Cognitive control of sequential knowledge in 2-year-olds: Evidence from an incidental sequence learning and generation task

Andrew J. Bremner[1,2,3], Denis Mareschal[2], Arnaud Destrebecqz[3], & Axel Cleeremans[3]

[1]Goldsmiths, University of London

[2]Centre for Brain and Cognitive Development, Birkbeck, University of London

[3]Cognitive Science Research Unit, Université Libre de Bruxelles

Address for correspondence: Dr Andrew J. Bremner, Department of Psychology, Whitehead Building, Goldsmiths, University of London, New Cross, London, SE14 6NW. United Kingdom. Tel. +44 (0) 20 7078 5074. Fax. +44 (0) 20 7919 7873. Email: a.bremner@gold.ac.uk

Word Count: 3855

References: 22

Running Head: Cognitive control in 2-year-olds

Acknowledgements: AC is a research director of the National Fund for Scientific Research (Belgium). This research was supported by European Commission grants HPRN-CT-2000-00065 and NEST-516542. Additional support was provided by an ESRC postdoctoral fellowship (T026271357) awarded to AB and by an institutional grant from the Université Libre de Bruxelles to AC. We thank the parents and children who took part in this study; Peter Willatts, Gethin Hughes, Robert French and the BMLF project members for helpful discussion, and also Sarah Fox and Ágnes Volein for assistance in data collection.

Abstract

Thirty-eight two-year-olds were trained under incidental instructions on a six element deterministic sequence of spatial locations. Following training, participants were informed of the presence of a sequence and asked to either reproduce or suppress the learned material. Children's production of the trained sequence was modulated by these instructions. When asked to suppress the trained sequence they were able to increase generation of paths that were not from the training sequence. Their performance was thus dependent on active suppression of knowledge rather than on a random generation strategy. This degree of control in two-year-olds stands in stark contrast to 3-year-olds' failure to control explicitly instructed rule-based knowledge (as measured the Dimensional Change Card Sort Task). We suggest that this is because the incidental nature of the learning enables the acquisition of a more procedural form of knowledge with which this age-group have more experience prior to the onset of fluent language.

Cognitive control of sequential knowledge in 2-year-olds: Evidence from an incidental sequence learning and generation task

Research into cognitive control in infancy and early childhood is central to our understanding of the origins and development of cognition. As well as establishing that a given age-group has attained a certain level of knowledge or conceptual complexity, it is equally important to determine the extent of control that they have over this knowledge. Knowledge that cannot be controlled and used appropriately is of little value.

One popular test of cognitive control is the 'Dimensional Change Card Sort' (DCCS) task (e.g., Kirkham, Cruess & Diamond, 2004; Kloo & Perner, 2005; Munakata & Yerys, 2001; Zelazo, Frye & Rapus, 1996). In this task, children are asked to sort bivalent cards (e.g. red cars and blue rabbits) according to one of two dimensions (e.g. by colour). After successfully sorted the cards by the first dimension, they are asked to switch to sorting by a second dimension (e.g. by shape not colour). Despite responding correctly to questions concerning the game rules, 3-year-olds typically fail to switch the rule by which they sort. By 4 years, children are typically able to switch rule. Explanations of this developmental change encompass a wide variety of executive functions such as changes in an ability to inhibit attentional inertia (Kirkham et al., 2004), an ability to modulate the perspective one takes of a single object (Kloo & Perner, 2005), and an ability to integrate hierarchical rule structures (Zelazo, 2004). Importantly, all such developmental accounts address changes in children's ability to manipulate or inhibit mental representations of the stimulus features and of rules acquired through explicit instruction.

Given that success in the DCCS task depends on an ability to control knowledge acquired through explicit instruction, it is worth asking whether knowledge acquired under incidental instructions might follow a different developmental trajectory. Incidental learning paradigms are frequently used in the adult learning literature to examine the acquisition of putative 'implicit' knowledge (Cleeremans, Destrebecqz & Boyer, 1998) - knowledge that is in some way inaccessible to explicit report (Shanks & St. John, 1994). Moreover, incidental learning (e.g., through the observation of peer and adult activities) is a central form of early learning, prior to the onset of fluent language (Rogoff, 1990).

