Educational and Child Psychology, (2003), Vol. 20(3)

Verbal and non-verbal working memory and achievements on National Curriculum tests at 11 and 14 years of age

Helen L. Jarvis and Susan E. Gathercole
University of Durham

Abstract

Recent research has found links between working memory and performance on National Curriculum tests. However, these studies have employed only verbal complex span tasks as measures of the central executive component of working memory. In the present study, children at Key Stages 2 and 3 of the National Curriculum (aged 11 and 14 years) were tested on measures of the phonological loop and visuospatial sketchpad as well as both verbal and non-verbal central executive tasks. Confirmatory factor analysis identified two factors: one verbal working memory factor associated with both the phonological loop and verbal central executive tasks, and one non-verbal working memory factor associated with both the visuospatial sketchpad and non-verbal central executive tasks. Further analyses were then used to explore the relationship between working memory and National Curriculum test scores in English, mathematics and science. The results suggested that both verbal and non-verbal working memory are important predictors of National Curriculum test scores.

Introduction

The term 'working memory' refers to a limited capacity system responsible for the simultaneous storage and manipulation of information during the performance of cognitive tasks (Baddeley, 1986). In recent years there has been considerable interest in the role of working memory in the development of skills in key scholastic domains including vocabulary (Daneman & Green, 1986), literacy (e.g. de Jong, 1998; Swanson, 1994) and arithmetic (Bull & Scerif, 2001). However, the resources proposed to underlie performance on working memory tasks differ widely across alternative models.

According to the most widely accepted model (Baddeley & Hitch, 1974; Baddeley, 2000) working memory consists of four components. At the heart of working memory is the central executive system, a domain general limited capacity system capable of controlling resources and monitoring information processing (e.g. Baddeley, 1986, 1996; Baddeley & Hitch, 1974; Baddeley, Emslie, Kolodny & Duncan, 1998). The central executive system is supported by two domain-specific storage components: the phonological loop that is responsible for the maintenance of auditory information, and the visuospatial sketchpad that is specialised for dealing with visual and spatial information. Baddeley (2000) recently identified the episodic buffer as a further subcomponent of working memory, responsible for integrating information from the subcomponents of working memory and long-term memory.

Other theorists have conceptualised working memory as a limited capacity system where processing and storage operations compete for a limited pool of resources (e.g. Case, Kurland & Goldberg, 1982; Daneman & Carpenter, 1980; Just & Carpenter, 1992). The working memory in this theory corresponds approximately to the central executive system within the Baddeley and Hitch (1974) model of working memory (Just & Carpenter, 1992).

Working memory is often assessed using complex span tasks, which require simultaneous processing and storage of information. According to Baddeley and Logie (1999), complex span performance relies upon a combination of domain-specific and domain-general resources, the visuospatial sketchpad and the phonological loop supporting domain specific storage, with processing drawing upon domain general resources. In addition, the central executive supports the coordination of processing and storage operations required in these tasks (Baddeley, 1996).

The contrasting view (Case, Kurland & Goldberg, 1982; Daneman & Carpenter, 1980; Just & Carpenter, 1992) is that a single working memory system supports both the storage and processing requirements of complex span tasks. Consistent with this view, performance on complex span tasks with different processing requirements, for example involving either language or numbers, are highly correlated with one another (e.g. Engle, Cantor & Carullo, 1992; Kyllonen, 1993; Shute, 1991; Turner & Engle, 1989). Other studies, however, have found marked dissociations between verbal and spatial complex span tasks (Jurden, 1995; Shah & Miyake, 1996). For example, Shah and Miyake (1996) found no significant correlation between verbal and spatial complex span measures. They used the reading span task (Daneman & Carpenter, 1980) in which participants read aloud sentences, while remembering the last word of each sentence. The spatial task involved performing a spatial transformation while keeping track of spatial locations. Scores significantly correlated with spatial abilities but not with verbal abilities. Correspondingly, the verbal complex span measure was correlated with verbal ability, but not spatial abilities, suggesting separate pools of resources for the two domains.

The different subcomponents of working memory have close links with a number of aspects of learning. Within the Baddeley and Hitch (1974) model of working memory, phonological loop skills have been associated with vocabulary acquisition in children with learning difficulties (e.g. Hulme & Mackenzie, 1992; Hulme & Roodenrys, 1995; Jarrold & Baddeley, 1997; Jarrold, Baddeley & Hewes, 1999; Russell, Jarrold & Henry, 1996), children with specific language impairment (e.g. Gathercole & Baddeley, 1990), and also within typically developing children (Gathercole & Adams, 1994; Gathercole & Baddeley, 1989; Michas & Henry, 1994). Scores on measures on central executive functioning have also been associated with vocabulary acquisition (e.g. Henry, 2001), language comprehension (e.g. Swanson & Ashbaker, 2000; Yuill, Oakhill & Parkin, 1989), and reading (Siegel & Ryan, 1989; Swanson & Ashbaker, 2000). Furthermore, all three of the main components of working memory have been associated with mental arithmetic (e.g. Dark & Benbow, 1990; De Rammelaere, Stuyven & Vandieredonck, 2001; Furst & Hitch, 2000; Hitch, 1978; Reuhkala, 2001; Seitz & Schumann-Hengsteler, 2000).

Working memory skills have also been found to be closely associated with performance on National Curriculum tests in England, in which all children in state schools are classified according to nationally expected standards in terms of their academic achievement. Children are assessed at three key stages, at 7, 11 and 14 years of age. At Key Stage 1 (for children aged 7 years) children are assessed on tests of English and mathematics, and at Key Stages 2 and 3 (ages 11 and 14) children are formally assessed on tests of English, mathematics, and science. Gathercole and Pickering (2000) found that children with low levels of performance on National Curriculum tests at 7 years of age showed marked impairments on measures of central executive functioning and of visuospatial memory. Gathercole, Pickering, Knight and Stegmann (in press) also found strong links between working memory and attainment levels in English and mathematics at 7 years of age, and links between working memory and mathematics and science at 14 years of age. At 14 years, working memory showed little association with English assessments.

However, each of the complex span tasks employed by Gathercole and Pickering (2000) and Gathercole *et al.* (in press) to tap central executive capacity were predominantly verbal in nature, involving for example the recall of the final words in sentences or sequences of digits in reverse order. An outstanding issue is whether in the light of evidence that verbal and spatial complex span tasks tap distinct resources (Jurden, 1995; Shah & Miyake, 1996), non-verbal complex span tasks are uniquely related to school achievements.

