Working Memory Impairments in Children with Specific Arithmetic Learning Difficulties

Janet F. McLean and Graham J. Hitch

Lancaster University, Lancaster, United Kingdom

Working memory impairments in children with difficulties in arithmetic have previously been investigated using questionable selection techniques and control groups, leading to problems concluding where deficits may occur. The present study attempted to overcome these criticisms by assessing 9-year-old children with difficulties specific to arithmetic, as indicated by normal reading, and comparing them with both age-matched and ability-matched controls. A battery of 10 tasks was used to assess different aspects of working memory, including subtypes of executive function. Relative to age-matched controls, children with poor arithmetic had normal phonological working memory but were impaired on spatial working memory and some aspects of executive processing. Compared to ability-matched controls, they were impaired only on one task designed to assess executive processes for holding and manipulating information in long-term memory. These deficits in executive and spatial aspects of working memory seem likely to be important factors in poor arithmetical attainment. © 1999 Academic Press

Key Words: working memory; executive processes; arithmetic; children; learning difficulties.

There are many reasons children may fail to learn arithmetic. Examples include anxiety about mathematics, lack of experience and poor motivation (Ashcraft & Faust, 1994; Levine, 1987), reading difficulties (Muth, 1984; Richman, 1983), and neuropsychological damage (McCloskey, Harley, & Sokol, 1991). A growing body of evidence suggests that arithmetical learning difficulties can be associated with cognitive deficits (e.g., Bull & Johnston, 1997; Geary & Brown, 1991; Geary, Brown, & Samanayake, 1991; Hitch & McAuley, 1991; Rourke & Findlayson, 1978; Rourke & Strang, 1978; Siegel & Ryan, 1989; Temple, 1991). The present study builds on this work by focusing on cognitive deficits in a subgroup of children within the general population who have specific difficulties in arithmetic but normal reading.

This research was conducted in partial fulfillment of a Ph.D. by the first author. The cooperation of schools in the Lancaster area and the financial support of the Economic and Social Research Council are gratefully acknowledged.

Address correspondence and reprint requests to Janet McLean, Department of Psychology, Florentine House, University of Glasgow, Glasgow G12 8BQ, United Kingdom.
Describing the cognitive deficits of this subgroup can be problematic. Studies based on clinical samples have suggested three different types of cognitive deficit in arithmetic (see Badian, 1983, and Geary, 1993, for reviews). These are characterized as visuospatial deficits (Hartje, 1987; Rourke, 1993; Rourke & Findlayson, 1978; Rourke & Strang, 1978), difficulties with reading and writing of numbers (Kosc, 1974; Temple, 1989), and problems of retrieving arithmetic facts from long-term memory (Badian, 1983; Benson & Weir, 1972; Jackson & Warrington, 1986). White, Moffitt, and Silva (1992) found that arithmetic disabled children performed poorly not only on the visuospatial tasks used by Rourke and colleagues, but also on the Making Trails tasks (Reitan, 1958), the Grooved Pegboard (Klove, 1963), and the WISC-R Coding (Wechsler, 1974), which they interpreted as reflecting difficulties with visual–motor integration. Taken together, clinical studies suggest that difficulties with arithmetic are associated with a variety of deficits. However, it is difficult to extrapolate from these observations to the general population.

In a comprehensive review of cognitive deficits in arithmetic disabled children which included nonclinical studies, Geary (1993) identified two basic categories with different developmental trajectories. The first is manifested by the use of developmentally immature arithmetic procedures which usually regain normal levels after 2 or 3 years of schooling. This category has been associated with deficits in counting, computational skill, and working memory (Geary, 1990; Geary, Bow-Thomas, & Yao, 1992). The second category involves a more persistent difficulty with the representation and retrieval of arithmetic facts from long-term semantic memory. Fact-retrieval difficulties have been linked to deficits in counting speed and working memory (Garnett & Fleischner, 1983; Geary, 1990; Geary et al., 1991). It is interesting to note that both of these categories include working memory, the limited capacity system responsible for transforming and maintaining temporary information (see, e.g., Baddeley & Hitch, 1974).

