

Animal memory: A review of delayed matching-to-sample data

Johan Lind ^{a,b*}, Magnus Enquist ^{a,b}, Stefano Ghirlanda ^{a,c}

November 10, 2014

^a Center for the Study of Cultural Evolution, Stockholm University, S-106 91, Stockholm, Sweden, *corresponding author

^b Department of Zoology, Stockholm University, S-106 91, Stockholm, Sweden

^c Department of Psychology, Brooklyn College, Brooklyn, NY 11210, USA

Abstract

We performed a meta-analysis of over 90 data sets from delayed matching-to-sample (DMTS) studies with 25 species (birds, mammals, and bees). In DMTS, a sample stimulus is first presented and then removed. After a delay, two (or more) comparison stimuli are presented, and the subject is rewarded for choosing the one matching the sample. We used data on performance vs. delay length to estimate two parameters informative of working memory abilities: the maximum performance possible with no delay (comparison stimuli presented as soon as the sample is removed), and the rate of performance decay as the delay is lengthened (related to memory span). We conclude that there is little evidence that zero-delay performance varies between these species. There is evidence that pigeons do not perform as well as mammals at longer delay intervals. Pigeons, however, are the only extensively studied bird, and we cannot exclude that other birds may be able to bridge as long a delay as mammals. Extensive training may improve memory, although the data are open to other interpretations. Overall, DMTS studies suggest memory spans ranging from a few seconds to several minutes. We suggest that observations of animals exhibiting much longer memory spans (days to months) can be explained in terms of specialized memory systems that deal with specific, biologically significant information, such as food caches. Events that do not trigger these systems, on the other hand, appear to be remembered for only a short time.

1 Introduction

Animal memory is the object of enduring fascination and debate (Suddendorf and Busby 2003; Raby et al. 2007; Clayton and Dickinson 1998; Roberts 2002; Bouton 2007; Pearce 2013; Kendrick et al. 1986). Even a cursory look at the literature reveals that sometimes animals appear to remember events in detail, and for a long time, while sometimes they seem to forget surprisingly fast. For example, many jays (family Corvidae) can remember the location of hidden food for many months (Bossema 1979; Bednekoff et al. 1997), but have difficulty remembering simple color stimuli for more than 25–40 s (Olson et al. 1995). This is just one example of the great variation in memory performance that has lead scholars to widely different conclusions. Some maintain that non-human animals have no explicit memories of past events, that is, they cannot recollect them in the same way as we do when we think about, say, childhood vacations (Roberts 2002; Suddendorf and Busby 2003). Others, in contrast, conclude that we have evidence that animals have “episodic-like” memories perhaps not dissimilar from those of humans (Clayton et al. 2001a,b; Zentall 2005). Comparison with human memory is central to the debate about non-human memory. Humans, in fact, appear capable of remembering practically arbitrary events for very long times without appreciable decline in performance, notably in tasks in which most other species perform poorly (Overman and Doty 1980).

Here we review well-controlled experiments using the delayed matching-to-sample paradigm (DMTS, summarized below), one of a number of tasks that has been used to probe animals’ working memory (Bouton 2007; Pearce 2013). By “working memory” we mean information about an event that is maintained for some time in the absence of reinforcement, and that can be used to guide behavior at a later time (Pribram et al. 1960; Baddeley 1991). We

are particularly interested in events without specific biological significance, such as colored lights, visual patterns, or sounds, as typically used in laboratory studies. We ask whether we can discern any species differences in working memory, and whether memory can be improved by training. We conclude by suggesting that animal memory is best characterized as the sum of specialized memory capacities, different across species and potentially lasting a long time, and a general memory capacity that works in essentially the same way in all non-human species, and which has a limited span ranging between a few seconds and a few minutes.

2 Methods

2.1 The delayed matching-to-sample (DMTS) paradigm

In a typical DMTS experiment a *sample* stimulus is presented for a short time, typically a few seconds. The sample is then removed and, after a delay, two *comparison* stimuli are presented, of which one is identical to the sample and the other is different. The animal is rewarded for choosing the comparison stimulus that matches the sample (several procedural variations are possible, such as requiring the animal to perform an action to remove the sample and initiate the delay). This paradigm has been used for many decades, for example to study whether animals possess conceptual or “symbolic” abilities (Finch 1942; Yerkes and Nissen 1939; Weinstein 1941; Giurfa et al. 2001), or to study the effect of drugs or brain lesions on memory (Dunnett and Martel 1990; Stanhope et al. 1995; Sloan et al. 2006; Horel et al. 1984, and many others). Here we are interested in DMTS as a probe into animals’ capacity to retain information about arbitrary events.

2.2 Estimation of memory parameters

We acquired over 90 data sets from published DMTS studies conducted with 25 species, each study using one or more delay intervals. Data sources are given in the Appendix. Performance in DMTS is customarily measured in terms of % correct trials, i.e., % trials in which the animal chose the comparison stimulus matching the sample. Thus 50% represents chance performance and 100% perfect performance. We used DMTS data to estimate two performance parameters:

Zero-delay performance: Performance when comparison stimuli are presented immediately after the sample is removed. This condition imposes minimum requirements on memory.

Performance half-life: The delay for which performance has fallen halfway between zero-delay performance and chance performance (50% correct).