The serial reaction time (SRT) task involves teaching adult participants a sequence of motor responses under incidental instructions (Nissen & Bullemer, 1987; Cleeremans & McClelland, 1991). Participants respond to a series of visual cues by pressing a corresponding key as quickly as possible. Unknown to them, the material contains sequential structure. After training, knowledge of the sequential regularities is probed through direct and indirect measures (Jiménez, Méndez & Cleeremans, 1996). Here, we report on an adaptation of this paradigm that makes it possible to explore 2-year-olds' ability to control sequence knowledge acquired incidentally.

To assess cognitive control of knowledge learned in the SRT task, Destrebecqz & Cleeremans (2001) adapted Jacoby's (1991) "Process Dissociation Procedure" (PDP) for use with the SRT. The PDP compares performance in two separate tasks: (1) an *inclusion task*; in which learned material should be reproduced, and (2) an *exclusion task*; in which learned material should be suppressed. This involves asking trained participants to generate sequences of keypresses that either resembles (inclusion) or differs from (exclusion) the training sequence as much as possible. In exclusion, participants must first activate the learned response and then inhibit this and select another response. Unlike traditional cognitive control tasks such as the DCCS, learning the sequential regularities contained in the SRT material is not based on mastering explicit rule structures, but rather on intentional use of incidentally acquired knowledge. In the following experiment, 2-year-olds were first taught one of two six-element deterministic sequences of spatial locations. We then tested their ability to control the learned knowledge by comparing generation performance under inclusion and exclusion instructions.

Method

Design

Children were trained on one of two six-element sequences of spatial locations on a play-board (S1: A-C-B-D-A-B, or S2: C-A-D-B-A-B). The elements of the sequence were always consistent in their spatial relations to one another, but their locations with respect to the play-board were varied between children (see Figure 1).

--Figure 1 about here--

S1 and S2 were balanced for the frequency of individual elements (A and B occur twice while C and D occur only once in both S1 and S2), and for the number of predictable elements given one or two elements of context. The sequential differences between S1 and S2 make it possible to assess learning by comparing participants' generation of material from the training sequence with their generation of material from the control sequence.

Following training, the children were asked to perform a generation task under either inclusion or exclusion instructions. Generation conditions (inclusion or exclusion), sequence (S1 or S2), and element locations (arrangements 1, 2, 3 and 4) were counterbalanced between participants.

Participants

Sixty 2-year-olds took part in this study. Usable data were obtained from 38 participants (26 girls) with a mean age of 723 days (24.1 months); SD = 8.3 days). Of the 22 excluded participants, 9 refused to complete the task, 1 was excluded due to experimenter error, 2 were excluded due to interference by their parent, 7 failed to meet the minimum training requirement (5 sequence repetitions), and 3 failed to meet the minimum generation task requirement (to have visited each location at least once). The children were selected based on voluntary parent participation in the research programme.

Materials

The play-board (Figure 1) was 60 centimetres in diameter. Each of its four locations was marked with a picture of an object (a sofa, a chair, a hat and a table). Other material consisted of two toy cats, and six toy dogs. The child was seated on their parent's lap, with the play-board placed on the table directly in front of them. The experimenter sat across the table facing the child. Sessions were recorded on videotape for later coding.

Procedure

The experimenter explained that the study involved a chasing game in which the experimenter would move a cat from place to place on a play-board, and that the child's task was to chase the cat with a toy dog as quickly as possible. The parent was asked to encourage the child to chase the cat, but not to prompt them to move in any particular direction. Neither the parent nor the child were told that the task contained sequential structure.

Training phase

The experimental session began once the child was seated. To begin, the play-board was covered by a sheet of cardboard, on which all six toy dogs were placed. The child was encouraged to pick her favourite dog. Once she had picked up a dog, the other five were removed. The experimenter then introduced a toy cat, and explained that in the game, 'I (experimenter) will be the cat, and you (participant) will be the dog'.

