

# Self-Ordered Pointing as a Test of Working Memory in Typically Developing Children

Lucy Cragg <sup>1</sup>& Kate Nation

*Department of Experimental Psychology, University of Oxford, UK.*

This is the author's post-print of the article. The definitive version was published in *Memory*, Volume 15 Issue 5, 2007. This can be accessed here : <http://www.informaworld.com/smpp/content~db=all~content=a780327620>

## Abstract

The self-ordered pointing test (SOPT; Petrides & Milner, 1982) is a test of non-spatial executive working memory requiring the ability to generate and monitor a sequence of responses. Although used with developmental clinical populations there are few normative data against which to compare atypical performance. Typically developing children (5-11 years) and young adults performed two versions of the SOPT, one using pictures of familiar objects and the other hard-to-verbalise abstract designs. Performance improved with age but the children did not reach adult levels of performance. Participants of all ages found the object condition easier than the abstract condition suggesting that verbal processes are utilised by the SOPT. However, performance on the task was largely independent from verbal and nonverbal cognitive ability. Overall the results suggest that the SOPT is a sensitive measure of executive working memory.

---

<sup>1</sup> Address for correspondence:

Lucy Cragg  
Department of Experimental Psychology,  
University of Oxford,  
South Parks Road, OXFORD. OX1 3UD, UK.  
Tel: +44 (0)1865 271357  
[lucy.cragg@psy.ox.ac.uk](mailto:lucy.cragg@psy.ox.ac.uk)

## Introduction

The self-ordered pointing test (SOPT) was developed by Petrides and Milner (1982) as a test of working memory for patients with frontal lobe lesions. The task takes the form of a set of pictures of familiar objects or abstract designs, arranged in a grid. These are presented in a different spatial arrangement on each trial and the participant is required to point to a different picture every time. The test requires executive abilities in order to organise and carry out a sequence of responses as well as to retain and constantly monitor the responses made.

Given its reputation as an executive task, the SOPT has been used as a test of working memory with childhood clinical populations that demonstrate an executive deficit, such as children with phenylketonuria (Diamond, Briand, Fossella, & Gehlbach, 2004; Smith, Klim, & Hanley, 2000), attention deficit hyperactivity disorder (Geurts, Verté, Oosterlaan, Roeyers, & Sergeant, 2004; Scheres et al., 2004), oppositional-defiant disorder (van Goozen et al., 2004), and autism (Geurts et al., 2004; Joseph, Steele, Meyer, & Tager-Flusberg, 2005). The majority of these studies have administered the SOPT as part of a battery of executive tasks in order to determine which executive components are deficient in the population being studied. However, there are few developmental data available to interpret these results against the level of performance for typically developing children in different age groups.

Normative developmental data are available for a spatial version of the self-ordered pointing task, which relies on the same underlying principles as the SOPT but requires remembering a sequence of locations instead of a sequence of pictures. This is more widely available as a test of executive working memory in children (De-Luca et al., 2003; Hughes, Plumet, & Leboyer, 1999; Luciana & Nelson, 1998, 2002; Rhodes, Coghill, & Matthews, 2004) due to its inclusion in the Cambridge Neuropsychological Testing Automated Battery (CANTAB), a widely used measure of neuropsychological function. Normative data for the CANTAB were provided by Luciana and Nelson (2002) who showed that the executive working memory skills tapped by the spatial SOPT are not fully developed by 12 years of age. While the normative data from the spatial self-ordered task give us some information on the developmental trajectory of executive working memory skills, the non-spatial version

may not follow the same pattern (Conklin, Luciana, Hooper, & Yarger, 2007). Although the tasks may share domain-general processes involved in generating and organising the sequence of responses, there may be domain-specific processes involved concerning the aspect of the stimuli to be remembered (location vs. identity) which may develop at different rates. The non-spatial version of the task is also useful to use alongside the spatial version with clinical populations to determine if there is a general underlying deficit in monitoring and manipulation information in working memory, or a domain-specific problem dependent on the type of stimuli used.

Studies using the non-spatial SOPT in typically developing preschoolers (Hongwanishkul, Happaney, Lee, & Zelazo, 2005) and school-age children (Archibald & Kerns, 1999) have indicated that performance on the task improves with age. However, as these studies included the SOPT as part of a battery of tests, a detailed assessment of developmental changes is not provided.

Most researchers using this task with children have administered the task following Petrides and Milner (1982), presenting the stimuli on paper with three repetitions of the set sizes 6, 8, 10 and 12. However, the error score is often summed or averaged across set sizes and as a result, changes in performance as a function of task difficulty have not been addressed. The present study aimed to examine the effect of set size by including it as a factor in the analyses. Previous studies have also collapsed the results across the three repetitions. Unfortunately however, this may result in practice or interference effects being missed. We specifically examined the effect of task repetitions, labelled 'games', to determine if repeating the task had a beneficial or detrimental effect on performance. The reliability of performance across task repetitions was also investigated.