To estimate these parameters, we first transformed the percentage of correct responses into a measure of performance ranging between 0 and 1:

$$y = \frac{\% \text{ correct} - 50}{50} \quad (1)$$

We then fitted an exponential function to each data set:

$$y(t) = M \exp(-t/\tau) \quad (2)$$

where t is the delay interval and M and τ are the fitted parameters. According to equation (2), performance at zero delay equals $50(1+M)$, while its half-life equals $\tau \ln 2$. Thus fitting equation (2) to data enables us to estimate these parameters even if performance at the exact delays that define them (0 and $\tau \ln 2$) was not actually assessed. We used a simple exponential function for three reasons. First, it fits the data remarkably well. Across all experiments, the mean difference between observed and predicted performance is 5% per data point (median 4%, range 0–17%), with no detectable difference across species ($\chi^2_{16} = 10.06$, $p = 0.86$, Kruskal-Wallis test relating mean squared error in fit to species). Second, equation (2) can be recast as a linear relationship by taking logarithms:

$$\ln y(t) = \ln M - \frac{t}{\tau} \quad (3)$$

Linear fits are numerically more robust than nonlinear fits (Björck 1996). Lastly, while other functions have been used to fit working memory data (Rubin and Wenzel 1996; Wickens 1998), these have additional parameters and cannot be linearized, which makes them less suitable to fit data sets with few observations, typical of DMTS studies.

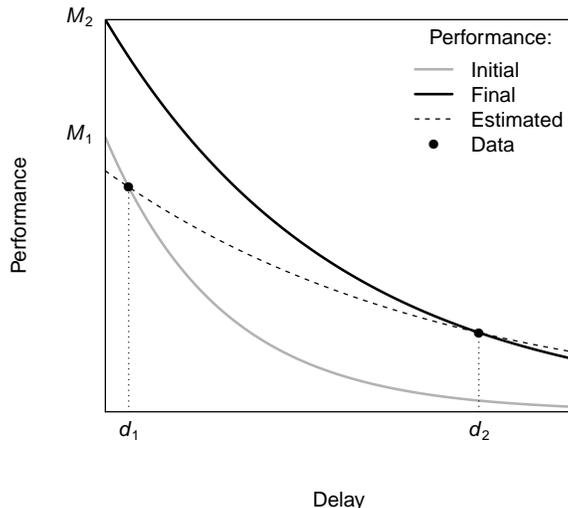


Figure 1: Estimation of memory parameters from delay titration studies, taking into account that, over the course of training, the initial performance function (gray line) may be lower, and may decay faster than the final performance function (black line, see Kangas et al. 2011, for empirical data). We would like to estimate the decay rate of final performance, but available data are only the *initial* performance at a short delay d_1 (black dot on gray line) and the *final* performance at a longer delay d_2 (black dot on solid black line). Using these data to fit the performance function 2 leads to the estimation that performance follows the dotted line. This underestimates the true zero-delay final performance (point M_2) and overestimates the final performance half-life (the dotted line representing estimated performance decays more slowly than black line representing final performance).

In fitting equation (3) to data, we encountered a few special circumstances. First, three studies observed performance slightly below 50% (in figure 2: bees 4, pigeons 5, macaques 3). We treated these observations as measurement error and considered all values ≤ 50 as equal to 51. We choose 51 rather than 50 because 50 would result in $\ln y(t) = -\infty$, which would prevent fitting. This choice very slightly overestimates DMTS performance at long delays. Second, some studies used only one delay interval, preventing estimation of performance half-life. For these studies, we used observed performance to estimate zero-delay performance, provided the delay interval was ≤ 1 s. Third, fitting equation (3) estimates an infinite half-life when performance over the probed range of delays is stable or increasing. This occurred in a black-capped chickadee study (maximum delay = 20s, birds 1 in figure 2), one dolphin study (maximum delay = 60s, dolphins 4), and two chimpanzee studies (maximum delay = 20s in both, chimpanzees 1 and 2). While the claim of infinite memory is clearly untenable, simply ignoring these studies would bias our estimates toward shorter memory spans. As a compromise, we estimated performance half-life as twice the longest probed interval. This choice is arbitrary yet appears generous in light of other data with the same or similar species.

2.3 Delay titration studies

While most studies reviewed below presented a set of delay intervals determined by the experimenter, some studies used a subject-driven titration procedure in which the delay interval is lengthened whenever a subject meets a pre-determined response criterion (e.g., two correct responses in a row). If the criterion is not met, the delay interval is shortened. These studies, although a minority, deserve special attention as they sometimes report striking long-delay performance. We detail here how we estimated memory parameters from titration data, and postpone to the Discussion their evaluation. Data from titration studies are not reported as delay-performance curves because performance is kept at criterion by adjusting the delay. Rather, the maximum delay achieved under a given performance criterion is reported. It is possible to estimate the M and τ parameters in equation (2) from such data exploiting the fact that,

typically, animals are first trained with a small or zero delay until a performance criterion is met that is more stringent than what is maintained during titration. For example, Kangas et al. (2010) trained pigeons to 85% correct at zero delay, and maintained them at 67% correct during titration. Let (d_1, y_1) be the initial delay and performance, and (d_2, y_2) the delay and performance at the end of titration. According to equation (2), we have:

$$y_i = M \exp(-d_i/\tau) \quad i = 1, 2$$

or, according to equation (3):

$$\ln y_i = \ln M - \frac{d_i}{\tau} \quad i = 1, 2$$

This is a system of two equations that can be solved exactly to yield

$$\begin{aligned} M &= \exp\left(\frac{d_2 \ln y_1 - d_1 \ln y_2}{d_2 - d_1}\right) \\ \tau &= \frac{d_2 - d_1}{\ln y_1 - \ln y_2} \end{aligned} \tag{4}$$

thus providing the desired estimates of M and τ . It should be borne in mind that these formulae assume that training does not change τ and M , a question we will return to when discussing whether working memory can improve by training. If memory improves through training, equation (4) underestimates M and overestimates τ , as shown in figure 1. We will see below that underestimation of M is not a specific feature of titration data, but rather a common occurrence. We will keep in mind that performance half-life may be overestimated from titration data, and we will highlight which data come from such studies.