A number of previous studies have compared measures of working memory in normal children, children with arithmetic difficulties, and children with enhanced arithmetical ability (e.g., Bull & Johnston, 1997; Dark & Benbow, 1990, 1991; Geary et al., 1991; Hitch & McAuley, 1991; Siegel & Ryan, 1989; Swanson, 1993, 1994). Although these studies differ in their assessment and aims, they agree in suggesting that working memory is related to differences in the ability to perform arithmetic. For example, Geary et al. (1991) investigated arithmetically disabled and normal children over 1 year. The disabled group received remedial education in mathematics and were below the 46th percentile on national achievement tests. Both forward and backward digit span were significantly reduced for the disabled group, consistent with a working memory deficit. Geary et al. (1991; see also Geary, 1990, 1993) proposed that although skills such as counting knowledge and strategy choice are the primary areas of deficit in the mathematically disabled child, working memory deficits may lead to a failure to develop long-term memory representations of basic facts. Geary et al. (1991)
suggest that if the representation of a problem’s integers is lost more rapidly from working memory it is less likely to be associated with the answer. However, their arithmetic disabled children also had poor reading skills. Their children’s average percentile ranking for reading was actually below that for arithmetic, and was much lower than the reading ranking for the control group. This is important because a significant amount of language is used in mathematics, and problems in reading and mathematics often co-occur (Muth, 1984; Richman, 1983; see Geary, 1993, for a review). Therefore, children with more general academic difficulties have been included in Geary and colleagues’ studies, and this leads to problems interpreting their findings with respect to arithmetic per se.

Evidence that working memory deficits differ as a function of the specificity of learning difficulties is presented by Siegel and Ryan (1989). They gave children two complex span tasks designed to assess working memory capacity. These were a counting span task (Case, Kurland, & Goldberg, 1982) and a sentence span task (adapted from Daneman & Carpenter, 1980). Children with specific arithmetic difficulties (i.e., WRAT Arithmetic ≤ 25th percentile and WRAT Reading ≥ 30th percentile) were impaired on counting span but not on sentence span. In contrast, children who were impaired on reading as well as on arithmetic had significantly lower counting and sentence spans. Siegel and Ryan suggested that the reading-and-arithmetic disabled group had a general working memory impairment, whereas the arithmetic disabled group had a domain-specific working memory deficit. In a related study, Hitch and McAuley (1991) confirmed that children with specific arithmetic difficulties were impaired on counting span but not on other complex span tasks. They also showed that these children had significantly lower digit spans than did age-matched controls.

A general criticism of the studies described above is that they have employed a design in which children with arithmetic difficulties are compared to normal achievers of the same chronological age. This type of design cannot determine the cause of the disability, as any cognitive deficits may be a consequence rather than a cause of low arithmetic attainment. Vellutino, Pruzek, Steger, and Meshoulam (1973) attempted to resolve this problem in the area of reading difficulties by suggesting that a younger ability-matched control group should be included. The logic is that if the poor achievers perform worse on some cognitive task than do younger, ability-matched controls, the deficit is unlikely to be a simple consequence of their low level of academic achievement. Given poor achievers’ greater age and experience, such a deficit would suggest they are arriving at similar levels of ability via different routes. However, Backman, Mamen, and Ferguson (1984) noted that age-matched controls remain useful for identifying differences in cognitive function at the same level of exposure to the skill in question.

A second concern about previous work is the sparsity of attempts to assess the various facets of working memory in children with poor arithmetic. If this system
is impaired, it is clearly important to know more precisely how it is affected. There are several different ways of viewing working memory (see, e.g., Baddeley & Hitch, 1974; Daneman & Carpenter, 1980; Just & Carpenter, 1992; Turner & Engle, 1989). The Baddeley and Hitch (1974) multicomponent model was used here, as it explains a wide range of experimental and neuropsychological evidence (see also Baddeley, 1986) and tasks have been developed for assessing some of its components. According to this model, working memory consists of a central executive processor which interacts with two slave subsystems, the phonological loop and the visuospatial sketch pad. The phonological loop is specialized for the storage and rehearsal of speech-based verbal information (Baddeley, 1992a, 1992b), whereas the sketch pad is specialized for holding visual and spatial material (Logie, 1986; Quinn & McConnell, 1996). In a recent development, Baddeley (1996) proposed a fractionation of the central executive into separate but overlapping functions of coordinating concurrent activities, switching retrieval plans, attending to inputs, and holding and manipulating information in long-term memory.

Evidence from studies of the normal population suggests that different components of working memory may have specialized roles in arithmetic (Ashcraft, 1995). For example, the phonological loop appears to be involved in counting (Logie & Baddeley, 1987) and in holding information in complex calculations (Logie, Gilhooly, & Wynn, 1994; see also Fürst & Hitch, in press). The visuospatial sketch pad appears to be involved in multidigit problems where visual and spatial knowledge of column positioning is required (Heathcote, 1994). Other researchers have noted the use of a visual number line (Dehaene, 1992) and spatial representations of individual numbers (Hartje, 1987), but without linking them to working memory. The role of the central executive has been noted several times (see, e.g., Ashcraft, Donley, Halas, & Vakali, 1992; Lemaire, Abdi, & Fayol, 1996; Logie et al., 1994). Ashcraft suggested that the executive is responsible for initiating and directing processing, comprehension, and retrieval from long-term memory. However, remarkably little is known. Accordingly, the following discussion generates tentative hypotheses about the role of different executive functions in arithmetic.

