APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **18**: 1–16 (2004)
Published online 31 October 2003 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/acp.934

Working Memory Skills and Educational Attainment: Evidence from National Curriculum Assessments at 7 and 14 Years of Age

SUSAN E. GATHERCOLE[1]*, SUSAN J. PICKERING[2], CAMILLA KNIGHT[2] and ZOE STEGMANN[2]

[1]*University of Durham, UK*
[2]*University of Bristol, UK*

SUMMARY

The relationship between working memory skills and performance on national curriculum assessments in English, mathematics and science was explored in groups of children aged 7 and 14 years. At 7 years, children's levels of attainment in both English and mathematics were significantly associated with working memory scores, and in particular with performance on complex span tasks. At 14 years, strong links persisted between the complex working memory test scores and attainments levels in both mathematics and science, although ability in the English assessments showed no strong association with working memory skill. The results suggest that the intellectual operations required in the curriculum areas of mathematics and science are constrained by the general capacity of working memory across the childhood years. However, whereas success in the acquisition in literacy (tapped by the English assessments at the youngest age) was also linked with working memory capacity, achievements in the higher-level skills of comprehension and analysis of English literature assessed at 14 years were independent of working memory capacity. Copyright © 2003 John Wiley & Sons, Ltd.

In the UK, it has become increasingly important to identify the factors that influence and predict pupils' achievements in key scholastic domains such as literacy and mathematics. This is due in part to the requirement placed by the government on all local education authorities in England to report their pupils' levels of attainment on standard national assessments of core areas of the curriculum. The implications of the resulting public scrutiny for the individual school are far-reaching, influencing directly both parental satisfaction and the popularity of the school as a choice by prospective parents.

Some of the environmental factors influencing children's achievements at the earliest assessment point, at 7 years of age (Key Stage 1), are now known. Pupils attending nursery before starting school tend to perform more highly at Key Stage 1 assessments, whereas pupils receiving free school meals and from ethnic minority backgrounds score relatively poorly (e.g. Lindsay & Desforges, 1999; Strand, 1999). Other evidence points to the importance of what Tymms (1999) has referred to as 'general developed abilities' at

*Correspondence to: Susan E. Gathercole, Department of Psychology, University of Durham, Science Laboratories, South Road, Durham DH1 3LE, UK. E-mail: S.E.Gathercole@durham.ac.uk

Contract/grant sponsor: Medical Research Council.

school entry as predictors of children's achievements at 7 years. For example, Key Stage 1 attainments in both English and mathematics are highly predicted by children's abilities when they start school to identify letters, to spell, and to display knowledge about reading, and to count (Singleton et al., 1999; Tymms, 1999). So, children who have already mastered the component skills that lie at the foundations of literacy and mathematics tend to progress well and achieve highly following several years of structured education at school in these domains.

It is notable that such predictors of future achievements in national curriculum and other related educational assessments typically tap embryonic forms of the complex skill domain tapped in subsequent national curriculum assessments of literacy (e.g. letter familiarity, spelling knowledge) and mathematics (e.g. counting, and other concepts of number). These abilities represent crystallized knowledge that has been built up on the basis of the experiences in the home, other social settings, and in some cases in the nursery school.

However, children's abilities at school entry are not determined by learning opportunities alone; they are constrained also by basic cognitive capacities to learn. Two children sharing similar learning environments may differ markedly in the knowledge and skills with which they start formal schooling, as a consequence of their different capacities to learn. Also, comparable levels of performance on knowledge-based assessments may be the product of quite different combinations of environmental opportunity and basic cognitive capacities to learn. Clearly, environmental factors and cognitive structures are likely to interact and cannot be fully separated. From an educational viewpoint, it would nonetheless be extremely valuable to distinguish to as great a degree as possible between the environmental and cognitive factors influencing a child's current level of attainment, as a means of identifying the child's true learning potential.

One way of reducing the confounding of previous experience and basic cognitive capacities to learn may be to provide fluid assessments of cognitive processing knowledge in addition to tests of crystallized knowledge. Fluid measures tap children's capacities to perform cognitive operations that are involved in many complex learning situations encountered in the course of acquiring important skills such as literacy and mathematics. They differ critically from knowledge-based assessments by using tasks and materials that are equally unfamiliar to all children, and that are relatively impervious to demographic factors such as ethnic background (Dollaghan et al., 1997) and maternal education level (Gathercole et al., in preparation, 2002). In this way, measures of cognitive processing skills provide relatively pure assessments of the child's future abilities to learn (Baddeley & Gathercole, 1999).

In the present article, we report findings from two studies that investigated the relationship between children's performance in national curriculum assessments and their performance in one particular area of fluid cognitive ability that appears to play an important role in learning complex skills, working memory. The term 'working memory' is used to refer to a mental workplace in which information can be stored and processed for brief periods of time in the course of demanding cognitive activities. According to the most widely accepted model, working memory consists of a number of separate but interacting temporary memory systems (Baddeley, 1986; Baddeley & Hitch, 1974). The central executive is a limited-capacity processing system that lies at the heart of this model of working memory. It is suggested to regulate the flow of information through working memory, to coordinate access to and retrieval from more durable knowledge systems such as long-term memory, to control action, and to schedule multiple cognitive activities (e.g. Baddeley, 1986; Baddeley et al., 1998). Other theorists have conceptualized complex working memory either as a limited capacity system whose resources can be flexibly allocated to both processing and storage (Case et al., 1982; Just & Carpenter, 1992), or portions of long-term memory activated by a limited attentional resource (e.g. Cantor & Engle, 1993; Cowan, 2001).

