Encoding strategy and not visual working memory capacity correlates with intelligence

Rhodri Cusack, Manja Lehmann, Michele Veldsman & Daniel J. Mitchell

MRC Cognition and Brain Sciences Unit.

Running head:

VWM and intelligence

Corresponding author:

Rhodri Cusack
MRC CBU
15 Chaucer Road
Cambridge
CB2 7EF

rhodri.cusack@mrc-cbu.cam.ac.uk

Tel: +44 1223 273689

Fax: +44 1223 359062

Abstract

There is conflicting evidence on whether or not the capacity of visual working memory (VWM) reflects a central capacity limit that also influences intelligence. We propose that encoding strategy, and more specifically attentional selection, underlie the correlation of some VWM tasks and IQ, and not variations in VWM itself. Change detection measures of VWM were found to be contaminated by some cognitive process that depressed performance at higher set-sizes, so that fewer items were remembered when eight were presented than when just four were presented. Measuring VWM instead using whole report gave a less variable estimate that was higher, particularly for larger set-sizes. Non-verbal IQ did not correlate with this estimate of VWM capacity, but instead with the additional factor that contaminates change detection estimates, which we propose reflects a lack of selection during encoding. A second experiment investigated the role of rehearsal using articulatory suppression and showed this could not account for the key differences between the procedures.

There is intense debate about the relationship between visual working memory (VWM) and intelligence, with conflicting evidence on whether or not VWM capacity reflects a central limit that is a bottleneck in higher-order cognition. A procedure commonly used to measure VWM is change detection (CD, . In a typical experiment, a memorandum comprising a number of coloured squares is presented briefly, to minimise chunking or the piecemeal transfer of items into phonological working memory. This is followed by a delay period sufficiently long (around 1s) to ensure the decay of sensory representations responsible for iconic memory, and then a probe display. The participant must determine whether the items in the probe had the same colour and position as those in the sample. When measured in this way, VWM has been found to predict fluid intelligence and educational achievement in adults and children and to predict top-down control assessed using electroencephalography (EEG) , suggesting it taps into a central capacity limit.

In contrast to this is evidence from other studies that the capacity of VWM and measures of top-down control vary independently across individuals. Like the classic studies by Sperling , Finke et al and Peers et al both used a whole report (WR) paradigm to measure VWM. The memoranda were a set of letters presented briefly. Participants were asked to orally report all of the letters they could remember. This measure of VWM was compared to a measure of the efficiency of top-down control, derived from the ability to report target letters from amongst distractors of a different colour (see the "Theory of Visual Attention", . It was concluded that VWM was independent of the efficiency of top-down control, both across a group of healthy controls and a group of patients with parietal and frontal damage .

What might be the cause of these discrepant results? We focus on two possible explanations. The first is the different memoranda used in the VWM tasks: colour & position versus letters . A second is the different experimental procedures used: CD versus WR .

Experiment 1

Method

The influence of procedure and memorandum type on estimates of VWM, and their relationship to intelligence and top-down control were investigated. 48 participants were tested (16 male, aged 19-63, mean 47). Approval was given by Cambridge Psychological Research Ethics Committee.

VWM was measured with a single WR and two CD procedures that were closely matched. All tasks required memory for letter identity, and one additionally required memory for spatial location. Identical memoranda displays were presented, with set-sizes of 2,3,4,6&8 letters, followed by a maintenance period (Figure 1). In the WR task for letters (VWM-WR-L), after the delay period a box appeared and participants were asked to type all the letters they remembered using a standard keyboard and then press enter. In one of the CD tasks (VWM-CD-L), a probe display was presented that comprised a single letter at the location of one of the letters in the memorandum display (chosen at random), and participants were asked to press one of two buttons to indicate whether the probe letter had been present anywhere in the memorandum, irrespective of location. There was a 50% chance that it had been present. A response could be made at any point after the beginning of the probe display and the next trial did not commence until after it had been made. In the second CD task (VWM-CD-LS), again a single probe letter appeared at the location of one of the letters in the memorandum, and participants had to indicate whether or not it had been present at the same location. There was a 50% chance that the letter was the same at the same location, a 25% chance the letter was present at a different location, and a 25% chance of it being a novel letter. The button mapping was counterbalanced across subjects.

The number of items remembered at each set-size was estimated for each procedure, correcting for guessing. To calculate the number of letters remembered in the CD tasks, a double-high-threshold model was used, which assumes that K_{CD} items were remembered perfectly and the remainder were guessed ($K_{CD}=N(H-F)$ where N=number of items in set, H=hits, F=false alarms,). To calculate the number of letters remembered in the WR trials, the number of letters correctly reported was counted and from this subtracted an estimate of the guessing rate. This used the conservative assumption that the participants had learnt the subset of 15 letters that could appear in the stimuli (hence possibly overestimating the guessing rate and slightly underestimating capacity): $K_{WR}=C-E*N/(15-N)$ where E=number of letters reported that were not in the display, C=correctly reported letters.