The first executive function identified by Baddeley (1996) is the capacity to coordinate performance on two or more separate tasks. Arithmetic can be considered a multiple task when it involves subtasks such as calculating partial totals and keeping track of other information. However, given that the subtasks are parts of an integrated skill, the requirement for coordination is presumably low relative to performing multiple independent tasks. The second executive function is switching retrieval strategies. This is clearly necessary for problems such as multidigit multiplication, which typically involves both multiplying and adding. Experimental evidence suggests that switching is also required for carrying operations (Fürst & Hitch, in press). The third executive function is attending selectively to different inputs. This is clearly a feature of multidigit
problems carried out as a series of subtasks, where attention is paid to selected parts of a problem at different times. The fourth executive function is activating and manipulating information in long-term memory. This seems likely to be involved in operations such as using the equivalence of two relationships (e.g., $5 + 2 = 4 + 3$) in order to simplify a calculation. In summary, therefore, nearly all the components of working memory seem likely to be involved in arithmetical calculation, each playing a somewhat different role.

In light of the foregoing considerations, the present study had two aims. The first was to gain a more complete picture of working memory deficits in children with specific arithmetic difficulties. This was achieved by using a battery of tasks assessing different components of working memory. The second aim was to consider whether working memory deficits are responsible for children's difficulties in arithmetic. This was investigated by applying the age-matched and ability-matched design described above. To ensure the selectivity of arithmetical deficits, all three groups were matched on reading. That is, in all three groups, children's reading was at normal, age-appropriate levels. An incidental advantage of using reading in this way is that it avoids controversial issues associated with matching groups on general intelligence (Siegel, 1988).

The battery for assessing working memory comprised two phonological loop tasks, Digit Span and Nonword Repetition (Gathercole, Willis, Baddeley, & Emslie, 1994); two visuospatial sketch pad tasks, Visual Matrix Span (Wilson, Scott, & Power, 1987) and Corsi Block Span (Milner, 1971); and five central executive tasks. The central executive tasks consisted of three Making Trails tasks (Reitan, 1958), to assess switching between retrieval plans; a Crossing Out task (Moran & Mefford, 1959), to measure selective attention; and a novel “Missing Item task,” designed as a measure of the ability to hold and manipulate information in long-term memory. All the tasks used numerical materials whenever possible, in order to maximize chances of observing deficits in the poor arithmetic group. A general assessment of working memory in the arithmetic domain was also made. This was provided by Addition Span (Adams & Hitch, 1997), a measure of how complex a mental addition children could perform accurately. More detailed rationales for these tasks are given in the Method section.

Expectations for comparisons between children with poor arithmetic (Poor Arithmetic) and their age-matched peers (Age-Matched) were as follows. Following previous work it was expected that children with Poor Arithmetic would be impaired on Digit Span (see, e.g., Geary et al., 1991; Hitch & McAuley, 1991; Siegel & Linder, 1984). Nonword Repetition should be similarly impaired if the reduction in digit span is due to a deficit associated with the phonological loop. If Rourke’s ideas generalize to the normal population, the Poor Arithmetic group should be impaired on the two visuospatial tasks, Corsi Blocks Span and Visual Matrix Span. Least clear of all, but perhaps most interesting, were expectations for the central executive tasks. Previous work demonstrating a deficit in counting
span (Siegel & Ryan, 1989) could be interpreted as suggesting an executive deficit. However, this conclusion may not be justified (Hitch & McAuley, 1991; see also Towse & Hitch, 1995). Furthermore, if the executive can be fractionated, then the possibility arises of deficits in specific executive functions. Of those considered, activating long-term memory seemed particularly interesting given previous work suggesting impaired access to long-term memory representations in arithmetic disabled children (Koontz & Berch, 1996).

No predictions were made for comparisons between the Poor Arithmetic group and the Ability-Matched controls, as this comparison has not previously been made. However, any deficit in the Poor Arithmetic group would be particularly interesting, as it would suggest a cognitive impairment that is not simply a consequence of low attainment.