In the Baddeley and Hitch (1974) model, the central executive is supplemented by two slave systems specialized for the maintenance of information in particular informational domains. The phonological loop is a temporary storage mechanism capable of storing limited amounts of information in terms of its phonological form. This subsystem is subject to rapid decay, although the contents of the phonological store can be maintained to a limited extent by a process of subvocal rehearsal (Baddeley, 1986). The second slave system, the visuo-spatial sketchpad, represents information in terms of its visuo-spatial features (e.g. Logie, 1995). A further subcomponent of working memory recently identified by Baddeley (2000) is the episodic buffer. It is suggested that the episodic buffer is responsible for integrating information from a variety of sources in the cognitive system, including both temporary and long-term memory systems. The detailed structure of the episodic buffer and methods of assessing its capacity have yet to be identified.

Close associations between individual differences in the central executive and phonological loop components of working memory and in capacities to acquire new knowledge and complex skills are well established. The phonological loop appears to play an important role in supporting the learning of the phonological structure of new words, both in native and second language acquisition (see Baddeley et al., 1998, for review). Accordingly, weak phonological loop skills are associated with poor vocabulary acquisition (e.g. Gathercole & Baddeley, 1989; Gathercole et al., 1997; Service, 1992), and severe impairments of the phonological loop may contribute to the profound language learning difficulties that characterize the developmental language pathology, Specific Language Impairment (Gathercole & Baddeley, 1990).

The impact of central executive capacity on learning may be even more pervasive. Poor central executive functioning has been suggested to compromise learning in key scholastic domains including literacy (e.g. Swanson, 1994; de Jong, 1998), arithmetic (Bull & Scerif, 2001; Passolunghi & Siegel, 2001; McLean & Hitch, 1999) and vocabulary (Daneman & Green, 1986), as well as more general aspects of cognitive capacity as indexed by college entrance scores (Daneman & Carpenter, 1980; Jurden, 1995) and occupational success (Kyllonen & Christal, 1990).

Gathercole and Pickering (2000a) provided evidence for a direct link between working memory abilities and pupils' performance on national curriculum assessments. Children failing to reach expected levels of attainment in English and mathematics performed poorly on central executive tasks involving both processing and storage of verbal material for brief periods of time. These children also showed some deficits, although less marked, on phonological loop measures. A combination of central executive and phonological loop measures proved highly effective at distinguishing children with poor performance on the national curriculum assessments from those with normal levels of achievement. Children with learning difficulties sufficiently severe to be classified by their schools as having special educational needs also perform extremely poorly on the same working memory measures (Gathercole & Pickering, 2001).

In the present article, we report further data on the relationship between working memory skills and pupils' attainments in national curriculum assessments. One group of children were aged seven or eight years of age, and had recently completed Key Stage 1 assessments. A second group of participants were 14 and 15 years of age, and were tested a few months after completing Key Stage 3 assessments. Both groups completed a selection of tests from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001) that had been found to discriminate most effectively pupils failing to achieve expected attainment levels in English and maths in the earlier study (Gathercole & Pickering, 2000a). The assessments included tasks that impose a combined storage and processing load widely used as measures of central executive capacities (backward digit recall, and listening recall) and storage-only measures of the phonological loop (word list recall, and word list matching). A further phonological loop measure, nonword repetition (Gathercole & Baddeley, 1996), was included for the Key Stage 1 group. On the basis of earlier findings that visuo-spatial working memory scores were not uniquely associated with performance on National Curriculum assessments (Gathercole & Pickering, 2000a), no visuo-spatial measures were included in the present study.

The study had two principal aims. The first aim was to establish whether earlier findings of close links between individual differences in working memory function and in national curriculum assessments at the end of the infant school years (Gathercole & Pickering, 2000a) could be replicated. The second was to determine whether associations between working memory and national curriculum performance extend to the final key stage of assessment, at 14 years of age.

METHOD

Participants

The younger group consisted of 40 children, 19 males and 21 females, from Year 3 of a state primary school in the southeast of England. The mean age of the group, who were tested at the beginning of the school year between September and October 2000, was 7 years 10 months ($SD = 3.58$ months, range 7 years 3 months to 8 years 6 months). All children were native English speakers.

The older group were 43 children (18 males and 25 females) attending Year 10 of a state secondary school in the southeast of England. Their mean age was 14 years 8 months ($SD = 3.20$ months, range 14 years 2 months to 15 years 2 months), when tested within the first four months of the school year, between October and December 2000. All children were native English speakers.

