# Withstanding the test of time: multisensory cues improve the delayed retention of incidental learning


| | |
|---|---|
| Journal: | *Developmental Science* |
| Manuscript ID | DS-11-17-0449-SR.R1 |
| Manuscript Type: | Short Report (under 4000 words) |
| Keywords: | Multisensory, Cognitive Development, Incidental Learning, Retention, Audiovisual, Education |
| Subject Area: | Childhood cognitive development |
| | |

# Withstanding the test of time: multisensory cues improve the delayed retention of incidental learning



Research highlights

- Delayed retention of learning following exposure to multisensory compared to unisensory cues was examined using a novel category learning task
- A greater depth of learning was found with multisensory than unisensory information in primary school children
- Findings have implications for the use of multisensory tools in education

**Abstract**

Multisensory tools are commonly employed within educational settings (e.g., Carter & Stephenson, 2012), and there is a growing body of literature advocating the benefits of presenting children with multisensory information over unisensory cues for learning (Baker & Jordan, 2015; Jordan & Baker, 2011). This is even the case when the informative cues are only arbitrarily related (Broadbent, White, Mareschal, & Kirkham, 2017). However, the delayed retention of learning following exposure to multisensory compared to unisensory cues has not been evaluated, and has important implications for the utility of multisensory educational tools. This study examined the retention of incidental categorical learning in five-, seven- and nine-year-olds (N=181) using either unisensory or multisensory cues. Results found significantly greater retention of learning following multisensory cue exposure than with unisensory information when category knowledge was tested following a 24-hour period of delay. No age-related changes were found, suggesting that multisensory information can facilitate the retention of learning across this age range.

**Keywords:** Multisensory; Cognitive Development; Incidental Learning; Retention; Audiovisual; Education

## 1. Introduction

Educational tools often require children to effectively utilise inputs to multiple sensory modalities. Indeed, 'multisensory' tools are commonly employed within educational environments as a means to support learning, and have a long history in pedagogy (Montessori, 1912; Singleton, 2009). Most notably, multisensory techniques have been advocated for children with learning difficulties (Carter & Stephenson, 2012; Ogden, Hindman, & Turner, 1989), or as an instructional approach to reading remediation (Gillingham & Stillman, 1997).

Despite a profusion of cognitive research on the development of multisensory integration across early development (e.g., Bahrick & Lickliter, 2000, 2004, 2012; Brandwein et al., 2011; Bremner, Lewkowicz, & Spence, 2012; Burr & Gori, 2012; Ernst, 2008; Gori, Del Viva, Sandini, & Burr, 2008), little is understood regarding the efficacy of multisensory learning tools during childhood. That is, although multisensory information has been shown to support perceptual learning as early as infancy (Bahrick & Lickliter, 2000; Flom & Bahrick, 2010; Kirkham, Wagner, Swan, & Johnson, 2012; Lewkowicz, 2000; Richardson & Kirkham, 2004; Wu & Kirkham, 2010), our understanding of the extent to which children use multisensory information flexibly during classroom learning is relatively limited. Furthermore, the depth and retention of learning following exposure to multimodal compared to unimodal cues has been relatively overlooked, and may have important implications for the utility of multisensory tools for formal learning.

Multisensory components used in educational tools may only moderately resemble the stimuli used to address cross-modal sensitivity in laboratory studies. Addressing the gap between our understanding of mechanisms supporting multisensory integration and the role of multisensory tools in children's learning is therefore an important challenge to pursue. For instance, do changes in ability to integrate bimodal cues across development (e.g., Burr & Gori, 2012; Gori et al., 2008; Nardini, Jones, Bedford, & Braddick, 2008) relate to changes across childhood in the ability to use multisensory cues in children's basic learning tasks?

Recent research has shown that, compared to the use of unimodal cues, synchronous multimodal information can facilitate immediate learning in 6- to 10-year-olds (Broadbent et al., 2017). These authors found that the simultaneous presentation of complementary visual and auditory information, in which both features were informative to category membership, resulted in superior immediate incidental learning of categories than with unisensory cues. This was particularly the case for 6-year-olds, given that learning with auditory cues alone resulted in particularly poor performance. In a further study, Broadbent et al. (2018) found that an auditory concurrent task resulted in significant detriment to multisensory learning in 6-year-olds, suggesting that at this young age children may have difficulties in using information to different sensory modalities as effectively as older children. Other studies have also reported the beneficial use of synchronous multisensory information on explicit learning tasks, such as teaching reading (Joshi, Dahlgren, & Boulware-Gooden, 2002) and the acquisition of numerical abilities in children (Baker & Jordan, 2015; Jordan & Baker, 2011; Jordan, Suanda, & Brannon, 2008; Moyer-Packenham et al., 2013; Thornton, Jones, & Toohey, 1983). However, it remains unclear whether children experience a greater depth of learning from exposure to multisensory cues than unisensory cues, and which would result in superior *retention* of learning, especially since synchronous audio-visual cues can interfere with each other (Thomas, Nardini, & Mareschal, 2017). Resolving this issue is of fundamental importance for multisensory learning practices in primary school settings.