To measure top-down control, we used a partial report (PR) procedure in which the stimuli were identical to the WR with set-size six, except that three of the letters were white and three black. After a delay period that matched the VWM tasks, a box appeared and participants were asked to type all of the letters in a target colour. For half of the participants the target colour was fixed as white and for the other half black. As a summary measure of top-down control we used alpha-prime , which gives an estimate of "distractibility" (i.e., higher scores correspond to worse top-down control). It contrasts performance for WR with 3 and 6 items to PR in which 3 target items had to be selected from 3 distractors. If selection were perfect, PR would yield the same performance as 3 item WR; if there were no selection at all, performance per item would be the same as in the 6 item WR.

The order of the four blocks VWM-WR-L, VWM-CD-L, VWM-CD-LS and PR was counterbalanced across participants. Across all of these blocks, the sample displays were 183 ms, the screen background was mid-grey, letters were randomly chosen from the set ABDEFGHJKMNQRTY, and their colour (black or white) in the VWM blocks was matched to the target colour in the partial report. Letters were presented in upper case Arial with height ~3.8° visual angle at a location randomised within a rectangle ~28° by 37° visual angle centred at the middle of the screen, subject to the constraint that the centres of two letters could not be closer than ~6.7°. As report takes longer than making a same/different response, there were fewer trials per block for WR and PR (10 practice+120 main block, total duration including instruction period around 18'40'') than for CD (20+240, 19'30'').

Non-verbal intelligence was estimated using Cattell's Culture Fair .

Results

As can be seen in Figure 2, there were considerable differences in estimates from the three VWM tasks. A task (3 levels) by set-size (5 levels) ANOVA on K confirmed effects of task ($F(2,94)=35.4$, $p<0.001$), set-size ($F(4,188)=103.5$, $p<0.001$) and an interaction ($F(8,376)=8.39$, $p<0.001$).

Strikingly, there was a clear difference between the VWM-WR-L and VWM-CD-L conditions, which required memory for identical information, but assessed it in different ways. Three key findings were salient:

(1) *WR-L gave higher estimate than CD-L.* This pattern, seen in Figure 2, does not provide evidence for output interference in the WR procedure. Instead, there appears to be some kind of interference that specifically affects CD. An

ANOVA with factors of task (VWM-WR-L & VWM-CD-L) and set-size confirmed an effect of task ($F(1,47)=23.1$, $p<0.001$) and an interaction ($F(4,188)=6.48$, $p<0.002$). This difference was driven by worse performance in CD, particularly at higher set-sizes (Table 1a).

(2) *WR-L was less variable.* Although the block was faster to acquire, WR gave smaller variance (Table 1b)

(3) *Decrease in the number of items remembered at higher set-sizes for CD but not WR.* The set-size dependence in CD cannot be an effect of a capacity limit alone, and instead suggests the measure is contaminated by some additional interfering cognitive process. At lower set-sizes, as more items are presented, more are remembered, for all procedures. But, above four items, while the WR condition stays approximately constant at higher set-sizes (consistent with a full VWM buffer), in the CD conditions there is a *decrease* in the number of items remembered with increasing set-size (4 to 8: WR, $t(47)=0.47$, NS; CD-L, $t(47)=2.4$, $p<0.05$; CD-LS, $t(47)=5.0$, $p<0.001$). This decline is significantly larger for CD than for WR (task by set-size (4 or 8) interaction CD-L, $F(1,47)=5.0$, $p<0.05$; CD-LS, $F(1,47)=21.4$, $p<0.001$).

These features suggest that the VWM-WR provides a more canonical measure of visual working memory capacity.

To understand how intelligence affects VWM, we conducted a median split by Cattell's Culture Fair score (mean 34, range 25-42). Performance in the three VWM procedures for the participants with lower (mean 31) and higher (mean 38) scores is shown in Figure 3. The first striking feature is that in the VWM-WR condition, there was no effect of intelligence, suggesting no fundamental relationship between VWM and aptitude (main effect of IQ $F(1,46)=0.00$, NS; set by IQ interaction $F(1,46)=0.45$, NS). Secondly, and importantly, there was an effect of intelligence in both of the CD conditions, but selectively for the higher set-sizes (WR vs. CD-L, task by set by IQ interaction $F(4,184)=3.1$, $p<0.05$; task by IQ interaction $F(1,46)=3.8$, $p=0.06$; WR vs. CD-LS, task by set by IQ interaction $F(4,184)=1.66$, NS; task by IQ interaction $F(1,46)=3.0$, $p=0.09$). Using IQ as a continuous measure rather than performing a median split gave a similar result (WR vs. CD-L ANCOVA- task by IQ interaction $F(1,46)=7.41$, $p<0.01$, Table 1c-e). This suggests that it is the additional cognitive process that depresses performance at higher set-sizes in CD that couples with intelligence, and not VWM *per se.*