2.4 Software and statistics

Data were analyzed with R, version 3.0.0 (Core Team 2013). The `nls` function was used to fit equation (3) to data while enforcing the constraints $0 \leq M \leq 1$ and $\tau \geq 0$. Functions `kruskal_test` and `spearman_test` from the `coin` package were used to perform Kruskal-Wallis tests and assess the significance of Spearman correlations.

3 Results

All data reviewed are shown in figure 2, and figure 3 shows our estimates of zero-delay performance and performance half-life for the featured species. Our unit of analysis is a “data set,” i.e., data collected from the same individuals under the same conditions. Note that several data sets may come from the same individuals, tested more than once under different conditions. Dolphin data, for example, come from just two individuals, and capuchin data from four. In both cases, the animals were tested over several years in the same laboratory. We display these data separately because many variables are known to affect DMTS performance, so that even data from the same individuals can exhibit substantial variation. For example, the three capuchin monkeys tested by D’amato and O’Neill (1971) performed worse under moderate illumination than in the dark (Capuchins 1 and 2 in figure 2). Even when data are available from a large number of naïve animals tested independently, such as for rats, rhesus macaques, and pigeons, it is hard to disentangle procedural differences from species differences. For example, pigeons have performed, overall, worse than rats (figure 3, left), but they have also been tested under more challenging conditions. The lowest pigeon scores in figure 3, left, for example, are 77% and 82% correct, and come from a study by Roberts (1980) in which trials were presented only 1 s apart. Longer intertrial intervals are known to be beneficial for memory (Nelson and Wasserman 1978), and indeed the reviewed rat studies used intertrial intervals of 5 s or more. In general, it seems that pigeons have often been used in fundamental memory research, sometimes aimed at stretching memory abilities, while rats have mainly been used in pharmacological studies in which good baseline performance is sought prior to drug administration (from these studies, we only included data from drug-free control groups). For these reasons we limit ourselves to discussing major qualitative features of the data.

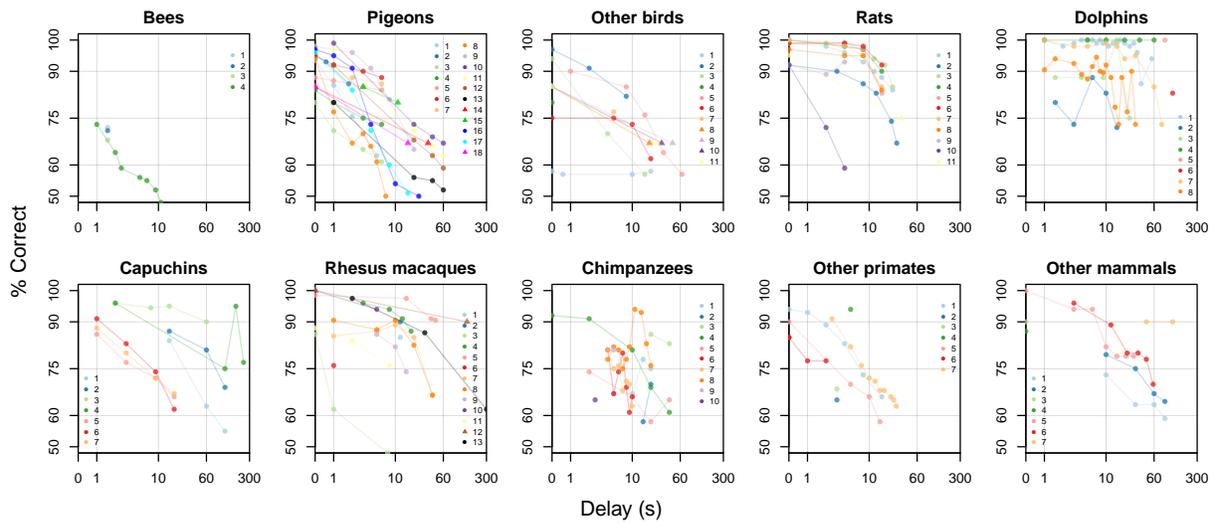


Figure 2: Performance in delayed matching-to-sample as a function of delay length across species. Axes are identical in all panels to enable direct comparison. For clarity of display, the horizontal axis is logarithmic with an added point for 0s delay. Species in multi-species panels are given in the Appendix, as are sources for all studies. Circles indicate studies using pre-determined delay intervals; triangles indicate studying in which delay interval was titrated (see Methods).

