Evidence from reaction times for an anticipatory process in symbolic delayed matching-to-sample

David W. Dickins
University of Liverpool

Benjamin J. A. Dickins
Center for Comparative Genomics and Bioinformatics, Pennsylvania State University

It was predicted that a delay between sample ofset and comparison onset in symbolic delayed matching-to-sample (DMTS) would allow time for the anticipatory retrieval of the correct comparison so that the match/no match status of the presented comparison could be decided more swiftly. Te relation between delay and reaction time (RT) was explored in participants after they had been similarly tested, as a control, on identity DMTS using the same stimuli. In most participants there was the predicted inverse relation between delay and RT in symbolic DMTS, but no such relation in identity DMTS. Subsequently an arithmetic test, designed to allow a simple calculation before or after presentation of the comparison, was used to demonstrate an analogous efect of a delay on RT.

Keywords: delayed matching-to-sample (DMTS), single-comparison alternate response procedure, reaction time (RT), human adults; anticipatory retrieval of comparison.

Te principal proposal in the present study[1] is that diferences in reaction time (RT) as a function of delay duration can provide chronometric evidence of anticipatory processes in symbolic (or arbitrary) delayed matching-to-sample (DMTS), but that these processes would not occur in identity DMTS. It was hypothesised that in arbitrary matching a process of retrieval of the stimulus with which the sample had been paired in training could begin, and perhaps reach completion during the delay. Tis would make ready for the process of matching between this retrieved (correct) comparison with the actual comparison when the latter subsequently appeared. Terefore over some range of delays from zero upwards the RT (since this is defned as the time between comparison onset and the response) should be a decreasing function of increased delay. In identity DMTS the stimulus to be matched would simply be the sample itself, with no other intermediate stimulus needing to be retrieved. No savings in RTs as a function of increased delay would be expected.

To test this hypothesis participants were exposed frst to identity DMTS and then to symbolic DMTS, using the same stimuli. In the first experiment identity matching was rapidly established using a single-comparison alternative-response (SCAR) DMTS procedure and the efects on performance of 4 levels of delay –zero, 500ms, 1000ms, and 1500ms - were compared on 8 separate blocks of 12 trials, 4 blocks using 12 abstract shapes as stimuli, and 4 blocks using 12 nonsense syllables (consonant-vowel-trigrams). In the second experiment 12 shape-syllable trained relations were taught to

a criterion of 12/12 correct trials per test block, using the same test procedure, and then performance was further tested on 24 further 12-trial blocks (using a sequence of six delays for successive blocks repeated 4 times in an alternately decreasing and increasing series) to reveal any systematic relations between RT and delay duration. In a third experiment, using an analogous DMTS procedure, a series of 48 trial-unique addition sums were given to the participants, where either the sample consisted of two numbers to be added $(x + y)$ to be checked against the putative total (z) as comparison, or vice versa $(z = x + y)$. Under each condition one 12-trial block was tested with a zero and one with a 1200 ms delay. Tis arithmetic experiment was designed as a demonstration of savings on RT when a delay provided an opportunity to perform an addition prior to the onset of the comparison, simulating the postulated anticipatory efects in Experiment 2.

Previous studies of the efects of delay in DMTS have been addressed to deleterious efects upon accuracy of performance of increasing delays, and have concerned primarily identity DMTS (e.g., Roberts, 1972; Roberts & Grant, 1974; Weavers, Foster, & Temple, 1998; White, 1985; Williams, Johnston, & Saunders, 2006). Typically in DMTS on a given trial one of a set of stimuli serves as a sample stimulus. Tis is presented for a fnite period the termination of which is followed by a delay of a given duration prior to the onset of one or more comparison stimuli. If there are two or more comparisons the response is to choose one of them. In identity matching the correct stimulus will be the same stimulus that was used as sample. If there is only one comparison one of two alternative responses is usually required, one response if the comparison matches (viz. is the same as or identical to the sample), and the other response if it does not match (when sample and comparison are diferent). On correct trials the sample stimulus somehow acts across the gap to set the occasion for the appropriate response, but the likelihood of this happening (= the accuracy, measured as the percentage of correct responses) commonly diminishes as the duration of the delay is increased.

In symbolic (or arbitrary) matching-to-sample, unlike identity MTS, participants have to learn a consistent relation between each specifc sample stimulus and a specifc comparison. (Wright, 2001). Te two stimuli have been arbitrarily paired for training, and a 'match' is when sample and comparison belong to the same training pair and a non-match response is required when sample and comparison belong to diferent trained pairs.

In cognitive terms DMTS may be construed as demonstrating a 'short-term' or 'working memory' for the sample stimulus (Baddeley, 2003). When the sample is no longer present on the screen the participant somehow maintains a 'representation' (Roitblat & von Fersen, 1992) of the sample (or its discriminant function is somehow retained in the brain) that at the end of the delay can be used to gauge a match or non-match with the comparison. In cognitive psychology such terms may be seen as metaphors based upon a computer-like analogy of psychological functions. Alternatively they can be seen as hypothetical entities corresponding to events in the brain potentially susceptible to demonstration (e.g., Funahashi, Bruce, & Goldman-Rakic, 1989; Fuster, 1973). Liebe, Logothetis, and Rainer (2008) were able to show that a local feld potential in the prefrontal cortex of macaque monkeys carried performance-related information, which made it possible to decode the behavioral choice as well as the reaction time in a DMTS task.

Behaviour analysts on the other hand seek to delineate a more functional account of the behavioural phenomena of DMTS. Wixted (1989) for example, sees MTS as a set of conditional discriminations, and endeavours to account for the forward efect of the sample (and its attenuation with an increasing delay) in terms of the temporal parameters of sample duration, delay, and intertrial interval. Such accounts may be seen as conficting with cognitive type descriptions, or they may simply be explanations at a diferent level.

In non-human studies of DMTS it has been suggested that features of sample stimuli may be processed in two ways before the onset of comparison stimuli. Information from the

sample may simply be carried over the delay, after which it is said to be 'retrospectively' processed, or it may be processed 'prospectively' (Zentall, 2001). Evidence for a prospective or 'anticipatory' process sought in pigeons by examining the efects of delays on error rates is ambiguous (Grant & Kelly, 2001).

Nevin, Davison, Odum, and Shahan (2007) describe the human subject, during the presence of the sample, as "Attending, discriminating, encoding features". During the delay (in either identity or symbolic DMTS), the participant is "attending to sample as coded". Tis, they say, "corresponds to the notion of rehearsal in more cognitive accounts of memory processes, and may best be conceptualized as attending to any sample-related behavior, measured or unmeasured, that the participant may emit during the retention interval." Also during the delay the participant is engaged in covert "Orienting, emitting observing behavior" in relation to the forthcoming comparison (they posit similar behaviour, presumably in the inter-trial interval, in relation to the sample before it appears). In the presence of the comparison the participant is "Attending, discriminating, emitting (the response)". Tis is reminiscent of an intuitive minimal list of 'sub-processes' involved in DMTS, devised by the frst author (Dickins, 2005) but, unlike this 'model', characterizations such as those of Wixted (1989) and Nevin et al. (2007) relate to studies solely of the accuracy of responding as a function of the delay, and seem almost universally to ignore reaction times. In the present study RT was the measure of choice, and human participants were trained and tested on a DMTS procedure that generally yielded high accuracy of responding irrespective of the delay over the range of delays investigated.

Te time a reaction takes may be related to concomitant neurophysiological processes which inherently require a certain time to take place, a principle exploited by the classic 'chronometric' approach to the study of experimental performances (Posner, 1978, 2005; Sternberg, 1969a, 1969b, 2001). Te principal proposal in the present study is that RT diferences may provide chronometric evidence of anticipatory processes in symbolic DMTS.

EXPERIMENT 1

Te purpose of Experiment 1 was to investigate the efects of a range of delays from zero to 1500ms on RT and accuracy in identity DMTS. It was expected that with delays of such short duration there would be little efect on accuracy, unlike the delays of many seconds commonly associated with diminished accuracy in much of the DMTS literature, such as that cited above. Nor were RTs, over the same range of delays, expected to vary signifcantly. Tis experiment serves as a kind of control for the second experiment.

Method

Participants

Twelve frst year psychology undergraduates enlisted as part of the Research Participation Scheme of the School of Psychology, University of Liverpool, which is a requisite part of the course. Participants were anonymised by being numbered S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, and S14.