A supplementary analysis was performed to investigate whether the lack of relationship between WR and IQ might be because the WR measure is *in general* less sensitive to individual differences. As reported above, WR is less variable than CD, and across subjects it has a smaller range (mean across set-size WR: 2.32-3.81; CD-L: 1.71-3.78; CD-LS, 1.51-3.55). The critical issue is whether this extra variance in CD reflects increased signal, or merely increased noise. To test this we used factor analysis to retrieve a single factor from the three measures (presumably underlying memory capacity), using maximum likelihood estimation on the mean across set-sizes. This single factor explained a substantial proportion of the variance (67.2%). Importantly, the factor loading for WR (0.85) was higher than for CD-L (0.60) or CD-LS (0.70). This indicates that WR is, if anything, more sensitive than CD to underlying individual differences, and was more tightly related to each of the CD measures than they are to each other (this was also confirmed with correlation, WR with CD-L 0.51; WR with CD-LS 0.59; CD-L with CD-LS 0.42). Similar results were obtained if IQ was partialled out prior to analysis (factor analysis 67.2%, weights WR: 0.92; CD-L: 0.58; CD-LS: 0.65; partial correlations WR with CD-L 0.53; WR with CD-LS 0.60; CD-L with CD-LS 0.38). WR is at least as sensitive as CD to individual differences and the absence of a relationship between IQ & WR cannot be attributed to lower sensitivity.

As the measure of selection, α', includes a term from VWM-WR, an unbiased assessment of individual differences in VWM (as measured this way) and top-down selection could not be calculated. There were no significant correlations between α' and VWM-CD-L or VWM-CD-LS at any set-size. It was also possible to examine what individual differences were associated with an ability to remember spatial information. As might be expected, there was a cost in having to also remember spatial position (CD-LS vs. CD-L, $F(1,47)=12.4$, $p<0.001$). To best estimate VWM with CD while ameliorating the influence of the drop at high set-sizes, for each person the maximum across set-size in the number of items remembered was calculated. These did not correlate with IQ or α', but the cost of also remembering spatial information (max(CD-L)-max(CD-LS)) did correlate with selection α' ($\rho(48)=0.30$, $p<0.05$; partial correlation with IQ removed $r(45)=0.30$, $p<0.05$) suggesting that selection and remembering spatial position share something over and above a dependency on IQ.

Discussion

WR gave a higher and less variable estimate of VWM than CD. At high set-sizes, an intriguing effect of IQ on performance was found in CD, with lower-Cattell participants remembering fewer items as more were presented beyond capacity. This deficit was absent in WR despite it requiring exactly the same memory load. What might be the cause of this depression in CD? The sensitivity of the effect to the number of items presented, even when above the capacity limit, suggests that it cannot be due to something that happens at the time of maintenance or comparison (e.g., the requirement to maintain the sample in the presence of the probe) as by the time these phases arrive, items above capacity should already have fallen out of VWM. Instead, we propose that the critical difference is at the time of encoding: that CD encourages a strategy of attempting to hold a visual snapshot of the entire display, but in WR the covert naming, or the ability to control which items are recalled on any trial, encourages more selective encoding. That fewer items were remembered when eight were presented than four in CD (particularly for those lower in aptitude) is consistent with a maladaptive lack of selection. It has been shown previously that selection is less effective in those that perform worse on VWM tasks . We propose that this maladaptive encoding strategy contaminates the CD measure at higher set-sizes, and that it is this, not VWM capacity that correlates with aptitude.

Experiment 2

Experiment 1 found that WR encourages better selection than CD. This might be because during the maintenance period of WR items were shuffled from VWM to the phonological loop, and the imminence of covert naming encouraged selection at encoding. Alternatively, in WR, participants can control which items they report and this might have encouraged more selection than in CD, in which random items are probed. It was the aim of experiment 2 to contrast these two accounts, by examining the effect of suppressing verbalisation.

Method

Memory capacity was assessed again using WR for letters, but with the addition of articulatory suppression, in which overt rehearsal of an interfering stimulus was used to prevent covert or overt rehearsal of the memoranda. Articulatory suppression is robust to the exact items being uttered (Baddeley, 1990, pp.79-80). Participants have been asked to rehearse a single word (e.g., "the-the-the…" Baddeley, 1990), numbers (two digits, overtly, Vogel, Woodman & Luck, 2006; covertly, Todd & Marois, 2004), or a list of items (over-learnt series, "a-b-c-d", Woodman & Luck, 2004). To ensure fully effective suppression, the current experiment used a hard task, in which five spoken random digits were rapidly presented (SOA 300ms) and participants asked to overtly rehearse them at this rate for the remainder of the trial. The digits were windowed to be 250ms in duration with linear onset & offset ramps to prevent spectral splatter, and presented in free field at a comfortable listening level. Between the end of the last digit and the sample display there was a delay of 1550ms to allow participants to enact rehearsal. Compliance was ensured by recording participants' voices using a portable digital recorder.