The primary aim of the present study was to examine whether there is evidence for separate verbal and non-verbal central executive resources in children and if so, if the two have distinguishable dissociations with performance on National Curriculum tests. A further aim was to establish whether previous findings of links between working memory and National Curriculum test scores (Gathercole & Pickering, 2000; Gathercole *et al.*, in press) could be extended to Key Stage 2 (11 years of age) of the National Curriculum as well as Key Stage 3 (14 years of age).

Method

Participants

The participants were 55 children (23 boys and 32 girls) with a mean age of 11 years and 6 months (SD = 3.24 months, range 10 years 11 months to 11 years and 9 months) and 73 children (35 boys and 38 girls) with a mean age of 14 years and 5 months (SD = 3.00 months, range 13 years 10 months to 14 years 9 months). The two groups of children were from two different local education authority schools in a suburban area of a city in north-east England. The socioeconomic background of the pupils at both the schools was mixed, but well above average. Percentages of pupils achieving level 4 or above on National Curriculum tests at 11 years of age were 81 per cent in English, 61 per cent in mathematics and 74 per cent in science, higher than percentages achieved nationally (65 per cent, 59 per cent and 68 per cent respectively). The percentages of pupils at 14 years of age achieving level 5 or above were 62 per cent in English, 67 per cent in mathematics and 68 per cent in science, in excess of the national percentages of 56 per cent, 59 per cent and 60 per cent, respectively. Both the working memory assessments and the National Curriculum tests were conducted during the summer term of the school year.

Procedure

All children took part in one testing session in which eight working memory tasks were administered. Two tasks tapping each putative aspect of working memory were used in the study. There were two measures of the phonological loop, taken from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001), and two of the visuospatial sketchpad. Two verbal complex span tasks, listening recall and backwards digit recall, were used along with two nonverbal complex span tasks, the spatial span task (Shah & Miyake, 1996) and the odd-one-out task, to tap the central executive component of working memory.

Each child was tested individually in a quiet area of the classroom. The order of presentation of the tasks was held constant with phonological loop tasks administered first. Visuospatial sketchpad tasks were administered second, followed by verbal and then nonverbal central executive tasks. A fixed order of testing across all children was employed in order to minimise individual variation due to differences in testing sequences.

Phonological loop tasks.

Each participant completed the digit recall test from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001). Participants are asked to recall, in the same order, sequences of digits spoken aloud by the experimenter. The digits are presented at the rate of one per second. Testing begins with three trials at a list length of two digits. The number of digits then increases by one every three trials until two lists of a particular length are recalled incorrectly. The score given is the maximum list length at which three sequences are recalled correctly. Test-retest reliability for digit recall is .82 for children aged 9 to 11 years (Pickering & Gathercole, 2001).

In the word recall test (Working Memory Test Battery for Children, Pickering & Gathercole, 2001) participants are asked to recall, in the same order, sequences of monosyllabic words spoken aloud by the experimenter. The structure of testing is identical to that for the digit recall task, but the score given is the maximum list length at which at least two out of three trials are recalled correctly, with an extra half a point if one out of three is correct at the next list length. Test-retest reliability for word recall is .64 for children aged 9 to 11 years (Pickering & Gathercole, 2001).

Visuospatial sketchpad tasks

The Visual Patterns Test (Della Sala, Gray, Baddeley, Allamano & Wilson, 1999) was administered to all children. Although it was originally developed for use with adults, it is suitable for use with children and can be normed alongside the Working Memory Test Battery for Children (Pickering & Gathercole, 2001). Participants are required to remember and recall chequerboard patterns. Each pattern is created by filling in half of the squares in a given grid. Following a practice trial, there are three trials at each grid size, from a 2 × 2 matrix to a 5 × 6 matrix. Each pattern is presented for three seconds, but there is no time limit for responding. The score given is the level of complexity (the number of filled squares contained by the grid) at which at least one of the three patterns is recalled correctly.

The dynamic matrices task is a computerised version of the Corsi blocks task, developed for the purpose of this study. The test was presented using Microsoft PowerPoint on a personal computer with a 33 cm monitor. Matrices increasing in size in the same manner as for the visual patterns test were presented in the centre of the screen. Squares within the matrices changed from white to black for one second in sequence. The participant was then asked to recall the sequence. The sequences were random with no location being highlighted more than once within a trial. The level of difficulty was increased by increasing the number of squares that went from white to black in a trial. Following a practice trial there were three trials at each level of difficulty. The score given was the longest sequence at which at least two of the three sequences were correctly reproduced.

Verbal central executive tasks

In the listening recall task (Working Memory Test Battery for Children, Pickering & Gathercole, 2001) participants hear a series of sentences and are asked to judge the veracity of each. At the end of each trial they are asked to recall the final word from each sentence. After two practice trials, each participant is given four trials with two sentences. After each four trials the number of sentences is increased by one. When two trials at any list length are incorrectly recalled, then the test ends. Each participant is given a score of the maximum list length at which they are correct on at least three out of four trials, and an additional half a point if correct on two trials at the next list length. Test–retest reliability for listening recall is .38 for 9- to 11-year-old children (Pickering & Gathercole, 2001).

The backwards digit recall test (Working Memory Test Battery for Children, Pickering & Gathercole, 2001) requires each participant to recall a sequence of spoken digits (between one and nine) in reverse order. The structure of the testing includes discontinuation and scoring criteria the same as for the digit recall test outlined above. Test–retest reliability for backwards digit recall is .71 for children aged 9 to 11 years (Pickering & Gathercole, 2001).

Non-verbal central executive tasks

The spatial span task (Shah & Miyake, 1996) was modified for the purposes of the present study in order to eliminate any involvement of long-term memory or verbal working memory. The test stimuli were thus five nonsense shapes presented either in a normal view or as a mirror image, in one of eight spatial orientations. Each participant was required to state whether each shape presented was 'normal' or a 'mirror image' of an original shape, while keeping track of the orientation of each shape. After each trial at list lengths of two to seven shapes, the participant was asked to recall the position of the top of each shape. Each shape was shown for only two seconds to minimise the possibility that participants delay the mental rotation in order to rehearse the orientations. Participants were given a score of the longest list length at which they were correct on at least two out of the three trials. They were given half a point extra if they were correct on one out of three trials at the next list length

The odd-one-out task (based on the procedure used by Russell *et al.* 1996) consisted of sets of three shapes. Two of the shapes were identical and one was different. The participant's task was to indicate the odd one out. Each set of three shapes was shown for only two seconds, then immediately followed by another set, to minimise the possibility that participants delayed the judgement of the odd one out to rehearse the spatial locations. Following each trial (in list lengths of two to seven) the participant was asked to recall the spatial location of all the odd-one-out shapes. The participant was given a score of the longest list length at which they were correct on at least two out of three trials. An extra half a point was awarded if the child made a correct response on one out of the three trials at the next list length.

