Author(s) Neil Stewart

Article Title: Order of acquisition in learning perceptual categories: A laboratory analogue of the age-of-acquisition effect?

Year of publication: 2008

Link to published version: http://dx.doi.org/10.3758/PBR.15.1.70

Publisher statement: None

Running Head: LABORATORY AoA

Order of Acquisition in Learning Perceptual Categories: A Laboratory Analogue of the Age of Acquisition Effect?

Neil Stewart

University of Warwick, England

Andrew W. Ellis

University of York, England

Stewart, N., & Ellis, A. W. (2008). Order of acquisition in learning perceptual categories: A laboratory analogue of the age of acquisition effect? *Psychonomic Bulletin & Review, 15,* 70-74.

main text + references = 3,819 words

neil.stewart@warwick.ac.uk

+44 (0) 24 7657 3127

Department of Psychology

University of Warwick

Coventry

CV4 7AL

England

Abstract

In the age of acquisition (AoA) effect, an advantage for recognition and production is found for items learned early in life compared to items learned later. In this laboratory analogue, participants learned to categorize novel random checkerboard stimuli. Some stimuli were presented from the onset of training; others were introduced later. At test, when early and late stimuli had equal cumulative frequency, early stimuli were classified significantly more quickly. Because stimuli were randomly assigned to be introduced either early or late, we can conclude that early stimuli were categorized more quickly because of their order of acquisition. This finding suggests that age, or order, of acquisition effects are a general property of any learning system.

Order of Acquisition in Learning Perceptual Categories:

A Laboratory Analogue of the Age of Acquisition Effect?

Stimuli learned earlier in life are processed more quickly and/or more accurately than stimuli learned later in life. The opportunity for such an age of acquisition (AoA) effect is found when a set of representations are not learned simultaneously but instead sequentially and cumulatively over time, with new items being added to an ever-growing ‘vocabulary’ of older items (see Johnston & Barry, 2006, and Juhasz, 2005, for recent reviews). The literature on AoA effects focused initially on the processing of words. Carroll and White (1973) showed that early learned words are produced faster in an object naming task than later acquired words, an observation that has since been replicated and extended in several different languages (Bates, Burani, D'Amico, & Barca, 2001; Bonin, Chalard, Méot, & Fayol, 2002; Caroll & White, 1973; Ellis & Morrison, 1998; Ghyselinck, Lewis, & Brysbaert, 2004; Snodgrass & Yuditsky, 1996). AoA effects have subsequently been reported in other lexical processing tasks, such as lexical decision, semantic categorization, written word naming, and eye fixations in reading (e.g., Brysbaert, Van Wijnendaele, & De Deyne, 2000; Juhasz & Rayner, 2003; Monaghan & Ellis, 2002; Morrison & Ellis, 2000).

But AoA effects are not only found in lexical processing tasks. For example, AoA effects occur when participants are required to discriminate real objects from invented non-objects (Holmes & Ellis, 2006; Moore, Smith-Spark, & Valentine, 2004), in distinguishing famous from unfamiliar faces (Moore & Valentine, 1999), and in categorizing actors' faces into television program categories (Lewis, 1999). The ubiquitous nature of AoA effects has led to the suggestion that they might be a general property of learning under conditions where sets of items are learned gradually and incrementally, and that they may be distributed throughout the cognitive system (e.g., Catling & Johnston, 2006; Ellis & Lambon Ralph, 2000; Moore & Valentine, 1999). If so, the study of AoA effects should reveal something of system-wide cognitive principles in learning and memory. So why, then, is there an advantage for early acquired items in these different tasks and situations?

Cumulative Frequency

If two words have equal frequencies, but one was learned before the other, then the total, cumulative frequency of exposure to the early acquired word will be greater than the frequency of exposure to the later acquired word. The cumulative frequency hypothesis proposes that AoA effects occur because of the differences in total number of life-time exposures to early and late acquired stimuli (Carroll & White, 1973; Lewis, 1999; Zevin & Seidenberg, 2002, 2004). AoA effects would then be real but theoretically uninteresting, because frequency effects (and thus cumulative frequency effects) are already well accommodated by many types of cognitive model (e.g., Lewis, 1999; Murray & Forster, 2004; Seidenberg & McClelland, 1989).