It was not clear *a priori* the degree to which phonological rehearsal would effect performance on trials with a short encoding display (183 ms) and so to ensure an assay of the effect of articulatory suppression, trials with a longer sample display were also included (1000 ms). However, this condition will overestimate VWM as it will allow time for visual chunking, and the piecemeal transfer of items from VWM to phonological memory (Cowan, 2001).

There were fourteen participants (7 female). All were informed prior to initial consent that their voice would be recorded during the experiment. Their mean age was 49 and Cattell score 38. There were two blocks, one with articulatory suppression (WR-AS), and one without, with the order counterbalanced across participants. A factorial design of sample display duration (183 and 1000ms) and set-size (2,3,4,6,8) yielded 10 conditions each of 24 trials per block, intermixed in random order. Apart from the addition of the articulatory suppression task and the modification of encoding duration in some trials, the paradigm and analysis were as in experiment 1.

Results

Figure 4 shows means and standard errors across subjects. An ANOVA found main effects of set-size ($F(4,52)=79.7$, $p<0.001$) and encoding duration ($F(1,13)=90.6$, $p<0.001$). The articulatory suppression task reduced performance ($F(1,13)=10.2$, $p<0.01$), but had no interaction with encoding duration ($F(1,13)=0.05$, NS).

For comparability with experiment 1, subsequent analyses then focussed on the short encoding condition. Importantly, even in the presence of articulatory suppression there was no suggestion of a downturn at high set-sizes (WR-AS, short encoding, set-size 4 vs. set-size 8, paired-sample $t(13)=0.37$, NS). This drop-off was statistically smaller than for lower-Cattell CD-L ($t(36)=2.08$, $p<0.05$), and similar to higher-Cattell CD-L ($t(36)=-0.7$, NS). These results show that rehearsal cannot account for the absence of a drop-off at higher set-sizes in WR.

A prediction from Cowan et al. (2005) is that in the presence of articulatory suppression, WR will yield a purer estimate of capacity and so reveal a positive correlation with intelligence. There was no evidence of such a relationship, although limited conclusions should be drawn from this due to the sample size (one-tailed test of Spearman's correlation between mean K & IQ, $\rho(13)=-0.28$, NS; Table 1f).

Overall performance in the presence of articulatory suppression (WR-AS, short encoding, mean K 3.04 ± 0.12) was reduced to the same level as the higher-Cattell CD-L condition from experiment 1 (mean $K=2.98\pm0.09$, $t(36)=0.40$, NS). However, as expected it was better than the lower-Cattell group in the CD-L condition (mean $K=2.76\pm0.09$, $t(36)=1.82$, $p<0.05$ one-tailed).

Discussion

There was no suggestion that articulatory suppression impacted higher set-sizes beyond capacity more than those around capacity. This suggests rehearsal of the memoranda is not required to encourage the selective strategy seen in WR. Instead, we propose the key characteristic is that the paradigm allows participants to choose which items they report, in contrast to the random probing of CD.

Performance of a hard articulatory suppression task reduced performance in WR. This might be a general dual-task decrement. Rehearsing digits has been shown before to interfere with general reasoning (Baddeley, 1986), and it may impact on executive control to the detriment of many aspects of the memory task. Alternatively, there might have been a more specific impact of articulatory suppression on maintenance. The bottleneck in WR for letters is primarily VWM (e.g., Sperling, 1960; Bundesen, 1990; Cowan, 2001), but phonological rehearsal may contribute to maintenance for the remainder of the memory delay, in which an occasional item may drop from VWM. The articulatory suppression task will have prevented such rehearsal.

General Discussion

The WR procedure gives a less variable and more easily interpretable estimate of VWM capacity than CD. CD appears to be contaminated by some additional cognitive process that depresses performance at higher set-sizes. The magnitude of this contamination was found to be greater in people that performed worse on an intelligence test, but there was no relationship between VWM and intelligence *per se.* Experiment 2 showed that rehearsal cannot account for the absence of a drop-off at higher set-sizes in WR.

In experiment 1, overall performance in WR was better than for CD. In experiment 2, the presence of a difficult articulatory suppression task reduced overall performance on WR to a similar level as for CD in experiment 1. It is not clear whether this reduction is because of general dual-task interference, or more specifically, suppression of rehearsal.

The inference that there is a difference in encoding strategy between CD & WR (see Discussion of experiment 1) has been confirmed in a subsequent experiment from our laboratory. In this, CD-L & WR-L trials were randomly intermixed, so that participants did not know until probed how they would have to respond. This led to a marked reduction in the difference between the procedures. This suggests that worse CD in participants with lower-Cattell is more likely due to the adoption of a sub-optimal strategy of trying to encode everything, rather than the result of a fundamental inability to select appropriately.

