of MD. One is through the investigation of deficits in domain-specific skills (Geary, 1993; Gersten et al., 2005). The other is through the investigation of deficits in domain-general cognitive skills (e.g., Johnson, Humphrey, Mellard, Woods, & Swanson, 2010; Peng & Fuchs, 2016; Swanson & Jerman, 2006).

Regarding the domain-specific skill deficits among MD, there is a consensus that MD is related to deficits in the numerical processing system. For example, Geary (1993) conducted a literature review of MD research with a focus on conceptual and procedural competencies of arithmetic and suggested that MD is related to deficits in the representation of arithmetic facts from semantic memory, the execution of arithmetic procedures, and visuospatial presentation of numerical information. The synthesis by Gersten et al. (2005) is consistent with Geary (1993), suggesting that the deficits in numerical-related skills, such as magnitude comparison, counting strategies, identification of numbers, and numerical working memory, are important markers for MD.

Regarding the domain-general skill deficits among MD, previous reviews focused on the memory system. Specifically, Swanson and Jerman (2006), and Peng and Fuchs (2016) used meta-analyses to examine the memory deficits of MD among elementary school children. Both reviews indicated that children with MD, regardless of comorbidity in mathematics and reading difficulties (MDRD), suffer from memory deficits. Compared to children with MDRD, children with MD only seem to have more severe visuospatial and numerical memory deficits. Johnson et al. (2010) also studied the deficit profiles of working memory, short-term memory, executive functions, and processing speed of MD. However, due to a limited number of effect sizes (less than 8 effect sizes for each skill), they could not draw a conclusion on these deficit profiles of MD or systematically examine moderators that influence the profiles.

Although previous reviews have advanced our understanding of domain-specific and domain-general deficits of MD, several important questions remain unanswered, which we aimed to address in the present study. Specifically, we examined (1) whether MD is related to deficits in other cognitive skills besides memory deficits, (2) whether the cognitive deficit profiles of MD are influenced by comorbidity (i.e., MDRD vs. MD only), types of MD screening (i.e., calculation vs. problem solving vs. comprehensive...
Running head: COGNITIVE DEFICIT PROFILES OF MD

mathematics), severity of MD, domains of task (i.e., verbal vs. numerical vs. visuospatial), and age, and (3) whether deficits in low-level cognitive skills (i.e., processing speed and short-term memory) explain the deficits in high-level cognitive skills (e.g., attention, working memory, and executive functions) among MD.

Investigating these questions is both theoretically and practically important. From the theoretical perspective, the deficit profiles in cognitive skills help further understand the nature of MD. That is, besides numerical processing and memory deficits, is MD also related to deficits in other cognitive skills? (Geary, 2005; Mazzocco, 2005). If yes, would this relation reflect the numerical-deficit nature of MD such that cognitive deficits among MD are related to numerical processing specifically? Is there heterogeneity among MD? If yes, does this heterogeneity reflect different cognitive deficit profiles among different types of MD or does this heterogeneity also reflect the changing cognitive deficit profiles as a function of time? Moreover, does MD have deficits in both low-level and high-level cognitive skills? If yes, do deficits in low-level cognitive skills explain the deficits in high-level cognitive skills, as suggested by the bottleneck theory (Salthouse, 1996; Swanson & Sachse-Lee, 2001; Peng, Sun, Li, & Tao, 2012)?

From a practical perspective, answering these questions has implications for MD identification and intervention. Specifically, if the cognitive deficit profiles vary among different types of MD or vary with time, there should be more accurate identification of MD subtypes and thus more individualized instruction for a MD subtype at a specific period of time. Moreover, an increasing number of studies in recent years have investigated whether training high-level cognitive skills (e.g., working memory) improves cognition and whether such improvement transfers to mathematics outcomes (e.g., Holmes, Gathercole, & Dunning, 2009; Kroesbergen, Van’t Noordende, & Kolkman, 2014). Although some studies have found training effects on cognitive skills, most have failed to find the transfer effects (e.g., Jacob & Parkinson, 2015; Melby-Lervåg & Hulme, 2013; Morrison & Chein, 2011; Peng & Miller, 2016; Shipstead, Redick, & Engle, 2012). The lack of far transfer has often been discussed as being related to: 1) variability in the types of cognitive training (Jacob & Parkinson, 2015), 2) variability in the domains of task training materials (Peng & Fuchs, 2016), and 3) variability in the population receiving the training (Shipstead et al., 2012). Investigating the moderation effects of types of MD, severity of MD, domains of task, and age on the cognitive deficit profiles of MD can provide empirical evidence that may help address the conflicting findings in the cognitive intervention research.