Procedure

Key Stage assessments

Each school provided transcripts of their pupils' levels of achievement on the relevant Key Stage assessments completed several months earlier, in July 2000. The Key Stage 1 assessments completed by the younger group were based on the following measures: (i) teacher assessments conducted between January and June 2000; (ii) tasks (reading, writing, and mathematics) administered between January and June 2000; and (iii) tests (reading comprehension, spelling, and mathematics) given in May 2000. All assessments used methods and materials developed and validated by the Qualifications and Curriculum Authority. Teacher assessments were classified at the following levels: W (working towards level 1), 1 (below nationally expected standard), 2 (nationally expected standard), and 3 (above nationally expected standard). The spelling and reading comprehension tests were only administered to children reaching at least level 2 on the reading task, so there

Table 1. Percentage of children achieving each level in relevant national curriculum assessments, by Key Stage group

	Key Stage 1 level/ stage					
Curriculum area	W	1	2c	2b	2a	3
English	2 (3)	5 (13)	10 (13)	20 (20)	12 (20)	50 (29)
Mathematics	2 (2)	2 (7)	12 (15)	22 (24)	27 (23)	32 (28)
	Key Stage 3 level					
Curriculum area	3	4	5	6	7	8
English	—	8 (20)	33 (22)	34 (22)	26 (8)	0 (1)
Mathematics	5 (9)	15 (18)	15 (24)	39 (23)	22 (17)	5 (3)
Science	15 (7)	17 (20)	7 (32)	36 (26)	24 (7)	2 (1)

Note: National data for 1999 are shown in parentheses.

were no W or level 1 classifications on either of the tests. Spelling performance was classified either as L (falling below level 2), level 2 or level 3. For the remaining tasks and tests, level 2 performance was divided into three grades in order to provide finer discrimination within the largest ability band. Grade 2a represents performance at the higher end, and grade 2c represents performance at the lower end, of the grade. Administration and marking of tests was audited and verified by the local education authorities. The percentages of children in the younger group achieving each grade/level on the tasks and tests at Key Stage 1 are shown in Table 1, alongside national figures for 1999 for comparison.

For the purposes of the present study, the children's attainments were classified on the basis of the standardized task and test results, rather than the teacher assessments whose psychometric properties are unknown. Each child was given a score ranging from 1 to 6 for English and mathematics. Individual attainment levels were assigned the numerical values shown in parentheses: W (1), 1 (2), 2c (3), 2b (4), 2a (5), and 3 (6). The mathematics score was based on the single mathematics test. The English score was based on each child's highest score on three assessments: the reading task, the reading comprehension test, and the writing task. The multiple assessments for English are likely to enhance the reliability of this derived score relative to the single mathematics measure. Attainment levels for the third curriculum area tested at Key Stage 1, science, were not analysed as they were based on teacher assessments only, and for all participating children except one were equivalent to the mathematics test attainment level.

Each child was also categorized as having low, average or high levels of attainment in English and mathematics. Children achieving levels W or 1 on the mathematics test were classified as low ability ($n = 2$), children obtaining level 2 were classified as average ability ($n = 25$), and those achieving level 3 ($n = 13$) were classified as high ability. Membership of the English ability groups was based on the highest attainment level achieved by the individual child on three of the English measures: reading, reading comprehension, and writing. The spelling test scores were not included in this classification as they lacked sufficient fine-grained discrimination within level 2 categories. There were three children in the low English achievement group, 17 children in the average group, and 20 children in the high achieving group. There was a high degree of overlap between the membership of the English and mathematics ability groups, as shown in Table 2.

Table 2. Co-membership of ability groups, by age and curriculum area

Age	Area of assessment	Ability group	Mathematics			Science		
			Low	Average	High	Low	Average	High
6 and 7	English	Low	2	1		—	—	—
		Average		17		—	—	—
		High		7	13	—	—	—
14 and 15	English	Low	1	2		1	2	
		Average	4	15	6	8	11	6
		High		5	5		5	5
	Science	Low	8	4		—	—	—
		Average		16	2	—	—	—
		High		2	9	—	—	—

The older group completed their Key Stage 3 assessments in July 2000, at 14 years of age. The school supplied the attainment levels of the participating children in the curriculum areas of English, mathematics and science. English assessment levels were not available for four of the children, and the mathematics attainment level was also missing for one of these children. In a small number of cases, assessment levels had not been assigned to the children in individual subjects due to school absences. Attainment levels at this stage range from 3 to 8, with levels 5 and 6 representing standards normally expected of this age group. Thus, levels 3 and 4 represent low achievements in comparison to the national norm, whereas levels 7 and 8 represent high achievements.

The achievement level of each child in each curriculum area was classified on this basis. In English, there were 5 low-achieving children, 22 average children and 12 high-achieving children (four children had no recorded scores in one of more of the English assessments, and so could not be classified). In mathematics, there were 8 low achieving children, 22 average children, and 12 high achieving children. In science, there were 9 low achieving children, 22 average children, and 12 high achieving children. The ability levels to which children were assigned in mathematics and science differed only for one child, due to the absence of a score on the mathematics test. Overlap in the composition of the three ability groupings in each curriculum area is shown in Table 2.

Working memory assessments

All children were tested individually in a quiet place in school in two sessions within a single week. Both age groups were given two central executive tests from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001): Backwards Digit Recall, and Listening Recall. They were also tested on two phonological loop measures from the same test battery: Digit Recall, and Word List Matching. The younger group also completed the Children's Test of Nonword Repetition (Gathercole & Baddeley, 1996), a further measure of phonological loop function.

Backwards Digit Recall

In this test, children hear sequences of spoken digits, and are asked to repeat them in backwards order. Testing commences with two-digit sequences, with the length of the sequences increasing by one digit over successive blocks of trials until the child makes errors on three lists at a particular list length. At this point, testing stops. Thus, the child must complete at least four trials at each list length (from a maximum of six trials) to proceed to the next level. The number of lists correctly recalled is scored, with credit given to the untested lists (i.e. the sixth and/or the fifth lists) if the child reaches the criterion for progression to the next sequence level before these lists are administered. This raw score was converted into a standard score for each child. The test-retest reliability of the test is 0.53 for children aged 5 to 7 years, and 0.71 for 9- to 11-year old children.