Studies using verbal memoranda have found that complex measures combining storage and processing correlate more strongly with intelligence than simple measures of storage alone (Baddeley, 1996). Both Miyake et al. (2001) and Cowan et al. (2005) investigated whether the same is true for visuospatial memory, and found it was not, as both simple and more complex visuospatial measures correlated with intelligence to the same degree. They concluded that there is an asymmetry between verbal and visuospatial storage, with the latter fundamentally dependent on central processes related to intelligence. Our findings suggest a possible reinterpretation: perhaps even in their simple tasks intelligence had an effect on strategic choices, rather than being related to VWM *per se.* Interestingly, Miyake et al's tasks focussed more on spatial memory, and our findings may have implications for this related field of study.

Given the parsimonious explanation of maladaptive selection, a surprising feature of the existing dataset is that our measure of top-down selection in experiment 1 did not correlate with performance in the CD procedures. It might be that as discussed above, the effect of intelligence on CD is mediated by differences in strategy, rather than in ability to select *per se*, and that when the task instructions explicitly require selection, the relationship is removed. This is consistent with both Gold et al (2006) and Cowan et al (in press), who found a trend of drop off in CD at high set-sizes but good performance on an explicit selection task. The partial report selection measure did show associations in other respects: better α' was correlated with less of a decline in performance when space also had to be remembered in a CD task. This might be because both measures tap into spatial representations, and may have further contributed to the association between selection and VWM measures observed in tasks that required memory for spatial position (e.g., Vogel, McCollough & Machizawa, 2005).

Many CD experiments in the literature neither require that subjects use the position information (as in CD-LS) nor require that they ignore irrelevant changes in position (as in CD-L). However, a drop in CD at high set-sizes has been observed in experiments where a single probe was always presented at fixation (e.g., Chee & Chuah, 2007), where a single item is probed but its position always kept similar to the sample (e.g., Song & Jiang, 2006) or when all of the items are probed and their position kept constant (e.g., Gold et al, 2006).

Although this report has focussed on the important differences between procedures, it should be noted that there were significant commonalities across the measurements, as evident from the good correlations and strong single factor model. Reassuringly, both procedures reflect VWM to a substantial degree.

In summary, the results of this study suggest that VWM capacity in itself is not related to cognitive aptitude, but that CD may tap into maladaptive encoding strategies that impact upon performance at higher set-sizes.

Acknowledgements

We thank Alejandro Vicente-Grabovetsky for assistance with computer programming, John Duncan and Daniel Bor for insightful comments, and Ian Nimmo-Smith and Peter Watson for statistical advice.

References

Baddeley, A. (1996). *The fractionation of working memory* (Vol. 93, pp. 13468-13472). National Acad Sciences.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523-547.

Chee, M.W.L & Chuah Y.M.L. (2007). Functional neuroimaging and behavioral correlate of capacity decline in visual short-term memory after sleep deprivation. *Proc Natl Acad Sci.*; 104, 9487-9492.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*, 87-114; discussion 114-185.

Cowan, N., Elliott, E.M., Scott Saults, J., Morey, C.C., Mattox, S., Hismjatullina, A., et al. (2005). On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology, 51*, 42-100.

Cowan, N., & Morey, C.C. (2006). Visual working memory depends on attentional filtering. *Trends in the Cognitive Sciences, 10*, 139-141.

Cowan, N., Morey, C.C.., AuBuchon, A.M., Zwilling, C.E., Gilchrist, A.L. (in press). Seven-year-olds allocate attention like adults unless working memory is overloaded. *Developmental Science.*

Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Chavda, S., & Shibuya, H. (1999). Systematic analysis of deficits in visual attention. *Journal of Experimental Psychology: General, 128*, 450-478.

Finke, K., Bublak, P., Krummenacher, J., Kyllingsbaek, S., Muller, H.J., & Schneider, W.X. (2005). Usability of a theory of visual attention (TVA) for parameter-based measurement of attention I: evidence from normal subjects. *Journal of the International Neuropsychology Society, 11*, 832-842.

Gold, J.M., Fuller, R.L., Robinson, B.M., McMahon, R.P., Braun, E.L., Luck, S.J. (2006). Intact attentional control of working memory encoding in schizophrenia. *J Abnorm Psychol*;115, 658-73.

Luck, S.J., & Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279-281.

Measuring Intelligence with the Culture Fair Tests. (1973). Champaign, IL: Institute for Personality and Ability Testing.

Miyake, A., Friedman, N.P., Rettinger, D.A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology. General*, *130*, 621-40.

Peers, P.V., Ludwig, C.J., Rorden, C., Cusack, R., Bonfiglioli, C., Bundesen, C., et al. (2005). Attentional functions of parietal and frontal cortex. *Cerebral Cortex, 15*, 1469-1484.