But there is now considerable evidence against the cumulative frequency hypothesis. First, AoA effects occur when cumulative frequency is controlled (Lewis, Chadwick, & Ellis, 2002; Pérez, 2007). Second, when the contributions of frequency and AoA are compared in the same experiments, AoA effects are greater in magnitude than cumulative frequency differences would predict (Ghyselinck et al., 2004). Third, though relative differences in cumulative frequency reduce with age, AoA effects do not (Barry, Johnston, & Wood, 2006; Lewis et al., 2002; Morrison, Hirsh, Chappell, & Ellis, 2002).

Neural Network Accounts

A second class of theories is based on the behavior of artificial neural networks required to learn sets of items over time in a cumulative fashion. Steyvers and Tenenbaum (2005) offered an account of AoA in terms of the growth of a semantic network. Concepts which are acquired earlier in the growth of a semantic network will be more connected within the network, and thus are retrieved more easily when the cognitive system searches through the network.

Ellis and Lambon Ralph (2000) demonstrated effects of the order of acquisition of items in a simple feed-forward connectionist network involving two layers of weights trained by back propagation of error. If the network was trained on one set of 'early' items before a second set of 'late' items were added into training alongside the early ones, then performance on the early items was superior to than on the late items, even when the cumulative frequencies of the early and late items were equated. Learning in a distributed network depends on adjustments to the strength of connections or weights between processing units. Items introduced into training at the outset have the opportunity to adjust the weights in directions optimal for their own representation. Later items may prefer a different weight structure but their attempts to reconfigure the weight space are resisted by the early items which continue to be experienced alongside them.

Ellis and Lambon Ralph's (2000) network account can be considered at a more general level. The learning system sets parameters (e.g., weights in a connectionist network) to represent early items. When the late items are introduced, the system is already biased towards regions of the parameter space that favor early items. If the learning algorithm cannot escape this local minimum, as is likely to be the case in high-dimensional parameter spaces, then the result is an AoA effect.

Differential Processing

Moore (2003) suggests that AoA effects result from using early novel stimuli to set up a specialized processing mechanism for the new stimulus class. Explicit processing of early items is used to set the parameters of the system before automatized processing of later items. Though Moore's account seems similar to the network account, with early items setting the system parameters, the model differs from the network accounts in assuming differential processing of early and late stimuli as a result of different affective responses to early and late stimuli. Thus the model accounts for AoA effects because early items are processed differently from later items and not because the learning algorithm is stuck in a local minimum.

A Laboratory Analogue of AoA Effects?

AoA has natural, real-world correlations with factors like frequency and imageability that make it very difficult to manipulate AoA while controlling possible confounding variables. An alternative approach would be to attempt to simulate age of acquisition effects in the laboratory under conditions which allow greater control over the nature of the stimuli and the conditions of learning (Lewis, 2006). Demonstrations of AoA effects for faces acquired after childhood (Lewis, 1999; Moore & Valentine, 1999) and for second language vocabularies acquired in late childhood or adulthood (Izura & Ellis, 2002; 2004) gave hope that if young adults were required to learn novel items in a cumulative fashion, those items might also show AoA effects.

In this article, we report an experiment where we induce a laboratory AoA effect. The experiment avoids the problem of natural confounds with other factors by using previously unseen artificial stimuli. Young adult participants learned to categorize novel checkerboard stimuli. The mapping of stimuli and categories was entirely arbitrary. Testing was done over five sessions. Some *early* checkerboards were trained from Session 1. The introduction of the remaining *late* checkerboards was delayed until Session 2. Training frequencies were adjusted so that early and late checkerboards had been seen equally often by the end of Session 4. The crucial question was whether the early checkerboards would show a processing advantage over the late ones in Session 5. Because performance was compared when early and late boards had equal cumulative frequently, an advantage for early items cannot be attributed to cumulative frequency. Because stimuli were randomly assigned to early and late sets, if there is an advantage for early boards over late boards, then this cannot be attributed to some uncontrolled intrinsic stimulus property.