Once the child had successfully followed the cat to two successive practise locations on either side of the midline, the play-board was revealed, and the experimenter exclaimed, 'Look at all these places where the cat can hide from the dog!'. 'Can the dog catch the cat here?' The experimenter would then place the cat on the first location in the training sequence. Once the child had placed her dog in the same location on the play-board, the experimenter would move the cat to the next location in the sequence. This was repeated until the child had chased the cat to all locations in the sequence. The sequence was repeated a minimum of 5 times and a maximum of 12 times. To keep the child interested in the game for as long as possible, she was given the opportunity to chase with different toy dogs. This change in dogs always occurred between repetitions of the sequence.

Finally, we used a participant-controlled variable training procedure. The training phase was terminated if the child became too disinterested to continue the training, or if a maximum of 12 repetitions of the sequence (blocks) had been reached.

Generation phase

The experimenter introduced the generation phase directly following the training phase: 'this time you are going to be a cat and I'll be a dog'. They were then told that during the first game 'the cat was always running away in a special way, from place to place'. They were then prompted with the first two locations that the cat had visited; the first two elements of the training sequence (S1: A-C, S2: C-A). At this point, the procedure for children in the inclusion and exclusion conditions differed.

Children in the inclusion condition were asked, 'do you think you can remember which way the cat went next?' 'Can you go the same way as the cat was going before?' The experimenter then encouraged the child to pick up the cat (placed at the second of the two prompt locations) and move it to a new location. Throughout the session the experimenter reminded them to go 'the same way as the cat was going before'. Only children who visited each of the locations at least once were included in analysis.

Children in the exclusion condition were shown a new cat. The experimenter told the child, 'this is a different cat, and this cat goes a different way than the other cat'. The experimenter then prompted the child with two locations comprising a transition that had not been present in the training sequence (S1: A-D, S2: C-B). Throughout the session the experimenter reminded them to go 'a different way than the other cat'. Only children who visited each of the locations at least once were included in analysis.

Children were encouraged to generate a minimum of 4 and a maximum of 18 transitions. The sequences generated by each child were coded from video records. If children visited the same location consecutively (e.g. A-A) only one visit to that location was included in the scored generation sequence. Inter-observer reliability was estimated by comparing the coded generation sequences of 12 randomly selected children (6 from the inclusion condition and 6 from the exclusion condition) with those of a second observer. This was achieved by aligning the sequences with reference to the largest continuous string of agreements. Cohen's κ was then calculated, yielding a satisfactory reliability of .85.

Results

Children were trained on a mean of 7.2 (SE=.23) repetitions of the training sequence (Inclusion: M=7.4, SE=.35; Exclusion: M=7.1, SE=.30. $t(36)=.69$, ns.). The proportion of generated pairs (e.g. A-C), triplets (e.g. A-C-B) and quadruplets (e.g. A-

C-B-D) that were part of the training sequence was then calculated for each child by dividing the number of generated pairs, triplets and quadruplets from the training sequence, by the total number of pairs, triplets and quadruplets generated. The mean proportions are shown in Figure 2.

--Figure 2 about here--

We conducted a mixed-design ANCOVA on children's scores for the proportion of generated chunks from the training sequence, with 1 within-subjects factor (Length of Chunk: pair, triplet or quadruplet), 1 between-subjects factor (Instructions: inclusion or exclusion), and one covariate (Number of training blocks received). This analysis revealed a significant effect of 'Length of Chunk' ($F(2,70)=7.3$, $p_{rep}=.99$, $\eta_p^2=.172$). Children produced fewer long chunks from the training sequence than short chunks. This is because the probability of making an error increases with the increasing length of the chunk. There was also a significant effect of Instructions ($F(1, 35)=4.0$, $p_{rep}=.88$, $\eta_p^2=.102$). Children produced less of the training material under exclusion than under inclusion instructions. No other effects or interactions reached significance ($Fs<1$).