Phillips, W.A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics, 16*, 283-290.

Sobel, K.V., Gerrie, M.P., Poole, B.J., & Kane, M.J. (2007). Individual differences in working memory capacity and visual search: the roles of top-down and bottom-up processing. *Psychonomic Bulletin and Review, 14*, 840-845.

Sperling, G. (1960). The information available after brief visual presentation. *Psychological Monographs, 74*, 1-29.

Vogel, E.K., McCollough, A.W., & Machizawa, M.G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature, 438*, 500-503.

Tables

Table 1. Statistics for experiments 1 & 2 for individual set-sizes.

	Experiment 1					Experiment 2
Set-size	(a) Mean WR-L vs. CD-L (t(47))	(b) Variability WR-L vs. CD-L (F(47,47))	(c) Correlation IQ & WR (spearman's ρ(48))	(d) Correlation IQ & CD-L (spearman's ρ(48))	(e) Correlation IQ & CD-LS (spearman's ρ(48))	(f) Correlation IQ & WR-AS (spearman's ρ(13))
2	-0.11	0.53	0.01	0.17	0.08	-0.31
3	3.0 (**)	3.7 (***)	0.01	-0.09	-0.07	-0.30
4	3.9 (***)	1.6 (*)	-0.12	-0.43	0.06	-0.54
6	2.8 (**)	2.1 (**)	-0.03	0.30 (*)	0.29 (*)	-0.21
8	3.6 (***)	3.8 (***)	0.00	0.27 (p=0.07)	0.19	-0.26

* p<0.05 ** p<0.01 *** p<0.001

Figure legends

Figure 1. The structure and timing of a trial for the change detection (CD, left) and whole report (WR, right) tasks.

Figure 2. The number of items remembered as a function of set-size for three different VWM tasks (WR-L = whole report for letter identity; CD-L = change detection for letter identity; CD-LS = change detection for letter identity and spatial position).

Figure 3. Performance in experiment 1 on three VWM tasks for lower- (lighter colours, diamonds) and higher- (darker colours, squares) intelligence groups created by a median split on a non-verbal IQ measure, Cattell's Culture Fair (WR-L = whole report for letter identity; CD-L = change detection for letter identity; CD-LS = change detection for letter identity and spatial position).

Figure 4. Estimated memory capacity as a function of set-size in experiment 2 for a whole report task with either long (1000 ms) or short (183 ms) encoding durations, and with or without articulatory suppression (AS).