Te experiments were part of a suite of experiments all approved by the University of Liverpool Ethics Committee.

Stimuli

Twenty-four stimuli (Figure 1) served both as samples and as comparison stimuli. 12 were abstract shapes (Stimuli A1-A12) and 12 were nonsense syllables (consonant-vowel-consonant trigrams) (Stimuli B1-B12).

Setting and Apparatus

Te sequence of 3 experiments required an uninterrupted session of about 1.5 hours, singly or in a group of 2 or 3 in a small laboratory, each participant sitting at a table in front of an Apple iMac computer, in the presence of the experimenter (frst author). Te experiments were run and all data were recorded using a script written in RunRev (Runtime_Revolution), Version 2.6, by Mr. Phil Jimmieson of the Department of Computer Science, University of Liverpool.

Instructional screens (see Appendix 1, 2, &

3) guided the participant through the program. When these had been read mouse-clicking a button or pressing Return or the space bar on the keyboard moved the participant on to the next stage of the program. In the 'response phases' (see below) responses were made by pressing either the C or the M key on the keyboard, and participants were encouraged to operate these with a fnger from each hand and to move on through the program by pressing the space bar with either thumb.

Instructions

Participants were shown the locations on the computer screen of the sample stimulus and the subsequent comparison stimulus. At the onset of the comparison stimulus the word SAME appeared in a box in the bottom left of the screen and the word DIFFERENT simultaneously appeared in a box in the bottom right.
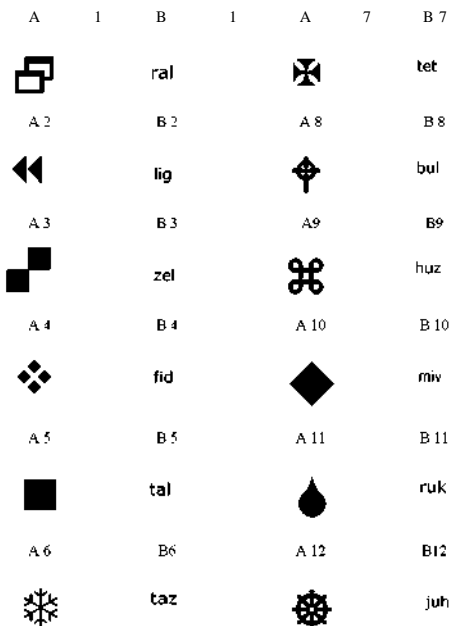
A	1	B	1	A	7	B 7
⌐		ral		✠		tet
A 2		B 2		A 8		B 8
◀◀		lig		✙		bul
A 3		B 3		A 9		B9
◣		zel		⌘		huz
A 4		B 4		A 10		B 10
❖		fid		◆		miv
A 5		B 5		A 11		B 11
■		tal		◖		ruk
A 6		B6		A 12		B12
❄		taz		☸		juh

Figure 1. Te stimuli used in Experiments 1 and 2. Te A stimuli were the shapes in columns 1 and 3 (from the left) and the B stimuli were the nonsense syllables in columns 2 & 4. Labels such as A8 or B8 shown in the fgure are for explanation only, and to show the correct pairings in the relations that were trained: these labels were not visible to the subject at any time.

Participants were verbally instructed that if the sample and the comparison were the same they should press the C key on their keyboard corresponding to SAME, but if they were not the same they should press the M key corresponding to DIFFERENT. Participants were given no informational feedback on trials, nor were there any other programmed consequences. After they were shown how to enter their name and details and how to move through the program by pressing the return key at the appropriate points the program was started and they received on screen a reiteration of the above instructions (see Appendix 1)

Procedure

Each participant was given 8 successive test blocks, each consisting of 12 identity MTS trials. Te frst 4 'AA' blocks deployed the 12 'A' stimuli, (see Figure 1) each of which served once as a sample and once as a comparison in each block in which the program paired half of the samples with the same stimulus as comparison, viz. constituting 6 'same-correct' trials, and shufed the pairings of the other 6 sample stimuli so that each was paired with a diferent comparison from that set of six, constituting 6 'diferent-correct' pairs. Pairs were presented in random order. New randomisations occurred for each block of trials. Te same procedure continued with the 12 'BB' trials in each of the last 4 blocks.

On each trial a sample stimulus was presented in a box on the left side of the computer screen for 200ms. Ofset of the sample began the programmed delay (see below), which terminated with the presentation of a comparison stimulus in a similar box on the right. Te comparison stimulus was exposed for 200ms. Te SAME and DIFFERENT buttons appeared (denoting the time during which either a 'C' or an 'M' response on the keyboard was efective) at the same time as the onset of the comparison, but lasted for 2s longer, or until a response was made. An intertrial interval of 1s followed.

Te delay values (ms) were changed between trial blocks, and were as follows: AA 500, 1500, 1000, 0; BB 1000, 500, 1500, and 0.

Te program automatically recorded the following data for each trial: trial number (1-12); RT; type of trial - C or M (same-correct or diferent-correct); response ('c' or 'm'); 'correct' or 'incorrect' (or 'timeout' if no response); sample stimulus; comparison stimulus.

Te data were analyzed on a trial-by-trial basis: fles generated by the RunRev script were parsed for relevant parameters using further custom-built scripts written and executed in Python 2.6.2. Tab-delimited fles thus generated were analysed in R 2.9.1 (RDevelopmentCoreTeam, 2009) using custom-built R scripts and employing functions from the 'coin' (Hothorn, Hornik et al. 2008) and 'SAGx' packages (Broberg, 2009). All scripts are available on request.

Results and Discussion

Accuracy

Without preliminary training participants responded at a high level of accuracy from the outset which was sustained with never more than 2 and on average less than 1 incorrect responses for each block of 12 trials, with the exception of participants S1 and S14 who,

on the zero delay test blocks only, performed poorly with numerous incorrect responses or timeouts (see Table 1). Tis can be explained post hoc by the brevity of the comparison exposure time of 200ms that, without a pause between sample ofset and comparison onset, may have been difcult for these participants to register. Te overall average percentage of correct responses was 94.3 %, with an SD of ± 4.2.

Reaction times

Figure 2 shows the box plots of RTs for correct responses on identity matching (AA and BB combined) for each of the 4 values of delay for all 12 participants tested. For each of the 4 delay levels the blue box on the left represents RTs of same-correct (C) responses, and the pink box on the right of each pair shows the RT distribution of diferent-correct (M) responses. Te medians for same-correct responses can be seen for each delay to be lower than those for diferent-correct responses, but these local diferences are not signifcant since the notches in adjacent boxes overlap. For both kinds of response medians can be seen to be lower for zero delay than they are for the other 3 delays.

Table 1: Number of responses correct /12 for successive response phases during identity DMTS testing in Experiment 1.

	AA trials				BB trials				Summation			
Delay (s)-»	0.5	1.5	1.0	0	1.0	0.5	1.5	0	Total correct	Total errors	TO	% correct
Participant												
S1	10	12	12	4	12	12	12	7	81	5	10	84.4
S2	12	12	11	10	12	12	11	12	92	2	2	95.8
S3	12	12	12	10	12	11	12	11	92	3	1	95.8
S4	12	12	11	11	12	11	11	10	90	6	0	93.7
S5	12	12	12	12	12	12	11	12	94	1	1	97.9
S6	12	10	12	11	12	11	12	11	91	4	1	94.8
S7	11	12	11	12	12	12	12	11	93	3	0	96.9
S8	10	11	12	11	12	11	11	11	89	4	3	92.7
S9	11	12	10	11	12	12	12	10	90	4	2	93.7
S10	11	12	12	11	12	12	12	12	94	0	2	97.9
S11	12	12	12	11	12	12	12	12	95	0	1	99.0
S14	12	11	12	5	12	12	12	9	85	7	4	88.5

Mean correct = 94.3% ± 4.2

Data analysis

Reaction times are typically distributed in a highly non-normal manner and are therefore often refractory to ANOVAs or analyses based on medians. Ideally one would like to draw inferences based on the parameters of the reaction time distribution itself (Whelan, 2008). However, in this experiment, while the overall sample sizes were large, those at the intersections of the variables of interest (delay, same versus diferent trials) and the relevant blocking factor (participant identity) were small, precluding robust estimation of parameters.