Despite the effect of Instructions, we cannot conclude from this evidence alone that participants could control their expression of the training sequence based on knowledge acquired during the training phase. The inclusion and exclusion scores might also reflect controlled expression of non-sequential information, such as the frequencies of the different locations (which are unequal in the training set) or simple spatial patterns, such as the frequencies of reversals (which are rare in the training set; S1: A-B-A, S2: B-A-B).

To determine whether the effect of Instructions was due to differential expression of genuine sequential knowledge, we compared the conditional probabilities associated with particular generated pairings of elements (Jiménez et al., 1996). We examined the probabilities that, prior to a specified target element, the children had generated more often the specific context element corresponding to the training sequence rather than the context element that corresponds to the other sequence or the context element that corresponds to both sequences. Moreover, the target pairings were selected such that their grammaticality differed with respect to training sequences S1 and S2 (e.g., they were grammatical in S1 but not S2). Generation scores computed according to S1 or S2 can then be used as a control for each other. This data also provides a measure of the degree to which the sequences have been learned by comparison to a baseline chance level of performance.

We thus compared the probabilities of children having generated the elements C and D prior to generation of B (denoted as C|B and D|B respectively)[1]. As element A appears before element B in both S1 and S2 we did not compare A|B between conditions. Because it is possible to generate one of three elements prior to B (repetitions were not allowed), the baseline probability for C|B and D|B is 0.33. S1 children were trained on element C appearing before B (A-C-B-D-A-B), whereas S2 children were trained on element D appearing before B (C-A-D-B-A-B). Thus, if children in the inclusion condition have learned the sequence, we would expect higher probabilities associated with C|B than with D|B for those children trained on S1, and higher probabilities associated with D|B than with C|B for those children trained on S2. We would also expect generation of C|B and D|B to be above chance for children trained on those particular transitions (children trained on S1 and S2 respectively). Moreover, children in the exclusion conditions should avoid generating strings that are grammatical in their taught sequence and, thus, should be more likely to generate sequences that are in fact grammatical in the alternative (untaught) sequence. In other words, those in S1 should show higher D|B than C|B generation, whereas those in S2 should show higher C|B than D|B generation. Figure 3 shows that this is the case.

--Figure 3 about here--

We analysed the conditional probabilities of children's generation of C|B and D|B using a mixed-design ANCOVA. The ANCOVA included 1 within-subjects factor (Context Element; C or D), 2 between-subjects factors (Training Sequence: S1 or S2; Instructions: inclusion or exclusion), and one covariate (Number of training blocks received). This analysis revealed a significant 3-way interaction of Context Element X Training Sequence X Instructions ($F(1,33)=9.7$, $p_{rep}=.98$, $\eta_p^2=.227$). No other effects or interactions were significant ($Fs<2$). This interaction confirms that children were able to control their expression of sequential information learned during training according to instructions. To explore this further, we conducted separate conditional probability analyses within the inclusion and exclusion instruction conditions.

Inclusion performance

A mixed-design ANCOVA with 1 within-subjects factor (Context Element: C or D), 1 between-subjects factor (Training Sequence: S1 or S2) and 1 covariate (Number of training blocks received) revealed a significant interaction of Context Element X Training Sequence ($F(1,16)=4.6$, $p_{rep}=.88$, $\eta_p^2=.224$). This interaction describes the difference in conditional probabilities associated with generation of C|B and D|B following training on S1 (where C|B is grammatical) and S2 (where D|B is grammatical). The children in the inclusion condition were more likely to have generated a grammatical path in both conditions. This indicates that the 2-year-olds had at least partially learned and were able to express the sequence that they had been trained on. No other effects were significant ($Fs<1$). The probabilities associated with production of C|B by S1 children and D|B by S2 children (see Figure 3a) were, as predicted, both significantly greater than chance (.33) ($t(8)=1.8$, $p_{rep}=.88$, $d=.59$; and $t(9)=2.0$, $p_{rep}=.89$, $d=.64$).