Methods

Literature Search

Articles for this meta-analysis were identified in three ways. First, a computer search of the Education Resources Information Center (ERIC), ProQuest, and PsycINFO was conducted. We used the earliest possible start date (1920) till June, 2017. The following terms were used to search in full text: (math*AND difficult* OR disabilit*) OR discalcul*. The terms math*, difficult*, disabilit*, discalcul* allow for inclusion of mathematics, difficulties / difficulty / disability / disabilities, dyscalculia / dyscalculic, and so forth. Second, we searched for unpublished literature through Dissertation and Masters Abstract indexes in ProQuest, Cochrane Database of Systematic Reviews, relevant conference programs (e.g., Conference of Society for Research in Child Development; National Council of Teachers of Mathematics Annual Meeting;
Psychology of Mathematics Education Annual Conference. Third, previous relevant reviews (Geary, 1993; Gersten et al., 2005; Peng & Fuchs, 2016; Swanson & Jerman, 2006) were reviewed to include extra articles not identified in the first two search steps. The initial search yielded 908 studies. Two authors of this study then reviewed all studies by titles and abstracts. After excluding the duplicate 22 articles, the remaining 882 articles were closely reviewed using the specific criteria. First, the study had to include a group of individuals with MD and a group of age-matched TD individuals. That is, the study had to report information showing that the MD group scored at least below the 35th percentile or below one standard deviation on mathematics screening measures, which is a common learning disability identification criterion (Swanson & Harris, 2013). Also, the study had to provide information showing that the IQ (non-verbal IQ, verbal IQ, or combination of non-verbal and verbal IQ) of individuals with MD were in the normal range (standardized score 80-120), and the age-matched TD group was comparable to the MD group. We excluded 499 studies that did not meet these criteria. Then, we reviewed the remaining 383 studies to select studies that compared the MD to TD group on measures that tap at least one of the following cognitive skills: phonological processing, processing speed, short-term memory, working memory, attention, visuospatial skills, and executive functions. We excluded another 308 studies that did not meet this criterion. The final sample included 75 studies, including 4 dissertations, 1 book chapter, and 70 peer-reviewed articles.

**Coding Procedure and Inter-Rater Reliability**

These 75 studies were coded according to the characteristics of participants, mathematics and IQ screening tasks, tasks used to measure phonological processing, processing speed, short-term memory, working memory, visuospatial skills, attention, and executive functions. Not all studies provided sufficient information on the variables of interest for the present study. In case of insufficient information, authors were contacted to obtain the missing information. Variables were discussed until a consensus was reached between the first and the second author. Then, both authors used the coding system to conduct the final coding of all studies. Across the total variable matrix, the mean inter-rater agreement (i.e., the percentage of agreement indexed by the data points coded in agreement divided by total coded data points) was .98, with the coefficient above .93 for all moderators of interest and all skills investigated in this study. Any disagreements between the raters were resolved by consulting the original article or by discussion.

**Missing data.** Not all studies provided sufficient information on the variables of interest for the present study. In case of insufficient information, authors were contacted to obtain the missing information. However, if missing data could not be obtained for a particular moderator variable, the study was excluded from that particular moderator analysis but was included in other moderator analyses for which the data were present.