Children were selected from a large sample attending normal schools by applying cutoffs to arithmetic and reading scores. This was considered more appropriate than using discrepancy scores (e.g., Swanson, 1994) or regression criteria, given the aim of investigating children with poor arithmetic in comparison to the chronological age norm. It also overcomes problems with poor approximation to the normal distribution. (For a review of selection techniques see Fletcher et al., 1989). The cutoff procedure selects children that fall below a set criterion for arithmetic ability and above one for reading ability. Children in the poor arithmetic group were chosen from the 4th year of schooling, as it was felt that by this age differences between reading and arithmetic would have emerged and it would be possible to find younger ability-matched controls.

METHOD

Participants

One hundred twenty-two children (64 females and 58 males) participated in the initial part of the study. Twenty-two were in their 3rd year of schooling (mean age 7 years 11 months), and 100 were in their 4th year of schooling (mean age 9 years 0 months). They attended five local primary schools in the Lancaster area. All children were administered the Graded Arithmetic–Mathematics Test (GAM; Vernon & Miller, 1976), the Primary Reading Test (France, 1979), and six 1-min written tests of speeded calculation (single- plus single-digit addition, single- plus two-digit addition, two- plus two-digit addition, subtraction, multiplication, and division; adapted from Hitch, 1978).

The Graded Arithmetic–Mathematics Test is a standardized group test for children aged 6–12 years. It has a time limit of 30 min and contains 70 written computation problems. These begin with arithmetical items and end with mathematical items. In the present study only arithmetic was assessed, as children never got as far as the mathematical items in the time limit. The Primary Reading Test (PRT) is another standardized group test suitable for 6- to 12-year-old children. It comprises 10 picture–word matching items and 38 sentence completion items and has a time limit of 30 min. Level 2 of the test was used in the
present study, as this was suitable for children aged above 7 years. Given that the standardized scores for the two assessment tests were calculated more than 20 years ago, raw scores were used to allocate children to groups.

Twelve children (7 girls and 5 boys) were selected for the Poor Arithmetic group on the basis that they fell in the bottom 25% of raw scores on the Graded Arithmetic–Mathematics Test and the middle 50% of raw scores on the Primary Reading Test. This number constituted 8.5% of the 4th-year sample. Twelve Age-Matched controls (6 girls and 6 boys) were selected whose arithmetic and reading raw scores were in the middle 50% of the sample. A further 12 Ability-Matched controls (8 girls and 4 boys) were selected from the 3rd-year sample on the basis that their arithmetic raw scores were similar to those of the Poor Arithmetic group and their reading was normal for their age. Descriptive statistics for each group are shown in Table 1.

$t$ tests were performed to ensure that the groups were adequately matched. These showed that the Poor Arithmetic group and Age-Matched controls did not differ on reading, $t(22) < 1$, but did on arithmetic, $t(22) = 10.1, p < .001$. The Poor Arithmetic group and Ability-Matched controls did not differ on arithmetic, $t(22) < 1$, but did on reading, $t(22) = 2.59, p < .05$.

Working Memory Battery

**Phonological Loop**

**Auditory digit span.** The forward auditory Digit Span task from the WISC-R subtest (Wechsler, 1974) was used as a standard measure of phonological short-term memory. Children were given the test twice, and digit span was calculated as the mean of the two scores. The reported reliability for 8- to 16-year-old children on this task is 0.85 (Weschler, 1974).

**Nonword repetition.** The Children’s Test of Nonword Repetition (Gathercole et al., 1994) was used as a second index of phonological short-term memory. This test involves the immediate repetition of auditorily presented nonwords and thus requires temporary storage of unfamiliar phonological sequences. Its reliability for 7-year-old children (the eldest reported) is 0.80 (Gathercole et al., 1994). Nonword repetition scores correlate reasonably well but not highly with digit span (Gathercole, Willis, Emslie, & Baddeley, 1992), suggesting that the two tasks tap a mixture of common and separate abilities.

The test consists of 40 nonwords, with 10 items at each length of from two to five syllables. The child was asked to listen to a tape recording of a female speaker reading aloud the nonwords in a neutral English accent and to repeat each one immediately after its presentation. Items were presented in a random order and there was a 3-s interval between the end of one item and the onset of the next. Credit was given for each correct repetition, with immediate self-corrections credited as a correct response.
Visuospatial Sketch Pad

Visual matrix span. A matrix span task adapted from Wilson, Scott, and Power (1987) was used to measure visual working memory. This task assesses immediate recognition memory for random visuospatial patterns of increasing complexity. Stimuli are presented relatively briefly in order to minimize the possibility of using verbal descriptions.