Listening Recall

On each trial of this test, children listen to a series of sentences, judge the veracity of each in turn, and then recall the final word of the sentences in sequence. The structure of the testing including discontinuation criteria is the same as the Backwards Digit Recall test outlined above. Both raw and standard scores were recorded. The test-retest reliability of the test is 0.83 for children aged 5 to 7 years, and 0.38 for 9- to 11-year old children.

Digit Recall

This test has the same structure as the Backwards Digit Recall test, except that the children are asked to recall the digits in the same sequence to the one spoken to them (i.e. forwards) rather than in backwards order. Both raw and standard scores were recorded. The test-retest reliability of the test is 0.81 for children aged 5 to 7 years, and 0.82 for 9- to 11-year old children.

Word List Matching

On each trial of this test, the child listens to a spoken sequence of familiar one-syllable words. Following a brief interval, the same words are presented again, either in the identical sequence, or with the position of two of the words within the sequence reversed. The child has to judge whether the sequences are the same or different. The number of words in each list increases over successive blocks of trial, following the same structure as outlined for the Backwards Digit Recall test. Both raw and standard scores were recorded. Test-retest reliability is 0.45 for children aged 5 to 7 years, and 0.42 for 9- to 11-year old children.

Non-word Repetition

In the Children's Test of Non-word Repetition (Gathercole & Baddeley, 1996), the child hears 40 spoken non-words, which range in length from two to five syllables. Following the presentation of each non-word, the child attempts to repeat the item. The total number of correct repetition attempts is scored, and a standard score calculated. Test-retest reliability assessed on a sample of five-year old children was 0.77. This test was administered to the younger age group only, as scores approach ceiling levels beyond nine years of age.

Results

The mean standard scores for the working memory measures for each group are shown in Table 3. The mean standard scores for the younger group were generally slightly higher than average, with the exception of the non-word repetition measure, which was rather low. The older group had scores on the backwards digit recall, listening recall, and digit recall tests that fell at the low end of the normal ability range. This group did, however, score highly on the word list matching test. As chance levels of performance on individual trials on this test would yield a 50% success rate, it is possibly that these high scores were inflated by guessing strategies.

Table 3. Descriptive statistics for each age group: age and working memory scores

	7 and 8 years				14 and 15 years			
	Raw score		Standard score		Raw score		Standard score	
Measure	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age in months	93.9	3.6	—	—	176.0	3.2	—	—
Phonological loop								
Digit recall	28.1	5.1	106.3	20.0	30.1	6.6	89.7	17.8
Word list matching	23.0	6.2	104.2	15.8	37.6	6.9	119.1	14.52
Non-word repetition	27.5	5.2	89.9	21.2	—	—	—	—
Pl composite			100.2	14.8			104.5	13.5
Central executive								
Backwards digit recall	11.8	3.7	102.1	22.9	14.1	2.7	89.0	13.1
Listening recall	12.3	4.8	106.1	18.8	14.3	5.0	95.0	14.6
CE composite			104.1	17.8			91.8	11.9

Table 4. Correlations between principal measures, for the younger group (lower triangle) and older group (upper triangle)

	1	2	3	4	5	6	7	8
1 Digit recall	—	**0.37**	—	**0.58**	0.20	−0.04	0.28	**0.32**
2 Word list matching	**0.51**	—	—	**0.32**	0.23	**0.32**	**0.41**	**0.49**
3 Non-word repetition	**0.44**	0.25	—	—	—	—	—	—
4 Listening recall	**0.50**	**0.35**	0.26	—	**0.46**	**0.32**	**0.54**	**0.50**
5 Backwards digit recall	0.26	0.15	0.26	**0.47**	—	**0.35**	**0.45**	**0.39**
6 English level	**0.41**	0.13	**0.36**	**0.49**	**0.44**	—	**0.63**	**0.68**
7 Mathematics level	**0.48**	0.17	0.25	**0.53**	**0.41**	**0.81**	—	**0.90**
8 Science level	—	—	—	—	—	—	—	—

Note: Coefficients printed in bold are significant at the $p < 0.05$ level.

The correlations between the working memory measures and the attainment levels in the curriculum areas are shown in Table 4. For the younger group, attainment levels in English and mathematics were highly correlated with one another, $r = 0.81$, $p < 0.001$. Both measures were significantly associated with digit recall ($r = 0.41$, $p = 0.01$, for English; $r = 0.481$, $p < 0.005$, for mathematics) and listening recall ($r = 0.49$, $p = 0.005$, for English; $r = 0.53$, $p < 0.001$, for mathematics). English level was also significantly correlated with non-word repetition scores ($r = 0.36$, $p < 0.05$). Word list matching was not significantly correlated with either attainment level. For the older age group, correlations were moderately high between attainment levels in English and mathematics, $r = 0.63$, $p < 0.001$, and English and science, $r = 0.68$, $p < 0.001$, and were very high between mathematics and science, $r = 0.90$, $p < 0.001$. Correlations were particularly high between mathematics and science attainment levels and three of the four working memory measures: word list matching ($r = 0.45$ and 0.39, $p < 0.001$, respectively), backwards digit recall ($r = 0.41$ and 0.49, $p < 0.001$, respectively), and listening recall ($r = 0.54$ and 0.50, $p < 0.001$, respectively). The correlations between attainment levels in English and the same three working memory scores were less strong although still significant (with *r*s ranging from 0.32 to 0.35, $p < 0.05$).