National Curriculum tests

The schools supplied attainment levels in English, mathematics and science for each pupil. These levels were based on standardised tests taken in the summer term, and were independent of teacher assessments of ability. At Key Stage 2, English test scores incorporate measures of reading, writing, spelling and handwriting. Two mathematics papers and a mental arithmetic test are used to generate a mathematics score, and there are two science papers. Each test has high reliability, with Cronbach's alpha for each subtest ranging from .86 to .89 (QCA, 2001). Attainment levels provided for each child range from 3 to 5, with level 4 indicating nationally expected standards. At Key Stage 3 English assessments differ somewhat from at Key Stage 2, assessing more complex abilities. For example, within the 'reading' subtest of English, children have to demonstrate their understanding of literature and make comments on reader–writer relationships within text. Again, a mathematics score is generated from mathematics and mental arithmetic tests, and there are two science papers. Cronbach's alpha for subtests range from .85 to .94 (QCA, 2001). The levels of attainment at this Key Stage range from 3 to 8, with levels 5 and 6 indicating nationally expected standards.

Results

Descriptive statistics for all the working memory measures and National Curriculum test levels are presented in Table 1. Within both age groups, skew and kurtosis for all measures met criteria for multivariate normality (Kline, 1998). No univariate or multivariate outliers were identified.

Table 1. Descriptive statistics of working memory measures and National Curriculum test scores for English, mathematics and science

	Key Stage 2		Key Stage 3	
Measures	Mean	SD	Mean	SD
Working memory measures				
Digit recall	4.82	0.92	5.21	0.93
Word recall	3.28	0.64	3.75	0.72
Visual patterns	7.24	1.32	8.41	1.31
Dynamic matrices	3.29	0.92	3.56	0.76
Listening recall	2.81	0.74	3.12	1.60
Backwards digit	4.00	0.92	4.56	0.85
Spatial span	2.50	0.59	3.08	0.40
Odd-one-out	3.48	0.80	3.60	0.77
National Curriculum tests				
English level	4.44	0.74	5.89	1.14
Mathematics level	4.20	0.78	5.88	1.46
Science level	4.53	0.60	5.42	1.10

Table 2. Zero order correlation coefficients between the working memory measures and national curriculum attainment levels; upper triangle displaying coefficients for Key Stage 2, lower triangle displaying coefficients for Key Stage 3

		1	2	3	4	5	6	7	8	9	10	11	12
1	Age	–	.32*	.36**	.14	.40**	.26	.21	.37**	.20	.06	.19	.26
2	Digit recall	.45**	–	.49**	.76**	.63**	.22	.37**	.40**	.31*	.34*	.44**	.21
3	Word recall	.34**	.53**	–	.47*	.64**	.30*	.46**	.32*	.36**	.40**	.53**	.37**
4	Backwards digit	.36**	.56**	.39**	–	.61**	.32*	.35**	.26	.38**	.35**	.49**	.27*
5	Listening recall	.27*	.46**	.34**	.43**	–	.47**	.59**	.51**	.47**	.46**	.58**	.46**
6	Visual patterns	.18	.30*	.37**	.42**	.37**	–	.56**	.51**	.72**	.29*	.46**	.42**
7	Dynamic matrices	.08	.19	.20	.35**	.17	.56**	–	.55**	.62**	.33*	.57**	.39**
8	Odd-one-out	.19	.43**	.37**	.57**	.27*	.33**	.25*	–	.50**	.42**	.54**	.46**
9	Spatial span	.19	.18	.28*	.33**	.26*	.41**	.44**	.35**	–	.26	.47**	.47**
10	English level	.17	.36**	.47**	.50**	.52**	.36**	.25*	.45**	.37**	–	.71**	.60**
11	Mathematics level	.05	.23*	.40**	.47**	.31**	.42**	.32**	.48**	.47**	.73**	–	.68**
12	Science level	.09	.21	.44**	.50**	.24*	.34**	.22	.48**	.41**	.65**	.85**	–

*$p < .05$ ** $p < .01$

Correlational analyses

The correlation matrix of the working memory tasks and the National Curriculum attainment levels is presented in Table 2. Age is also included. The upper triangle displays correlation coefficients at age 11, and the lower triangle displays correlation coefficients at age 14.

At 11 years of age, age was significantly correlated with digit recall, $r(53) = .32$, $p < .05$, word recall, $r(53) = .36$, $p < .01$, listening recall, $r(53) = .40$, $p < .01$ and odd-one-out scores, $r(53) = .37$, $p < .01$. At 14 years of age, age was significantly correlated to digit recall, $r(71) = .45$, $p < .01$, word recall, $r(71) = .34$, $p < .05$, listening recall, $r(71) = .27$, $p < .05$ and backwards digit recall, $r(71) = .36$, $p < .01$.

The majority of the scores on the working memory measures were also significantly correlated with each other. The correlation coefficients between scores on the two tasks aimed at tapping the phonological loop were significant at age 11, $r(53) = .49$, $p < .01$ and age 14, $r(71) = .53$, $p < .01$, as were those between the two tasks aimed at tapping the visuospatial sketchpad, $r(53) = .56$, $p < .01$, $r(71) = .56$, $p < .01$. Scores on the two verbal central executive tasks were also significantly correlated in both age groups, $r(53) = .61$, $p < .01$, $r\ (72) = .43$, $p < .01$, as were those on the two non-verbal central executive tasks, $r(53) = .50$, $p < .01$, $r(71) = .35$, $p < .01$.

The correlation coefficients between scores on phonological loop and verbal central executive tasks were also highly significant in both age groups, the highest correlation being between digit recall and backwards digit recall, $r(53) = .76$, $p < .01$, $r(71) = .56$, $p < .01$. Scores on visuospatial sketchpad and non-verbal central executive tasks were also significantly correlated in both age groups, the highest correlations being between scores on the visual patterns test and spatial span, $r(53) = .72$, $p < .01$, $r(71) = .41$, $p < .01$.

Correlational analyses also revealed that at 11 years of age, odd-one-out scores correlated highly with scores on non-verbal tasks such as dynamic matrices, $r(53) = .55$, $p < .01$, and the visual patterns test, $r(53) = .51$, $p < .01$. However, at 14 years of age, the odd-one-out task correlated more highly with tasks within the verbal domain, such as backwards digit span, $r(71) = .57$, $p < .01$.