The present version was administered using a laptop computer. Stimuli consisted of a set of 15 matrices made up of different numbers of black boxes and white boxes. Each box was 2 cm square, and was drawn on a dark background.
The first matrix had two boxes, one black and one white, and appeared in the top left-hand side of the screen for 2000 ms. This was followed by a blank screen for 2000 ms before a similar matrix appeared. The second matrix differed from the first in that the black box had changed to white. Children were asked to point to the square which had changed. Once the child had given a response, the matrix increased by two more squares either down or across the screen. The boxes were filled black and white randomly, and the procedure was repeated as before. Trials continued up to a $30 \times 30$ matrix. Span was measured as the largest matrix remembered correctly before two successive failures. Each child completed the task twice and the two values of span were averaged.

*Corsi blocks span.* The Corsi blocks task (Corsi, 1972), described by Milner (1971), was administered to measure spatial span. The apparatus for this task is a set of identical blocks glued to random positions on a board. On each trial the participant observes the experimenter tap a sequence of blocks and then attempts to reproduce the sequence. Span is measured by gradually increasing sequence length to find the point where performance breaks down. Evidence that performance on this spatiotemporal task dissociates from verbal serial recall is presented by Isaacs and Vargha-Khadem (1989) and by Pickering, Gathercole, and Peaker (1998).

The apparatus consisted of nine 4.4-cm square wooden blocks which were painted black and fastened in random positions on a black board. To aid the experimenter, the numbers 1–9 were painted on the side of each block facing away from the child. Blocks were tapped at a rate of one per second and no block was tapped more than once in a trial. The first trial was a sequence of length 2. If this was repeated correctly, then a sequence of length 3 was tapped. If an error was made the child was given a second chance at sequence length 2. If the child succeeded on the second attempt, the task continued at length 3. If not, the task was stopped. This incremental procedure was continued to determine span, the longest sequence correctly repeated before two successive failures. Each child completed the task twice and the two values of span were averaged.

**Central Executive**

*Making Trails tasks.* Following Baddeley (1996), the ability to switch retrieval strategies was assessed using variants of the Making Trails task (Reitan, 1958). This task involves generating a stream of responses by alternating between two different sequences. The dependent variable was the time taken to complete the trail (or as much as could be done). There were three versions of the basic task, as follows.

For *Trails Written* the material consisted of 22 circles on a page of A4 paper. Half the circles had a number in the center (1–11), and half had a letter (A–K). Children were asked to start at number 1 and make a trail with a pencil so that each number alternated with its corresponding letter (e.g., 1-A-2-B-...-11-K; see Fig. 1). They were not prompted if a mistake was made in any of the sequences.
Reported reliabilities for this version of the Making Trails task lie between 0.60 and 0.90 (Lezak, 1995).

*Trails Verbal* was similar to Trails Written except that the children were asked to do the task orally. Again they began at 1 and finished with K, or (in case of error) the 22nd response.

*Trails Color* also was similar and was designed to minimize any effects of reading ability. It was adapted for use with children from an adult version (D’Elia & Satz, 1989). The materials consisted of 22 circles, half of which were pink and half yellow, on a page of A4 paper. Each circle contained a number from 1 to 11 such that each number appeared once in a pink circle and once in a yellow one. Children were asked to make a written trail starting with yellow 1, then pink 1, and so on, alternating colors with each number.

*Crossing Out task.* A Crossing Out task adapted from Moran and Mefford (1959) was used to measure selective attention. It is similar to a standard cancellation task except that the target is repeatedly shifted. Demands on selective attention are considered high because ignoring nontargets involves inhibiting responses to items previously designated as targets.

Each child was given a sheet of paper with 10 rows of digits. The first digit in each line was colored red and identified the target; the rest were black. The child was asked to cross out all the digits in the line that were the same as the red one.
Each line contained four randomly positioned targets. After the first line, which was used as practice, the child was instructed that the number to be crossed out was different for each line. The dependent variable was the time taken to complete the 9 remaining rows.

**Missing Item task.** This was a novel task developed to measure the capacity to hold and manipulate information accessed in long-term memory. It was a paper-and-pencil task with items consisting of one addition and a second, incomplete addition (e.g., \(2 + 3 = 4 + ? = ?\)). The child was asked to complete the equivalence and then the total. Three practice trials were given in which the experimenter talked through how to do the task. After this the child completed 15 more problems under instructions to be as accurate as possible. No prompting was given for incorrect responses. The dependent variable was the time taken to complete all 15 problems. The rationale for this task was that the child had to access long-term memory to complete the first relationship \((2 + 3 = ?)\) and then maintain this information and access long-term memory again to complete the second relationship \((5 = 4 + ?)\).