One aspect of the standard administration of the non-spatial SOPT which may be problematic when used with children is that when the pictures are arranged in a grid, a high score can be obtained simply by repeatedly choosing the same location. This is prevented in adults using a verbal warning if the strategy is adopted, however this may be confusing to young children. To avoid this scenario, we presented the pictures in random locations which changed each time a response was required. This meant that it was not possible to consistently choose the same location.

Another factor that may influence task performance is the availability of verbal encoding and verbal rehearsal strategies. To explore this, our task compared performance in an object condition (where pictures were easy to name) and an abstract condition where pictures were very hard to label. Performance should be better in the object version of the task if children use verbal encoding to help remember the objects.

The level of verbal ability required by the task is an important factor to take into consideration when studying developmental populations who may have concomitant or comorbid language problems in addition to other deficits. Joseph, Steele, Meyer and Tager-Flusberg (2005) used the SOPT to test the hypothesis that children with autism are impaired in using verbal encoding and rehearsal strategies to aid working memory. Their results showed that the typically developing group (aged 5;10-13;10 years) found a condition with line drawings of objects significantly easier than an abstract condition, suggesting that verbal encoding was being used to help remember the objects. Furthermore, it appears that language ability is correlated with performance on the object condition of the SOPT, such that children with better language skills are more successful. Joseph et al. (2005) found that language level, measured by the Expressive Vocabulary Test (Williams, 1997) and the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997) was significantly correlated with performance on the object condition, but not the abstract condition of the SOPT, once age had been controlled for. The same relationship was shown by Hongwanishkul et al. (2005) in preschoolers and we predicted a similar pattern of results in our own experiment.

The role of language abilities in task performance may become more important with age as children become more reliant on verbal strategies such as rehearsal. We predicted that the older children in our experiment would benefit more from using verbal strategies in the object condition of the SOPT than the younger children, and therefore that we would find a greater difference in performance between the object and abstract conditions in the older children. This is supported by evidence that verbal rehearsal strategies are not used to aid memory for pictorial stimuli until after the age of 8 years (Halliday, Hitch, Lennon, & Pettipher, 1990; Hitch & Halliday, 1983).

In summary, the present study aimed to provide a more detailed analysis of SOPT performance in typically developing children. Prior to testing the children we tested a sample of

adults to ensure that our modified task produced the expected pattern of results. This sample also acted as a comparison group to help determine the age at which children reach adult levels of performance. As well as examining changes over development, the effects of set size and task repetition manipulations were specifically examined. We predicted that performance would be better in the object condition than the abstract condition, due to the use of verbal labelling. Furthermore, we hypothesised that if the use of verbal strategies increases with age then the difference in performance between object and abstract conditions would also increase. On the basis of previous research (Hongwanishkul et al., 2005; Joseph et al., 2005) we predicted that language ability would correlate with performance on the object, but not abstract condition of the task.

## Method

### *Participants*

Ninety children and 15 young adults participated in this study. Data were collected from 15 children in each of the following British school year groups: Year 1 (5-6yrs); Year 2 (6-7yrs); Year 3 (7-8yrs); Year 4 (8-9yrs); Year 5 (9-10yrs) and Year 6 (10-11yrs). All of the children attended state primary schools and were selected at random by class teachers. Informed parental consent was received for all children that participated. Bilingual children and those with a statement of Special Educational Needs were excluded from the study. The young adults who participated were all students at Oxford University, some of whom received course credits for taking part. Background information for all participants is presented in Table 1.

The Matrices and Vocabulary subtests of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) were completed by all children. As shown in Table 1, performance was close to average for all groups of children suggesting that the sample was representative of typically developing children. The adults completed only the Matrices subtest and as a group, they achieved scores in the high-average range.

### *Apparatus*

The experimental task (available from [www.psy.ox.ac.uk/lcd](http://www.psy.ox.ac.uk/lcd)) was created and controlled using E-Prime software and run on a Dell laptop computer. The participants responded using an ELO touchscreen with a screen size of 304mm by

**Table 1** Participant characteristics

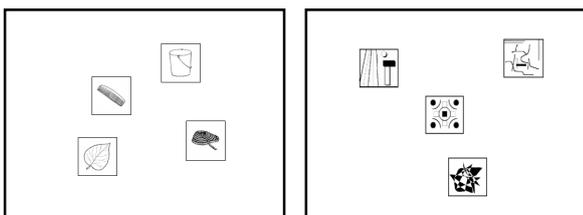
Age Group	Gender Male:Female	Age (yrs)		WASI Matrices <sup>1</sup>		WASI Vocabulary <sup>1</sup>	
		Mean	SD	Mean	SD	Mean	SD
Year 1	7:8	6.01	.33	53.7	10.7	50.0	8.04
Year 2	7:8	6.81	.25	56.3	11.0	50.6	7.99
Year 3	8:7	8.13	.28	52.1	12.5	52.5	13.3
Year 4	9:6	9.15	.24	52.2	9.10	52.1	10.6
Year 5	7:8	10.1	.32	49.9	7.68	59.4	8.68
Year 6	7:8	11.0	.24	48.2	8.14	55.2	11.7
Adults	5:10	19.4	1.35	58.6	5.82	---	---

Notes. <sup>1</sup>T scores: M=50, SD=10.