Highly significant correlations were also found between a number of the working memory measures and National Curriculum attainment levels. At Key Stage 2 the strongest associations were found between English levels and listening recall, $r(53) = .46$, $p < .01$, and mathematics levels and listening recall, $r(53) = .58$, $p < .01$. Science scores were highly correlated with listening recall, $r(53) = .46$, $p < .01$ as well as odd-one-out scores, $r(53) = .46$, $p < .01$, and spatial span, $r(53) = .47$, $p < .01$. At Key Stage 3 the strongest associations were found between English levels and backwards digit recall, $r(71) = .50$, $p < .01$, and listening recall, $r(71) = .52$, $p < .01$. Mathematics levels were most strongly associated with odd-one-out scores, $r(71) = .48$, $p < .01$, and spatial span, $r(71) = .47$, $p < .01$, as well as backwards digit recall, $r\ (71) = .47$, $p < .01$. Science levels were most highly correlated with backwards digit recall, $r(71) = .50$, $p < .01$, and odd-one-out scores, $r(71) = .48$, $p < .01$.

In order to evaluate the extent to which unique associations were obtained between working memory tasks and National Curriculum attainment levels, composite scores were calculated for the phonological loop, visuospatial sketchpad, verbal central executive and non-verbal central executive by averaging the *z* scores on the associated tasks. Partial correlations between each construct and English, mathematics and science scores were calculated, eliminating variance related to age and the other working memory constructs in each case. Corresponding correlational analyses were conducted partialling out variance associated with age and the constructs in the opposite domain, for example, in the partial correlations involving the phonological loop the visuospatial constructs (visuospatial sketchpad and non-verbal executive) were partialled out. However, the high correlations between the odd-one-out task with tasks in the verbal domain at age 14 suggest that the task tapped a verbal rather than a non-verbal construct. Scores on the odd-one-out task were therefore excluded from further analysis of the data at 14 years of age, with the spatial span task used as a single measure of non-verbal executive processes. This procedure is highly conservative, given the high degree of intercorrelations between the variables but does provide a very stringent test of the specificity of the relationships between the components of working memory and attainment. The partial correlation coefficients are presented in Table 3.

The partial correlations revealed that when the other three working memory constructs were taken in to account, the only unique link between working memory and National Curriculum attainment levels at Key Stage 2 was that between nonverbal central executive scores and science, $r(49) = .29$, $p < .05$. The few significant partial correlations, however, were likely to be a result of the high intercorrelations between simple and complex span tasks within the same domain. When partialling out only the constructs in the opposite domain, i.e. verbal or non-verbal as appropriate, a stronger pattern of associations emerged. The verbal constructs (phonological loop and verbal central executive) were highly correlated with English levels, $r(50) = .33$, $p < .01$, $r(5) = .33$, $p < .01$, and mathematics levels, $r(50) = .41$, $p < .01$, $r(50) = .39$, $p < .01$. The visuospatial constructs (the visuospatial sketchpad and non-verbal central executive) were significantly related to mathematics levels, $r(50) = .39$, $p < .01$, $r(50) = .44$, $p < .01$, and science levels, $r(50) = .30$, $p < .01$, $r(50) = .40$, $p < .01$. At Key Stage 3 even when all other working memory constructs were partialled out, verbal central executive scores were significantly correlated with levels in English,

Table 3. Partial correlation coefficients between working memory constructs and attainment levels in English, mathematics and science

Assessment	Working memory construct					Working memory construct				
	PL	VSSP	VCE	NVCE	SS	PL	VSSP	VCE	NVCE	SS
Constructs partialled out:	VSSP VCE NVCE/ SS	PL VCE NVCE/ SS	PL VSSP NVCE/ SS	PL VSSP VCE	PL VSSP VCE	VSSP NVCE/ SS	PL VCE	VSSP NVCE/ SS	PL VCE	PL VCE
Key Stage 2										
English	.14	-.00	.13	.17	–	.33**	.16	.33**	.23	–
Mathematics	.19	.17	.15	.20	–	.41**	.39**	.39**	.44**	–
Science	-.03	.02	.12	.29*	–	.09	.30**	.15	.40**	–
Key Stage 3										
English	.21	.00	.41**	–	.20	.39**	.09	.51**	–	.21
Mathematics	.17	.13	.25*	–	.32**	.29*	.26*	.34**	–	.38**
Science	.19	.02	.23	–	.28*	.31*	.14	.33**	–	.31*

Note: PL = phonological loop; VSSP = visuospatial sketchpad; VCE = verbal central executive; NVCE = non-verbal central executive; SS = spatial span.

$^{*}p < .05$ $^{**}p < .01$

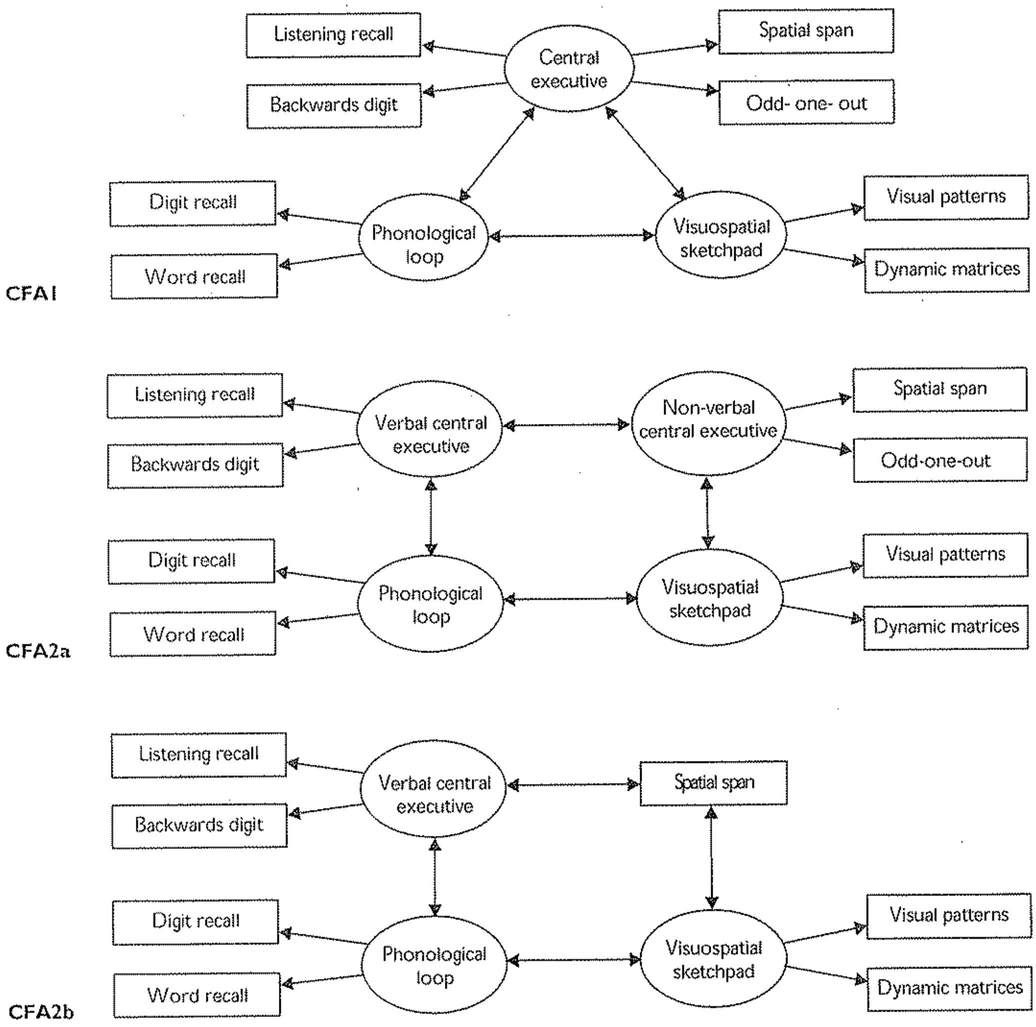
$r(67) = .41$, $p < .01$, and mathematics, $r(67) = .25$, $p < .05$, and non-verbal central executive scores were significantly correlated with mathematics levels, $r(67) = .32$, $p < .01$, and science levels, $r(67) = .28$, $p < .05$. When controlling for constructs in the other domain phonological loop scores were also significantly correlated with levels in English, $r(68) = .39$, $p < .01$, mathematics, $r(68) = .29$, $p < .05$ and science, $r(68) = .31$, $p < .05$ and visuospatial sketchpad scores were correlated with mathematics levels, $r(68) = .26$, $p < .05$.