Exclusion performance

A mixed-design ANCOVA with 1 within-subjects factor (Context Element: C or D), 1 between-subjects factor (Training Sequence: S1 or S2) and 1 covariate (Number of training blocks received) revealed a significant interaction of Context Element X Training Sequence ($F(1,16)=4.8$, $p_{rep}=.88$, $\eta_p^2=.232$). This interaction describes the difference in conditional probabilities associated with generation of C|B and D|B following training on S1 (where C|B is grammatical) and S2 (where D|B is grammatical). Children in the exclusion condition were more likely to generate an ungrammatical path in both conditions. We conclude that the children tested above were able to suppress the expression of the training sequence by reference to their knowledge of that sequence. No other effects were significant ($Fs<2$). We made no specific predictions concerning the generation of grammatical pairs with respect to baseline (.33) under exclusion instructions. However, it is interesting to note that the probabilities associated with generation of grammatical paths for both S1 and S2 trained children were significantly below chance ($t(8)=2.1$, $p_{rep}=.90$, $d=.72$; and $t(9)=1.9$, $p_{rep}=.88$, $d=.61$).

Discussion

Following incidental training on a sequence of spatial locations, 2-year-olds were asked to either reproduce or suppress their knowledge of the sequence. Analyses revealed that: (i) 2-year-olds' production of the trained sequential material is modulated by these instructions, and (ii) those asked to suppress the trained material were able to increase their generation of sequence paths that were not part of the training sequence. Thus, exclusion instructions resulted in active suppression of knowledge of the training sequence rather than in a random generation strategy. These results provide evidence of incidental sequence learning in 2–year-olds, and add to the growing evidence of cognitive flexibility in early childhood (Deák, 2003).

Our findings contrast strikingly with 3-year-olds' performance on other measures of cognitive control, which is characteristically inflexible (e.g. Zelazo et al., 1996). One potential explanation of the relative ease with which children control their knowledge in the current task is the dimensional complexity involved in the task switch. Perner & Lang (2002) found that 3- and 4-year-olds are more successful at a version of the DCCS which requires intradimensional or 'reversal' switches (e.g. changing from sorting red to red and blue to blue, to sorting red to blue and blue to red) rather than interdimensional switches. Nevertheless, some intra-dimensional shift control tasks remain a significant challenge to children under 4 years (Russell, Mauthner, Sharpe & Tidswell, 1991; Russell, Hala & Hill, 2003; Russell, Jarrold & Potel, 1994).

The most salient difference between the current task and those which preschool children find difficult is that the former involves control of incidentally acquired knowledge acquired through a motor schema, rather than control of declarative, rule-like knowledge acquired through explicit instruction. Our results suggest that young children have more skill in manipulating the former rather than the latter. Intuitively, this is congruent with the fact that learning that occurs before the onset of fluent language tends to be incidental rather than instructed.

An ability to control knowledge in an inclusion/exclusion task is generally taken as an indication that the relevant knowledge is explicit (Jacoby, Toth, & Yonelinas, 1993). However, this need not be the case. Indeed, in this study as in others using the PDP, children can base their generated responses on a feeling of "familiarity" (Richardson-Klavehn, Gardiner & Java, 1996) rather than on any explicit knowledge of the sequence structure. Familiarity could take the form of sensitivity to the trained transitions themselves or to the motor responses associated with the trained transitions. Thus, the 2-year-olds may have favoured specific transitions in the inclusion task (and avoided those transitions in the exclusion task) simply because these were more familiar. We suggest that children's ability to control this less explicit form of knowledge (Dienes, Altmann, Kwan & Goode, 1995) can help explain why our results depart from those of previous studies of cognitive control in 3-years olds (e.g., Kirkham et al., 2004; Kloo & Perner, 2005; Munakata & Yerys, 2001; Zelazo et al., 1996). Incidental learning and control tasks may thus provide an important addition to the executive control literature, as they allow control of sub-explicit knowledge to be measured.