**Analytic Strategies**

**Effect size.** Hedges’ g, corrected for sample size bias, was used as the measure of effect size. We chose Hedges’ g as it provides a better estimate of effect sizes than Cohen’s d on small sample sizes (MD profiling research tends to have small sample sizes) (Grissom & Kim, 2005). For studies reporting means, standard deviations, and sample size, the following formulae were used:

\[
g'^n = g \left(1 - \frac{3}{4(N_{MD} + N_{TD} - 2)}\right)
\]
Running head: COGNITIVE DEFICIT PROFILES OF MD

With \( g = \frac{\bar{X}_{MD} - \bar{X}_{TD}}{s} \) and \( g = \sqrt{\frac{N_{MD}-1}{N_{MD}+N_{TD}-2} \cdot \frac{S_{MD}^2 + S_{TD}^2}{N_{MD}+N_{TD}-2}} \)
in which \( g^u \) is the unbiased estimate of Hedges’ \( g \), \( g \) is Hedges’ \( g \) as traditionally defined, \( N_{MD} \) is the number of participants in the MD group, \( N_{TD} \) is the number of participants in the TD group, \( \bar{X}_{MD} \) is the mean of the target skill (e.g., working memory) scores for participants in the MD group, \( \bar{X}_{TD} \) is the mean of the target skill scores for participants in the TD group, \( S \) is the pooled standard deviation of \( S_{MD}^2 \) is the variance of target skill scores for the participants in the MD group, and \( S_{TD}^2 \) is the variance of target skill scores for the participants in the TD group. Because Cohen’s benchmark for effect size is often adopted in educational research (Hill, Bloom, Black, & Lipsey, 2008), especially in the profiling research on children with learning disabilities (Johnson et al., 2010; Peng & Fuchs, 2016), we used Cohen’s benchmark (.02 ~ .05 means small effects, .50 ~ .80 means medium effects, and > .80 means large effects, Cohen, 1988) to interpret the magnitude of effect sizes.

Effect sizes of deficits on all cognitive skills were estimated for MD in comparison to the TD group. Next, meta-regression analyses were used to examine the moderation effects of types of MD (i.e., comorbidity and types of MD screening), severity of MD, age, and domains of task on the deficit profile for each cognitive skill. For the moderation analyses, each moderator was examined with other moderators simultaneously controlled in one meta-regression model. For moderators that were categorical, we created different sets of dummy codes and entered them into the meta-regression model (Cohen, Cohen, West, & Aiken, 2013). To examine the “bottleneck theory”, we controlled for the group differences between MD and TD on processing speed or short-term memory for the group comparison between MD and TD on working memory, attention, or executive functions. Specifically, we first identified studies that provided correlations among at least one low-level cognitive skill and one high-level cognitive skill. Then, controlling for the group difference on the low-level cognitive skill and the correlation between low-level and high level cognitive skills, we calculated the group difference between MD and TD on the high-level cognitive skill, which would be synthesized to indicate the group differences on high-level cognitive skills, partiaing out the effects of low-level cognitive skills.

**Nested structure of effect sizes.** We considered all eligible effect sizes in each study. That is, studies could contribute multiple effect sizes as long as the sample for each effect size was independent. For studies that reported multiple effect sizes from the same sample (e.g., two effect sizes based on two working memory measures were calculated for MD vs. TD in one study), we accounted for the statistical dependencies using the random effects robust standard error estimation technique developed by Hedges, Tipton, and Johnson (2010). This analysis allows for the clustered data (i.e., effect sizes nested within samples) by correcting the study standard errors to take into account the correlations between effect sizes from the same sample. The robust standard error technique requires that an estimate of the mean correlation (\( \rho \)) between all the pairs of effect sizes within a cluster be estimated for calculating the between-study sampling variance estimate, \( \tau^2 \). In all analyses, we estimated \( \tau^2 \) with \( \rho = .80 \); sensitivity analyses showed that the findings were robust across different reasonable estimates of \( \rho \).

Because we included studies from a wide age span and on different mathematics and cognitive skills, we hypothesized that the research body is reporting a distribution of effect sizes with significant between-studies variance, as opposed to a group of studies...
attempting to estimate one true effect size. Thus, we used a random-effect model for the current study (Lipsey & Wilson, 2001). Weighted, random-effects meta-regression models using Hedges et al.’s (2010) corrections were conducted with ROBUMETA in STATA (Hedberg, 2011) to summarize effect sizes and to examine potential moderators.