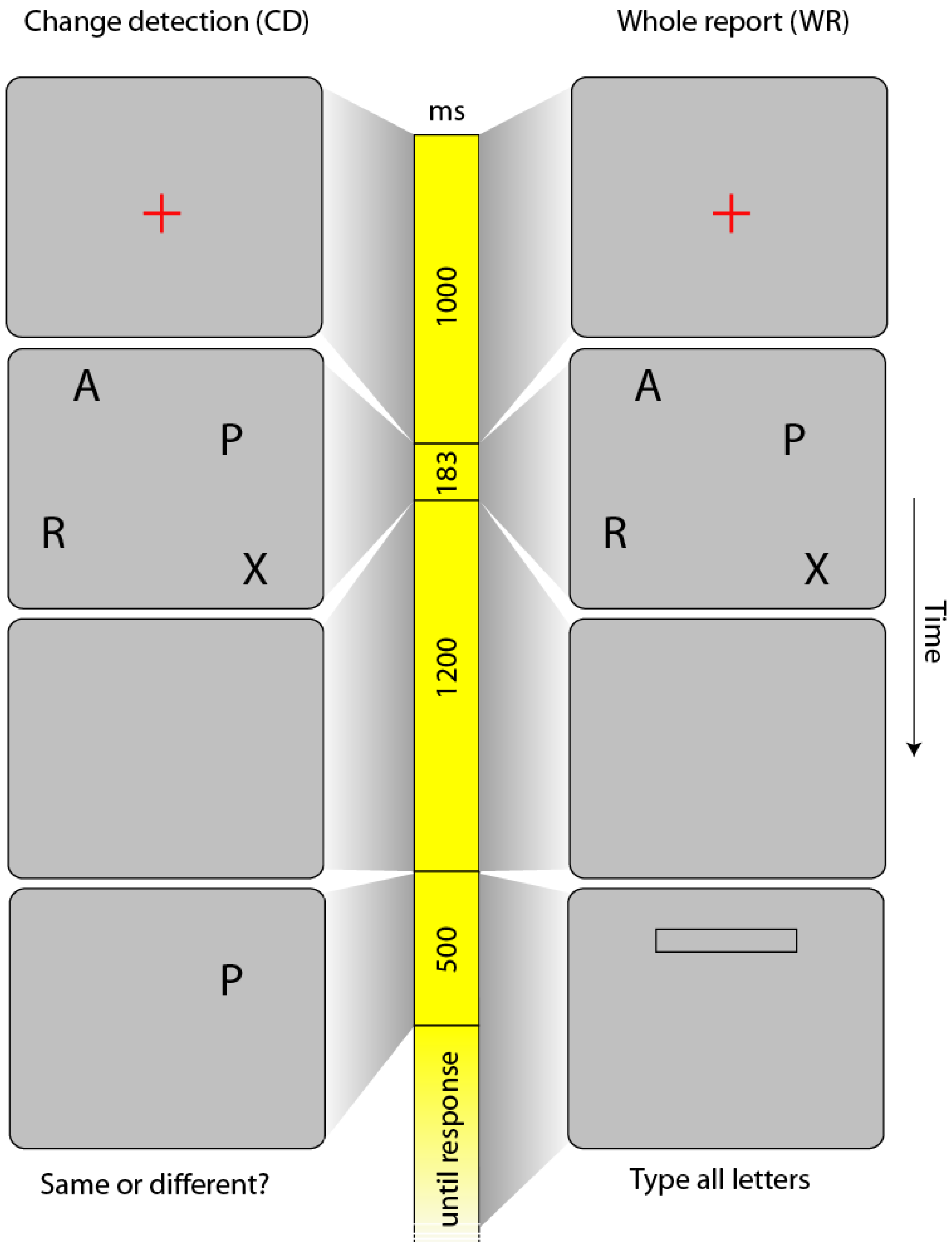
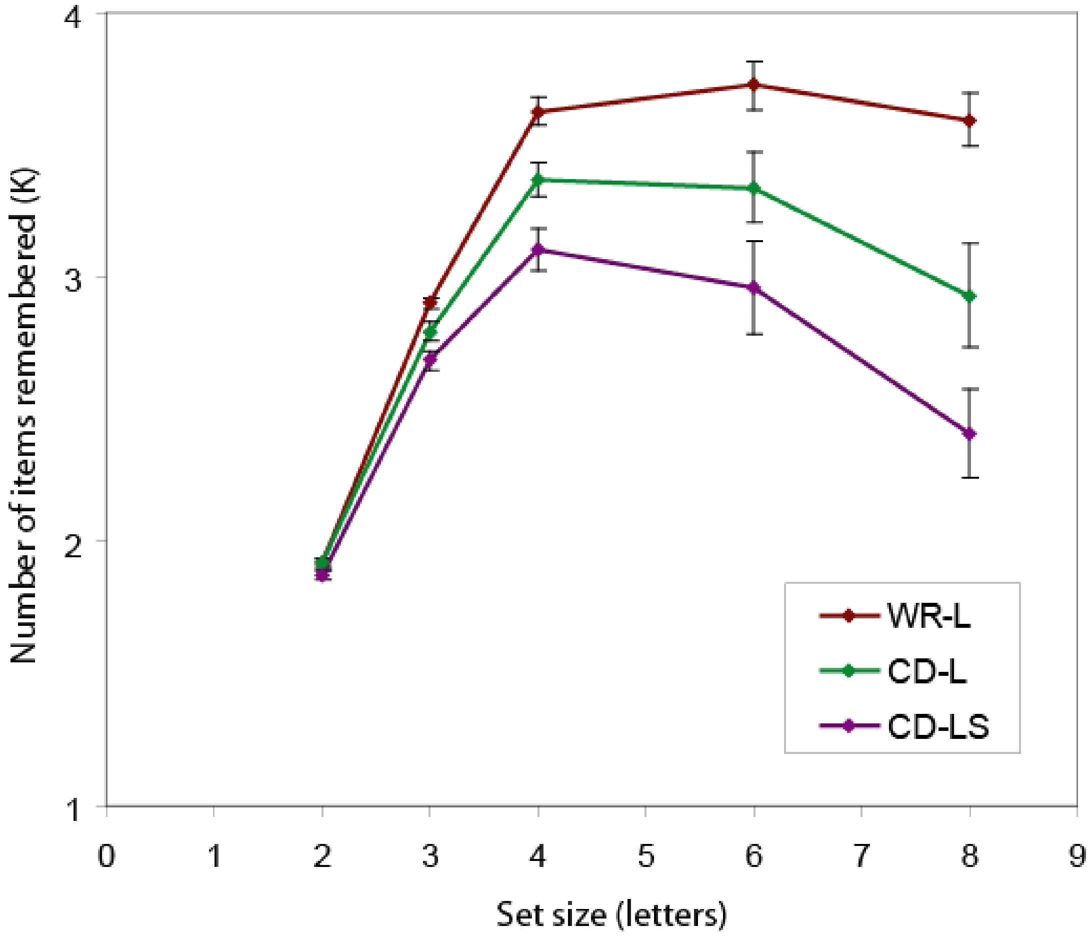
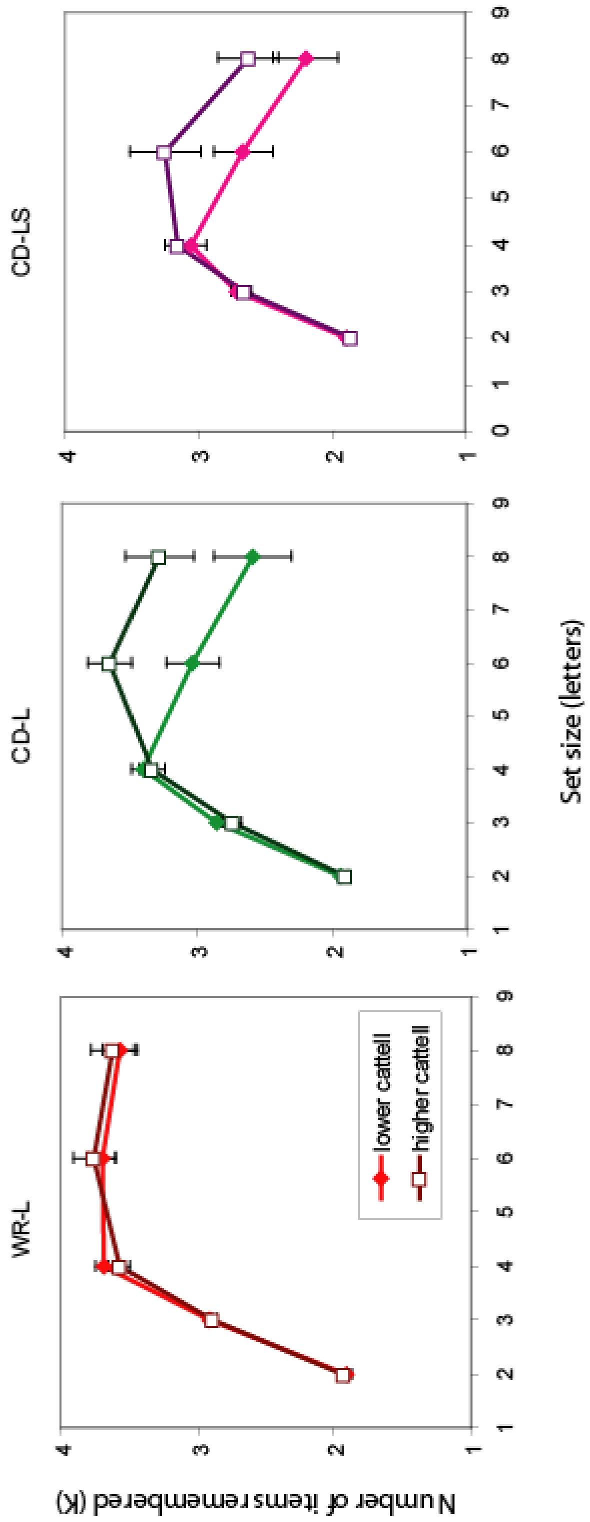