Instead we therefore adopted a non-parametric approach based on ranking of the dependent variable (and conditional inference by permutation: Hothorn, Hornik, van de Wiel, & Zeileis, 2006) to lessen the efect of extremely long RTs which might distort measures of central tendency. Linear-by-linear association (LBL) tests were carried out in which reaction times were ranked, delays were coded as an ordered factor, and participant identity was introduced as a stratifying factor for permutation. A quadratic-form test statistic was used and these tests were instantiated in

the 'kruskal_test' convenience function of the 'coin package': (Hothorn, Hornik, van de Wiel, & Zeileis, 2008). Variables in Wilcoxon Mann-Whitney (WMW) tests were similarly treated, with stratifcation by subject, except that the binary independent variable was dummy-coded and a scalar test statistic was used (Wilcox test).

Te chief inference to be drawn from the data was whether or not there was a relation and, if so, in what direction, between delay and RT. While ranks showed a positive relation, that is increased RTs with increasing delay (Figure 2), this failed to reach signifcance in same-correct or diferent-correct trials or when all data were combined (two-tailed $p > 0.10$ for all LBL tests). (Jonckheere-Terpstra tests also failed to confrm signifcant trends in overall or individual data with one-tailed tests, not corrected for multiple comparisons, with an uncorrected $p < 0.05$ identifed in only one subject, S14, who showed the highest median RT of all subjects at zero delay). One signifcant diference did emerge: same-correct trials showed lower RTs than diferent-correct trials (one-tailed $p < 0.0001$ for WMW tests). (Tis was confrmed in several participants separately - S3, S4, S5, S8, S9 and S14 - with uncorrected $p < 0.05$).

Te diference between M and C RTs, with M being longer than C, is interesting, but the main fnding relevant to our main hypothesis was the lack of any signifcant negative relation between increasing delay and decreasing RT. (Instead a positive but non-signifcant relation was observed in the data.) Tis was predicted as a corollary of the main hypothesis: in Identity MTS there was no time to be saved in any delay by anticipatory retrieval of an absent stimulus because this was not needed, (and at this stage there were no trained relations between stimuli to be evoked). Te sample stimulus only needed to be registered, and then presumably held over any delay in some kind of working memory (WM) store (Baddeley, 1986). If it had any efect a delay might be expected to lead to the attenuation of WM, perhaps associated with a longer RT. Te suggestion of a positive relation between delay and RT fts this notion.
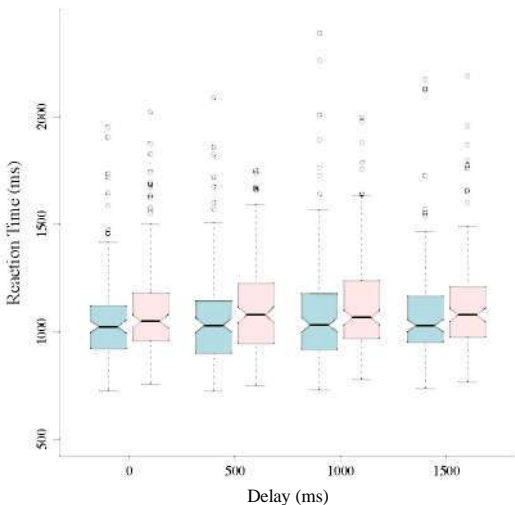

Figure 2. Box plots of RTs for identity DMTS (AA and BB combined) for each of the 4 values of delay for all 12 subjects tested. For each of the 4 delay levels the blue plot on the left of each pair represents RTs of same-correct (C) responses, and the pink plot on the right shows the RT distribution of diferent-correct (M) responses.

EXPERIMENT 2

Te purpose of this experiment was to test the efects of a range of delays, up to 3s, on RT and accuracy in symbolic DMTS. In symbolic DMTS, a sample stimulus and a comparison stimulus are deemed to 'match' if an arbitrarily assigned relation between them has previously been trained. A non-match would be between two stimuli belonging to diferent trained relations. Te main hypothesis of the study was that the appearance of the sample enables a prospective process of retrieval of the comparison with which it has a trained relation to be initiated. Tis comparison, once it had been retrieved, could itself be matched (in a similar way to that of identity matching) with the comparison, when it appears, after the delay. If this hypothesised process takes a discernible fnite time, and has time to be completed during a delay, the subsequent RT, timed from the onset of the comparison, should be proportionately shorter than RTs with zero delays after a minimal sample time.

Method

Participants

The 12 participants from Experiment 1 subsequently participated in the symbolic MTS procedure in Experiment 2.

Stimuli

Te same stimuli were used as in Experiment 1 (see Figure 1) except that only the 'A' stimuli served as samples and only their corresponding 'B' stimuli served as comparisons. Te stimuli are shown in their correct (trained) pairs side by side in Figure 1, for example 'A1' and 'B1', or 'A5' and 'B5'.

Procedure

In this single-comparison alternative-response (SCAR) procedure the correct AB pairings were acquired by observation learning in the frst instance, in "study phases", in which the 12 correct pairs were simply presented once each in random order in a Pavlovian manner without any overt response being required. Te participant was simply shown an instructional screen stating: "Memorize the following pairs." (Appendix1).

Study phases alternated with "response phases" which were blocks of 12 trials in each of which either a SAME or a DIFFERENT response was required, depending respectively upon whether the stimuli composing the sample-comparison pairs now presented were combined as in the study phase, or rearranged. On individual trials participants were given no informational feedback, nor was there any other programmed consequence. Te only feedback was the outcome, which followed completion of the block of 12 response trials. If at least one of the responses (unspecifed) had been incorrect (or too late) the participant was returned to another iteration of the study phase. If all the responses had been correct, the participant was moved on to testing. Testing consisted of a sequence of 24 12-trial response phases, with no feedback whatsoever, with no further study phases being given whether or not errors or late responses occurred. A specifc delay was interposed between sample ofset and comparison onset in each test block. Tere were 6 delays, in steps of 400ms from zero to 2s in Group 1 (S1, S14, S2, and S3), and in steps of 600ms from 0-3s in Group 2 (S4-S11), and these delays were given 4 times in alternating descending, ascending, descending and ascending order. (Owing to a programming error the planned second 1600ms delay for the 4 participants in Group 1 was actually set at 2200ms and the data for this delay were omitted from the analyses.)

Sample and comparison stimuli were presented as in Experiment 1. In the study phases on each trial an 'A' sample stimulus was presented for 1s, followed after a zero delay by the presentation for 2s of the correctly corresponding 'B' comparison. In the study phases there were no SAME or DIFFERENT boxes and no opportunity to respond. Comparison ofset was followed by an intertrial interval of 1s.

In the response phases on each trial an 'A' sample stimulus was presented for 200ms. Tis was found to be of sufcient duration to support accurate identity and symbolic DMTS in an earlier study (Dickins, 2003) of the current

hypothesis using a multiple choice DMTS procedure. Ofset of the sample was followed by a delay, which during training, for the 7 participants S1-S6 and S14, was zero, but for the 5 participants S7-S11 was 1500ms. T is was followed immediately by the presentation of a 'B' comparison stimulus. Exposure of the comparison stimulus lasted for 2s, or until a response was made. During this time the word SAME appeared in the bottom left of the screen and the word DIFFERENT in the bottom right, as in Experiment 1. On each trial the participant had to make a SAME or a DIF-FERENT response, depending upon whether sample and comparison were paired as in the preceding study phase (to which the correct response would be to select the 'C' key, SAME), or had been interchanged, (in which case the correct response would be to select the 'M' key, DIFFERENT). For each run of a response phase half of the sample-comparison pairs were chosen at random for mutual rearrangement and the other half were as in the study phase. An intertrial interval of 1s followed.

At the start, before the frst study phase proper, an exemplar study phase of 3 stimulus pairs was given using other shape and non-sense syllable stimuli not used in the main study. Study and response phases with these alternated, with appropriate instructions (see Appendix 1) until the participant got all 3 response trials right, when the training proper began.

Training continued with alternating study and response phases until a response phase was reached in which all 12 responses were cor-rect. A congratulatory instruction screen (see Appendix 1) then introduced the testing part of the program, which consisted of a series of response phases uninterrupted by any further study phases.