Factor analysis

Factor analysis and structural equation modelling were then conducted using the EQS 6 structural equation package (Bentler, 2001). The purpose of this approach was to test, formally, different theoretical models of the relationships between latent constructs tapped by a number of measures. Each model assessed in structural equation modelling generates coefficients for the paths between constructs and variables, indicating the strength of relationships. A number of statistics are produced that indicate the goodness of fit of the model to the input correlation matrix. By comparing the fit indices across competing models it is possible to find the best theoretical account of the data.

Consider first the data from the children at 11 years of age. Four models of the structure of working memory were tested. The first model (CFAI) corresponded to the standard Baddeley and Hitch (1974) working memory model with three factors representing the central executive, phonological loop and visuospatial sketchpad. The second model (CFA2a) fractionated the central executive in to distinct verbal and non-verbal components (Shah & Miyake, 1996), and was therefore composed of both verbal and non-verbal central executive components in addition to the phonological loop and visuospatial sketchpad. The third model (CFA3a) eliminated the distinction between the central executive and the two domain-specific storage systems and consisted of one verbal factor

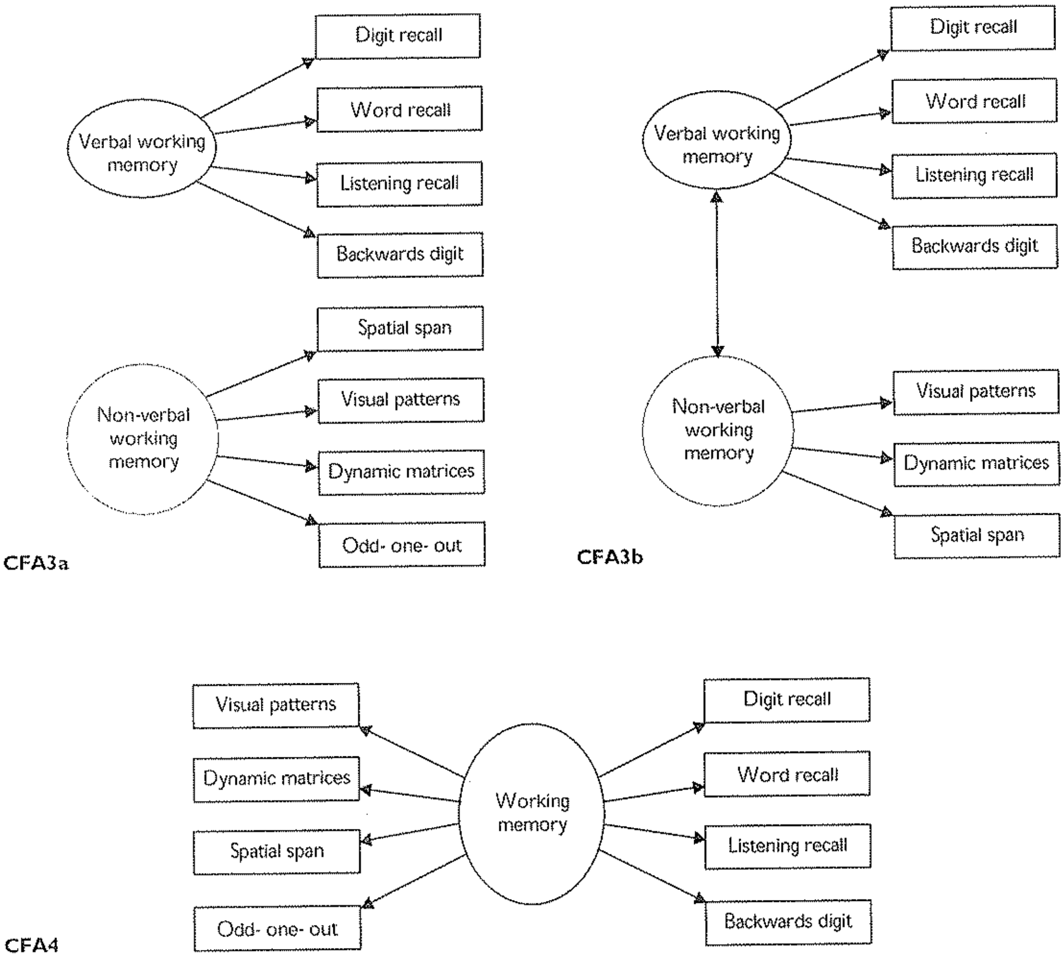
Figure 1a. Diagrammatic representation of the confirmatory factor analysis models



incorporating the phonological loop and verbal complex span measures, and one non-verbal factor including both the visuospatial sketchpad and non-verbal complex span measures. In the final model (CFA4), all of the working memory tasks were associated with a single common factor. A diagrammatic representation of these models is shown in Figures 1a and 1b.