Change detection (CD)
Whole report (WR)
ms
1000
183
1200
500
until response
A
P
R
X
P
Time
Same or different?
Type all letters

Figure 1



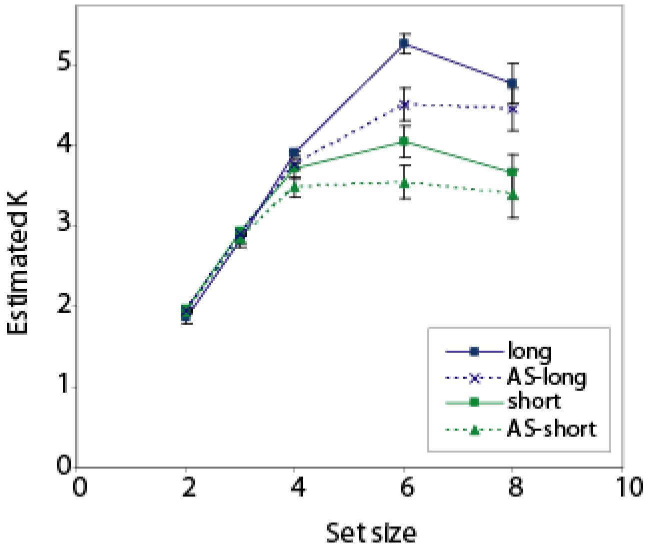
Number of items remembered (K)
4
3
2
1
0
1
2
3
4
5
6
7
8
9
Set size (letters)
WR-L
CD-L
CD-LS

Figure 2



WR-L
CD-L
CD-LS
lower cattell
higher cattell
Number of items remembered (K)
Set size (letters)

Figure 3



Estimated K
Set size
long
AS-long
short
AS-short
0
1
2
3
4
5
0
2
4
6
8
10

Figure 4