In testing, an array of 6 delays was then scheduled for all participants in the same reiter-ated sequence as described above so that each subject experienced 6 x 4 x 12 = 288 test trials in all, half of which on each test block were same-correct and half diferent-correct in ran-dom sequence (Excepting S14 who missed the last 3 blocks and did 252 trials and S11 who missed the last 2 blocks and did 264 trials.)

As in Experiment 1 the program automati-cally recorded the following data for each trial: trial number (1-12); RT; type of trial - C or M (same-correct or diferent-correct); response (c or m); correct or incorrect (or timeout if no response); sample stimulus; comparison stimulus.

Te same methods of analysis were used as in Experiment 1 for the purposes of efective contrast and because samples sizes were again

Table 2: Number of responses correct /12 for successive response phases during symbolic DMTS training in Experiment 2.

| | Response phases -* | | | | | | | | | | | | | | |
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
S4	8	11	10	**12**											
S11	6	11	11	**12**											
S2	7	9	10	11	**12**										
S3	4	9	10	10	**12**										
S8	8	9	9	11	**12**										
S7	8	8	6	8	10	11	**12**								
S5	5	8	9	9	10	11	**12**								
S7	8	8	6	8	10	11	**12**								
S9	8	8	11	10	11	10	**12**								
S10	3	7	5	8	10	11	**12**								
S14	8	4	7	9	10	10	11	11	**12**						
S6	5	7	7	10	11	11	10	10	**12**						
S1	5	7	9	9	9	9	8	7	9	8	7	11	7	11	**12**

Table 3a. Overall accuracy of individual subjects during symbolic DMTS testing in Experiment 2.

Participant	2 correct responses	2 errors	2 timeouts	Σ errors	% accuracy
S1	226	37	13	50	81.9
S2	274	2	0	2	99.3
S3	264	12	0	12	95.7
S4	245	43	0	43	85.1
S5	246	38	4	42	85.4
S6	280	7	1	8	97.2
S7	268	18	2	20	93.1
S8	274	10	4	14	95.1
S9	282	5	1	6	*97.9*
S10	280	6	2	8	97.2
S11	235	28	1	29	89.0
S14	<u>222</u>	17	1	18	92.5
				Mean correct = 92.4% ± 5.8	

small at the intersection of delay, same-correct/diferent-correct, and participant identity (for the 12 participants the parameters were: range 11-24, mean 21.5; SD ± 2.87).

Results and Discussion

Accuracy

All participants attained the training criterion after varying numbers of study phase/response phase cycles (range 4-15 cycles, median 7 cycles: see Table 2).

Tey then carried out the suite of tests with an overall accuracy of 92.4 %, SD ± 5.8. (see Table 3a).

All but two participants (S11 and S14) completed the full set of 24 test blocks and the numbers of errors (incorrect choices + timeouts) per test block for each participant are shown in Table 3b.

Te participants here are arranged in order of the number of training cycles required to reach the training criterion, as in Table 3a. It can be seen that accuracy was sustained at a high level for most subjects irrespective of training history and despite the total lack of any programmed consequences for continuing to respond correctly. In statistical terms, within a sample of one block, binomial signifcance levels are p<0.003 for 11/12 correct, and p<0.019 for 10/12. Out of a total of 283 12-trial blocks, including timeouts, there were only 26 blocks in which there were 3 or more errors. Half the participants (S2, S3, S8, S9, S10, S6) had no such blocks. S1, who was the slowest to reach criterion, was the exception since his performance deteriorated to chance level over most of the last ten test blocks. T is may have been due to fatigue because it had taken him 15 cycles of study and response phases to achieve the training criterion of 12/12 responses correct, twice the median of 7 cycles. Te re is no suggestion, apart from S1, that there is any relationship between the number of training cycles required to reach criterion and subsequent accuracy during testing.

Table 3b. Block-by-block accuracy of individual subjects (errors/12 trials) during symbolic DMTS testing in Experiment 2.

	Decreasing delays						Increasing delays						Decreasing delays						Increasing delays					
	5	4	3	2	1	0	0	1	2	3	4	5	5	4	3	2	1	0	0	1	2	3	4	5
54	3	2	3	2	1	1	4	3	3	1	1	1	1	2	0	1	2	2	1	1	2	3	2	1
S11	1	3	2	0	1	2	1	2	1	2	1	3	1	1	0	0	1	1	0	2	2	2		
52	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
53	2	1	0	0	1	2	0	1	0	0	0	0	0	0	0	2	0	1	0	0	0	1	0	0
58	0	1	2	0	0	1	2	1	0	0	2	1	0	1	0	1	0	0	1	1	0	1	0	1
55	1	2	1	1	2	0	0	2	1	2	0	1	2	1	3	4	3	1	3	1	3	1	1	3
S7	1	0	0	3	1	1	0	0	0	1	0	1	1	1	1	2	1	1	1	2	0	0	1	0	1
59	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0
S10	1	0	0	0	0	0	1	1	1	0	1	1	0	0	0	0	1	0	1	0	1	0	0	0
S14	0	1	2	0	2	1	1	0	3	1	0	2	1	0	1	0	0	0	0	1	2			
56	2	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
S1	0	0	0	0	2	0	0	2	0	0	0	3	0	3	2	5	0	3	4	4	6	6	6	4

Reaction times

Te data were collated on a trial-by-trial basis as in Experiment 1. Figure 3 shows the box plots of RTs for symbolic MTS for each of the 6 values of delay for (a) the 4 participants in Group 1 (delays 0-2s); and (b) the 8 participants in Group 2 (delays 0-3s).

For each of the delay levels the blue box on the left represents RTs of same-correct (C) responses, and the pink box on the right shows the RT distribution of diferent-correct (M) responses. As in Experiment 1 for both groups the medians for same-correct responses can be seen for each delay to be lower than those for diferent-correct responses, but these local differences are not signifcant since the notches in adjacent boxes overlap. For the 8 participants in Group 2 (Figure 3b) both kinds of response medians can be seen to be higher for zero delay than they are for the other 5 delays. Inspection suggests a general tendency for the medians of both kinds of response to become lower with increasing delays up to 2400ms for same-correct and up to 1800ms for diferent-correct. In Fig.3a, the Group 1 participants show a similar fall in median RTs from zero delay to 1200ms, except that the median for same-correct responses at

800ms is higher than that for both 400ms and 1200ms. For longer delays in both groups there is a tendency for median RTs to rise again.

Analysis

Because of the diferent range of delays (0-2s in Group 1 and 0-3s in Group 2 - see above) the data from the 2 groups were frst analysed separately.

Decreases in RT were visible in both groups of participants over the two overlapping delay intervals and signifcant diferences overall were noted in both groups (two-tailed $p < 0.01$ for both; LBL tests). When analysed separately for same-correct versus diferent-correct responses, however, this trend is most evident in same-correct (two-tailed $p < 0.05$ for both groups), but $p < 0.10$ for group 1 and $p > 0.10$ for group 2 for diferent-correct responses).

Since analysing the two groups of participants separately reduced power we undertook an analysis in the two groups combined of zero versus 1200 ms delay, the only 2 delays shared by all participants. T i s analysis yielded a signifcant diference due to delay for both same-correct and diferent-correct responses (one-tailed $p<0.01$ for both).
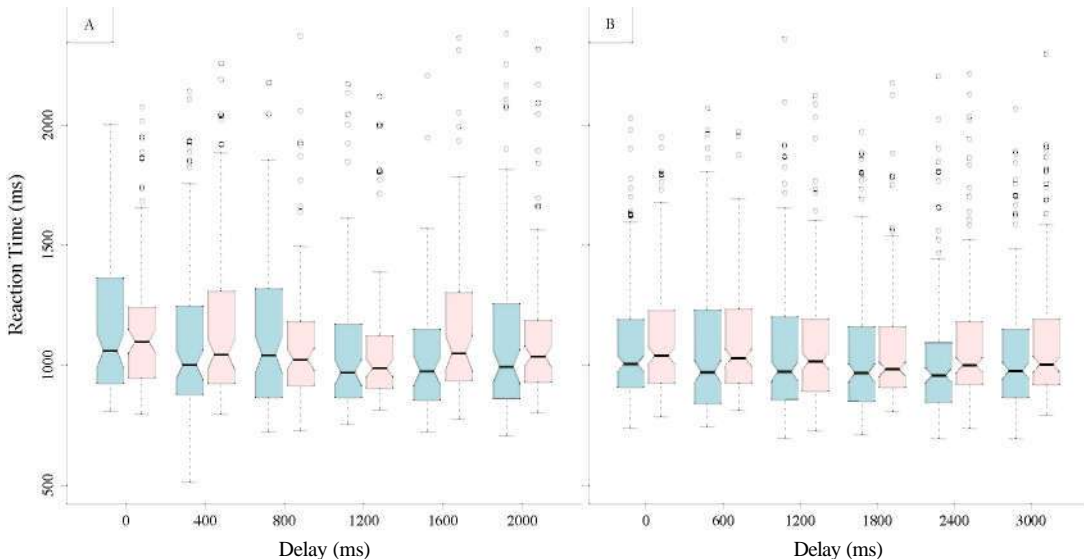

Figure 3. Box plots of RTs for symbolic DMTS for each of the 6 values of delay for (a) the 4 subjects in Group 1 delays 0-2s; and (b) the 8 subjects in Group 2 delays 0-3s. For each of the delay levels the blue plot on the left represents RTs of same-correct (C) responses, and the pink plot on the right shows the RT distribution of diferent-correct (M) responses.