Experiment

Method

Participants. Twenty-seven University of Warwick students (19 female and 8 male) aged between 20 and 35 participated. Pay (between £30 and £38) comprised an hourly rate and a bonus proportional to the sum of the reciprocals of correct reaction times (RTs), to reward fast, accurate responding.

Stimuli. Sixteen different 12 x 12 checkerboards were randomly generated for each participant. Half were assigned to Category A and half to Category B. Each square within a checkerboard was set to be either black or white with probability .5. Each checkerboard square measured 6 x 6 pixels. Checkerboards were presented on a uniform gray background. The checkerboards were presented in the center of the monitor with a 40 cm viewing area diagonal, a 1024 x 768 resolution, and a 87Hz refresh rate. The viewing distance was approximately 50 cm. Feedback was presented below the checkerboards in 1 cm high, white, sans-serif text.

Design. There were five experimental sessions, each containing eight blocks of 96 trials. For each participant, 8 of the 16 checkerboards (4 from Category A and 4 from Category B) were designated early, and the remaining 8 designated late. Only early boards were presented in Session 1. In Sessions 2-4, late boards were presented more often than early boards, so that by the end of Session 4, early boards and late boards had been presented equally often (see Table 1). The comparison of interest is between performance on early and late boards in Session 5, by which time the early and late boards had been presented equally frequently.

Procedure. Participants were tested individually in a quiet cubicle. Each session lasted about 45 minutes. Participants were allowed to schedule two sessions in a day, in which case one session occurred in the morning and the other in the afternoon, with at least one hour between sessions. Across participants the five sessions were distributed over a minimum of 3 and a maximum of 5 consecutive days.

Participants were instructed to categorize the checkerboards as quickly as possible without making mistakes, and were told that their performance would determine their bonus. The introduction of the additional late boards was drawn to participants' attention at the beginning of Session 2.

Each trial began with a 500 ms blank screen. A checkerboard was then selected at random from only the early boards for Session 1, or from both early and late boards in the other sessions. The board was presented at the beginning of a vertical retrace (Stewart, 2006a). Participants responded by pressing one of two push-to-make buttons (labeled 'A' and 'B') on a button box connected to the parallel port (Stewart, 2006b). The button pressed and the RT from stimulus onset (to the nearest ms) were recorded. The correct answer (either 'A' or 'B') was then displayed immediately below the checkerboard for 1 s before the next trial began. At first participants had to guess, but by paying attention to the feedback they could learn the correct category for each checkerboard. To motivate participants to perform quickly and accurately, a summary table was displayed at the end of each block. The table showed the percentage of correct responses and the mean RT for those responses for each block completed in the session.

Results and Discussion

Figures 2 and 3 show how the accuracy and speed (10% trimmed mean correct RT) of categorization changed over experimental sessions for the early and late boards. For the early boards, accuracy and speed improved quickly over Session 1. Performance on the early boards dropped between Session 1 and Session 2 when the late boards are introduced because participants had to learn to discriminate the early boards from the late boards. In Sessions 2 to 4, performance on both the early and late boards improved, with accuracy reaching about 93% correct for both early and late boards and RT dropping to about 600 ms.

The critical comparison is between early and late boards in Sessions 5, by which time early and late boards had been presented equally often. A stimulus type by block ANOVA was run on the 10% trimmed mean correct RTs in Session 5. The impact of oder of acquisition was revealed in a significant main effect of stimulus type, $F(1, 26) = 5.13$, $p = .032$, with faster responses to the early boards (mean = 562 ms) than to the late boards (mean = 578 ms). There was also a main effect of block, $F(7, 182) = 4.29$, $p = .0002$, showing a reduction in RT over the eight blocks of Session 5. The stimulus type x block interaction was not significant, $F(7, 182) = 1.13$, $p = .35$, indicating that the advantage for early stimuli was stable throughout Session 5.