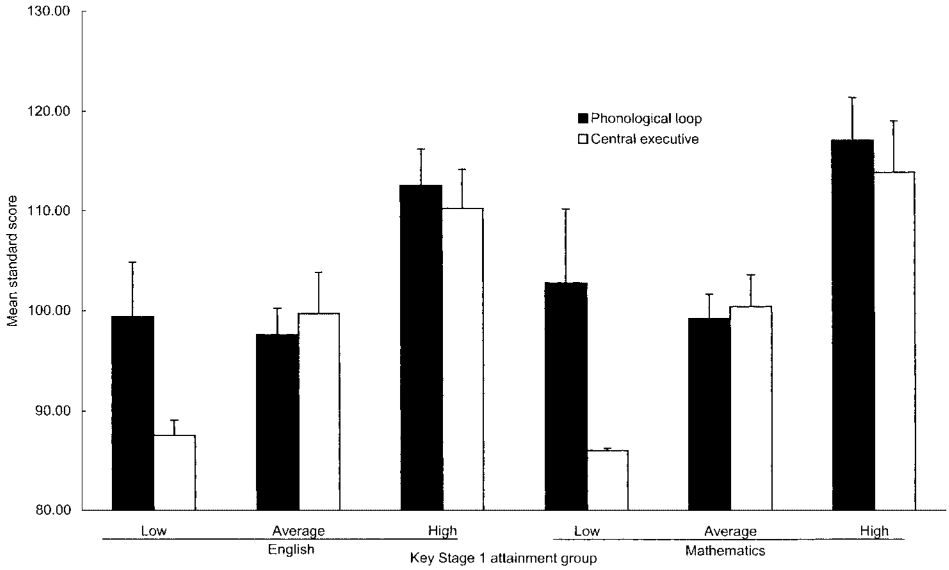

Figure 1. Mean standard scores on phonological loop and central executive measures as a function of English and mathematics ability groups in the younger group, with standard error bars

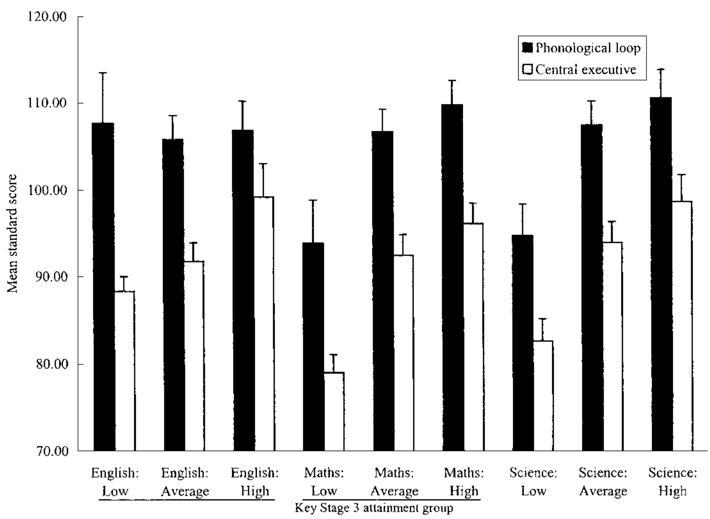

Figure 2. Mean standard scores on phonological loop and central executive measures as a function of English, mathematics and science ability groups in the older group, with standard error bars

A principal aim of the study was to assess the extent to which scores on the working memory tests effectively discriminated the children allocated to different ability groups on the basis of their national curriculum assessments. Figures 1 and 2 summarize the working memory profiles of the English and mathematics ability groups for the younger children based on their Key Stage 1 assessments, and of the English, mathematics and science ability groups for the older children on the basis of their Key Stage 3 assessments. The phonological loop and central executive standard scores were calculated from the mean values for each of the relevant measures. For the younger age group (Key Stage 1), central executive scores increased across the three ability groups in both English and mathematics. Phonological loop scores did not differ between the low and average ability groups in the two curriculum areas, but were substantially greater for the high than the average ability groups. For the older children, both central executive and phonological loop scores increased sizeably between the low and average ability groups, and between the average and high ability groups. There was, however, no consistent association between working memory scores and ability in English.

A series of MANOVAs were carried out in order to determine whether scores on the various working memory tests differed significantly across groups. Consider first the analyses of the data from the younger age group. There was a significant effect of English ability group (low, medium, and high) on working memory scores, $F(10, 64) = 2.07$, $p < 0.05$, using Hotelling's *T*. Univariate analyses established significant effects of group on three of the measures: non-word repetition, $F(2, 40) = 3.85$, $p < 0.05$; digit recall, $F(2, 40) = 5.09$, $p < 0.05$; listening recall, $F(2, 40) = 4.90$, $p < 0.05$. The effect of group on word list matching was marginally non-significant, $F(2, 40) = 3.21$, $p = 0.052$. Planned contrasts established significant differences between both the low and the average ability groups with the high ability group on digit recall, listening recall, and non-word repetition ($p < 0.05$, in each case), but no significant differences between the low and the average ability groups.