References

Cleeremans, A., Destrebecqz, A., & Boyer, M. (1998). Implicit learning: News from the front. Trends in Cognitive Sciences, 2, 406-416

Cleeremans, A. & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R.M. French & A. Cleeremans (Eds.), Implicit Learning and Consciousness. Hove, UK: Psychology Press, pp. 1-40.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of sequence events. Journal of Experimental Psychology: General, 120, 235-253.

Deák, G. O. (2003). The development of cognitive flexibility and language abilities. In R. Kail (Ed.), Advances in child development and behavior, Vol. 31 (pp. 271-327). San Diego: Academic Press.

Destrebecqz, A. & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. Psychonomic Bulletin & Review, 8, 343-350.

Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. Journal of Experimental Psychology: Learning, Memory and Cognition, 21, 1322-1338.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. Journal of Memory and Language, 30, 513-541.

Jacoby, L. L., Toth, J.P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. Journal of Experimental Psychology: General, 122, 139-154.

Jiménez, L., Méndez, C. & Cleeremans, A. (1996). Direct and indirect measures of sequence learning. Journal of Experimental Psychology: Learning, Memory and Cognition, 22, 948-969.

Kirkham, N. Z. , Cruess, L. M., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. Developmental Science, 6, 449-467.

Kloo, D., & Perner, J. (2005). Disentangling dimensions in the dimensional change card sorting task. Developmental Science, 8, 44-56.

Munakata, Y., & Yerys, B. E. (2001). All together now: When dissociations between knowledge and action disappear. Psychological Science, 12, 335-337.

Nissen, M. J., & Bullemer, P. T. (1987). Attentional requirements for learning: Evidence from performance measures. Cognitive Psychology, 19, 1-32.

Perner, J., & Lang, B. (2002). What causes 3-year-olds' difficulty on the dimensional change card sorting task? Infant and Child Development, 11, 93-105.

Richardson-Klavehn, A., Gardiner, J. M., & Java, R. I. (1996). Memory: Task dissociations, process dissociations and dissociations of consciousness. In G. Underwood (Ed.), Implicit cognition (pp. 85-158). Oxford, UK: Oxford University Press.

Rogoff, B. (1990) Apprenticeship in thinking. Oxford, UK: Oxford University Press.

Russell, J., Hala, S., & Hill, E. (2003). The automated windows task: the performance of preschool children, children with autism, and children with moderate learning difficulties. Cognitive Development, 18, 111-137.

Russell, J., Jarrold, C., & Potel, D. (1994). What makes strategic deception difficult for children – the deception or the strategy? British Journal of Developmental Psychology, 12, 301-314.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The “windows task” as a measure of strategic deception in preschoolers and autistic subjects. British Journal of Developmental Psychology, 9, 331-349.

Shanks, D. R., & St. John, M. F. (1994). Characteristics of dissociable human learning systems. Behavioral and Brain Sciences, 17, 367-447.

Zelazo, P. D. (2004). The development of conscious control in childhood. Trends in Cognitive Sciences, 8, 12-17.

Zelazo, P. D., Frye, D. & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. Cognitive Development , 11, 37-63.

Footnotes

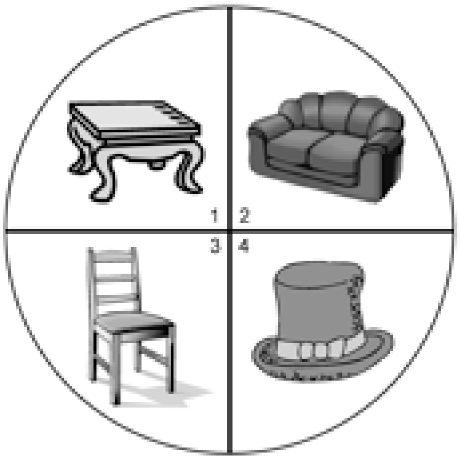
Footnote 1: This particular analysis was chosen because it is unique in testing children's learning of a pairing which: (a) was not presented across a gap in training (as would be the case if, for example, comparing the conditional probability of generating D or C after B), and (b) was not explicitly taught in the prompted pair at the beginning of the test phase (as would be the case if comparing the conditional probability of generating C or D before A). Choosing A or B as the target element in the conditional pairing is the most suitable analysis of control performance because the only possibility for participants in the exclusion condition is to generate an element from the alternate (control) sequence.