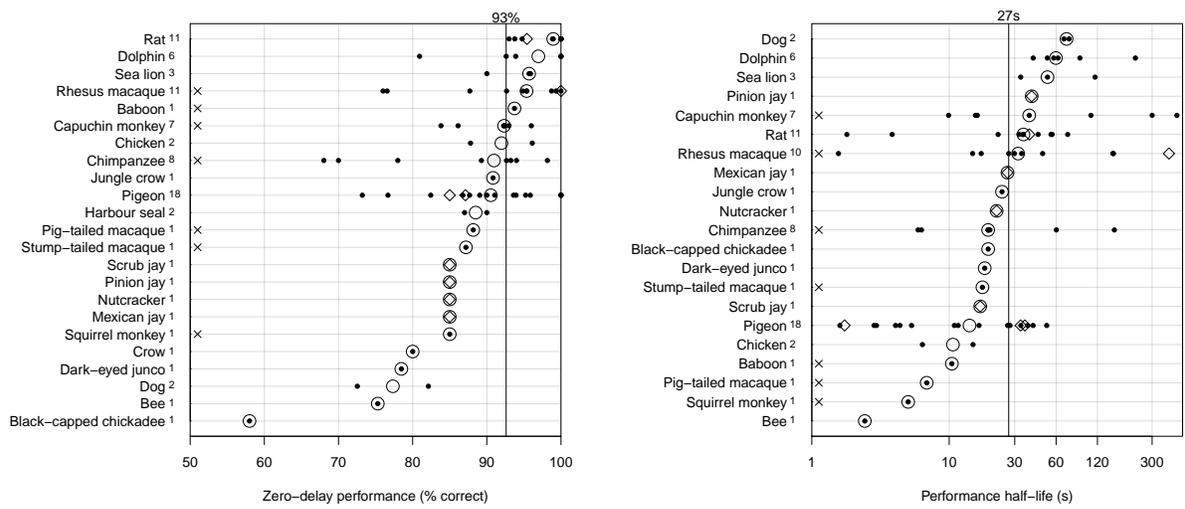


Figure 3: Estimated zero-delay performance (left) and performance half-life (right) in delayed matching-to-sample across species. Performance half-life is defined as the delay for which performance drops from its zero-delay value to a value halfway toward 50% (chance performance). Estimates from individual studies (see figure 2 and Appendix) are given as open diamonds (titration studies) or filled circles (other studies). Open circles are species medians. Black vertical lines are medians of all data points. Species are arranged by median value, from top to bottom in decreasing order. Number of data sets per species is indicated near the species name. Crosses highlight primate species.

3.1 Zero-delay performance

In figure 3, left, estimated zero-delay performance of different species is shown. Median estimates vary between 58% (black-capped chickadee) and 99.5% correct (rat), with a median across all data sets of 92% correct. While this variation appears substantial at first sight, we cannot exclude that, in fact, all species may be capable of practically perfect matching at zero delay. The reason is that only species that are rarely studied have been reported to perform poorly (figure 3, left; the Spearman rank correlation between maximum zero-delay performance and number of data sets is $r_s = 0.80$, $N = 23$, $p < 0.001$). Many details of training procedures are important for optimal performance (see the above-mentioned effects of ambient illumination and inter-trial interval), so that experimenters' familiarity with a species can influence performance greatly. This point is well illustrated by comparing rat and dog data. Rats are alongside pigeons, the most common species in laboratory studies of learning, for which training procedures have been developed for many decades. Dogs, on the other hand, have been studied only sporadically after Pavlov. Although we have little reason to believe that rats should perform much better than dogs, estimated zero-delay performance in eight rat experiments was at least 92%, with a median of 99.5%. In contrast, two dog data sets (from the same eight individuals, Kuśmierk and Kowalska 2001) yielded only 73% and 82%. Another reason why many estimates of zero-delay performance are almost certainly lower than what a species can achieve is that few studies seek to maximize zero-delay performance. Rather, animals are often trained for a predetermined number of trials or until a given criterion is met, which can be as low as 70% or 80% correct. For example, we estimated a zero-delay performance of only $\sim 75\%$ for bees based on data in Giurfa et al. (2001). The bees, however, were trained for fewer than 100 trials, at the end of which performance was still improving (Fig. 1 in Giurfa et al. 2001). For comparison, rats and pigeons commonly receive thousands of training trials. It seems likely that bees could be trained to perform better.

3.2 Performance half-life

The right panel of figure 3 shows the estimated performance half-life of different species. Median estimates vary between 2.4 s (bees) and 71 s (dogs), with a grand median of 27 s. Half-life data are more suggestive of species differences than zero-delay performance data. For example, all species yielding at least one half-life estimate over a minute are mammals. Pigeons, in particular, fall short of the one minute mark despite being a laboratory species for which training techniques are well developed. The cautionary notes put forward in the previous section, however, continue to apply. Similarly to the best zero-delay performance, the longest half-life estimated for each species correlates with the number of available data sets (figure 3, right, Spearman rank correlation $r_s = 0.74$, $N = 21$, $p < 0.001$; among birds, pigeons are the most studied and also have the longest half-life estimate). It appears therefore possible that seldom studied non-mammal species could perform at the level of mammals given improved experimental procedures.

4 Discussion

4.1 Are there species differences in DMTS performance?

The main conclusion of our review is that many (perhaps all) non-human animals are capable of near-perfect performance in zero-delay match-to-sample, and that in all species performance degrades rather quickly when a non-zero delay is used, with good performance having been reported at intervals ranging from a few seconds to a few minutes. For comparison, Overman and Doty (1980) report that humans easily attain errorless performance after a 48 hour delay.¹ We could not reach firm conclusions about species differences in memory spans, although performance half-life over a minute has been observed, to date, only in mammals. One interesting observation is that primates do not appear to stand out from other mammals (or other species in general) in either zero-delay performance or performance half-life (figure 3).

¹Overman and Doty (1980) also report above-chance performance in pigtailed macaques after a 48 hour delay. This result, however, is confounded by the fact the correct comparison stimulus was a familiar stimulus, responding to which had been reinforced 9 times during its initial presentation, while the incorrect comparison stimulus was a novel stimulus. Thus the animals' above-chance performance may derive from standard instrumental learning. Similarly, Prusky et al. (2004) report above-chance performance in rats after a 16 minutes delay, but approaching the sample stimulus had been previously reinforced.

4.2 Can memory span be improved by training?

The most straightforward way to answer this question would be to compare performance half-life across studies using different amounts of training. There are several reasons why this cannot be done. A practical reason is that amount of training is not always reported accurately, or at all. A more substantive problem is that it is hard to disentangle the effect of amount of training from other procedural differences known to affect performance half-life.