**General Resources**

**Addition Span task.** This task was developed by Adams and Hitch (1997), and involved mentally adding a pair of numbers. The size of the numbers was gradually increased to find a span beyond which performance breaks down. Addition Span shares the requirement to combine temporary information storage and ongoing processing with complex span tasks such as reading span and listening span (Daneman & Carpenter, 1980). However, given evidence that mental arithmetic loads on more than one component of working memory (Furst & Hitch, in press; Logie et al., 1994), it was considered here as a nonspecific measure of working memory resources.

The test was conducted orally and began with a single-digit plus single-digit addition. If the answer was correct, a one-digit plus two-digit addition was given next. If this was also answered correctly, the size of the next addition was increased by one digit according to the protocol described by Adams and Hitch (1997). Span was recorded as the largest problem answered correctly before two successive failures. Each child was given the task twice and a mean span was calculated from the two estimates. None of the additions involved carry operations, as pairs of integers making up the additions always had totals between 5 and 9.

**Procedure**

The task battery was administered 3 months after the initial screening, and toward the end of the school year. Each child was tested individually in a quiet area and completed the 10 tasks in two sessions about a week apart. Tasks were given in a different random order for each child, with the exception that the three Trail Making tasks were given together.
RESULTS

Differences between groups were analyzed using one-way analysis of variance (ANOVA) for each task supplemented by planned comparisons between the Poor Arithmetic and Age-Matched groups, and between the Poor Arithmetic and Ability-Matched groups. Table 2 shows means and standard deviations together with the results of the ANOVAs. There were a number of significant main effects, but there were no significant effects for Nonword Repetition and Visual Matrix Span, and only marginal effects for the Digit Span and Crossing Out tasks.

Planned comparisons revealed that the Poor Arithmetic group had significantly lower scores than did the Age-Matched controls on Corsi Span, $F(1, 33) = 13.35$, partial ($\eta^2$) = .29; Trails Written, $F(1, 33) = 5.73$, partial ($\eta^2$) = .15; Trails Verbal, $F(1, 33) = 7.82$, partial ($\eta^2$) = .19; Trails Color, $F(1, 33) = 13.04$,...
partial ($\eta^2$) = .28; Missing Item, $F(1, 33) = 32.15$, partial ($\eta^2$) = .49; and Addition Span, $F(1, 33) = 13.70$, partial ($\eta^2$) = .29. The Poor Arithmetic group also had lower digit spans than did the Age-Matched controls and this difference approached significance $F(1, 33) = 4.09$, $p = .051$. Comparisons against the Ability-Matched controls revealed that the Poor Arithmetic group were significantly impaired on the Missing Item task, $F(1, 33) = 9.17$, partial ($\eta^2$) = .22, but not on any of the other tasks.

Error rates were recorded for the Crossing Out and Missing Item tasks. Generally, these were low. The overall mean percentage was 0.6% for the Crossing Out task and 6.2% for the Missing Item task. Error rates on the Trail Making tasks were not recorded because of difficulties scoring and interpreting them.

Table 3 shows partial correlations (to control for effects of age) between the 10 working memory tasks and scores on the Graded Arithmetic–Mathematics Test, the Primary Reading Test, and the Calculation Test. Five of the working memory measures correlated significantly with arithmetic ability. These were the Missing Item task, Addition Span, Trails Verbal, Trails Color, and Corsi Span. The Calculation Test also showed a high correlation with arithmetic ability. It is interesting to note that none of the measures correlated with reading ability.

DISCUSSION

The results show that children with specific difficulties in arithmetic are impaired in several, but not all, aspects of working memory. However, as explained above, tasks on which the Poor Arithmetic group were impaired relative to Ability-Matched controls were considered the most likely to be informative about the basis of their learning difficulty. Accordingly, this aspect of the results is discussed first.