There was also a significant effect of mathematics ability group on working memory performance, $F(10, 64) = 2.30$, $p < 0.05$. By univariate analysis, significant group effects were found on digit recall, $F(2, 40) = 8.67$, $p < 0.005$, and on listening recall, $F(2, 40) = 5.19$, $p < 0.05$. Planned contrasts established no significant differences between the low and average ability groups on the working memory measures, although significantly better scores were found for the high than the average ability groups on digit recall, listening recall, and non-word repetition ($p < 0.05$, in each case).

In order to assess the extent to which working memory scores can identify children belonging to low-ability groups, a discriminant function analysis was performed in which the working memory scores were used to predict the three children with low abilities in English and mathematics. Two of these children belonged to the low ability groups in both English and mathematics, and the third child was classified as low ability in English but average ability in mathematics. Due to the high degree of overlap between membership of the English and mathematics ability groups, the groupings were not analysed separately by curriculum area. The best discrimination of group membership was achieved by using two of the working memory measures, listening recall and non-word repetition, with discrimination function coefficients of 0.49 and 0.78, respectively. The χ^2 value of this function was 6.50 with 2 df, $p < 0.05$. All three low-achieving children were correctly assigned to the low-ability group, plus 7 of the 37 children in the normal ability group. Thus, 82.5% of the original grouped cases were correctly classified. Inaccuracies in classification were due to a relatively high false positive rate (normal ability children misclassified as being of low ability), with all of the low ability children being correctly classified.

Corresponding analyses were performed on the data for the older age group, classified according to their Key Stage 3 assessments in English, mathematics, and science. In MANOVAs, significant effects of ability group on working memory scores were found in both science and mathematics: for mathematics, $F(8, 68) = 2.58$, $p < 0.05$, and for science,

$F(8, 70) = 3.47$, $p < 0.005$. Univariate analyses established significant effects of mathematics grouping on the two central executive measures: listening recall. $F(2, 41) = 5.93$, $p < 0.01$, and backwards digit recall, $F(2, 41) = 6.08$, $p = 0.005$. All pairwise differences between the three groups on these two measures were significant by planned contrasts ($p < 0.05$, in each case). Significant effects of science group on univariate analyses were obtained for listening recall, $F(2, 42) = 7.68$, $p < 0.05$, and word list matching, $F(2, 41) = 7.07$, $p < 0.005$. Planned contrasts established significantly lower scores of the low than average ability groups on the word list matching and digit recall measures, and between the average and high ability groups on the word list matching, listening recall, and backwards digit recall measures ($p < 0.05$, in each case).

In the corresponding MANOVA in which participants were grouped on the basis of English ability, however, no significant group effect was obtained on working memory scores, $F(8, 62) = 1.64$, $p > 0.05$. In univariate tests, a significant effect of group was found on one working memory measure: backwards digit recall, $F(2, 38) = 3.90$, $p < 0.05$.

The power of working memory scores to identify children belonging to low-ability groups in these curriculum areas was tested in a series of discriminant function analyses. For the purpose of these analyses, children originally classified as high ability were combined with the average ability children to form a single 'normal' ability group, for each curriculum area. For mathematics, best discrimination of the low and normal ability children was achieved by using the two central executive measures, backwards digit recall and listening recall, with standardized discrimination function coefficients of 0.58 and 0.65, respectively. The χ^2 value of this function was 11.29 with 2 df, $p < 0.005$. All 8 children belonging to the low-ability mathematics group were correctly assigned to this category, and 25 of the 33 children in the normal ability group were also correctly classified. Thus, 80.5% of the children were correctly classified using these two measures: all of the low-ability children, and around 75% of the normal ability group.

In the corresponding analysis of science ability groups, best group discrimination was achieved using listening recall and word list matching scores, with standardized discrimination function coefficients of 0.70 and 0.66, respectively. The resulting function had a χ^2 value of 19.93 with 2 df, $p < 0.001$. Eleven of the 13 children classified as low-ability in science were correctly grouped according to this function, and 23 of the 28 normal ability children. In all, 83% of the children were correctly grouped—85% of the low ability children, and 82% of the normal ability group. Finally, the discrimination function analysis using English ability groups (low and average) failed to generate a function with a significant χ^2 value, $p > 0.05$.

DISCUSSION

Close associations were found between children's scores on working memory measures and their national curriculum assessments. The results replicate earlier findings that children failing to achieve nationally expected levels in curriculum assessments at 7 years (Gathercole & Pickering, 2000a) have poor working memory function, and extend these findings to the national assessments of secondary school children at 14 years of age.

The detailed nature of the associations of scholastic achievements with working memory in this study did, however, differ at 7 and 14 years of age. At 7 years, children with high abilities in both English and mathematics scored better on working memory measures than children of low or average ability. Also, working memory scores were effective at discriminating children with low ability from the rest of the sample. At this age, the children's abilities in English and mathematics were very closely associated with one another.

A more specific pattern of associations between working memory and curriculum assessments emerged at 14 years. Attainment levels in English and mathematics were less closed linked to another than at Key Stage 1, although abilities in mathematics and science remained highly associated with one another. Achievements in mathematics and science were strongly correlated with working memory scores, with test scores differing significantly between both the low and average ability groups, and the average and high ability groups. In contrast, the children's performance on the school assessments in English was only weakly associated with working memory scores, and failed to differ significantly across the three English ability groups.