A direct comparison of 10% trimmed mean correct RTs for early and late boards averaged over all blocks of Session 5 shows that the early boards were categorized 16 ms faster than the late boards (the 95% confidence interval on this mean is 1 - 31 ms). Consistent with the significant main effect in the ANOVA, the difference averaged over blocks is significant, $t(26) = 2.21$, p = .036, $\eta^2 = .16$. Because the RT data were positively skewed, a Wilcoxon signed rank test was run and confirmed the significance of the difference, $p = .015$. Alternative analyses based on Ratcliff's (1993) suggestions of either mean RT, mean inverse RT, and median RT using both t-tests and Wilcoxon signed rank tests all give significant differences.

In Session 5, accuracy was virtually identical for the early and late boards (mean proportion correct = .93 for early and late boards), $t(26) = 0.42$, $p = .68$, showing the RT difference between early and late boards cannot be attributed to a speed-accuracy tradeoff.

General Discussion

The problems of assessing causality in studies with natural stimuli, where word frequency and AoA are highly correlated, led Lewis (2006) to suggest that studies might be conducted with novel stimuli randomly assigned to be introduced early or late. We have presented the first such laboratory analogue of the AoA effect. Participants categorized early and late checkerboards to a high degree of accuracy yet they classified the early boards significantly faster than the late ones. The results cannot be explained in terms of differences in simple frequency in Session 5 or cumulative frequency across sessions. The use of randomly generated stimuli randomly assigned to be learned early or introduced later rules out other uncontrolled intrinsic stimulus properties as explanations.

The study of order of acquisition effects in incremental learning offers the prospect of experimental studies which could test a variety of accounts of how and why AoA effects arise. For example, network accounts predict larger AoA effects when early and late mappings are inconsistent and smaller AoA effects for consistent mappings (Ellis & Lambon Ralph, 2000; Lambon Ralph & Ehsan, 2006; Smith, Cottrell, & Anderson, 2001; Zevin & Seidenberg, 2002). This prediction is supported by the observation of substantial AoA effects for reading Japanese kanji characters (Havelka & Tomita, 2006), minimal effects for reading the highly regular words of Italian (Bates et al., 2001), and greater effects for irregular, exception words than for regular, consistent words in English (Ellis & Monaghan, 2002; Monaghan & Ellis, 2002). Experimentally manipulating the consistency of mappings learned for early and late items will provide a concrete test of this hypothesis. Moore (2003) proposed that AoA effects result from using early novel stimuli to set up a specialized processing mechanism for the new stimulus class. The differential processing of early and late stimuli is driven by differences in affective responses to early and late stimuli. It should be possible to manipulate experimentally affective responses and therefore to test that aspect of the theory.

The kind of task and stimuli deployed in the present experiment have not previously been employed in studies of AoA effects but have been widely used in studies of perceptual categorization and learning. Studies of categorization and concept formation have not, however, analyzed the effects of incremental training. Our results indicate that theories in those areas will need to be adapted to account for order of acquisition effects. Exemplar models are perhaps the most successful models of perceptual categorization (e.g., Nosofsky, 1986), but as currently constituted they do not predict AoA effects beyond those attributable to cumulative frequency (Lewis, 1999). There are, however, several ways in which they might be adapted to do so. In exemplar models, stimuli are classified according to their similarity to stored category exemplars. Early exemplars could be weighted more heavily in similarity calculations, although there is no precedent for such a primacy assumption. Representations of early stimuli may be more distinct (cf. Ellis & Lambon Ralph, 2000) and thus less confusable, or there may be less generalization between early acquired items. Learned attentional weights (Kruschke, 1992) or learned orders of feature sampling (Lamberts, 2000) may favor early items, and fail to be relearned for later items.