**Publication bias.** Publication bias (the problem of selective publication, in which the decision to publish a study is influenced by its results) was examined using the method of Egger, Smith, Schneider, and Minder (1997) and funnel plot. We did not find a significant publication bias based on Egger et al.’s (1997) publication bias statistics (i.e., the standard errors of correlations did not significantly predict correlations among studies with ROBUMETA in Stata, $ps > .05$), except for the comparisons on phonological processing and working memory, $ps < .05$. Further funnel plot analyses showed reasonable symmetry in all reported comparisons (the significant Egger’s test for the comparisons on phonological processing may be due to four outliers and two outliers for working memory). Taken together, Egger’s test and funnel plot suggested that there was little influence of publication bias in the data and thus the original dataset was used in all reported analyses. That said, we controlled for the publication type in our moderation analysis to account for any potential effects of publication bias (Rothstein, Sutton, & Borenstein, 2006).

**Results**

The 75 studies included in the meta-analysis represented a total of 13,001 children (5251 for MD, and 7750 for TD) obtained from 126 independent samples. There were 846 effect sizes that indicated the comparison between MD and TD on phonological processing (108 effect sizes, with 24 effect sizes for manipulation and 84 effect sizes for retrieval), processing speed (47 effect sizes), short-term memory (192 effect sizes), working memory (286 effect sizes), attention (49 effect sizes), visuospatial skills (20 effect sizes), and executive function (144 effect sizes, with 114 effect sizes for inhibition, 11 effect sizes for updating, 19 effect sizes for switching).

The Deficit Profiles of MD

The estimated average effect size indicating the deficits of MD as compared to TD was as follows: phonological processing: Hedges’ $g = -.91$, CI95[-1.24, -.57] (Manipulation: Hedges’ $g = -1.31$, CI95[-1.82, -.79]; Retrieval: Hedges’ $g = -.61$, CI95[-.80, -.43]), processing speed: Hedges’ $g = -.90$, CI95[-1.08, -.72], short-term memory: Hedges’ $g = -.56$, CI95[-.67, -.46], working memory: Hedges’ $g = -.76$, CI95[-.88, -.64], visuospatial skills: Hedges’ $g = -.43$, CI95[-.63, -.23], attention: Hedges’ $g = -.72$, CI95[-.92, -.52] (Subjective: Hedges’ $g = -.44$, CI95[-.73, -.14]; objective: Hedges’ $g = -.86$, CI95[-1.09, -.63]), and executive function: Hedges’ $g = -.50$, CI95[-.60, -.39] (Inhibition: Hedges’ $g = -.37$, CI95[-.48, -.26]; Updating: Hedges’ $g = -.76$, CI95[-1.04, -.48]; Switching: Hedges’ $g = -.75$, CI95[-.87, -.63]). To sum, compared to the TD group, the MD group showed severe deficits in phonological processing and processing speed, and moderate deficits in phonological processing-retrieval, short term memory, working memory, visuospatial skills, and executive functions.

Factors that Influence the Deficit Profiles of MD

Next, we examined whether comorbidity, types of MD screening, severity of MD, age, or domains of task influenced the deficit profiles of each cognitive skill of MD. Our findings show that 1) regarding the comorbidity, the MD group showed severe deficits in phonological processing-manipulation and processing speed, and moderate deficits in
phonological processing-retrieval, short-term memory, working memory, visuospatial skills, and executive functions. In contrast, the MDRD group showed more severe deficits than MD only group on phonological processing, processing speed, short-term memory, visuospatial skills, and executive functions; 2) Types of MD screening affected the deficit profiles of phonological processing, working memory, and executive functions such that MD identified with word-problem solving difficulties seemed to have less severe deficits on these cognitive skills than those identified with calculation difficulties or comprehensive mathematics difficulties. For MD identified with comprehensive mathematics difficulties, they had more severe numerical working memory deficits than their visuospatial working memory deficits, and they also had more severe phonological processing-manipulation deficits than their retrieval deficits; 3) Severity of MD was only related to the severity of processing speed deficits; 4) Age only affected deficits in phonological processing and attention such that younger individuals with MD suffer from more severe deficits in these skills than older individuals with MD; 5) Domains of task mostly affected the deficit profiles of processing speed and working memory such that deficits on these skills among MD are most severe in the numerical domain; 6) Low-level skills (short-term memory and processing speed) significantly accounted for some (not all) variance in the group differences between MD and TD on high-level skills (working memory, attention, and executive functions).