These results indicate that the contribution of working memory to levels of scholastic attainment during the school years varies considerably according to curriculum area. Consider first the English assessments. Strong links with working memory scores were present at 7 but not 14 years, and these were both greater in magnitude and more consistent at the earlier age for the central executive than the phonological loop tasks, replicating the findings of Gathercole and Pickering (2000a). Detailed consideration of the English tests administered at Key Stages 1 and 3 indicates that the apparent diminution in the link between central executive skills and proficiency in English may not reflect a genuine developmental change. At Key Stage 1, the three tests that were used to classify the children's attainment levels in the present study—reading, reading comprehension, and spelling—directly tap literacy abilities. In contrast, the tests administered at Key Stage 3 consisted of a reading and writing paper in which children answered questions on two written passages, and a second paper requiring an essay answer to a question relating to a scene from a Shakespeare play that had been studied by the children in class. So whereas the English tests at 7 years focused strongly on literacy skills, the assessments at 14 years provided much broader evaluations of the quality of the children's comprehension and interpretation of the literary pieces of work, and of the complexity and maturity of their written language. Although the two types of assessment are entirely appropriate for the age groups for which they were designed, they are clearly not equivalent in any formal sense.

The present findings of close links between English test achievements at 7 but not 14 years of age and performance on central executive tasks suggest a close relationship between working memory and children's emerging literary skills during the early years of learning to read and write. This conclusion fits well also with other recent evidence that children with learning difficulties in the area of literacy show marked deficits in complex working memory (de Jong, 1998; Siegel & Ryan, 1989; Swanson & Alexander, 1997; see also, Gathercole & Pickering, 2000b), and suggests that the capacity to process and store material simultaneously may crucially constrain the successful acquisition of literacy skills during the early years of learning to read.

The current study provides little evidence, however, that working memory contributes to the development of the higher-level conceptual and analytic abilities that are tapped by the English assessments at 14 years. By this age, it can reasonably assumed that as the majority of children have achieved a basic functional level of literacy, differences in the quality of their written language work reflect variation in intellectual skills affecting analysis and interpretation rather than literacy *per se*. The absence of a link between the storage and processing capacities associated with the central executive and higher-level language and reasoning abilities such as comprehension, inference and interpretation is somewhat surprising on the basis of previous findings (e.g. Oakhill et al., 1986).

The nature of the assessments administered in mathematics varied much less than the English assessments across Key Stage 1 and 3, although the later assessment stage obviously incorporated more advanced computations. At Key Stage 1, the mathematics test comprised some tests administered orally and some written questions, and was designed to tap the children's abilities in number and algebra, and in shape, space, and measurement. At Key Stage 3, the children completed two written tests of mathematical ability (one of which allowed the use of the calculator), and included some context-free, short mathematical computations employing the four basic arithmetic operations.

The consistently high associations between scores on the central executive measures and mathematical ability at 7 and 14 years fit well with other recent findings. In both children and adults, mental arithmetic activity has been shown to be selectively impaired by concurrent activities that demand the general working memory resources associated with the central executive (Adams & Hitch, 1997; Logie et al., 1994). It has also been found that children with poor mathematical abilities show deficits in complex working memory span tasks (Bull & Scerif, 2001; Mayringer & Wimmer, 2000; Siegel & Ryan, 1989). It is notable that in the present study (see also Passolunghi & Siegel, 2001), low levels of performance on complex working memory span tasks occurred irrespective of whether the task involved the manipulation of number (as in backwards digit recall) or language more generally (listening recall). This feature of the results indicates that the central executive problems of these children related to general capacity, rather than more specifically from difficulties in processing numerical information.

An important question is why this strong link between mathematics ability and working memory exists. At a basic level, the child must have adequate abilities to process number information, to count, to apply arithmetic rules, and to retrieve arithmetic facts from long-term memory (e.g. Geary, Hoard & Hamson, 1999). A suggestion by Geary et al. (1991) is that some children fail to acquire long-term memory representations for basic number facts because information in working memory decays too rapidly for the associations to be formed. Other authors have suggested that problems in inhibitory control in the central executive may be crucial (e.g. Bull & Scerif, 2001, Passolunghi & Siegel, 2001), although it is not entirely clear how such difficulties may compromise performance on complex span tasks. Another possibility consistent with the present findings is that the general-purpose workspace provided by working memory to support the many simultaneous cognitive demands of processing and storage placed by mathematical computations (Gathercole & Pickering, 2001). By this account, poor working memory capacity limits an individual's ability to meet these demands, and leads to errors even if basic mathematical competences are intact.

The associations between central executive capacities and attainments in science have not to our knowledge been previously reported. The very high associations between scores on the mathematics and science assessments at both ages raises the possibility that it is skills or processes common to both domains, such as number processing and problem solving, that benefit from central executive support.

One potential limitation of the present study is its use of a cross-sectional design, involving comparison of two samples of children differing be seven years of age. As the samples were randomly selected, we assume that they are representative of the larger population and therefore that group differences in associations between curriculum performance and working memory reflect genuine developmental change. It would be extremely valuable now to extend the current evidence of important shifts in the nature of the contribution of the central executive to learning and attainments in a larger-scale longitudinal study tracking the progress of a single sample of children across all school assessment stages, in order to provide a fine-grained analysis of the nature of the changing contributions of working memory to learning.