If AoA effects were confined to a specific level of a particular system (e.g., phonological mappings, semantic mappings) as some authors suggest, our finding of AoA effects for categorization of novel checkerboard stimuli would be very surprising. Instead, our results are consistent with the suggestion that AoA effects might be a general property of learning, distributed throughout the system (e.g., Catling & Johnston, 2006; Ellis and Lambon Ralph, 2000; Moore & Valentine, 1999) and offer a way forward for studies of AoA effects freed from the confounds that plague studies using natural stimuli.

References

Barry, C., Johnston, R. A., & Wood, R. F. (2006). Effects of age of acquisition, age, and repetition, priming on object naming. *Visual Cognition, 13,* 911-927.

Bates, E., Burani, C., D'Amico, S., & Barca, L. (2001). Word reading and picture naming in Italian. *Memory & Cognition*, *29*, 986-999.

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*, 89-114.

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age of acquisition effects in semantic tasks. *Acta Psychologica*, *104*, 215-226.

Carroll, J. B., & White, M. N. (1973). Word frequency and age-of-acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology, 25,* 85-95.

Catling, J. C., & Johnston, R. A. (2006). The effects of age of acquisition on an object classification task. *Visual Cognition, 13,* 968-980.

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1103-1123.

Ellis, A. W., & Monaghan, J. (2002). Reply to Strain, Patterson, and Seidenberg (2002). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 215-220.

Ellis, A. W., & Morrison, C. M. (1998). Real age of acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 515-523.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica*, *115*, 43-67.

Havelka, J., & Tomita, I. (2006). Age of acquisition in naming Japaneses words. *Visual Cognition, 13,* 981-991.

Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition, 17,* 884-910.

Izura, C., & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica*, *23*, 245-281.

Izura, C., & Ellis, A.W. (2004). Age of acquisition effects in translation judgment tasks. *Journal of Memory and Language*, *50*, 165-181.

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition, 13,* 789-845.

Juhasz, B. J. (2005). Age of acquisition effects in word and picture identification. *Psychological Bulletin, 131,* 684-712.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 1312-1318.

Kruschke, J. K. (1992). ALCOVE: An exemplar based connectionist model of category learning. *Psychological Review, 99,* 22-44.

Lamberts, K. (2000). Information-accumulation theory of speeded classification. *Psychological Review, 107,* 227-260.

Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: empirical and computational evidence. *Visual Cognition, 13,* 928-948.

Lewis, M. B. (1999). Age of acquisition in face categorization: Is there an instance-based account? *Cognition*, *71*, B23-B39.

Lewis, M. B. (2006). Chasing psycholinguistic effects: A cautionary tale. *Visual Cognition, 13,* 1012-1026.

Lewis, M. B., Chadwick, A. J., & Ellis, H. D. (2002). Exploring a neural-network account of age-of-acquisition effects using repetition priming of faces. *Memory & Cognition, 30,* 1228-1237.

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 183-206.

Moore, V. (2003). An alternative account for the effects of age of acquisition. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 67-93). New York: Nova.

Moore, V., Smith-Spark, J. H., & Valentine, T. (2004). The effect of age of acquisition on object perception. *European Journal of Cognitive Psychology, 16,* 417-439.

Moore, V., & Valentine, T. (1999). The effects of age of acquisition in processing famous faces and names: Exploring the locus and proposing a mechanism. In *Proceedings of the twenty-first annual meeting of the Cognitive Science Society* (pp. 416-421). Mahwah, NJ: Erlbaum.

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology*, *91*, 167-180.

Morrison, C. M., Hirsh, K. W., Chappell, T., & Ellis, A. W. (2002). Age and age of acquisition: An evaluation of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology , 14 ,* 435-459.

Murray, W. S., & Forster, K. I. (2004). Serial mechanism in lexical access: The rank hypothesis. *Psychological Review, 111,* 721-756.