The current study aimed to examine the retention of category learning over a 24-hour delay using either unimodal or bimodal cues to inform category membership. Categorization and concept learning are central to many academic skills (Mareschal, Quinn, & Lea, 2013) and can be examined using basic, educationally-relevant tasks (e.g., Broadbent et al., 2017). A delay of 24-hours was selected to represent a typical period of retention between learning and recall used in a classroom environment and typical of other delayed-retention tasks in children (e.g., Vellutino, Steger, DeSetto, & Phillips, 1975). In addition, incidental learning is core to education; with learning objectives often encompassing the development of knowledge that has not transpired from explicit instruction. It is, therefore, important to examine the role of unisensory and multisensory cues on incidental learning across childhood. Given that the presentation of multisensory cues leads to superior immediate learning in children (Baker & Jordan, 2015; Jordan & Baker, 2011; Joshi et al., 2002), and particularly the incidental learning of categorical information (Broadbent et al., 2017), we hypothesized that when categorical information is presented simultaneously to more than one modality, this would result in greater retention of incidental learning than with unisensory cues. In light of previous findings of a protracted development of multisensory integration (Burr & Gori, 2012; Gori et al., 2008; Nardini, Bales, & Mareschal, 2015) and ability to utilize multisensory information (Broadbent et al., 2017), developmental differences across middle childhood in the extent to which multisensory information is supportive of learning were also anticipated.



## 2. Methods

### 2.1 Participants

Data from 181 children were included in the study. Participants were selected from three separate age groups; ‘5-year-olds’, N=60, mean age (years) = 5.63, SD= .28, (N= 31 males); ‘7-year-olds’, N=61, mean age= 7.69, SD= .26 (N= 30 males); and ‘9-year-olds’, N=60, mean age = 9.76, SD = .32 (N= 28 males). Participants in each group were randomly allocated to one of three learning conditions: Visual (unisensory), Auditory (unisensory) or Audiovisual (multisensory), in a between-subjects design; N=20 per condition (except for 7-year-olds, N=21 in the Auditory condition). Children from each school were allocated across the three conditions, to control for school differences. Sample sizes for each group, per condition, were determined by power analysis for ANOVA with df =1, f = 0.40.

Children were recruited from local schools and informed written parental consent was obtained for each participant, in accordance with the University ethics committee guidelines. All participants had normal hearing and normal (or corrected-to-normal) vision, and no known developmental or neurological disorder, as assessed on the parental consent form. All testing sessions were conducted in a quiet room within the participant’s school and children were rewarded for participating with a certificate and stickers. Testing sessions for each participant were conducted over two consecutive days. Session 1 on day 1 lasted approximately 15 minutes. Session 2 was conducted on day 2 to examine the retention of incidental learning and lasted approximately 5 minutes. Mean time (hours) between test sessions = 23.04 (SD = 1.13).

**2.2 Stimuli**

The Multisensory Attention Learning Task (MALT; Broadbent et al., 2017) is a computerised continuous performance task, developed to examine the role of unimodal and multimodal information on attentional vigilance and incidental learning of categorical information. For full details of stimuli and methods, see Broadbent et al. (2017). In brief, visual stimuli consisted of seven different animal line drawings; one target animal ('frog') and six non-target animals ('owl', 'dog', 'goat', 'pig', 'elephant', and 'cat'), presented on a 15" laptop screen approximately 50cm in front of the participant. Auditory stimuli consisted of congruent animal sounds, consistent with the different visual animal stimuli.