The best *prima facie* evidence that memory span can be extended through training comes perhaps from titration procedures (see Methods). With such a procedure, Mello (1971) trained three rhesus macaques to an average estimated performance half-life of about 7 minutes, considerably longer than the next longest estimate of 2.5 minutes (from data in Horel et al. 1984), who trained without titration. Further indications that memory may be trained comes from the common practice of starting animals on a zero-delay task and shifting them to progressively longer delays as performance improves. This procedure is similar to titration, with the difference that the delay progression is controlled by the experimenter rather than by the animal. The dolphin trained by Herman and Gordon (1974), for example, scored initially only 73% correct with a mere 3 s delay, but through training eventually achieved excellent performance with a 2 minute delay. Kangas et al. (2011), indeed, provide direct evidence that training may cause an increase in both zero-delay performance and performance half-life. They fitted equation (2) to pigeon data from different stages of training, finding that the best fitting function changes in the way depicted in figure 1. While the hypothesis that memory span increases with training is compatible with these findings, we should not forget that the performance function quantifies behavior and not directly memory. Another way to account for improvements in performance would be to assume that, although memory abilities do not change, the animal learns to cope with lower signal-to-noise ratios in memory contents. For example, we could assume that a longer delay interval results in a fainter memory trace of the sample stimulus (Kendrick et al. 1986). What improves during training, then, may be the ability to recover information from fainter memory traces, rather than the strength of memory traces. Whether these two hypotheses can be distinguished based on purely behavioral data is an interesting topic for future research (we also note that the hypotheses are not mutually exclusive).

4.3 How long can animal memories last?

We started this paper recalling seemingly contradictory data about animal memory spans, which have been observed to vary from seconds to months even within the same species (Olson et al. 1995). We think that these data point to the existence of two kinds of memory systems, which we may call *specialized* and *generalist*.

Specialized memory systems reflect adaptations to process particular kinds of information, and may be expected to differ across species in different ecological niches (Olson et al. 1995). The memory span of specialized memory systems should be appropriate to the animal's ecological needs. The best known examples of such systems are probably memory for locations where food has been hidden and for various features of food. The first may support memories that last for months, as appropriate in the case of winter caches mentioned in the Introduction. Memories for eaten food may last for hours, as appropriate to learn about the possible toxic effect of novel foods, which may be apparent only hours after ingestion. Taste learning is triggered by specific features of foods (Garcia et al. 1955, 1966), which may differ across species (e.g., rats use flavor cues and birds visual cues, Denny 1986). Various kinds of imprinting may also be considered the product of a specialized memory system. Recognition of potential sexual partners based on sexual imprinting, for example, may result in life-long memories (Immelmann 1972; Oetting et al. 1995).

The generalist memory system, in contrast, is possibly remarkably similar across species. It would be able to retain information about many kinds of events, but with a limited memory span of a few seconds to a few minutes. DMTS and related laboratory procedures (Kendrick et al. 1986) would typically tap into the generalist memory system, given that animals are required to remember biologically neutral stimuli such as abstract pictures or colored lights. These tasks could in principle also recruit specialized memory systems, given appropriate stimuli. For example, the fact that rats appear to have, on average, a longer performance half-life than most primates (figure 3, right) may derive simply from the common practice of requiring rats to match the *location* of the sample stimulus (e.g., Sloan et al. 2006; Dunnett and Martel 1990), a task that taps into well-developed spatial abilities. Steele and Morris 1999, for example, report that rats can remember locations for up to two hours after a single learning trial. In other settings, rats may perform significantly worse. Wallace et al. (1980), for instance, studied DMTS in rats using visual stimuli, resulting in a performance half-life of under 2 s (rats 10 in figure 2, and the worst half-life for rats in figure 3, right).

5 Acknowledgement

This study was supported by the Swedish Research Council.

References

- Baddeley, A., 1991. Human memory theory and practice. Lawrence Erlbaum Associates, publishers, London.
- Bednekoff, P. A., Balda, R. P., Kamil, A. C., Hile, A. G., 1997. Long-term spatial memory in four seed-caching corvid species. *Animal Behaviour* 53 (2), 335–341.
- Berriman, R., Cumming, W. W., Nevin, J. A., 1963. Acquisition of delayed matching in the pigeon. *Journal of the experimental analysis of behavior* 6 (6), 101–107.
- Björck, Å., 1996. Numerical methods for least squares problems. SIAM: Society for Industrial and Applied Mathematics.
- Blough, D. S., 1959. Delayed matching in the pigeon. *Journal of the Experimental Analysis of Behavior* 2 (2), 151–160.
- Bossema, I., 1979. Jays and oaks: an eco-ethological study of a symbiosis. *Behaviour*, 1–117.
- Bouton, M. E., 2007. Learning and behavior: A contemporary synthesis. Sinauer Associates.
- Brodbeck, D. R., Shettleworth, S. J., 1995. Matching location and color of a compound stimulus: Comparison of a food-storing and a nonstoring bird species. *Journal of Experimental Psychology: Animal Behavior Processes* 21 (1), 64.
- Clayton, N. S., Dickinson, A., 1998. Episodic-like memory during cache recovery by scrub jays. *Nature* 395 (6699), 272–274.
- Clayton, N. S., Griffiths, D. P., Emery, N. J., Dickinson, A., 2001a. Elements of episodic-like memory in animals. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356 (1413), 1483–1491.
- Clayton, N. S., Yu, K. S., Dickinson, A., 2001b. Scrub jays (*Aphelocoma coerulescens*) form integrated memories of the multiple features of caching episodes. *Journal of Experimental Psychology: Animal Behavior Processes* 27 (1), 17.
- Core Team, R., 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL <http://www.R-project.org>
- D’amato, M. R., O’Neill, W., 1971. Effect of delay-interval illumination on matching behavior in the capuchin monkey. *Journal of the Experimental Analysis of Behavior* 15 (3), 327–333.
- D’Amato, M. R., Worsham, R. W., 1972. Delayed matching in the capuchin monkey with brief sample durations. *Learning and Motivation* 3 (3), 304–312.
- Davachi, L., Goldman-Rakic, P. S., 2001. Primate rhinal cortex participates in both visual recognition and working memory tasks: functional mapping with 2-DG. *Journal of neurophysiology* 85 (6), 2590–2601.
- Davenport, R. K., Rogers, C. M., Russell, I. S., 1975. Cross-modal perception in apes: Altered visual cues and delay. *Neuropsychologia* 13 (2), 229–235.
- Dayer, C. A., Baron, S., Light, K. E., Wenger, G. R., 2000. Effects of ethanol on working memory and attention in pigeons. *Journal of Pharmacology and Experimental Therapeutics* 293 (2), 551–558.
- Denny, M. R., 1986. Retention of SR in the midst of the cognitive invasion. (Eds Kendrick, DF and Rilling, ME and Denny, MR) *Theories of animal memory*, Psychology Press, 35–50.