228mm. The touchscreen was placed approximately 270mm from the edge of the table with an ergonomic mouse mat centred in front of it which acted as a hand-rest. The experiment was carried out in a quiet area in the school or university. The participants sat within comfortable reaching distance of the touchscreen and were asked to begin by placing their dominant hand on the hand-rest.

### Materials and Procedure

Participants were shown a set of pictures and were required to touch a different picture on each trial, until all of the pictures had been touched once. There were two versions of this task, one using line drawings of objects, and one using black and white abstract patterns. The line drawings were pictures of objects taken from the online database of the International Picture-Naming Project, Center for Research in Language, University of California, San Diego (Szekely et al., 2004). The objects were high frequency words with an early age of acquisition. The abstract pictures were kindly donated by Dr Louise Phillips at Aberdeen University following her use of the SOPT with older adults (Philips et al., 2002). They were chosen as they were hard to verbalise. Examples of both sets of stimuli are shown in Figure 1.



**Figure 1** Examples of the object and abstract stimuli used

Each picture measured 43 by 43 mm and was presented on a blue background. Set sizes of 4, 6, 8 and 10 pictures were used with a unique set of pictures for each set-size. The task was repeated

three times at each set-size level to create three ‘games’, which differed only in the location of the pictures on the screen. To distinguish between these, each game began with a brightly coloured screen, displayed for 2000ms to tell the participant whether it was Game 1, 2 or 3 and another screen displayed ‘game over’ for 1000ms when they had touched the required number of pictures.

The children completed all conditions of the task in a fixed order. The participants were first shown a demonstration using four pictures of objects. They were then asked to perform the task themselves using first 4, then 6, 8 and 10 pictures of objects. This was then repeated for the abstract pictures but without the demonstration. Set size 4 was used as a practice and was therefore excluded from data analysis. No feedback was given to the participant at any stage of the task except to remind participants that they should not touch a picture which they had already touched. There were no time restrictions yet all children completed the task in approximately ten minutes.

### Results

Performance on this task was assessed in two different ways. First, the number of errors was calculated, defined as touching a picture already selected. Second, following Joseph et al. (2005) span was also measured, defined as the number of consecutive novel responses prior to the first error. A one-way ANOVA showed that there was no effect of gender on the total error score ( $F < 1$ ) for either children or adults. Therefore, gender was not included as a factor in further analyses. Due to unequal variance between groups, the Greenhouse-Geisser correction was used in all ANOVAs and the Games-Howell test was used for post-hoc comparisons.

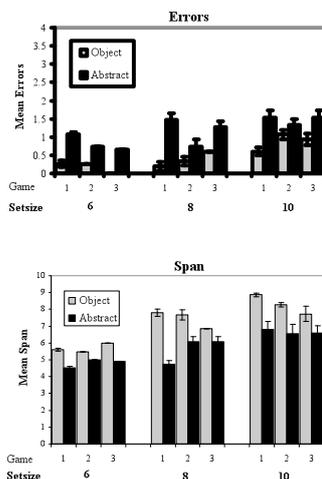
### Reliability

To examine the consistency between the object and abstract conditions of the task,

correlation and regression analyses were performed. The error rates for the two conditions correlated reasonably highly,  $R^2=.74$ ,  $p<.001$ . Once age was controlled, the number of errors on the abstract condition accounted for 24% additional variance in the object condition ( $F(1,102)=62.0$ ,  $p<.001$ ) and the number of errors in the object condition accounted for a similar amount of unique variance in the abstract condition ( $R^2=.26$ ,  $F(1,102)=62.0$ ,  $p<.001$ ). This suggests that the two conditions are tapping into some of the same processes. To assess internal reliability, Cronbach's alpha was computed for the error scores in Games 1, 2 and 3, collapsed across conditions. The reliability was acceptable with a value of  $\alpha=.88$ . This indicates that performance in the three games was consistent across individuals.