Relative to the Ability-Matched group, the Poor Arithmetic group were significantly impaired only in time to complete the Missing Item task. As no difference was noted in errors, this is unlikely to reflect a simple speed–accuracy trade-off. The Missing Item task was designed to assess the executive function of holding and manipulating information in long-term memory. However, identification of the critical feature of this task must be tentative bearing in mind its novelty and the current status of the proposed fractionation of executive function. Some information is provided by the observation that children with Poor Arithmetic did not differ from Ability-Matched controls on the Calculation Test and Addition Span. These tasks require access to similar arithmetical knowledge but do not appear to involve such demanding interactions with long-term memory. However, it is still open whether the children’s problem lies with resources for interactions with long-term memory or elsewhere. An alternative explanation is that the problem is one of understanding part–whole relations. Yet another is that the Poor Arithmetic group had weak conceptual and procedural knowledge of equivalence (see, e.g., Rittle-Johnson & Alibali, 1999). Thus, even though error
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>GAM Score</td>
<td>1.00</td>
<td>.28</td>
<td>.62***</td>
<td>.31</td>
<td>.06</td>
<td>.07</td>
<td>.42*</td>
<td>.33</td>
<td>-.52***</td>
<td>-.47**</td>
<td>-.25</td>
<td>.60***</td>
</tr>
<tr>
<td>2.</td>
<td>PRT Score</td>
<td>1.00</td>
<td>.20</td>
<td>.16</td>
<td>.16</td>
<td>.11</td>
<td>.15</td>
<td>.12</td>
<td>-.06</td>
<td>.07</td>
<td>.12</td>
<td>.12</td>
<td>.26</td>
</tr>
<tr>
<td>3.</td>
<td>Calculation Test 1</td>
<td>1.00</td>
<td>.36*</td>
<td>.30</td>
<td>-.05</td>
<td>.30</td>
<td>-.34</td>
<td>-.53**</td>
<td>-.54**</td>
<td>-.22</td>
<td>-.60***</td>
<td>.51**</td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Digit Span</td>
<td>1.00</td>
<td>.46**</td>
<td>.22</td>
<td>.31</td>
<td>-.23</td>
<td>-.07</td>
<td>-.30</td>
<td>-.06</td>
<td>-.31</td>
<td>.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Nonword Repetition</td>
<td>1.00</td>
<td>.34*</td>
<td>-.08</td>
<td>-.02</td>
<td>-.16</td>
<td>-.08</td>
<td>-.09</td>
<td>-.11</td>
<td>-.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Matrix Span</td>
<td>1.00</td>
<td>.04</td>
<td>-.24</td>
<td>-.02</td>
<td>-.12</td>
<td>.04</td>
<td>-.18</td>
<td>-.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>Corsi Span</td>
<td>1.00</td>
<td>-.40*</td>
<td>-.42*</td>
<td>-.44**</td>
<td>-.26</td>
<td>-.40*</td>
<td>.34*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>Trails Written (s)</td>
<td>1.00</td>
<td>.26</td>
<td>.53**</td>
<td>.40*</td>
<td>.38*</td>
<td>-.20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>Trails Verbal (s)</td>
<td>1.00</td>
<td>.37*</td>
<td>.21</td>
<td>.34*</td>
<td>-.38*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>Trails Color (s)</td>
<td>1.00</td>
<td>.30</td>
<td>.47**</td>
<td>-.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>Crossing Out (s)</td>
<td>1.00</td>
<td>.42*</td>
<td>-.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>Missing Item (s)</td>
<td>1.00</td>
<td>-.39*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>Arithmetic Span</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Correlation coefficients: df = 33 in all cases.

* $p < .05$.

** $p < .01$.

*** $p < .001$. 

---

**Working Memory and Arithmetic Difficulties**
rates were low, difficulty with the idea of equivalence could have reduced the speed with which appropriate procedures could be carried out. However, further (as yet unpublished) data argue against both of these interpretations. A somewhat different perspective is provided by Ericsson and Kintsch’s (1995) concept of long-term working memory as activated knowledge structures. These authors proposed that expertise within a domain is associated with more developed knowledge structures and an increase in the functional capacity of long-term working memory. In this framework, therefore, poor performance on the Missing Item task could be interpreted as reflecting an impairment in long-term working memory relative to children’s level of expertise.

More generally, a number of researchers have suggested ways in which learning difficulties in arithmetic might reflect links between working memory and long-term memory (e.g., Geary et al., 1991; Hitch & McAuley, 1991; Siegler & Shrager, 1984). To discuss all of these would be beyond the scope of the present study. However, the present results do speak to one particular suggestion, namely that some children fail to develop long-term memory for basic number facts because information in working memory decays too quickly for relevant associations to be formed (Geary et al., 1991). To the extent that simple and complex spans for verbal materials are influenced by decay (see, e.g., Baddeley, 1986; Towse & Hitch, 1995), this account is supported by the lower digit spans and addition spans for the Poor Arithmetic group compared with their Age-Matched controls. However, faster decay of information would not explain the absence of deficits in other retention tasks such as Nonword Repetition and Matrix Span. Furthermore, if faster decay were a causal factor, children with Poor Arithmetic might have been expected to perform significantly worse than Ability-Matched controls on all the above tasks, and this was not the case. Overall, therefore, the present results provide little support for the hypothesis that faster decay in working memory is a major factor in poor arithmetical attainment (Geary et al., 1991).