Again the same diference in RTs between same-correct and diferent-correct trials was found: same-correct trials showed lower latencies than diferent-correct trials (one-tailed p < 0.01 for group 1 and p < 0.0001 for group 2; WMW tests). (Tis was confrmed in most participants separately with uncorrected p < 0.05 for S2, S3, S4, S5, S6, S7, S9 and S10).

Te tentative conclusion from Experiment 2 is that though the predicted negative relation between increasing delay and decreasing RT was found as an overall efect, it can be discerned at the individual level only in some participants, and only over delays of up to 1200ms.

The individual differences could be explained post hoc by assuming that only certain participants take advantage of a delay to perform an anticipatory strategy. Perhaps others simply wait to pursue a retrospective analysis once the comparison has appeared or perhaps mixtures of strategies are adopted.

Te delays deployed in these experiments that seemed to be required for the most discernible reductions in RT (c.1200ms) were much longer than the reductions achieved (c.40ms). Te results are insufciently clear to calculate a frm value for the hypothetical optimal delay, defned as the shortest delay associated with the maximum shortening of RT, which, if it could be demonstrated, might indicate how long the hypothesised anticipatory processes took to occur.

EXPERIMENT 3

Te purpose of Experiment 3 was to test the rationale of the preceding experiments by substituting the familiar task of mental arithmetic for the postulated process of prospectively retrieving the correct comparison from prior experimental training in Experiment 2. Within the same SCAR procedure the participant would be faced in some test blocks by a sample stimulus consisting of two (1- or 2-digit) numbers to be added together, to be compared with the subsequently presented putative total (a single number which was either the correct total or not). Such trials were characterized as

'x + y = z' or AS (addition-as-sample). It was predicted that if the participant was given a short delay between sample and comparison this would allow time for the arithmetic computation of 'x + y' and enable a more rapid match or non-match to the putative total 'z' than if there were no delay. As a control participants were also given the converse 'z = x + y' or AC (addition-as-comparison) trials, in which the putative total appeared as sample, and could only be retained in WM until the terms to be added were presented as comparison and the mental addition could begin.

Method

Participants

Nine of the participants, S2, S3, S4, S5, S6, S7, S8, S9, and S10, participated in this experiment.

Stimuli

As shown in Table 4 the stimuli consisted of a series of one- or two-digit numbers, with the sample consisting of two numbers to add, and the comparison their presumptive total (x + y = z, or 'addition-as-sample': 'AS'), or vice versa (z = x + y, or 'addition-as-comparison': 'AC'). Te re was an "easy" condition and a "hard" condition, each consisting of two blocks of twelve AS trials and two blocks of twelve AS trials.

Procedure

Each trial resembled a response trial in symbolic DMTS with presentation and timing exactly as in Experiment 2 except one of the stimuli consisted of 2 numbers to be added together, x + y, and the other was a single number, z, which either was or was not the correct total of x + y. "Same" and "diferent" responses were required for correct and incorrect sums respectively.

In 2 blocks of 12 trials the sample stimulus consisted of the 2 numbers to add and the comparison stimulus was the presumptive total (x + y = z, or addition-as-sample - AS). In the other 2 blocks these relations were inverted (z = x + y, or addition-as- comparison - AC). In one block of each type a 1200 ms delay was interposed

between sample ofset and comparison onset, and in the other there was a zero delay. Te order of blocks was approximately counterbalanced across participants.

Participants S2, S3, S4, S5 & S7 were given relatively "easy" calculations (see Table 4(a)), with 48 single digit numbers and 96 double digit numbers). Participants S6, S8, S9, & 10 were given a diferent set of relatively "hard" calculations in which all the numbers were double digit (Table 4b).

As before the sample exposure time was 200ms, the maximum comparison exposure time and the response opportunity window were 2s simultaneously, and the intertrial interval was 1s.

Results and Discussion

Accuracy

Participants gave the correct answers on either SAME or DIFFERENT trials on an average of 78.7% responses, SD $\pm$ 9.7 (see Table 5a). When the scores of individuals are examined for each type of test (Table 5b), discounting timeouts, performance varies between 'easy' and 'hard', as expected, and between individuals. In a few instances performance is indistinguishable from chance (S5, S8, and particularly S10, but in the AS ('x + y + z') condition with the delay all participants, except S6, perform at their best, and signifcantly above chance. Te overall percentage accuracy for each of the four conditions shown in Table 5b places them, in decreasing levels of accuracy, in the order:
[AS+ delay] > [AC– delay]   [AC+ delay] > [AS – delay]

Reaction times

Figure 4 shows the box plots of RTs for the arithmetic task for zero versus 1200ms delay for AC (addition as comparison, z = x + y) on the left and AS (addition as sample, x + y = z) on the right for (a) all 9 participants ('hard' and 'easy' combined); (b) the 5 'easy' participants (c) the 4 'hard' participants. For each of the delay levels the blue box on the left represents RTs of same-correct (C) responses, and the pink box on the right shows the RT distribution of diferent-

Table 4. Te 4 blocks of 12 arithmetic trials in Experiment 3: same-correct trials in blue, diferent-correct in red.

(a): the 'Easy' condition

"EASY"

z = x + y	x + y = z	z = x + y	x + y = z
5=2+3	9+5=14	17=8+9	18+20=38
15=6+9	19+7=26	21=4+17	22+24=44
8=3+5	10+12=22	30=13+17	25+18=43
44=27+17	8+3=11	16=9+7	13+15=28
32=12+20	11+16=27	36=19+17	4+3=7
25=8+17	29+16=45	13=4+9	16+25=41
6=6+4	23+25=23	39=27+20	13+21=42
33=13+24	15+16=35	47=13+11	23+19=18
29=5+1	9+14=3	24=11+8	12+6=20
10=2+7	7+5=31	15=23+17	3+1=16
37=18+11	17+18=48	19=6+9	14+6=34
9=14+19	1+2=12	40=21+18	9+7=4

(b): the 'Hard' condition

"HARD"

z = x + y	x + y = z	z = x + y	x + y = z
48=22+26	35+39=74	29=16+13	49+32=81
49=17+32	18+69=87	31=16+15	28+37=65
58=33+25	51+31=82	34=15+19	36+41=77
53=27+26	46+42=88	40=23+17	35+22=57
43=12+31	35+44=79	30=13+17	39+32=20
51=25+26	38+37=75	37=25+12	41+26=67
55=21+33	22+47=78	33=13+23	33+29=64
54=18+29	29+39=86	35=18+21	28+31=63
47=24+37	58+28=69	39=14+18	29+34=62
52=14+36	45+48=85	38=19+16	48+24=73
50=29+26	47+38=68	36=11+27	26+47=59
61=13+39	25+53=93	32=17+16	27+37=72

correct (M) responses. Te overall median RTs for the 4 conditions (Figure 4a) places them, for both same-correct and diferent-correct trials, in the increasing order:
[AS+ delay] < [AS – delay] < [AC– delay] < [AC+ delay]

Analysis

Again a similar analysis was performed to that carried out for Experiments 1 and 2, for

similar reasons: in this case the intersections had the following parameters: Range 1-6; Mean 4.375, SD= 1.280. Te efects of delay on RT were found to be in opposite directions for the AC and the AS conditions. For AC (z = x + y): there was a small but non-signifcant increase in RT with the delay (p>0.1 both for same-correct and diferent-correct combined, and for each type of trial separately). For AS (x +y = z) there was a signifcant decrease in RT with the delay: (p<0.0001 for same-correct and diferent-correct together; p<0.0001 for same-correct responses; and p<0.01 for diferent-correct responses).