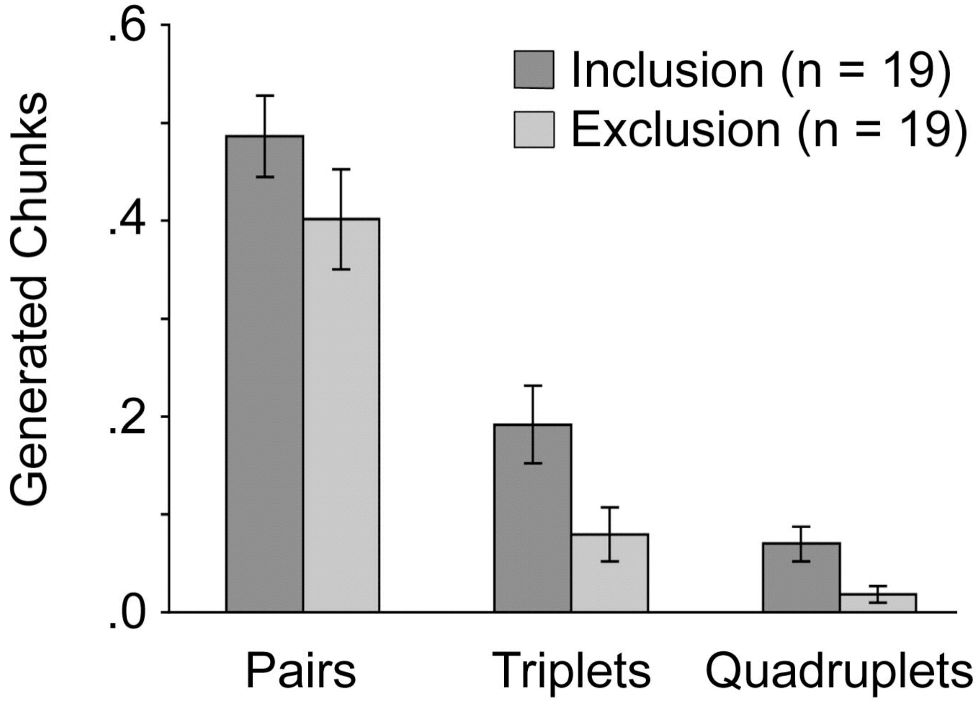
Figure Captions

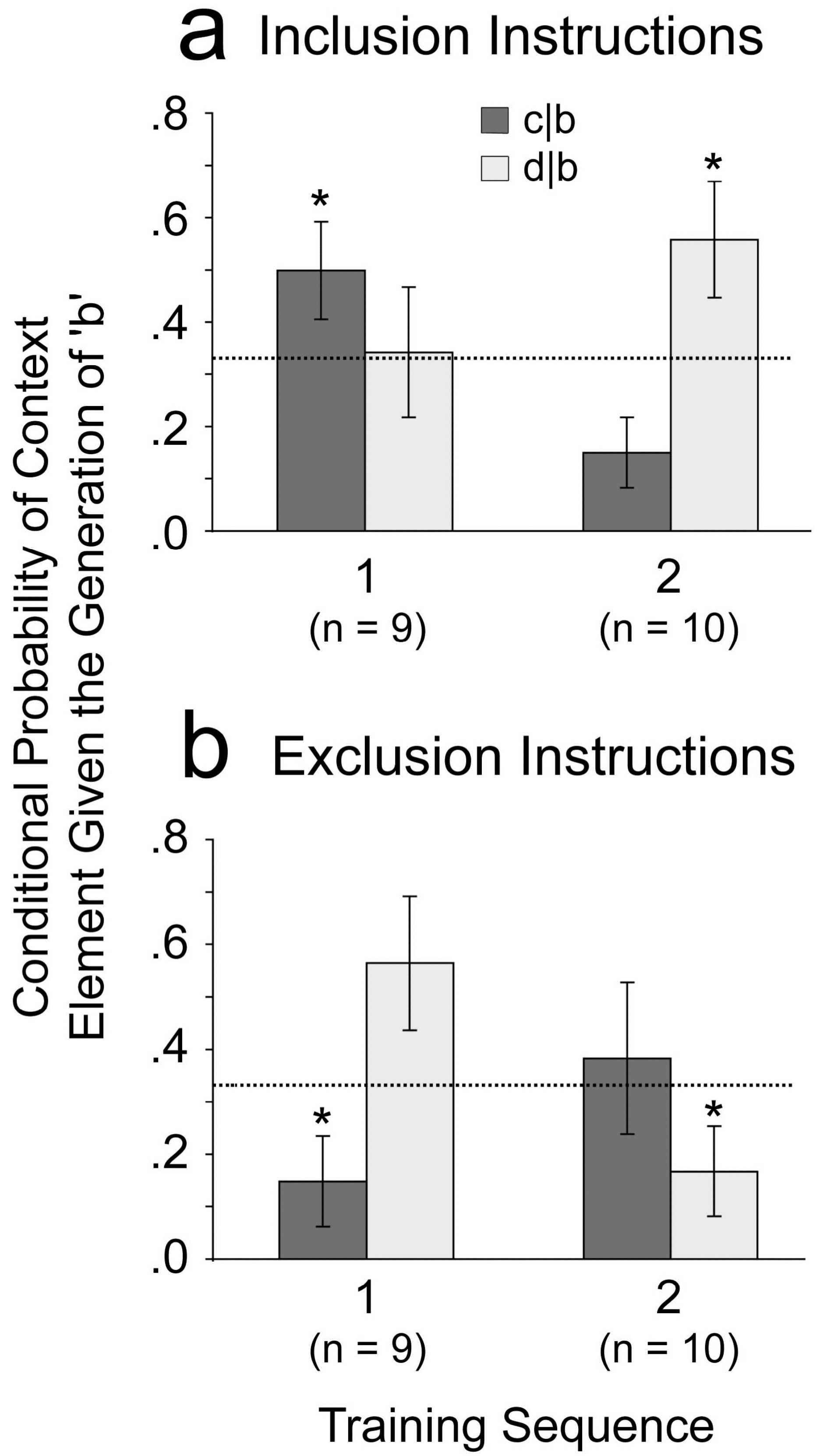
Figure 1: The play-board used in the task (number labels not visible to the participants). Children were assigned to one of four groups in which the elements of the training sequences were: (1) A;1, B;2, C;3, D;4, (2) A;2, B;4, C;1, D;3, (3) A;4, B;3, C;2, D;1, or (4) A;3, B;1, C;4, D;2.

Figure 2: Children's generation of the training sequence under both inclusion and exclusion instructions. Generation scores correspond to the number of generated pairs, triplets and quadruplets from the training sequence divided by the total number of pairs, triplets and quadruplets generated. Error bars correspond to SE.

Figure 3: Conditional probabilities associated with children's generation of grammatical and ungrammatical paths preceding 'B'. For children trained on S1 C|B is grammatical and D|B ungrammatical. For children trained on S2 D|B is grammatical and C|B ungrammatical. Error bars correspond to SE. The dotted line corresponds to the chance level of .33. *= p_{rep} >.88, when compared to 0.33.


1
2
3
4


.6
.4
.2
.0
Generated Chunks
Inclusion (n = 19)
Exclusion (n = 19)
Pairs
Triplets
Quadruplets


a Inclusion Instructions
c|b
d|b
.8
.6
.4
.2
.0
1
(n = 9)
2
(n = 10)
b Exclusion Instructions
Conditional Probability of Context Element Given the Generation of 'b'
Training Sequence