- Diekamp, B., Gagliardo, A., Güntürkün, O., 2002. Nonspatial and subdivision-specific working memory deficits after selective lesions of the avian prefrontal cortex. *The Journal of neuroscience* 22 (21), 9573–9580.
- Dunnett, S. B., Martel, F. L., 1990. Proactive interference effects on short-term memory in rats: I. Basic parameters and drug effects. *Behavioral neuroscience* 104 (5), 655–665.
- Emmerling, M. R., Gregor, V. E., Schwarz, R. D., Scholten, J. D., Callahan, M. J., Lee, C., Moore, C. J., Raby, C., Lipinski, W. J., Davis, R. E., 1994. PD 142676 (CI 1002), a novel anticholinesterase and muscarinic antagonist. *Molecular neurobiology* 9 (1-3), 93–106.
- Etkin, M., D'Amato, M. R., 1969. Delayed matching-to-sample and short-term memory in the capuchin monkey. *Journal of Comparative and Physiological Psychology* 69 (3), 544–549.
- Ferraro, D. P., Grilly, D. M., 1974. Effects of chronic exposure to Δ^9 -tetrahydrocannabinol on delayed matching-to-sample in chimpanzees. *Psychopharmacologia* 37 (2), 127–138.
- Finch, G., 1942. Delayed matching-from-sample and non-spatial delayed response in chimpanzees. *Journal of Comparative Psychology* 34 (3), 315–319.
- Garcia, J., Ervin, F. R., Koelling, R. A., 1966. Learning with prolonged delay of reinforcement. *Psychonomic Science* 5 (3), 121–122.
- Garcia, J., Kimeldorf, D. J., Koelling, R. A., 1955. Conditioned aversion to saccharin resulting from exposure to gamma radiation. *Science* 122, 157–158.
- Giurfa, M., Zhang, S., Jenett, A., Menzel, R., Srinivasan, M. V., 2001. The concepts of 'sameness' and 'difference' in an insect. *Nature* 410 (6831), 930–933.
- Goto, K., Watanabe, S., 2009. Visual working memory of jungle crows (*Corvus macrorhynchos*) in operant delayed matching-to-sample1. *Japanese Psychological Research* 51 (3), 122–131.
- Grant, D. S., 1976. Effect of sample presentation time on long-delay matching in the pigeon. *Learning and Motivation* 7 (4), 580–590.
- Hampson, R., Espana, R., Rogers, G., Porrino, L., Deadwyler, S., 2009. Mechanisms underlying cognitive enhancement and reversal of cognitive deficits in nonhuman primates by the ampakine CX717. *Psychopharmacology* 202 (1-3), 355–369.
- Hampson, R., Jarrard, L., Deadwyler, S., 1999. Effects of ibotenate hippocampal and extrahippocampal destruction on delayed-match and -nonmatch-to-sample behavior in rats. *The Journal of Neuroscience* 19, 1492–1507.
- Hampton, R. R., Shettleworth, S. J., 1996. Hippocampal lesions impair memory for location but not color in passerine birds. *Behavioral neuroscience* 110 (4), 831.
- Herman, L. M., Gordon, J. A., 1974. Auditory delayed matching in the bottlenose dolphin. *Journal of the Experimental Analysis of Behavior* 21 (1), 19–26.
- Herman, L. M., Hovancik, J. R., Gory, J. D., Bradshaw, G. L., 1989. Generalization of visual matching by a bottlenosed dolphin (*Tursiops truncatus*): Evidence for invariance of cognitive performance with visual and auditory materials. *Journal of Experimental Psychology: Animal Behavior Processes* 15 (2), 124–136.
- Herman, L. M., Thompson, R. K., 1982. Symbolic, identity, and probe delayed matching of sounds by the bottlenosed dolphin. *Animal Learning & Behavior* 10 (1), 22–34.
- Horel, J. A., Voytko, M. L., Salsbury, K. G., 1984. Visual learning suppressed by cooling the temporal pole. *Behavioral neuroscience* 98 (2), 310–324.
- Immelmann, K., 1972. Sexual and other long-term aspects of imprinting in birds and other species. *Advances in the Study of Behavior* 4, 147–174.