In all 9 participants the RT for the x + y = z condition is faster after the 1200ms delay than after a zero delay, with savings on average of 300ms. Tese results are consonant with the intention of the experiment which was to give participants, by means of the delay, extra time to perform the addition of x and y before checking the presented value of z to see whether or not it matched the calculated sum. Te highly signifcant efect of the delay in these AS trials, especially for same-correct responses, contrasts strongly with the slight efect in the opposite direction on AC trials. Participants have to register and remember 2 numbers (x and y) long enough to be able to add them together, hold the result of this calculation, and also to register and hold in memory the putative total (z) to see whether it is the same or a diferent number, and then

respond accordingly. Depending on the speed of a participant's mental arithmetic (and the difculty of the sum in each trial) the expected outcome was that if x and y are presented frst, and a delay is provided, the addition could be completed and the calculated sum held ready. T is would facilitate the subsequent comparison with the putative total z. Te addition process was used as an analogue of the postulated retrieval of the trained relation between sample and the matching comparison in Experiment 2.

With regard to accuracy, AS trials with a delay are markedly superior to the 3 other conditions, though AS without a delay showed the least accurate performance. In the AC condition the participant can work out the two sides of the sum 'at leisure', since the components to be added are longer on the screen. Note though that the RTs of correct AS responses without a delay are faster, probably because the addition can still be started earlier than in AC, during the brief sample period.

GENERAL DISCUSSION

Experiment 1 was designed as a control to show that with simple Identity matching, in which the sample itself may or may not be the

Table 5a. Overall accuracy on arithmetic test trials in Experiment 3.

Participant	Right	Wrong	Timeout	Σ error	% Accuracy
'Easy'					
S3	43	3	2	5	93.5
	43	4	1	5	91.5
S4	38	8	2	10	82.6
S5	31	13	4	17	70.4
S7	36	10	2	12	78.3
S6	31	10	7	17	75.6
S8	32	11	5	16	74.4
S9	34	9	5	14	79.1
S10	27	16	5	21	62.8
				SD	± 9.7
				mean	78.7%

Table 5b. Accuracy per test type on arithmetic test trials in Experiment 3.

Participant	x + y = z		z = x + y	
	Zero	1200ms	Zero	1200ms
'Easy'				
S3	11/12	12/12	9/12	11/11
S4	9/12	10/11	10/11	9/12
S5	8/12	10/12	8/11	5/9
S7	7/12	12/12	7/10	10/12
'Hard'				
S6	7/11	7/11	8/10	9/9
S8	8/11	10/12	6/8	8/12
S9	9/11	9/10	9/11	7/11
S10	5/10	11/12	5/10	4/9
Overall ratio	74/103	93/104	72/93	74/97
Overall %	71.8%	89.4%	77.4%	76.3%

same as the comparison, no benefit in terms of shorter RT results from a short delay, since no intervening stimulus need be retrieved, nor could it be, since none had at this stage been paired with the sample. We found if anything a positive relation between increasing delays and increased RT, though these efects did not quite reach statistical signifcance.

In Experiment 2 it was predicted that, after training on AB relations, participants might deploy a prospective or anticipatory strategy: as soon as they had registered the 'A' sample, they would retrieve the appropriate 'B' stimulus, and this 'B' could in turn be in efect identity-

matched with the forthcoming 'B' comparison. A sufcient delay would provide time for this anticipatory process to be completed. The subsequent RT would be no longer than an RT for identity matching, with or without a delay. Alternatively participants might process the relation between sample and comparison retrospectively, by waiting until the comparison appeared, and then retrieving for this actual 'B' comparison (by a 'derived' symmetrical relation) the 'A' with which it had been paired in training, and then matching this 'A' with the remembered 'A' sample. Here there would be no benefit from a delay, and no inverse relation
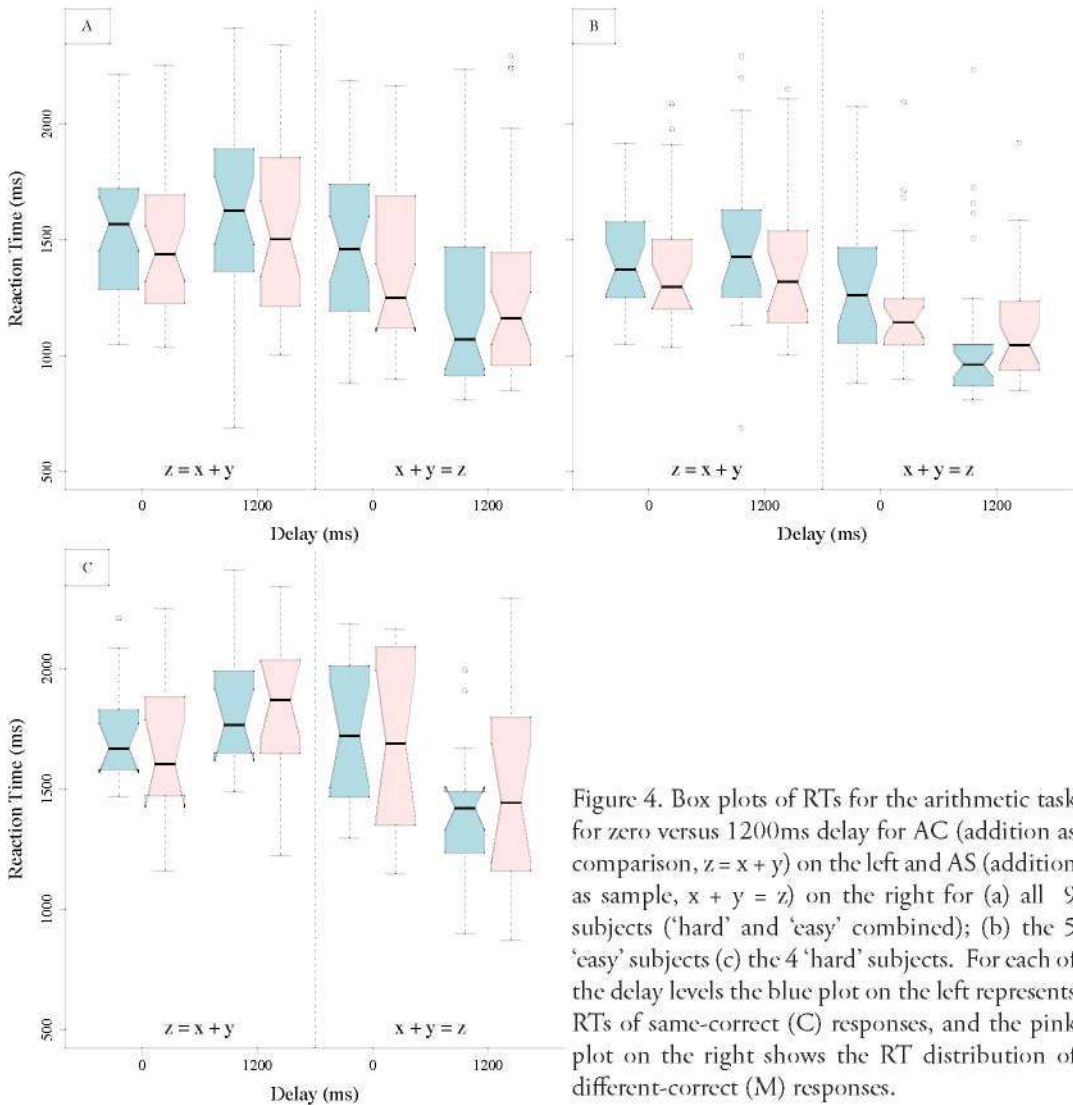

Figure 4. Box plots of RTs for the arithmetic task for zero versus 1200ms delay for AC (addition as comparison, z = x + y) on the left and AS (addition as sample, x + y = z) on the right for (a) all 9 subjects ('hard' and 'easy' combined); (b) the 5 'easy' subjects (c) the 4 'hard' subjects. For each of the delay levels the blue plot on the left represents RTs of same-correct (C) responses, and the pink plot on the right shows the RT distribution of different-correct (M) responses.

between delay and RT would be expected.