#### Adult data

To assess SOPT performance in adults a three-way repeated-measures ANOVA was run with either errors or span as the dependent variable. The within-subject factors were condition (object, abstract), set size (6, 8, 10) and game (1, 2, 3). The results are presented in Figure 2. Unsurprisingly, a significant main effect of set size, (errors:  $F(1.90, 26.6)=30.2$ ,  $p<.001$ ,  $\eta^2=.683$ ; span:  $F(1.90,26.6)=50.9$ ,  $p<.001$ ,  $\eta^2=.784$ ) showed that even though span length increased with set size, more errors were made as more pictures were added. Fewer errors and longer spans were observed in the object condition (errors:  $M=.47$ ,  $SD=.26$ ; span:  $M=7.14$ ,  $SD=.50$ ) than in the abstract condition (errors:  $M=1.15$ ,  $SD=.46$ ; span:  $M=5.70$ ,  $SD=1.02$ ), as shown by a significant main effect of condition (errors:  $F(1,14)=39.6$ ,  $p<.001$ ,  $\eta^2=.739$ ; span:  $F(1,14)=26.7$ ,  $p<.001$ ,  $\eta^2=.656$ ).



**Figure 2.** Mean (S.E.) errors and span for adult participants

The main effect of game was not significant in the adults; however there was a condition by game interaction (errors:  $F(1.64, 22.9)=5.29$ ,  $p<.05$ ,  $\eta^2=.274$ ; span:  $F(1.73,24.3)=4.11$ ,  $p<.05$ ,  $\eta^2=.227$ ). This was due to more errors and shorter spans in Game 1 than in Games 2 and 3 for the abstract condition, but no difference between games for the object condition. There was also a set size by game interaction for the errors ( $F(3.30, 46.2)=3.34$ ,  $p<.05$ ,  $\eta^2=.193$ ), however this was not significant when span was used as the dependent variable ( $F(2.99, 41.9)=1.64$ , ns).

In summary, these results show that as expected, adults made more errors as set size increased and found the object condition easier than the abstract condition. Performance across the three games differed according to the condition, with no difference between games when the pictures were of objects, but a higher error score and shorter span in the first game than in the following two for abstract pictures. These results replicate those of the original study by Petrides and Milner (1982) and as such provide a good background against which to consider the data from the children.

#### Child Data

A four-way mixed-measures ANOVA was used to analyse the children's data with the within-subject factors condition (object, abstract), set size (6, 8, 10) and game (1, 2, 3), and year group (Year 1, Year 2, Year 3, Year 4, Year 5, Year 6) as the between-subjects factor. Comparing the error rates of the six groups of children showed that performance on the SOPT improved with age resulting in a significant main effect of year group, (errors:  $F(5,84)=5.64$ ,  $p<.001$ ,  $\eta^2=.251$ ; span:  $F(5,84)=5.18$ ,  $p<.001$ ,  $\eta^2=.236$ ). Post-hoc tests revealed that Year 2 children made significantly more errors than the children in Years 4 to 6 and that Year 1 and Year 2 children had significantly shorter spans than children in Years 4 to 6.

This demonstrated that the younger children differed from the older children but that there were no differences within these groups. As a result of this the children were split into two groups: younger children (Years 1 to 3) and older children (Years 4 to 6) and the results reanalysed with age group (younger children, older children) replacing year group as the between-subjects factor. Reducing the data in this way allowed for greater clarity when interpreting the results.

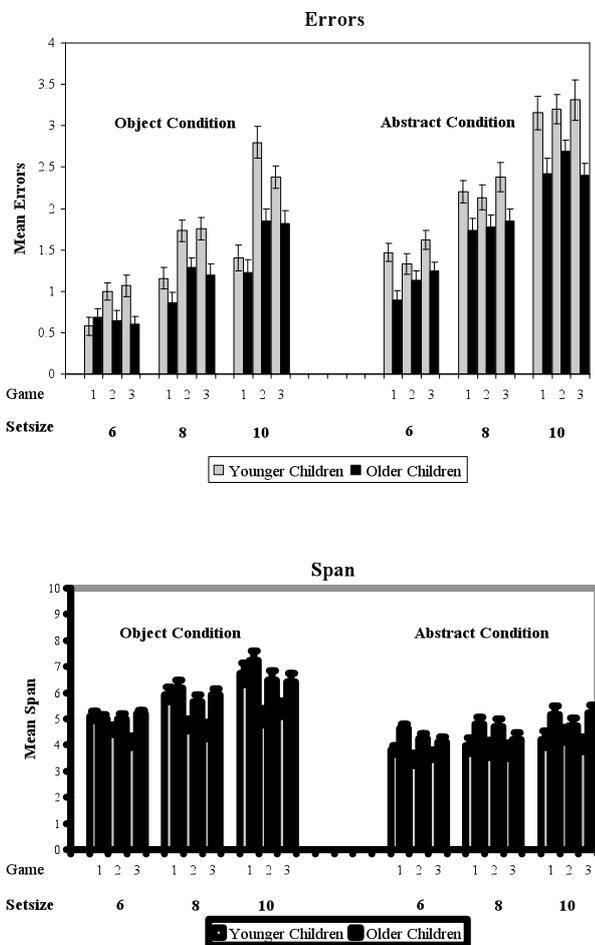
A four-way mixed measures ANOVA was performed with age group as the between-subjects factor and condition (object, abstract), set size (6,