- Jarrard, L. E., Moise Jr, S. L., 1970. Short-term memory in the stump-tail macaque: Effect of physical restraint of behavior on performance. *Learning and Motivation* 1 (3), 267–275.
- Johnston, K., Everling, S., 2006. Neural Activity in Monkey Prefrontal Cortex Is Modulated by Task Context and Behavioral Instruction during Delayed-match-to-sample and Conditional Prosaccade-Antisaccade Tasks. *Journal of Cognitive Neuroscience* 18 (5), 749–765.
- Kangas, B. D., Berry, M. S., Branch, M. N., 2011. On the development and mechanics of delayed matching-to-sample performance. *Journal of the Experimental Analysis of Behavior* 95, 221–236.
- Kangas, B. D., Vaidya, M., Branch, M. N., 2010. Titrating-delay matching-to-sample in the pigeon. *Journal of the Experimental Analysis of Behavior* 94, 69–81.
- Kendrick, D. F., Rilling, M. E., Denny, M. R., 1986. *Theories of animal memory*. Psychology Press.
- Kuśmirek, P., Kowalska, D. M., 2001. Effect of sound source position on learning and performance of auditory delayed matching-to-sample task in dogs. *Acta neurobiologiae experimentalis* 62 (4), 251–262.
- Mauck, B., Dehnhardt, G., 2005. Identity concept formation during visual multiple-choice matching in a harbor seal (*Phoca vitulina*). *Learning & behavior* 33 (4), 428–436.
- Mello, N. K., 1971. Alcohol effects on delayed matching to sample performance by rhesus monkey. *Physiology & behavior* 7 (1), 77–101.
- Nakagawa, S., Etheredge, R. J., Foster, T. M., Sumpter, C. E., Temple, W., 2004. The effects of changes in consequences on hens performance' in delayed-matching-to-sample tasks. *Behavioural processes* 67 (3), 441–451.
- Nelson, K. R., Wasserman, E. A., 1978. Temporal factors influencing the pigeons' successive matching-sample performance: Sample duration, intertrial interval, and retention interval. *Journal of the experimental analysis of Behavior* 30 (2), 153–162.
- Nissen, H. W., Riesen, A. H., Nowlis, V., 1938. Delayed response and discrimination learning by chimpanzees. *Journal of Comparative Psychology* 26 (2), 361–386.
- Oetting, S., Proeve, E., Bischof, H.-J., 1995. Sexual imprinting as a two-stage process: mechanisms of information storage and stabilization. *Animal Behaviour* 50 (2), 393–403.
- Olson, D. J., Kamil, A. C., Balda, R. P., Nims, P. J., 1995. Performance of four-seed caching corvid species in operant tests of nonspatial and spatial memory. *Journal of Comparative Psychology* 109 (2), 173.
- Overman, Jr., W. H., Doty, R. W., 1980. Prolonged visual memory in macaques and man. *Neuroscience* 5 (11), 1825–1831.
- Pack, A. A., Herman, L. M., Roitblat, H. L., 1991. Generalization of visual matching and delayed matching by a California sea lion (*Zalophus californianus*). *Animal Learning & Behavior* 19 (1), 37–48.
- Pearce, J. M., 2013. *Animal learning and cognition: an introduction*. Psychology Press.
- Porritt, M., Poling, A., 2008. Scopolamine effects under a titrating-delayed-nonmatching-to-position procedure. *Psychological Record* 58, 37–49.
- Pribram, K., Miller, G. A., Galanter E., 1960. *Plans and the structure of behavior*. Holt, Reinhart, and Winston, New York.
- Prusky, G., Douglas, R., Nelson, L., Shabanpoor, A., Sutherland, R., 2004. Visual memory task for rats reveals an essential role for hippocampus and perirhinal cortex. *Proceedings of the National Academy of Sciences of the United States of America* 101 (14), 5064–5068.
- Raby, C. R., Alexis, D. M., Dickinson, A., Clayton, N. S., 2007. Planning for the future by western scrub-jays. *Nature* 445 (7130), 919–921.