Using the non-parametric Jonckheere-Terpstra trend test we found in Experiment 2 an overall efect of delay of the kind indicative of anticipatory strategies. A signifcant negative relation between delay and RT was found for C (same-correct) responses and M (diferent-correct) responses combined, and for C (same-correct) responses separately, but not for M (diferent-correct) responses on their own. It should be noted that such a signifcant efect of delay could not be demonstrated in parametric tests when individual diferences were given due consideration. Te application of parametric tests to RT data is not straightforward however because of problems with the distribution of RTs (Whelan, 2008). Apart from this a tentative post hoc explanation of these results may be: (1) that an anticipatory strategy was only adopted by some participants; and (2) that M non-match responses entailed, perhaps in all participants, some extra, retrospective processing compared with the direct recognition of the matching comparison in C same-correct trials. Such a diference in RT between match and non-match trials has been found by other authors, e.g by Sternberg (1966), and for arithmetic, with neural concomitants, by (Menon, Adleman, White, Glover, & Reiss, 2001; Menon, Mackenzie, Rivera, & Reiss, 2002). A third factor (3) may have been increased noise in the RT data with longer delays, perhaps due to attenuation of WM. Inspection of the data suggested that the predicted efect was most discernible over the shortest 2 or 3 delays. When RTs solely for the delay of 1200ms were compared with those for zero delay over all 12 participants, all of whom had been subjected to these 2 delay values, a robust efect of delay in shortening RTs was found.

Experiment 3 was designed as a mathematical simulation of an opportunity for prospective processing in 'x + y = z' or 'addition-as-sample' (AS) trials contrasted with the obligation to perform retrospective processing in 'z = x + y' or 'addition-as-comparison' (AC) trials. In AS trials, after a 1200ms delay as compared with a zero delay, RTs were faster in all participants, especially on same-correct trials, whereas on AC trials RTs went up slightly after a delay for both same-correct and diferent-correct trials. Te results on this arithmetic task indicate the universal adoption by participants of a prospective strategy when the opportunity arose.

Te delays deployed in these experiments, which thus appear to have brought about a discernible reduction in RT, were longer than the reductions achieved. Te results are insuffciently clear to calculate a frm value for the hypothetical optimal delay, which, if it could be demonstrated, might indicate how long the hypothesised anticipatory processes took to occur. Approximately however, in Experiment 2, the saving is in the order of 40ms, an order of magnitude less than the delays, which were in hundreds of milliseconds. T i s value accords with a classic fnding (Sternberg, 1966) with regard to the time taken to retrieve a single item from a remembered list of stimuli. In one of Sternberg's experiments resembling DMTS a series of digits were presented (corresponding to a sequence of several sample stimuli) followed by a single digit as comparison. T i s digit did or did not appear in the sample list, and participants responded accordingly. For correct responses (the 8 participants averaged an accuracy of 98.7%) the relation between RT and the number of sample digits was linear, with a slope of 37.9 ± 3.8 ms, with a zero intercept of 397.2 ± 19.3 ms. Interestingly RT was not related to the position in the sample list of a matching digit: this suggests that as in a list with no matches participants searched the entire list on each trial. Subsequent research (McElree & Dosher, 1989) questioned the linear relationship indicative of purely serial processing through the list, suggesting parallel processing for all but the most recent item. Here we are concerned with only one item, and if an anticipatory process is making a contribution, linear processing is imposed by the sequential stimulus presentation. An alternative possibility is that under some circumstances processing might only begin, once the comparison has appeared, and might combine two processes, (a) prospective retrieval of the matching comparison from the sample and (b) retrospective determination of the corresponding or non-corresponding

sample from the comparison. Tese notional processes might run in parallel, though there is a suggestion, as in the arithmetic example of Menon et al (2002), that only in the case of a non-match resulting from process (a) will process (b) be instigated as a double check, which would explain the longer RTs often found, as here, for accurate diferent-correct compared with same-correct responses.

We have noted above in justifcation of the statistical tests used here that the application of parametric tests to RT data is not straightforward however because of problems with the distribution of RTs (Whelan, 2008). A more powerful way to construe such a parameter might be (a) to train individual participants over a long period until they reached an asymptote of short RTs in the manner of Tomonari and his colleagues and E. Arntzen personal communication, and then (b) to titrate the delay so that with a run of reduced RTs the program will shorten the delay, and with longer RTs it lengthens the delay, until an optimal delay is attained which is associated with the fastest RTs. Such manipulations may preclude difculties arising owing to the imprecision and complexity of RT distributions in the relatively unconstrained conditions of the present type of experiment.

Overall the data support the initial hypothesis and ofer a way forward for the analysis of component processes in matching-to-sample procedures. Whether such processes are conceptualized at the level of functional analysis, or in terms of computer-like operations that seem logically required to efect the behaviours concerned, the implication is that there are concomitant physiological events in the brain which constitute the proximate mechanism of how the brain 'does' behaviour, and which may increasingly become the focussed target of imaging and similar studies (Timberlake, Schaal, & Steinmetz, 2005).

In previous studies (Bentall, Dickins, & Fox, 1993; Bentall, Jones, & Dickins, 1999; Spencer & Chase, 1996) RTs have been used in the study of derived responses in stimulus equivalence. While RTs in these tests diminished with repeated testing, there was an enduring relative increase in RT associated with transitive relations (and increasing with increasing nodal distance), compared with directly trained relations or symmetry. Te frst author has also reported elsewhere (Dickins, 2003) results from a similar study to the present one, but using a multiple-choice DMTS procedure, in which ANOVAs revealed signifcant efects of delays on reducing RTs in tests not only of trained relations (AB and BC) but also of the derived relations of symmetry (BA and CB tests), transitivity (AC), and equivalence (CA) tests. Here the results indicated a longer anticipatory process in transitivity and equivalence. Tere can be little doubt that in transitive responding, especially when given a consistent sequence of transitivity trials, that a participant retrieves the (absent) nodal stimulus 'B' in the course of organizing the response to the comparison. Descriptions of such anticipatory processes in terms of 'covert' behaviours, such as 'naming' (Horne & Lowe, 1996) are no less speculative than the postulated 'processes' in studies of the present kind, and are likely to be partly confrmed and partly superseded by studies of RTs coupled with manipulation of temporal parameters and types of stimuli in behavioural experiments, and parallel investigations of concomitant physiological events in the brain (Barnes-Holmes, Regan et al., 2005; Barnes-Holmes, Staunton et al., 2005, Dickins et al, 2001, Timberlake et al, 2005).

REFERENCES

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

Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829-839.

Barnes-Holmes, D., Regan, D., Barnes-Holmes, Y., Commins, S., Walsh, D., Stewart, I., Smeets, P.M., Whelan, R., & Dymond, S. (2005a). Relating Derived Relations as a Model of Analogical Reasoning: Reaction Times and Event-Related Potentials. *Journal of the Experimental Analysis of Behavior, 84*, 435–451.

Barnes-Holmes, D., Staunton, C., Whelan, R., Barnes-Holmes, Y., Commins, S., Walsh, D., Stewart, I., Smeets, P.M., & Dymond, S.

(2005b). Derived stimulus relations, semantic priming, and event-related potentials: testing a behavioral theory of semantic networks. *Journal of the Experimental Analysis of Behavior, 84*, 417-433.

Bentall, R.P., Dickins, D.W., & Fox, S.R.A. (1993). Naming and equivalence: Response latencies for emergent relations. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 46B*, 187-214.

Bentall, R.P., Jones, R.M. & Dickins, D.W. (1999). Control over emergent relations during the formation of equivalence classes: Response error and latency data for 5-member classes. *Te Psychological Record, 49*, 93-116.

Broberg, P. (2009). SAGx: Statistical Analysis of the GeneChip. R package version 1.18.0. from http://bioconductor.org/packages/release/bioc/html/SAGx.html

Dickins, D.W. (2003). *Te efects of a non-zero delay on response times in matching-to-sample (MTS) tasks in a stimulus equivalence (SE) paradigm.* Paper presented at the European Association for Behaviour Analysis, Parma, Italy, July 2003.