8, 10) and game (1, 2, 3) as the within-subject factors. The results are illustrated in Figure 3. The older children ( $M=1.46$ ,  $SD=.43$ ) made significantly fewer errors than the younger children ( $M=1.93$ ,  $SD=.50$ ) and had significantly longer spans ( $5.27(.78)$  vs.  $4.53(.82)$ ) as shown by a significant main effect of age group (errors:  $F(1,88)=21.9$ ,  $p<.001$ ,  $\eta^2=.199$ ; span:  $F(1,88)=19.5$ ,  $p<.001$ ,  $\eta^2=.182$ ). As in adults, both span and the number of errors increased as the set got larger, shown by a main effect of set size (errors:  $F(1.92, 169)=294$ ,  $p<.001$ ,  $\eta^2=.770$ ; span:  $F(1.93,170)=44.3$ ,  $p<.001$ ,  $\eta^2=.335$ ). A significant set size x age group interaction was found for errors ( $F(1.92,169)=4.36$ ,  $p<.05$ ,  $\eta^2=.047$ ) due to the fact that the difference in error rate between the younger and older children increased as the set size got larger. However, this was not significant for span ( $F(1.93,170)=1.17$ , ns).

main effect of condition (errors:  $F(1,88)=212$ ,  $p<.001$ ,  $\eta^2=.707$ ; span:  $F(1,88)=168$ ,  $p<.001$ ,  $\eta^2=.657$ ), with fewer errors and longer spans in the object condition (errors:  $M=1.34$ ,  $SD=.52$ ; span:  $M=5.52$ ,  $SD=1.02$ ) than the abstract condition (errors:  $M=2.05$ ,  $SD=.62$ ; span:  $M=4.28$ ,  $SD=.96$ ). The interaction between age and condition was not significant, showing that performance improved with age to the same extent in both conditions. A significant interaction occurred between condition and set size (errors:  $F(1.85, 163)=7.52$ ,  $p<.001$ ,  $\eta^2=.079$ ; span:  $F(1.78,156)=5.81$ ,  $p<.01$ ,  $\eta^2=.62$ ) as there was a greater difference between conditions as the set size increased.

There was also a significant main effect of game, (errors:  $F(2.00, 176)=23.7$ ,  $p<.001$ ,  $\eta^2=.213$ ; span:  $F(1.99,175)=15.1$ ,  $p<.001$ ,  $\eta^2=.147$ ) which a test of simple effects showed was due to significantly fewer errors being made in Game 1 ( $M=1.48$ ,  $SD=.56$ ) than Games 2 ( $M=1.80$ ,  $SD=.58$ ) and 3 ( $M=1.80$ ,  $SD=.66$ ), which did not differ in error score. An identical pattern was found for span length with longer spans in Game 1 ( $M=5.24$ ,  $SD=1.04$ ) than in Games 2 ( $M=4.75$ ,  $SD=1.05$ ) and 3 ( $M=4.71$ ,  $SD=1.08$ ). There was an interaction between set size and game for errors ( $F(3.60, 317)=3.58$ ,  $p<.01$ ,  $\eta^2=.039$ ) but not span length ( $F(3.44, 303)<1$ ), due to a greater difference in error rate between Game 1 and the other two games as the set size increased.

Performance across the three games varied depending on the condition, as shown by an interaction between condition and game, (errors:  $F(1.81, 160)=15.2$ ,  $p<.001$ ,  $\eta^2=.147$ ; span:  $F(1.95, 172)=5.94$ ,  $p<.01$ ,  $\eta^2=.063$ ). In the object condition participants showed longer spans and made fewer errors in Game 1 than in Games 2 or 3 for pictures of objects, whereas in the abstract condition performance was similar across the three games. A significant age by condition by game interaction, (errors:  $F(1.81,160)=7.31$ ,  $p<.001$ ,  $\eta^2=.077$ ; span:  $F(1.95, 172)=5.24$ ,  $p<.01$ ,  $\eta^2=.056$ ) showed that this effect was larger in the younger children. This pattern of results was further illustrated by a significant condition by game by set size interaction for error rate,  $F(3.79, 334)=3.02$ ,  $p<.05$ ,  $\eta^2=.033$ . In the object condition fewer errors were made in Game 1 than Games 2 and 3, and this difference grew larger as set size increased, whereas for abstract pictures performance across games was similar for all three set sizes.