Discussion

The present study, to our knowledge, was the first meta-analysis that systematically and comprehensively examined the cognitive deficit profiles of MD and the factors that influence these deficit profiles. Findings have theoretical and practice implications for MD. Theoretically, the findings add to our understanding of MD. First, MD suffers from comprehensive cognitive deficits that are not specifically related to numerical processing, with exceptions that deficits in processing speed and working memory are more related to the numerical domain. Thus, the cognitive deficits of MD are generally not affected by domains of materials. Any individual who is identified with MD is likely to suffer from both domain-specific (numerical processing) and domain-general cognitive deficits. Second, MD is a group with heterogeneity, which lies in two aspects. One is that there is heterogeneity among MD subgroups. The MDRD group is different from the MD only group such that the MDRD group demonstrates more severe cognitive deficits. MD identified with different screening measures also differ from each other. Calculation difficulties are related to more severe deficits in phonological processing and working memory and comprehensive mathematics difficulties are related to more severe deficits in executive functions and working memory. The other aspect of heterogeneity of MD is reflected by age. That is, deficits in attention and phonological processing are more severe among younger individuals with MD. This age effect may be attributed to the characteristics of early mathematics learning (or instruction) that emphasizes automatized arithmetic facts retrieval.

Third, previous research suggests that different cut-off points on MD screening may result in MD with different profiles, indicating that MD may be a continuous construct (e.g., Branum-Martin et al., 2013; Murphy et al., 2007). However, the present review found that after controlling for comorbidity, types of MD screening, and other moderators, severity of MD only affected deficits in processing speed, but not in other cognitive skills. This finding is consistent with previous studies and reviews (Landerl et al., 2004; Peng
suggesting that the severity of MD, in general, does not affect the deficit profiles of MD, and MD may be a discrete construct. That said, the current study only focused on concurrent data. Future studies, especially those that focused on at-risk MD and persistent MD (Vukovic & Siegel, 2010), should consider longitudinal data to further investigate this issue. Fourth, the deficits in the high-level cognitive deficits are relatively independent of the deficits in low-level cognitive skills of MD. Compared to the deficits in low-level skills (except for processing speed), MD is more strongly and stably related to the deficits in high-level cognitive skills. Thus, it is likely that for individuals with MD, both their basic information processing system (e.g., numerical processing) and the complex information processing system (e.g., forming strategies and understanding conceptual knowledge) in mathematics are impaired.

Our findings also have implications for practice. Regarding the diagnosis of MD, our findings suggest the efficiency of MD diagnosis, especially in early grades, might be improved by conducting screenings on skills with salient cognitive deficits among MD. For example, in early grades (e.g., kindergarten and first grade) where traditional MD screening tests (e.g., calculations) may be insensitive (i.e., most children reach the floor on those tests), teachers/practitioners can use measures, such as processing speed, phonological processing, and attention ratings to help identify children at-risk for MD.

Interventions can be designed to compensate for the cognitive deficits of MD. For example, given that MD is related to distinct deficits in processing speed and working memory, especially in the numerical domain, instruction for MD should emphasize increasing numerical processing speed such as building fluency (e.g., improving arithmetic facts retrieval during calculations; using/selecting more efficient strategies during word problem-solving) to reduce working memory load in complex mathematic problems. This suggestion is also in line with the Load Reduction Instruction (LRI) theory (Martin, 2016). Researchers can design interventions that directly address these cognitive deficits among MD. Also, the cognitive interventions can be individualized for a specific type of MD. For example, whereas updating training may be crucial to remediate word-problem solving difficulties, phonological processing training may benefit MD identified with calculation difficulties most, especially those identified at a young age. Intervention materials can cover verbal, visuospatial, and numerical domains, but working memory training in the numerical domain seem most promising. With all being said, more research is needed to investigate whether and how direct cognitive training leads to and transfers to improved mathematics performance among MD.

References