Dickins, D.W. (2005). On aims and methods in the neuroimaging of derived relations. *Journal of the Experimental Analysis of Behavior, 84*(3: Special Issue on Te Relation of Behavior and Neuroscience), 453-483.

Dickins, D.W., Singh, K.D., Roberts, N., Burns, P. , Downes, J.J., Jimmieson, P. , & Bentall, R.P. (2001). An fMRI study of stimulus equivalence. *Neuroreport, 12*(2), 405-411.

Funahashi, S., Bruce, C.J., & Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61*, 331-349.

Fuster, J.M. (1973). Unit activity in prefrontal cortex during delayed response performance: neuronal correlates of transient memory. *Journal of Neurophysiology, 36*, 61-78.

Grant, D.S., & Kelly, R. (2001). Anticipation and short-term retention in pigeons. In R. G. Cook (Ed.), *Avian visual cognition*:

[On-line]. Available: http://www.pigeon.psy.tufts.edu/avc/.

Horne, P.J., & Lowe, C.F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior, 65*, 183-353.

Hothorn, T., Hornik, K., van de Wiel, M.A., & Zeileis, A. (2006). A Lego System for Conditional Inference. *Te American Statistician, 60*(3), 257-263.

Hothorn, T., Hornik, K., van de Wiel, M.A., & Zeileis, A. (2008). Implementing a Class of Permutation Tests: Te coin Package. *Journal of Statistical Software, 28*(8), 1-23.

Liebe, S., Logothetis, N.K., & Rainer, G. (2008). Prediction of Behavioral Choice and Reaction Time from Local Field Potential in Macaque Prefrontal Cortex. *AREADNE 2008: Research in Encoding and Decoding of Neural Ensembles 2, 70 (06 2008)*.

McElree, B., & Dosher, B.A. (1989). Serial position and set size in short-term memory: time course of recognition. *Journal of Experimental Psychology: General, 118*, 346-373.

Menon, V., Adleman, N.E., White, C.D., Glover, G.H., & Reiss, A.L. (2001). Error-Related Brain Activation during a Go/NoGo Response Inhibition Task. *Human Brain Mapping, 12*, 131-143.

Menon, V., Mackenzie, K., Rivera, S.M., & Reiss, A.L. (2002). Prefrontal cortex involvement in processing incorrect arithmetic equations: evidence from event-related fMRI. *Human Brain Mapping, 16*, 119-130.

Nevin, J.A., Davison, M., Odum, A.L., & Shahan, T.A. (2007). A Teory of Attending, Remembering, and Reinforcement in Delayed Matching to Sample. *Journal of the Experimental Analysis of Behavior, 88*(2), 285–317.

Posner, M.I. (1978). *Chronometric Explorations of Mind*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Posner, M.I. (2005). Timing the brain: Mental chronometry as a tool in neuroscience. *PLoS Biol, 3*(2), e51.

RDevelopmentCoreTeam. (2009). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation

for Statistical Computing, http://www.R-project.org.

Roberts, W.A. (1972). Short term memory in the pigeon: Efects of repetition and spacing. *Journal of Experimental Psychology, 94*, 74-83.

Roberts, W.A., & Grant, D.S. (1974). Short term memory in the pigeon with presentation time precisely controlled. *Learning and Motivation, 5*, 393-408.

Roitblat, H.L., & von Fersen, L. (1992). Comparative cognition: Representations and processes in learning and memory. *Annual Review of Psychology, 43*, 671-710.

Runtime_Revolution. from www.runrev.com

Spencer, T.J., & Chase, P.N. (1996). Speed analyses of stimulus equivalence. *J Exp Anal Behav, 65*(3), 643-659.

Sternberg, S. (1966). High-Speed Scanning in Human Memory. *Science, 153*, 652-653.

Sternberg, S. (1969a). Te discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276-273 I275.

Sternberg, S. (1969b). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*, 421-457.

Sternberg, S. (2001). Separate modifability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychologica, 106*, 147-246.

Timberlake, W., Schaal, D.W., & Steinmetz, J.E. (2005). Special Issue on the Relation of Behavior and Neuroscience. *Journal of the Experimental Analysis of Behavior, 84* (November), 305-692.

Tomanari, G.Y., Sidman, M., Rubio, A.R., & Dube, W.V. (2006). Equivalence classes with requirements for short response latencies. *Journal of the Experimental Analysis of Behavior, 85*, 349–369.

Weavers, R., Foster, T.M., & Temple, W. (1998). Reinforcer efficacy in a delayed matching-to-sample task. *Journal of the Experimental Analysis of Behavior, 69*, 77-85.

Whelan, R. (2008). Efective analysis of reaction time data. *Te Psychological Record, 58*, 475–482.

White, K.G. (1985). Characteristics of forgetting functions in delayed matching to sample. *Journal of the Experimental Analysis of Behavior, 44*, 15-34.

Williams, D.C., Johnston, M.D., & Saunders, K.J. (2006). Intertial Sources of Stimulus Control and Delayed Matching-to-sample Performance in Humans. *Journal of the Experimental Analysis of Behavior, 86*(2), 253–267.

Wixted, J.T. (1989). Nonhuman short-term memory: A quantitative reanalysis of selected fndings. *Journal of the Experimental Analysis of Behavior, 52*, 409-426.

Wright, A. (2001). Learning strategies in matching to sample. In R. G. Cook (Ed.), *Avian Visual Cognition*. www.pigeon.psy.tufts.edu/avc/wright: On-line.

Zentall, T.R. (2001). Te case for a cognitive approach to animal learning and behavior. *Behavioural Processes, 54*, 65-78.

Appendix 1.1: Instructions Experiment 1

This should be an easy task. If the first stimulus, called the sample, and the second stimulus, the comparison, are identical, as they will be on about half the trials, select the C key for SAME.
If they are not identical, select M, DIFFERENT.
Try to respond both as accurately and as quickly as you can. Your responses will be recorded but you will not be given trial-by-trial feedback.

Congratulations. You have reached the end of the experiment.
The experimenter may wish to ask you a few questions before you leave.
Thank you very much for your participation.

Appendix 1.2: Instructions Experiment 2

You start with a practise STUDY PHASE
You will see a series of pairs of items (stimuli) to learn and memorize. First you will see a SAMPLE STIMULUS on the left. Then a COMPARISON STIMULUS will appear on the right. Here is a simple example to give you the idea.
Study the following three pairs.

Now follows a "RESPONSE PHASE".
You will be presented with 3 sample-comparison TEST TRIALS.
If the stimuli are paired as they were in the STUDY PHASE press the C key: SAME.
If the stimuli have been rearranged press the M key: DIFFERENT.
If you press the wrong key, or are too late, on one or more of the trials, you will be returned to the Study Phase.
If you get them all right, you will continue to the start of the actual experiment.

You are back in the STUDY PHASE.
You must have made at least one incorrect choice or failed to respond in time.
Study the 3 practise pairs again, and then another RESPONSE PHASE will follow.

Well done: you understand the task.
Now the experiment proper starts with the first STUDY PHASE with 12 sample-comparison pairs to learn.

STUDY PHASE 1
Memorize the following pairs.
Do not press any keys during this phase.

RESPONSE PHASE 1
You will now be presented with some test pairs.
If the stimuli are paired as they were in the preceding Study phase press the C key: SAME.
If the stimuli have been rearranged press the M key: DIFFERENT.
If you press the wrong key, or are too late, on one or more of the trials you will be returned to the Study Phase.
If you get them all right, you will continue to a new, second Study Phase

---

Well done: you have successfully completed the training phase.
Now follows a series of further Response Phases with no returns to the Study Phase.
Please do your best to respond accurately and quickly.

Good.
Move on when you're ready.

Well done.
You have reached the end of the experiment.
Thank you very much for your participation.

Appendix 1.3: Instructions Experiment 3

MATHS TEST: SOME SIMPLE ARITHMETIC
Here one stimulus will represent the addition of two numbers, such as 3 + 2.
The other stimulus will be either the correct sum of these numbers, in this case 5, in which case
select SAME, or it will be incorrect, such as 6, in which case select DIFFERENT.
Again please try to get them all right as fast as you can.

Congratulations. You have reached the end of the experiment.
The experimenter may wish to ask you a few questions before you leave.
Thank you very much for your participation.