**Figure 3.** Mean (S.E.) errors and span for child participants

The children performed differently depending on the stimuli used, as shown by a significant

In summary, the children's performance was similar to the adults in that they made more errors and showed shorter spans in the abstract condition than in the object condition. They also made more errors as set size increased. However, a series of one-way ANOVAS comparing younger children, older children and adults showed that the children had not reached adults levels of performance on either the object condition (errors:  $F(2,102)=31.7$ ,  $p<.001$ ; span:  $F(2,102)=28.6$ ,  $p<.001$ ) or the abstract condition (errors:  $F(2,102)=27.2$ ,  $p<.001$ ; span:  $F(2,102)=23.1$ ,  $p<.001$ ). Post-hoc analyses showed significant differences in both errors and span between all three age groups (younger children, older children, adults) in both conditions.

The children and adults differed in terms of performance across the three games as shown by a significant age by condition by game interaction,  $F(3.59, 183.3)=3.87$ ,  $p<.01$ ,  $\eta^2=.071$ . Whereas the adults made more errors in Game 1 in the abstract condition but showed no difference between games in the object condition, the children showed no difference between games in the abstract condition but made fewer errors in Game 1 of the object condition. This effect was larger for the younger children.

#### *The Effects of Verbal and Nonverbal Ability in the SOPT*

A significant main effect of condition suggests that verbal factors are playing a role in SOPT performance, with verbal labelling improving performance in the object condition. To assess this further, individual differences in the children's verbal and nonverbal abilities were compared to performance on the SOPT. If verbal skills are important in the task then we would expect that children with a high verbal ability will perform well in the object condition of the SOPT. Correlations between the different measures were calculated and are presented in Table 2.

The error and span measures correlated highly, suggesting that both were measuring the ability to maintain and monitor items in working memory. Vocabulary and Matrices, measuring

verbal and nonverbal abilities respectively, correlated significantly with both error and span measures for both conditions. Once chronological age was controlled for however, the only relationship which remained significant was between Matrices and the object condition of the SOPT.

Two series of hierarchical multiple regressions were performed for each condition separately. Both the total error score and the total span score for each condition, summed across all games and set sizes, were used to enable comparisons with previous studies. The predictor variable chronological age (CA) was always entered in the first block, followed by raw scores on either WASI Vocabulary or Matrices in the second block. These showed that age was a significant predictor of SOPT performance, accounting for 13.4% of variance in error score in both the object condition ( $F(1,88)=13.6$ ,  $p<.001$ ) and the abstract condition ( $F(1,88)=13.6$ ,  $p<.001$ ). Age accounted for a similar amount of variance in span length (object condition:  $R^2=.15$ ,  $F(1,88)=15.6$ ,  $p<.001$ ; abstract condition:  $R^2=.14$ ,  $F(1,88)=14.4$ ,  $p<.001$ ). Once age had been accounted for, Vocabulary did not account for significant extra variance in either condition for either errors or span. Nonverbal abilities, as indexed by Matrices, accounted for additional variance in the object condition over and above age-related improvements (errors:  $R^2=.071$ ,  $F(1,87)=7.74$ ,  $p<.01$ ; span:  $R^2=.086$ ,  $F(1,87)=9.75$ ,  $p<.01$ ). However, Matrices accounted for no additional variance in the abstract condition once the contribution of age had been accounted for. Overall, these analyses show that the contribution of age is the same in both the object and abstract condition. Interestingly, nonverbal skills seemed to play a greater role in the object condition than the abstract condition while verbal abilities did not account for variance in either condition. Thus neither verbal nor nonverbal abilities are good predictors of performance on this working memory task.

**Table 2.** Correlations for the Different Measures: Above the Diagonal are Pearson Correlation Coefficients, Below the Diagonal are Partial Correlations Controlling for Chronological Age

	Errors-Object	Errors-Abstract	Span-Object	Span-Abstract	WASI Vocabulary	WASI Matrices
Errors-Object		.677**	-.868**	-.623**	-.368**	-.427**
Errors-Abstract	.627**		-.640**	-.845**	-.378**	-.340**
Span-Object	-.846**	-.581**		.581**	.346**	.461**
Span-Abstract	-.563**	-.820**	.510**		.363**	.286*
Vocabulary	-.145	-.162	.08	.127		.562**
Matrices	-.286*	-.172	.318*	.096	.239°	

Notes. \*\* $p<.001$ , \* $p<.01$ , ° $p<.05$

## Discussion

This study provides a detailed investigation of developmental changes in performance on the SOPT and how these are affected by various task manipulations, as well as the role of verbal skills in task performance. The results demonstrate that the SOPT is an appropriate task to use with a wide age-range of children. Randomising the spatial location of the pictures reduced the complexity of the instructions without changing the demands of the task and the choice of set sizes was of an appropriate difficulty level for the age range tested. The reliability of the SOPT was acceptable, with a Cronbach's alpha value above the recommended value of .8 (Coolican, 2004).