Nosofsky, R. M. (1986). Attention, similarity and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115,* 39-57.

Pérez, M. A. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *Quarterly Journal of Experimental Psychology, 60,* 32-42.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510-532.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Review, 96,* 523-568.

Smith, M. A., Cottrell, G. W, & Anderson, K. L. (2001). The early word catches the weights. In T. K. Leen, T. G. Dietterich, & V. Tresp (Eds.), *Advances in neural information processing systems* (Vol 13, pp. 52-58). Cambridge, MA: MIT Press.

Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, and Computers*, *28*, 516-536.

Stewart, N. (2006b). Millisecond accuracy video display using OpenGL under Linux. *Behavior Research Methods, 38,* 142-145.

Stewart, N. (2006a). A PC parallel port button box provides millisecond response time accuracy under Linux. *Behavior Research Methods, 38,* 170-173.

Steyvers, M. & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29,* 41-78.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47,* 1-29.

Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition, 32,* 31-38.

Author Note

Neil Stewart, Department of Psychology, University of Warwick; Andrew W. Ellis, Department of Psychology, University of York.

We thank James S. Adelman, Gordon D. A. Brown, and Stian Reimers for insightful discussion and Alan Garrec and Petko Kusev for their help running the experiment. Raw data are available from http://www.stewart.psych.warwick.ac.uk/.

Correspondence concerning this article should be addressed to Neil Stewart, Department of Psychology, University of Warwick, Coventry, CV4 7AL, UK. E-mail: neil.stewart@warwick.ac.uk.

Table 1

The Frequency of Early and Late Boards

Session	Early	Late
1	768	0
2	256	512
3	256	512
4	256	512
5	384	384

Figure Captions

Figure 1. An example of a checkerboard stimulus.

Figure 2. Mean proportion of correct responses by block. Blocks within the same session are joined. Error bars are standard error of the mean.

Figure 3. 10% trimmed mean correct reaction time by block. Blocks within the same session are joined. Error bars are standard error of the mean.