Other studies have attempted to explain the deficits of children with arithmetic difficulties in terms of speed of calculation (Garnett & Fleischner, 1983; Geary, Widaman, Little, & Cormier, 1987; Geary et al., 1991). These studies have found that arithmetic disabled children are slower than their normal-achieving peers in solving arithmetic problems (Garnett & Fleischner, 1983; Geary & Brown, 1991) and that speed of addition fact retrieval is related to performance on arithmetic tests (Geary & Burlingham-Dubree, 1989). However, in the present study there is no evidence to suggest that the Poor Arithmetic group were slower than their Ability-Matched peers in solving simple arithmetic problems. For example, there was no difference in the number of arithmetic problems they could do in 1 min on the Calculation Test (see Table 1). Thus, inclusion of Ability-Matched controls is valuable in showing that children with poor arithmetic do not have an unusually slow calculation speed relative to their level of attainment.

The second major feature of the results is that the Poor Arithmetic group were
significantly impaired relative to the Age-Matched controls on 6 of the 10 working memory measures: Corsi Span, Trails Written, Trails Verbal, Trails Color, Missing Item task, and Arithmetic Span (Digit Span was also near significance). Given that the Corsi task has no obvious arithmetical or numerical content, the first observation raises the possibility that a spatiotemporal working memory deficit contributes to poor arithmetic attainment. Interestingly, incidental observation of children’s behavior suggested that none of the Poor Arithmetic group had unusually poor handwriting or misaligned their arithmetic problems, implying that they did not have problems with spatial coordination. Neuropsychological studies of dyscalculia by Rourke and colleagues (e.g., Rourke & Findlayson, 1978; Rourke & Strang, 1978) had found visual–spatial deficits in tasks such as Picture Completion, Picture Arrangement, Block Design, and Picture Assembly from the WISC-R (Wechsler, 1974). Given that the Poor Arithmetic group were impaired on a test of spatial but not of visual working memory in the present study, Rourke and colleagues’ conclusions appear to generalize only partly to children with problems within the normal school population.

A second interesting set of differences between the Poor Arithmetic and Age-Matched groups are those on the Making Trails tasks. The visual versions, Trails Written and Trails Color, have been administered in previous studies of learning disabled children (e.g., McIntosh, Dunham, Dean, & Kundert, 1995; Share, Moffitt, & Silva, 1988; White et al., 1992; Williams et al., 1995). White et al. (1992) concluded that arithmetic disabled children were impaired on these tasks relative to normal and reading disabled children of the same age and interpreted this as reflecting a problem of visuomotor integration. However, although their basic finding was replicated, the observation that children were also impaired on the verbal Trails task argues against a visuomotor explanation. The idea of a deficit in switching between retrieval plans provides a more parsimonious interpretation. Moreover, given that the arithmetical content of the Trails tasks was minimal, it seems unlikely that children’s poor performance was a simple consequence their low arithmetic ability.

It is interesting that the Poor Arithmetic group were impaired on Digit Span relative to the Age-Matched controls, though the difference just missed significance. Other studies where groups with poor and normal arithmetic were matched on reading ability have shown small impairments in digit span (e.g., Hitch & McAuley, 1991). However, the picture is not entirely clear, as Bull and Johnston (1997) and Swanson (1994) found no evidence for lowered digit span in their groups of arithmetic disabled children when reading was controlled for statistically. As noted above, the deficit in digit span does not appear to reflect a problem associated with the phonological loop, as Nonword Repetition was unaffected. An alternative possibility is that the deficit arises from a specific difficulty in remembering numerical stimuli.

Results from the partial correlation matrix are useful in showing that all the
tasks which gave significant differences between the groups were correlated with Graded Arithmetic–Mathematics Test score. None of the measures correlated with reading ability. This was surprising with regard to Digit Span and Nonword Repetition given evidence for an association between verbal short-term memory and reading (Rugel, 1974). However, it should be noted that the way the groups were selected resulted in the sampling of a restricted range of reading abilities, reducing sensitivity of this aspect of the correlational analyses.

CONCLUSIONS

Children with specific difficulties in arithmetic were impaired relative to Age-Matched controls on a number of tasks tapping different aspects of working memory. Two of these, Corsi Span and Making Trails, suggest possible underlying deficits in spatiotemporal working memory and executive processes for switching retrieval plans. When compared with younger Ability-Matched controls, children with specific difficulties in arithmetic were significantly impaired on only one measure, a Missing Item task. This deficit is tentatively interpreted as reflecting a deficit in executive processes controlling interactions with long-term memory (Baddeley, 1996) or in long-term working memory (Ericsson & Kintsch, 1995).

REFERENCES


Received June 22, 1998; revised July 12, 1999