The fit statistics for these models for the Key Stage 2 data are presented in Table 4. The fit statistics used are chi squared (χ^2), the comparative fit index (CFI) (Bentler, 1990), Bollen's incremental fit index (IFI), the standardised root mean square of the model residuals (SRMR), and the root mean square error of approximation (RMSEA). The best-known index of fit is χ^2, which measures the degree to which the covariance's predicted by the model differ from the observed covariances. Small and insignificant χ^2 values indicate good fit. CFI and IFI indicate the extent to which the model is better than a baseline model with all

Figure 1b. Diagrammatic representation of the confirmatory factor analysis models



covariances set to zero. Values should equal or exceed .90 for adequate fit of model to the data. The SRMR is the square root of the averaged squared residuals, i.e. differences between observed and predicted covariances. A value of 0.08 or less represents acceptable goodness of fit (Hu & Bentler, 1999). The RMSEA is also a measure of the discrepancies between observed and predicted covariances, and values less than .05 correspond to a good fit and values less than .08 correspond to an acceptable fit.

Model 1, the standard three-factor working memory model, did not provide satisfactory fit to the data (both fit indices < .90). The model that yielded fit indices (CFI and IFI) in excess of .90 was the two-factor domain-specific model composed of one verbal and one non-verbal factor (CFA3a). It should, however, be noted that the fit of this model was not ideal. The χ^2 value was significant ($p = .03$) and the RMSEA value was .11. The factor loadings and item error terms for this model are presented in Figure 2a. All loadings and variances are significant at the .05 probability level.

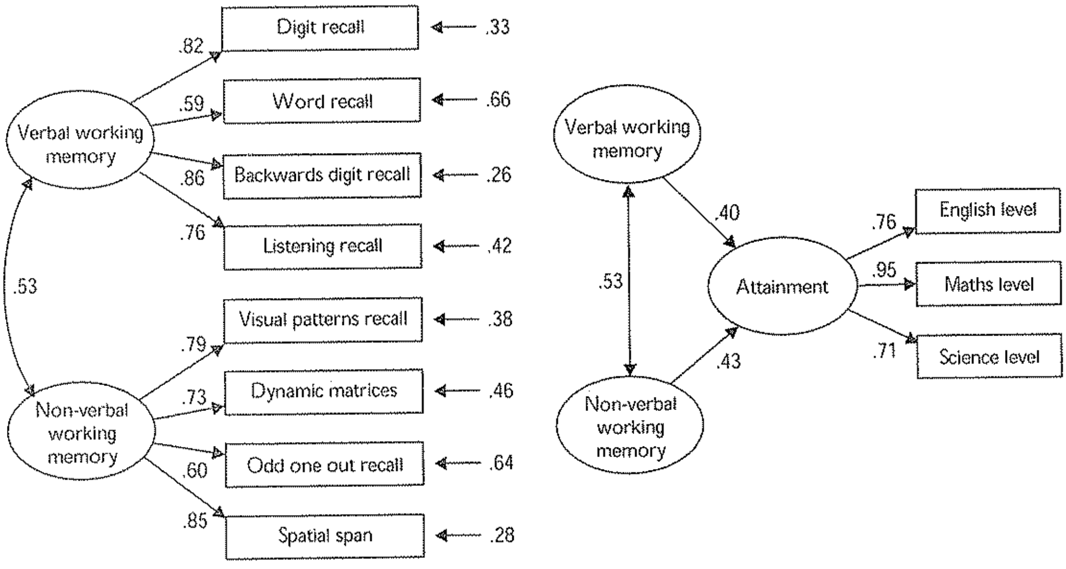
Table 4. Goodness of fit statistics for the estimated models

Model	d.f.	χ^2	p	CFI	IFI	SRMR	RMSEA
Key Stage 2							
CFA1	17	55.5	.00	.80	.81	.11	.21
CFA2a	16	41.3	.00	.87	.88	.24	.17
CFA3a	19	32.0	.03	.93	.94	.09	.11
CFA4	20	72.9	.00	.72	.73	.12	.22
SEM	49	56.3	.21	.97	.97	.09	.05
Key Stage 3							
CFA1	11	14.7	.20	.96	.98	.07	.07
CFA2b	9	22.8	.01	.86	.88	.16	.15
CFA3b	14	11.4	.58	1.0	1.0	.05	.00
CFA4	14	28.5	.01	.85	.86	.09	.12
SEM	39	47.2	.17	.97	.97	.08	.05

Note: CFI = Bentler's comparative fit index; IFI = Bollens incremental fit index; SRMR = standardised root mean squared residual; RMSEA = root mean square error of approximation

The factor loadings produced in the best fitting, two-factor confirmatory factor analysis model (CFA3) were then incorporated into a structural equation model (SEM) in which verbal and non-verbal working memory predicted English, mathematics and science scores. In the model, both verbal and non-verbal factors were causally linked with a single attainment factor associated with English, mathematics and science attainment levels. The fit indices produced for this model are shown in Table 4. Alternative models in which verbal and non-verbal working memory differentially predicted English, mathematics and science were also tested, but failed to satisfy statistical criteria for a good fit.

Figure 2. Confirmatory factor analysis model and full structural equation model for Key Stage 2



a) Confirmatory factor analysis model

b) Structural equation model

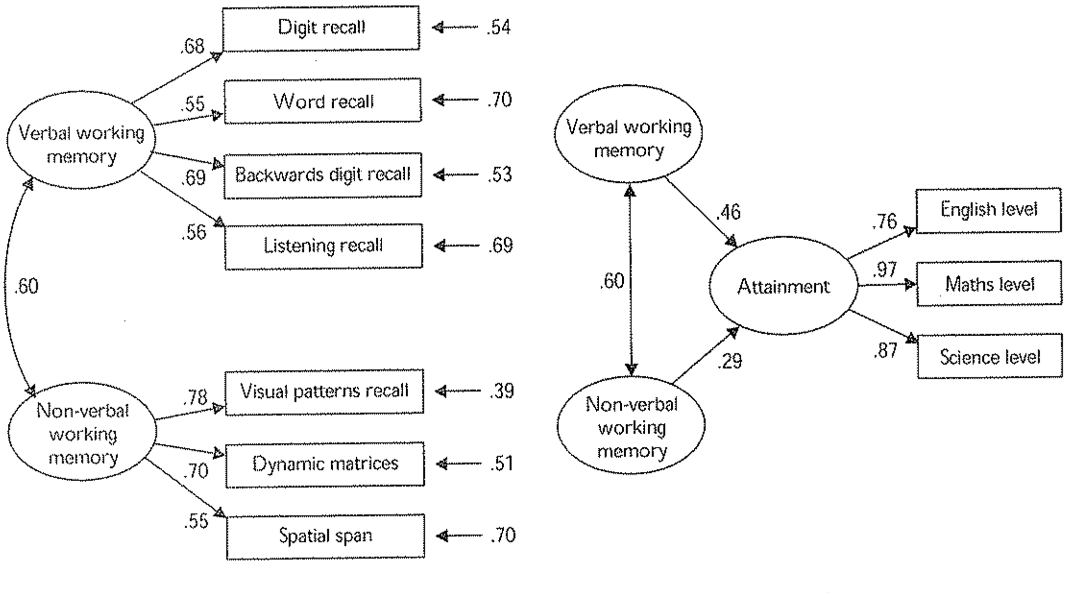
This model provided an excellent fit to the data, with a CFI of .97, an RMSEA value of .05, and a non-significant χ^2 value ($p = .21$). In this model highly significant paths existed between each working memory domain and attainment: for verbal working memory the standardised path coefficient was .40, $p < .05$, and for non-verbal working memory the standardised path coefficient was .43, $p < .05$. A structural equation model diagram of this model is presented in Figure 2b.