Performance on the SOPT improved with development with the older children making significantly fewer errors and creating longer spans than the younger children. The greatest difference between the age groups was found at larger set sizes when the task demands were higher. This was the first time children and adults have been directly compared on the non-spatial SOPT. While the findings should be treated with caution as the adult sample were students rather than from the general population, the results were consistent with Luciana and Nelson's (2002) comparison of 12-year-olds and adults on a spatial version of the self-ordered pointing task in showing that the older children did not reach adult levels of performance.

Despite developmental improvement, age only accounted for a relatively small portion of variance in our study, similar to the findings of Archibald and Kerns (1999). A large amount of unexplained variance remained which was not accounted for by either verbal or nonverbal abilities. This suggests that there are other within-group differences, namely in working memory ability, that are more important factors in predicting successful task performance. Comparisons with other complex working memory tasks should be carried out to confirm this.

An interesting difference in performance between the children and the adults was shown by comparing the task repetitions at each set-size. Repeating the task had a detrimental effect on children's performance in the object condition. While the adults showed no difference in performance across games, the children made fewer errors and longer spans in Game 1 compared to the following two games, which did not differ. This could be because the children lost

interest in the task as the games progressed, although if this were the case it would be expected that performance would be worse in the third game than the second. Good internal reliability between the games also suggests that the children were not losing interest in the task.

An alternative explanation is that the memory trace from the first game interfered in performance in the subsequent games. This only occurred in the object condition, possibly because there is more information about the stimuli to help encode them, and therefore a stronger memory trace. The difference between games also interacted with set size, implying that the interference had a larger effect when other task demands were high. The younger children showed the largest difference between games, suggesting that there was more interference from previous items in this age group than in the older children. In contrast, the adults showed no interference effects suggesting that the ability to inhibit the interference of memory traces from previous games develops with age.

An opposite pattern of results was found in the abstract condition. While there was no difference between games for the children, repeating the task led to improved performance in Games 2 and 3 for adults. This could be due to the fact that the adults were aware that abstract pictures are less memorable than objects and therefore consciously attempted to encode them. This would lead to a benefit in performance in the second game once the pictures have been encoded. If this is the case, it suggests an important development in meta-memory or strategy use.

As predicted, performance was better in the object condition than the abstract condition as the availability of verbal labels for objects made them easier to remember. Contrary to our prediction, it was surprising that the relationship found in previous studies (Hongwanishkul et al., 2005; Joseph et al., 2005) between verbal ability and performance on the object condition of the SOPT was not demonstrated in this study. This result is puzzling, especially since the difference in performance between object and abstract conditions suggests that verbal labelling is being used. The reason why no relationship was found is unclear, however we offer two possible explanations. The first relates to the vocabulary tests used to measure verbal ability in the different studies. Both Hongwanishkul et al. (2005) and Joseph et al. (2005) used the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997), a measure of receptive vocabulary which uses

pictorial stimuli and does not require a verbal response. The expressive vocabulary test used by Joseph et al. required the generation of verbal labels, whereas the expressive vocabulary measure used in this experiment entailed giving a definition. Arguably, the measures used by Hongwanishkul et al. and Joseph et al. are more similar to the skills involved in the SOPT, nevertheless, it is surprising that we found no relationship between verbal ability and SOPT performance. If it is indeed the case that general language ability is important in this task then we should expect to see a relationship between the two using any reliable measure of language ability.

An alternative explanation for the lack of association between verbal ability and the object condition is that we used pictures of objects which have an early age of acquisition. Therefore, all of the children should have found it very easy to name the stimuli. If more demanding vocabulary had been used in the object condition, we may have seen a relationship between SOPT performance in the object condition and vocabulary, as assessed by the WASI. Vocabulary difficulty has been shown to influence the recall of words from short-term memory in both children (Nation, Adams, Bowyer-Crane, & Snowling, 1999) and adults (Walker & Hulme, 1999).

Based on evidence that children under the age of 8-years-old do not use verbal rehearsal strategies to aid memory performance (Hitch & Halliday, 1983), we predicted that there would be a greater difference between the two conditions in the older children than the younger children. Our results did not support this conclusion, with a similar difference between the two conditions in all children. One possible explanation for these findings is that none of the children were using verbal rehearsal strategies because the task did not rely on a verbal output (cf. Hitch & Halliday, 1983). This is supported by Joseph et al. (2005) who found that a measure of verbal span was not associated with performance on the object condition of the SOPT. The two tasks differed in that the stimuli in the span task were presented auditorily, and as such were in a suitable format for verbal rehearsal. In contrast, in the SOPT the children needed to spontaneously adopt the strategy of recoding the picture stimuli into verbal form before they could verbally rehearse. Alternatively, it may be that the executive demands of the task prevented verbal rehearsal. Hitch and Halliday (1983) suggested that there may be a lack of rehearsal in young children

because naming pictures taxes the central executive of working memory, leaving insufficient capacity for rehearsal. However, it may be that it is not naming the pictures, but having to continuously generate and monitor responses that taxes the central executive and prevents a verbal rehearsal strategy from being adopted.