Finally, it should be acknowledged that although the working memory assessments effectively predict the attainment levels of the majority of children at Key Stage 1, and in the areas of science and mathematics at Key Stage 3, up to 25% of children achieving normal attainment levels for their age were incorrectly predicted to fall in low attainment categories on the basis of their working memory scores. The relationship between fluid cognitive ability (working memory scores) and crystallized knowledge (indexed by performance on the national curriculum assessments) is therefore close, but not perfect. Speculatively, these 'over-achieving' children may have basic learning capacities that place them at risk of educational under-achievement, but have been compensated either by other exceptional fluid abilities that were not assessed in the present study, or by beneficial environmental experiences that promoted learning.

ACKNOWLEDGEMENTS

This research was supported in part by a Medical Research Council programme grant on 'Working memory and learning disability' awarded to Alan Baddeley and Susan Gathercole.

REFERENCES

Adams, J. W., & Hitch, G. J. (1997). Working memory and children's mental addition. *Journal of Experimental Child Psychology, 67*, 21–38.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417–422.

Baddeley, A. D., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and the executive control of memory. *Quarterly Journal of Experimental Psychology, 51A*, 819–852.

Baddeley, A. D., & Gathercole, S. E. (1999). Individual differences in learning and memory: psychometrics and the single case. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences* (pp. 31–54). Washington, DC: APA.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–90). New York: Academic Press.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: inhibition, task switching, and working memory. *Developmental Neuropsychology, 19*, 273–293.

Cantor, J., & Engle, R. W. (1993). Working memory capacity as long-term activation: an individual differences approach. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19*, 1101–1114.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386–404.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour, 19*, 450–466.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25*, 1–18.
de Jong, P. F. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology, 70*, 75–96.
Dollaghan, C., Campbell, T., Needleman, H., & Dunlosky, G. (1997). Reducing bias in language assessment: a processing-dependent measure. *Journal of Speech, Language and Hearing Research, 40*, 519–525.
Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: a longitudinal study. *Journal of Memory & Language, 28*, 200–213.
Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: a study of young children learning new names. *British Journal of Psychology, 81*, 439–454.
Gathercole, S. E., & Baddeley, A. D. (1996). *The children's test of nonword repetition*. London: Psychological Corporation.
Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Short-term memory and long-term learning in children. *Developmental Psychology, 33*, 966–979.
Gathercole, S. E., & Pickering, S. J. (2000a). Working memory deficits in children with low achievements in the national curriculum at seven years of age. *British Journal of Educational Psychology, 70*, 177–194.
Gathercole, S. E., & Pickering, S. J. (2000b). Assessment of working memory in six- and seven-year old children. *Journal of Educational Psychology, 92*, 377–390.
Gathercole, S. E., & Pickering, S. J. (2001). Working memory deficits in children with special educational needs. *British Journal of Special Education, 28*, 89–97.
Geary, D. C., Brown, S. C., & Samanayake, 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, 47*, 175–192.
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., & Hamson, C. O. (1999). Numerical and arithmetical cognition: patterns of function and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology, 74*, 213–231.
Hitch, G. J., & McAuley, E. (1991). Working memory in children with specific arithmetic learning difficulties. *British Journal of Psychology, 82*, 375–386.
Jurden, F. H. (1995). Individual differences in working memory and complex cognition. *Journal of Educational Psychology, 87*, 93–102.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*, 122–149.
Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability (is little more than) working memory capacity. *Intelligence, 14*, 389–433.
Lindsay, G., & Desforges, M. (1999). Computerised baseline assessment of literacy. *Journal of Research in Reading, 22*, 55–66.
Logie, R. H. (1995). *Visuo-spatial working memory.* Hove: Lawrence Erlbaum.
Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory and Cognition, 22*, 395–410.
McLean, J., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240–260.
Mayringer, H., & Wimmer, H. (2000). Pseudoname learning by German-speaking children with dyslexia: evidence for a phonological learning deficit. *Journal of Experimental Child Psychology, 75*, 116–133.
Oakhill, J. V., Yuill, N., & Parkin, A. J. (1986). On the nature of the difference between skilled and less-skilled comprehenders. *Journal of Research in Reading, 9*, 80–91.
Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology, 80*, 44–57.
Pickering, S. J., & Gathercole, S. E. (2001). *Working memory test battery for children*. London: Psychological Corporation UK.

Service, E. (1992). Phonology, working memory and foreign-language learning. *Quarterly Journal of Experimental Psychology, 45A*, 21–50.

Siegal, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973–980.

Singleton, C., Horne, J., & Thomas, K. (1999). Baseline assessment in Scotland. *Journal of Research in Reading, 22*, 67–80.

Strand, S. (1999). Baseline assessment, value-added and the prediction of reading. *Journal of Research in Reading, 22*, 14–26.

Swanson, H. L. (1994). Short-term memory and working memory—Do both contribute to our understanding of academic achievement in children and adults with learning disabilities? *Journal of Learning Disabilities, 27*, 34–50.

Swanson, H. L., & Alexander, J. E. (1997). Cognitive processes as predictors of word recognition and residing comprehension in learning-disabled and skilled readers: revisiting the specificity hypothesis. *Journal of Educational Psychology, 89*, 128–158.

Tymms, P. (1999). *Baseline assessment and monitoring in primary schools*. London: David Fulton.

Hoard, Geary, Hamson (1999). (To be supplied).