Corresponding analyses were then performed on the data from the 14-year-old children. As with the younger children, four confirmatory factor analysis models of the structure of the working memory assessments were tested. Models CFA2b and CFA3b differed in one respect. As a result of the high correlations between the odd-one-out task and the verbal measures at Key Stage 3, the odd-one-out task was not included during modelling. Thus in CFA2b Spatial Span was used as a single indicator of nonverbal central executive capacity and in CFA3b only visual patterns, dynamic matrices and spatial span were used as nonverbal working memory indicators. For a diagrammatic representation of the models assessed for children aged 14 see Figure 1 (CFA1, CF2b, CFA3b, CFA4). The fit statistics for these models are presented in Table 4.

Model 1, with separate phonological loop, visuospatial sketchpad and central executive factors produced fit statistics indicative of a satisfactory fit (both fit indices > .90). However, all the fit statistics for model 3b with one verbal and one non-verbal working memory factor indicated improved indices from model 1. The fit indices (CFI and IFI) were both 1.0 and the RMSEA value was .00. Factor loadings and item error terms for this model are shown in Figure 3. All loadings and variances are significant at the .05 probability level.

The factor loadings from the two-factor confirmatory factor analysis model were then incorporated into a structural equation model (SEM) in which the verbal and non-verbal

Figure 3. Confirmatory factor analysis model and full structural equation model for Key Stage 3



a) Confirmatory factor analysis model

b) Structural equation model

factors predicted National Curriculum scores. In the model, the two working memory factors were both specified as predictors of a single attainment factor that was associated with English, mathematics and science. The fit indices for this model are shown in Table 4. Alternative models in which verbal and non-verbal working memory differentially predicted English, mathematics and science were also tested, but failed to satisfy statistical criteria for a good fit.

This model provided an excellent fit to the data, with a CFI of .97, an RMSEA of .05, and a non-significant χ^2 value ($p = .17$). In this model highly significant paths existed between each working memory domain and attainment: for verbal working memory the standardised path coefficient was .46, $p < .05$, and for nonverbal working memory the standardised path coefficient was .29, $p < .05$. A structural equation model diagram of this model is presented in Figure 3b.

Discussion

This study provides direct evidence for links between working memory and performance on National Curriculum tests at 11 and 14 years of age. Both verbal and non-verbal working memory predicted attainment at both ages.

The results build upon previous evidence of relationships between National Curriculum test scores and working memory at 7 years of age and at 14 years of age (Gathercole & Pickering, 2000; Gathercole *et al.*, in press) and findings of the involvement of the phonological loop, visuospatial sketchpad and central executive components of working memory in domains of skill related to education (e.g. Gathercole & Baddeley, 1993; Seitz & Schumann-Hengsteler, 2000; Reuhkala, 2001). It should, however, be noted that the present findings of close associations between working memory and attainment in English at Key Stage 3, are inconsistent with previous reports (Gathercole *et al.*, in press) and provide little evidence for developmental changes in the contribution of working memory to the acquisition of knowledge and skill in language.

Detailed analysis of interrelations between specific working memory tasks and attainment indicated that complex span tasks that are associated with the central executive in the Baddeley and Hitch (1974) model of working memory were most closely related to attainment in all curricular areas. The central executive was uniquely correlated to attainment at Key Stage 3. Verbal central executive tasks were significantly correlated with attainment levels in English and mathematics. Spatial span, an index of non-verbal central executive processes, was significantly correlated with mathematics and science levels. The two slave systems, however, were not uniquely related to levels in English, mathematics and science. This supports previous evidence that the central executive in particular plays a crucial role in the acquisition of complex cognitive abilities and skills such as literacy, comprehension and arithmetic (e.g. Bull, Johnson & Roy, 1999; Swanson, 1994; Yuill *et al.*, 1989).

It should, however, be noted that the data did not provide strong support for the specific Baddeley and Hitch (1974) model of working memory incorporating a domain-general central executive and subsidiary domain-specific storage systems. As tasks used in the present study all incorporated a storage element, it was not possible to assess the unique contributions of storage and processing to attainment in the present study. The finding of a dissociation between verbal and non-verbal working memory however, are consistent with those of Shah and Miyake (1996), who found evidence for separate pools of resources for verbal and spatial working memory, and further extend these findings to 11- and 14-year-old children.

Detailed analysis also revealed that the predictive relationships between verbal and non-verbal working memory abilities and school achievements reveal a marked degree of domain specificity (e.g. see also Daneman & Carpenter, 1980; Daneman & Tardiff, 1987; Shah & Miyake, 1996). Partial correlations revealed that verbal working memory tasks were uniquely associated with English and mathematics performance, whereas visuospatial tasks shared unique links with mathematics and science. One apparent developmental change in the associations between working memory and attainment concerns science. Verbal complex span tasks were uniquely correlated with science scores at 11 years of age, but not at 14 years of age. In addition, it is notable that non-verbal working memory contributes rather less to attainment at Key Stage 3 than at Key Stage 2.

A further important difference between the two age groups concerns the correlations between the odd-one-out task and the other working memory measures. At aged 11, odd-one-out scores correlated highly with other non-verbal measures but at aged 14 it correlated highly with verbal measures. In considering the demands made by the odd-one-out task it is plausible to suggest that in the older age group, where speed of processing is likely to be more efficient (e.g. Carella & Hale, 1994), participants were able to recode the spatial locations within the task in to a verbal format, such as 'left, middle and right'. In the younger age group, the time constraints imposed during the task may have prevented this recoding, resulting in visuospatial working memory being used to complete the task. Nonetheless, the lack of validity of the odd-one-out task as a non-verbal complex span task has implications for its further use.