Figure 1

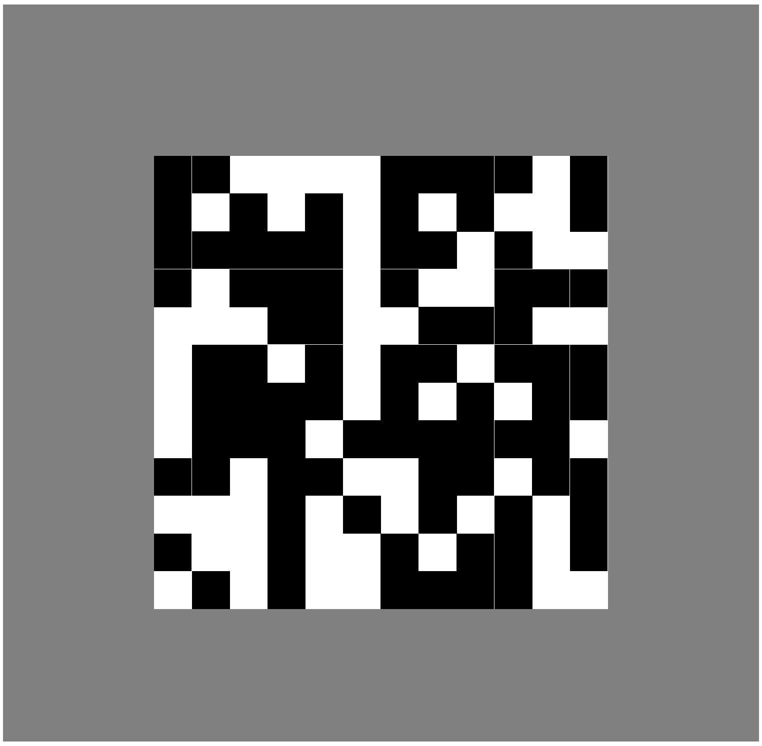


Figure 2

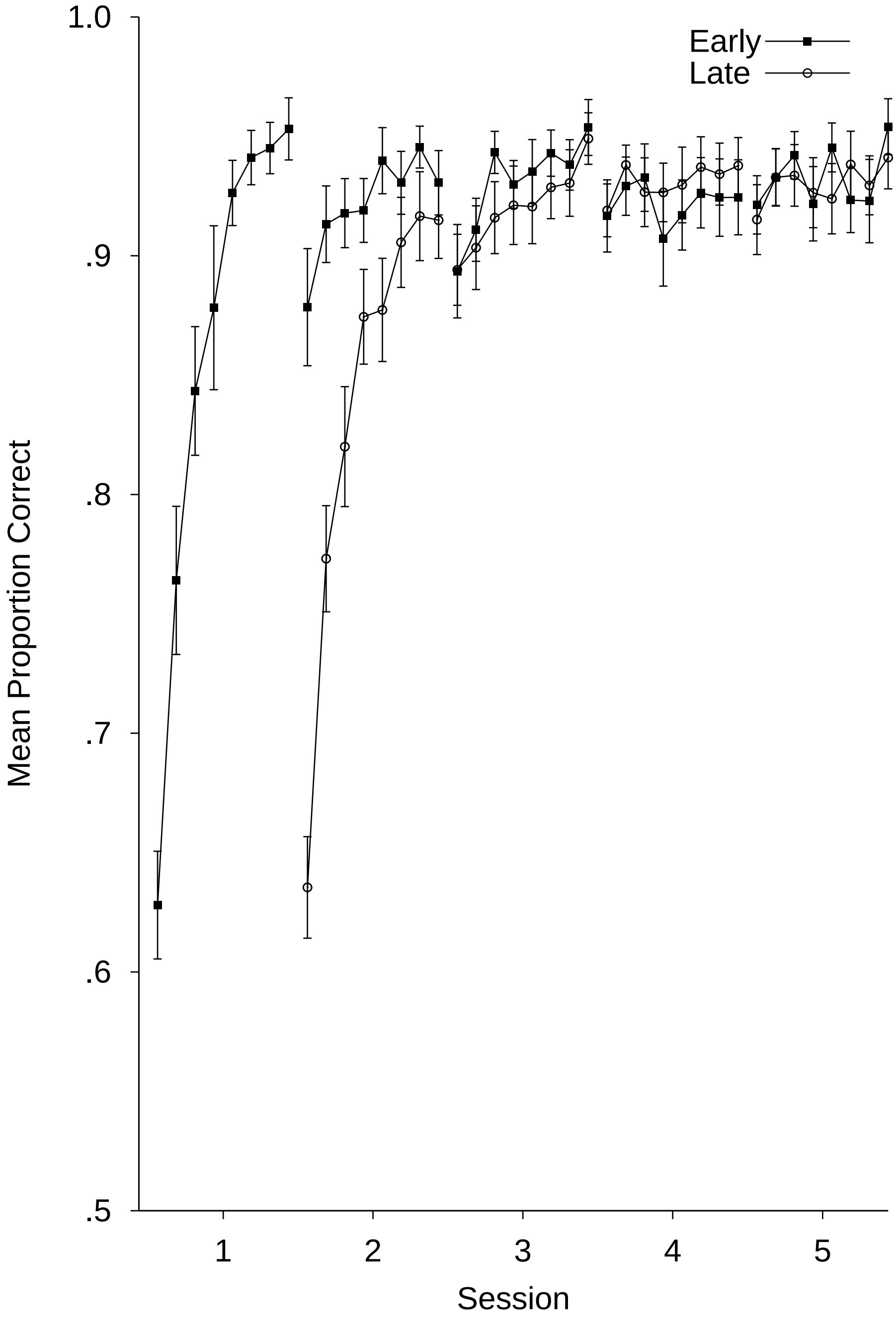
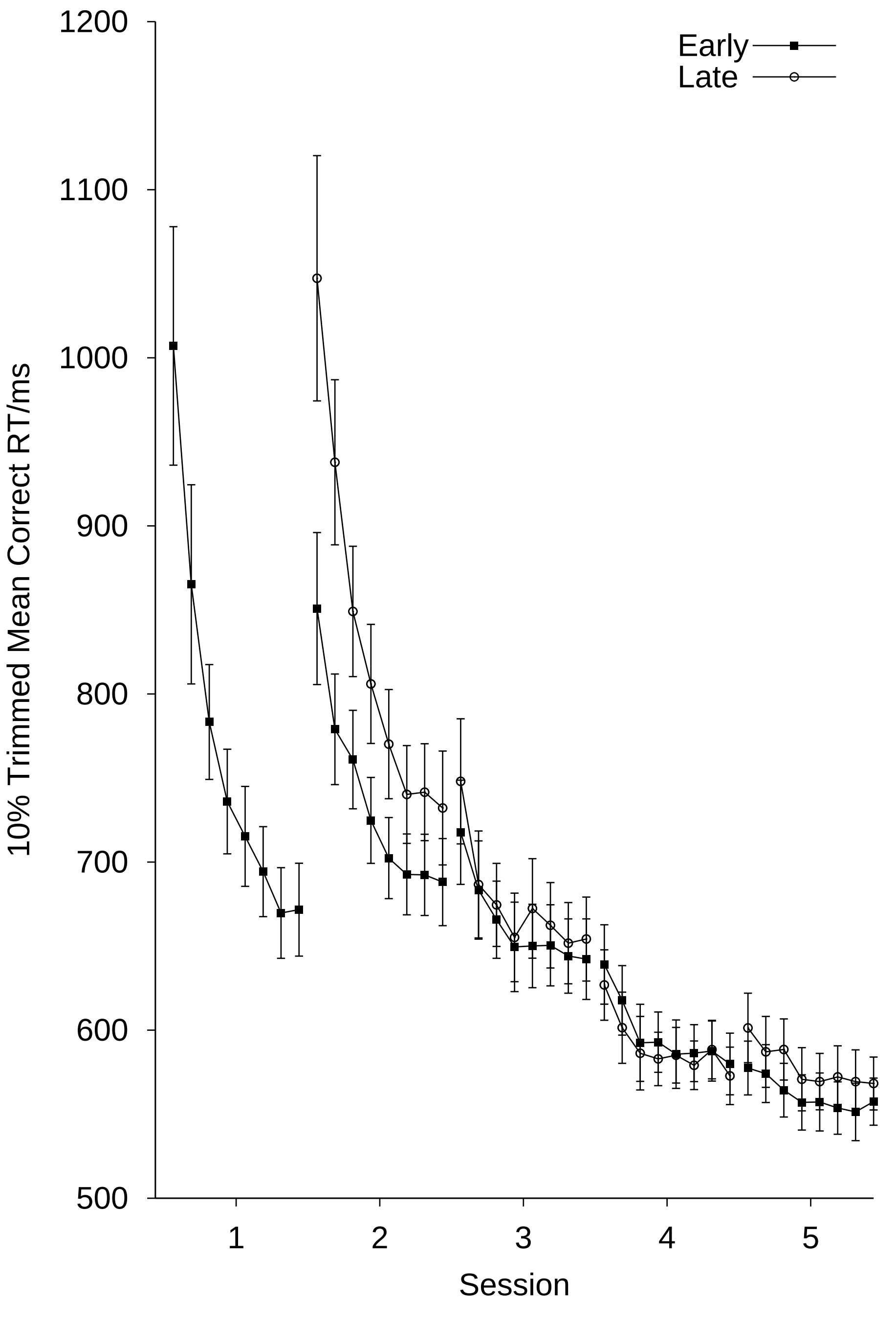


Figure 3


1200
1100
1000
900
800
700
600
500
10% Trimmed Mean Correct RT/ms
1
2
3
4
5
Session
Early
Late