In conclusion, the results of this study show that the SOPT is an appropriate measure of working memory for use with children across a wide age-range. Performance on this task showed improvement between the ages of 5 and 11 years, but had not reached adult levels by the end of this period. With increasing age, the children were able to remember a greater number of pictures and the ability to resist interference from previous memory traces also improved. Despite a difference between the two conditions suggesting use of verbal encoding, our predictions concerning the increasing use of verbal rehearsal strategies with age were not supported. Further study into the role of verbal factors in the SOPT in both adults and children, using techniques such as concurrent articulation and explicit questioning about strategy use may reveal if, and to what extent, verbal strategies are used in this task at different stages of development.

### Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council. We are grateful to the staff and pupils of the participating schools for their help and cooperation and to Dr Louise Philips from the University of Aberdeen for providing the stimuli used in the abstract condition. We would like to thank Jon Brock for helpful comments on an earlier draft.

### References

- Archibald, S. J., & Kerns, K. A. (1999). Identification and description of new tests of executive functioning in children. *Child Neuropsychology*, 5(2), 115-129.
- Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working Memory Performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology*, 31(1), 103-128.
- Coolican, H. (2004). *Research methods and statistics in psychology* (4th ed.). London: Hodder and Stoughton.
- De-Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J. A., Proffitt, T. M., Mahony, K., et al. (2003).

- Normative data from the Cantab. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology*, 25(2), 242-254.
- Diamond, A., Briand, L., Fossella, J., & Gehlbach, L. (2004). Genetic and neurochemical modulation of prefrontal cognitive functions in children. *American Journal of Psychiatry*, 161(1), 125-132.
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.
- Geurts, H. M., Verté, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism? *Journal Of Child Psychology And Psychiatry*, 45(4), 836-854.
- Halliday, M. S., Hitch, G. J., Lennon, B., & Pettipher, C. (1990). Verbal short-term memory in children: The role of the articulatory loop. *European Journal of Cognitive Psychology*, 2, 23-38.
- Hitch, G. J., & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society, Series B*, 302, 324-340.
- Hongwanishkul, D., Happaney, K. R., Lee, W. S. C., & Zelazo, P. D. (2005). Assessment of hot and cool executive function in young children: Age-related changes and individual differences. *Developmental Neuropsychology*, 28(2), 617-644.
- Hughes, C., Plumet, M. H., & Leboyer, M. (1999). Towards a cognitive phenotype for autism: Increased prevalence of executive dysfunction and superior spatial span amongst siblings of children with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40(5), 705-718.
- Joseph, R. M., Steele, S. D., Meyer, E., & Tager-Flusberg, H. (2005). Self-ordered pointing in children with autism; failure to use verbal mediation in the service of working memory? *Neuropsychologia*, 43, 1400-1411.
- Luciana, M., & Nelson, C. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia*, 36(3), 273-293.
- Luciana, M., & Nelson, C. A. (2002). Assessment of neuropsychological function through use of the cambridge neuropsychological testing automated battery: Performance in 4- to 12-year-old children. *Developmental Neuropsychology*, 22(3), 595-624.
- Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology*, 73(2), 139-158.
- Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, 20(3), 249-262.
- Philips, L. H., MacPherson, S. E. S., & Della Sala, S. (2002). Age, cognition and emotion: The role of anatomical segregation in the frontal lobes. In J. Grafman (Ed.), *Handbook of Neuropsychology* (2nd ed., Vol. 7, pp. 73-97). B.V.: Elsevier Science.
- Rhodes, S. M., Coghill, D. R., & Matthews, K. (2004). Methylphenidate restores visual memory, but not working memory function in attention deficit-hyperkinetic disorder. *Psychopharmacology*, 175, 319-330.
- Scheres, A., Oosterlaan, J., Geurts, H., Morein-Zamir, S., Meiran, N., Schut, H., et al. (2004). Executive functioning in boys with ADHD: Primarily an inhibition deficit? *Archives of Clinical Neuropsychology*, 19(4), 569.
- Smith, M. L., Klim, P., & Hanley, W. B. (2000). Executive function in school-aged children with phenylketonuria. *Journal of Developmental and Physical Disabilities*, 12(4), 317-332.
- Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, 51(2), 247.
- van Goozen, S. H. M., Cohen-Kettenis, P. T., Snoek, H., Matthys, W., Swaab-Barneveld, H., & van Engeland, H. (2004). Executive functioning in children: A comparison of hospitalised ODD and ODD/ADHD children and normal controls. *Journal of Child Psychology and Psychiatry*, 45(2), 284-292.
- Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(5), 1256-1271.
- Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. London: The Psychological Corporation.
- Williams, K. T. (1997). *Expressive vocabulary test*. Circle Pines, MN: American Guidance Service.