In conclusion, this study provides further evidence for a distinction between verbal and spatial working memory resources. It also demonstrates that working memory is a strong predictor of educational attainment, as measured by National Curriculum attainment levels. The impact of working memory capacities on performance on National Curriculum tests is likely to be a result of working memory being employed for storage, processing and integration of information during complex and demanding activities (Just & Carpenter, 1992). Such activities are common in the school classroom, for example writing while formulating the next part of a text, or engaging in mental arithmetic.

The strong links found here between working memory capacities and children's scholastic attainments have important practical implications for educational practice as well as cognitive theory. First, using measures of working memory in addition to more commonly used knowledge- based assessments may provide better estimates of a child's chance of future academic success. Secondly, one reason why children may fail to achieve expected levels in key curricular domains is that their performance on learning tasks in the classroom is constrained by their working memory capacities. There may be significant benefits from creating structured learning activities that reduce opportunity for failure due to inadequate working memory resources. One way of achieving this may be to decompose complex task sequences involving intermediate storage and concurrent processes in to component stages, supported where possible by external memory prompts rather than working memory.

Note

The authors would like to thank the pupils and staff of Gosforth Central Middle School and Gosforth High School in Newcastle upon Tyne for their time and assistance with this study.

References

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

Baddeley, A.D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology, 49A,* 5–28.

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 working memory. *Quarterly Journal of Experimental Psychology, 51A,* 819–852.

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

Baddeley, A.D. & Logie, R.H. (1999). The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* New York: Cambridge University Press.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246.

Bentler, P.M. (2001). EQS structural equation modelling software (version 6.0) [computer software]. USA: mvsoft.

Bull, R., Johnson, R.S. & Roy, J.A. (1999) Exploring the roles of the visuo-spatial sketchpad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental neuropsychology, 15,* 421–442.

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.

Caplan, D. & Waters, G. (1990). Short-term memory and language comprehension: A critical review of the neuropsychological literature. In G. Vallar & T. Shallice (Eds.), *Neuropsychological impairments of short- term memory* (pp.337–389). New York: Cambridge University Press.

Carella, J. & Hale, S. (1994). The rise and fall in information-processing rates over the life span. *Acta Psychologia, 86,* 109–197.

Case, R., Kurland, D.M. & Goldberg, J. (1982). Working memory capacity as long- term activation: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19,* 1101–1114.

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.

Daneman, M. & Tardiff, T. (1987). Working memory and reading skill re-examined. In M. Colthart (Ed.), *Attention and performance XII: The psychology of reading.* Hillsdale, NJ: Erlbaum.

Dark, V.J. & Benbow, C.P. (1990). Enhanced problem translation and short-term memory: Components of mathematical talent. *Journal of Educational Psychology, 82,* 420–429.

Della Sala, S., Gray, C., Baddeley, A.D., Allamano, M. & Wilson, L. (1999). Pattern span: A tool for unwelding visuo-spatial memory. *Neuropsychologia, 37,* 1189–1199.

de Jong, P. (1998). Working memory deficits of reading disabled children. *Journal of Experimental Child Psychology, 70,* 75–96.

De Rammelaere, S., Stuyven, E. & Vandieredonck, A. (2001). Verifying simple arithmetic sums and products: Are the phonological loop and central executive involved? *Memory and Cognition, 29,* 267–273.

Engle, R.W., Cantor, J. & Carullo, J.J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18,* 972–992.

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

Gathercole, S.E. & Adams, A. (1994). Children's phonological working memory: Contributions of long- term memory and rehearsal. *Journal of Memory and Language, 33,* 672–688.

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 and 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. (1993). *Working memory and language.* Hillsdale, NJ: Erlbaum.

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

Gathercole, S.E., Pickering, S.J., Knight, C. & Stegmann, Z. (in press). Working memory skills and educational attainment: Evidence from National Curriculum assessments at 7 and 14 years of age.

Henry, L.A. (2001). How does the severity of a learning disability affect working memory performance? *Memory, 9,* 233–247.

Hitch, G.J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology, 10,* 303–323.

Hu, L. & Bentler, P.M. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modelling, 6,* 1–55.

Hulme, C. & Mackenzie, S (1992). *Working memory and severe learning difficulties.* Hove, UK: Lawrence Erlbaum.

Hulme, C. & Roodenrys, S. (1995). Verbal working memory development and its disorders. *Journal of Child Psychology and Psychiatry, 36,* 373–398.

Jarrold, C. & Baddeley, A. D. (1997). Short-term memory for verbal and visuo-spatial information in Down's syndrome. *Cognitive Neuropsychiatry, 2,*102–122.

Jarrold, C., Baddeley, A.D. & Hewes, A.K. (1999). Dissociating working memory: Evidence from Down's and Williams Syndrome. *Neuropsychologia, 37,* 637–651.

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.

Kline, R.B. (1998). *Principles and practice of structural equation modelling.* New York: Guilford.

Kyllonen, P.C. (1993). Aptitude testing inspired by information processing: A test of the four- sources model. *Journal of General Psychology, 120,* 375–405.

Michas, I.C. & Henry, L.A. (1994). The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology, 12,* 147–163.

Pickering, S.J. & Gathercole, S.E. (2001). *Working Memory Test Battery for Children.* London: Psychological Corporation: UK.

Qualifications and Curriculum Authority (QCA) (2001). *Statutory tests standards and evaluations reports 1999-2001.* Retrieved from http://www.qca.org.uk/ca/tests/standardsmain.asp#evaluation

Reuhkala, M. (2001). Mathematical skills in ninth- graders: Relationship with visuo-spatial abilities and working memory. *Educational Psychology, 21,* 387–399.

Russell, J., Jarrold, C. & Henry, L. (1996). Working memory in children with autism and with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, 37,* 673–686.

Seitz, K. & Schumann-Hengsteler, R. (2000). Mental multiplication and working memory. *European Journal of Cognitive Psychology, 12,* 552–570.

Shah, P. & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125,* 4–27.

Shute, V.J. (1991). Who is likely to acquire programming skills? *Journal of Educational Computing Research, 7,* 1–24.

Siegel, 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.

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. & Ashbaker, M.H. (2000). Working memory, short- term memory, speech rate, word recognition, and reading comprehension in learning disabled readers: Does the executive system have a role? *Intelligence, 28,* 1–30.

Turner, M.L. & Engle, R.W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127–154.

Yuill, N., Oakhill, J. & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology, 80,* 351–361.

Address for correspondence: Helen Jarvis, Department of Psychology, University of Durham, Science Laboratories, South Road, Durham, DH1 3LE, UK.
E-mail: h.l.Jarvis@durham.ac.uk