- Roberts, W. A., 1980. Distribution of trials and intertrial retention in delayed matching to sample with pigeons. *Journal of Experimental Psychology: Animal Behavior Processes* 6 (3), 217–237.
- Roberts, W. A., 2002. Are animals stuck in time? *Psychological Bulletin* 128 (3), 473.
- Roberts, W. A., Kraemer, P. J., 1981. Recognition memory for lists of visual stimuli in monkeys and humans. *Animal Learning & Behavior* 9 (4), 587–594.
- Rodriguez, J. S., Zürcher, N., Bartlett, T. Q., Nathanielsz, P. W., Nijland, M. J., 2011. CANTAB delayed matching to sample task performance in juvenile baboons. *Journal of Neuroscience Methods* 196 (2), 258–263.
- Rubin, D. C., Wenzel, A. E., 1996. One hundred years of forgetting: A quantitative description of retention. *Psychological Review* 103 (4), 734–760.
- Skov Rackette, S. I., Miller, N. Y., Shettleworth, S. J., 2006. What-where-when memory in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes* 32 (4), 345–358.
- Sloan, H. L., Good, M., Dunnett, S. B., 2006. Double dissociation between hippocampal and prefrontal lesions on an operant delayed matching task and a water maze reference memory task. *Behavioural brain research* 171 (1), 116–126.
- Smirnova, A., Lazareva, O., Zorina, Z., 2000. Use of number by crows: investigation by matching and oddity learning. *Journal of the Experimental Analysis of Behavior* 73 (2), 163–176.
- Smith, L., 1967. Delayed discrimination and delayed matching in pigeons. *Journal of the Experimental Analysis of Behavior* 10 (6), 529–533.
- Soto, P. L., Dallery, J., Ator, N. A., Katz, B. R., 2013. A critical examination of best dose analysis for determining cognitive-enhancing potential of drugs: studies with rhesus monkeys and computer simulations. *Psychopharmacology* 228, 611–622.
- Stanhope, K., McLenachan, A., Dourish, C., 1995. Dissociation between cognitive and motor/motivational deficits in the delayed matching to position test: effects of scopolamine, 8-OH-DPAT and EAA antagonists. *Psychopharmacology* 122 (3), 268–280.
- Steele, R. J., Morris, R. G., 1999. Delay-dependent impairment of a matching-to-place task with chronic and intrahippocampal infusion of the NMDA-antagonist D-AP5. *Hippocampus* 9, 118–136.
- Suddendorf, T., Busby, J., 2003. Mental time travel in animals? *Trends in cognitive sciences* 7 (9), 391–396.
- Vonk, J., 2003. Gorilla (*Gorilla gorilla gorilla*) and orangutan (*Pongo abelii*) understanding of first-and second-order relations. *Animal Cognition* 6 (2), 77–86.
- Wallace, J., Steinert, P. A., Scobie, S. R., Spear, N. E., 1980. Stimulus modality and short-term memory in rats. *Animal Learning & Behavior* 8 (1), 10–16.
- Washburn, D. A., Hopkins, W. D., Rumbaugh, D. M., 1989. Video-task assessment of learning and memory in macaques (*Macaca mulatta*): Effects of stimulus movement on performance. *Journal of Experimental Psychology: Animal Behavior Processes* 15 (4), 393–400.
- Weinstein, B., 1941. Matching-from-sample by rhesus monkeys and by children. *Journal of Comparative Psychology* 31 (1), 195.
- Wenger, G. R., Kimball, K. A., 1992. Titrating matching-to-sample performance: Effects of drugs of abuse and intertrial interval. *Pharmacology, Biochemistry and Behavior* 41, 283–288.
- White, K. G., 1985. Characteristics of forgetting functions in delayed matching to sample. *Journal of the experimental analysis of behavior* 44 (1), 15–34.
- Wickens, T. D., 1998. On the form of the retention function: Comment on Rubin and Wenzel (1996): A quantitative description of retention. *Psychological Review* 105 (2), 379–386.

- Yerkes, R. M., Nissen, H. W., 1939. Prelinguistic sign behavior in chimpanzee. *Science* 89, 585–587.
- Zentall, T. R., 2005. Animals may not be stuck in time. *Learning and Motivation* 36 (2), 208–225.
- Zhang, S., Bock, F., Si, A., Tautz, J., Srinivasan, M. V., 2005. Visual working memory in decision making by honey bees. *Proceedings of the National Academy of Sciences of the United States of America* 102 (14), 5250–5255.
- Zimmerberg, B., Glick, S. D., Jarvik, M., 1971. Impairment of recent memory by marihuana and THC in rhesus monkeys. *Nature* 233, 343–345.

Data sources

Data sources are listed following the layout in figure 2. Studies marked * used titration of the delay interval (see Methods).

Bee. 1–3, Giurfa et al. (2001); 4 Zhang et al. (2005).

Pigeons. 1, Blough (1959); 2, Diekamp et al. (2002); 3–6, Roberts (1980); 7, Skov Rackette et al. (2006); 8, Smith (1967); 9, White (1985); 10–13, Grant (1976); 14, *Kangas et al. (2010); 15, *Dayer et al. (2000); 16, Berriman et al. (1963); 17, Kangas et al. (2011); 18, *Wenger and Kimball (1992).

Other birds. Black-capped chickadee: 1, Brodbeck and Shettleworth (1995). Chicken: 2–3, Nakagawa et al. (2004). Crow: 4, Smirnova et al. (2000). Dark-eyed junco: 5, Brodbeck and Shettleworth (1995); 6, Hampton and Shettleworth (1996). Jungle crow: 7, Goto and Watanabe (2009). Mexican jay: 8, *Olson et al. (1995). Pinion jay: 9, *Olson et al. (1995). Nutcracker: 10, *Olson et al. (1995). Scrub jay: 11, *Olson et al. (1995).

Rat. 1, Dunnett and Martel (1990); 2, Hampson et al. (1999); 3, Sloan et al. (2006); 4–8, Stanhope et al. (1995); 9–10: Wallace et al. (1980); 11, *Porritt and Poling (2008).

Dolphin. 1, Herman and Thompson (1982); 2–6, Herman and Gordon (1974); 7, Herman et al. (1989).

Capuchin monkey. 1–2, D’amato and O’Neill (1971); 3–4, D’Amato and Worsham (1972); 5–7, Etkin and D’Amato (1969).

Rhesus macaques. 1–2, Davachi and Goldman Rakic (2001); 3, Emmerling et al. (1994); 4, Hampson et al. (2009); 5, Horel et al. (1984); 6, Johnston and Everling (2006); 7–8, Washburn et al. (1989); 9, Weinstein (1941); 10–11, Zimmerberg et al. (1971); 12 *Mello (1971); 13, Soto et al. (2013).

Chimpanzee. 1–2, Davenport et al. (1975); 3, Ferraro and Grilly (1974); 4–5, Finch (1942); 6–9, Nissen et al. (1938); 10, Yerkes and Nissen (1939).

Other primates. Baboon: 1, Rodriguez et al. (2011). Gorilla: 2, Vonk (2003). Orangutan: 3, (Vonk 2003). Pig-tailed macaque: 4–5, Overman and Doty (1980). Stump-tailed macaque: 6, Jarrard and Moise Jr (1970). Squirrel monkey: 7, Roberts and Kraemer (1981).

Other mammals. Dog: 1–2, Kuśmierk and Kowalska (2001); Harbor seal: 3–4, Mauck and Dehnhardt (2005); Sea lion: 5–6, Pack et al. (1991).