Three different learning conditions were used, in a between-subjects design. In the visual condition, contrasting visual features were used to distinguish between two different categories ('families') of frogs. Frogs from family 1 had few spots (2 or 3), varying in size and colours across category members. Members within family 2 had many spots (7 or 8), varying in colours and sizes consistent with members from family 1. Non-target animals were similarly marked with spots of varying colours, size and number, for consistency across stimuli. In this condition, auditory stimuli remained consistent across exemplars so that there was no informative audio dimension. That is, for target stimuli (frogs), only one of the two auditory-cue 'families' (see below for further details) was used, counterbalanced across participants.

In the auditory condition, only auditory features were used to differentiate family members, distinguishable by two different frog croaks. Family 1 exemplars croaked with a 'high and long-short' sound, whilst family 2 exemplars croaked with a 'deep and short-long' croak. In this condition, visual stimuli remained consistent (only visual frogs from one category were used, counterbalanced across participants), so that there was no informative visual dimension.

In the audiovisual (multisensory) condition, both visual and auditory features could be used to discriminate category membership. Therefore, given that these cues were presented redundantly in this condition, participants, did not have to attend to both in order to succeed on the task or learn categorical information. For example, family 1 members had few spots (visual) and a long-short croak (auditory), whilst family 2 members had many spots and a short-long croak. The two possible combinations of categorizing audiovisual features were counterbalanced across participants.

The saliency and discriminability level of the visual and auditory features of target exemplars were examined using a discriminability task and pilot study, reported in Broadbent et al (2017). The authors found no significant differences in children's ability to discriminate across category boundaries in these two conditions, suggesting the visual and auditory tasks were of equal difficulty and salience.



## 2.3 Procedure

### 2.3.1 Session 1.

*2.3.1.1 Auditory working memory*

As a measure of auditory working memory, each participant initially completed the Digit Span Backwards (DSB) task from the British Ability Scales–II (BAS-II; Elliott, Smith, & McCulloch, 1996).

*2.3.1.2 Multisensory attention learning task (MALT)*

Before presentation of the MALT, a short audio and visual detection task was conducted to familiarise participants with task stimuli. Participants were shown one of each animal in turn and asked whether they could hear and see the exemplar. All participants answered affirmatively for each of the seven exemplars and continued with the task.

For the MALT task (Fig. 1), participants were instructed to press the space bar as quickly as possible whenever a frog (target) appeared on the screen, whilst inhibiting a response to any other animal stimuli. The task screen consisted of a white screen with an image of a lily pad in the top left-hand corner and an image of a log in the top right-hand corner. Stimuli were presented individually in the centre of the screen for 300ms. If the space bar was (correctly) pressed after the presentation of a target stimulus, the same frog reappeared in a net. The frog then immediately travelled to the top left- or top right-hand corner of the screen to the correct frog habitat; i.e., frog exemplars from one family consistently travelled to the lily-pad habitat, whilst frog exemplars from the other family travelled to the log habitat, counterbalanced across participants.

FIGURE 1 HERE



The task consisted of up to 200 trials, separated into four blocks by a motivation screen to allow for rest-breaks. Across the task, target stimuli (frogs) were presented on 40 percent of trials (80 trials; 40 exemplars from each family). Twenty of each non-target (distractor) stimuli were presented randomly throughout the task. Completion of the task was determined either by 50 correct responses to frog targets, or until the maximum 200 trials were completed. Participants were scored as having reached criterion or not, and data were only analysed from those who met the ‘50 correct’ criteria. Therefore, all participants included in the analyses had received the same number of category learning trials (having observed 50 frogs travelling to their correct habitat). All participants except for two 5-year-olds met response criteria.

**2.3.2 Session 2.**

*2.3.2.1 Retention of category learning.*

To examine the retention of incidental category learning on the MALT after a 24-hour delay, participants were asked to complete a category identification task the following day. Eight exemplars from each category of the given learning condition (total = 16 trials) were presented in a random order. Participants responded verbally to whether the frog had lived at the lily pad or the log during the game that had been presented on the previous day, and the researcher pressed the correct habitat image positioned on the keyboard on keys ‘z’ and ‘m’, respectively. Participants viewed each frog individually, and no feedback was given throughout the identification task. Total correct categorization responses were recorded.

## 3. Results

### 3.1 Session 1

**3.1.1 Auditory working memory.** Digit Span Backwards (DSB) raw scores were converted to standardised T-scores and compared across groups using a one-way analysis of variance (ANOVA). A significant difference was found between groups; 5 years: M = 62.97, 95% CI [60.73, 65.21]; 7 years: M = 54.75, 95% CI [52.53, 56.97]; 9 years: M= 57.58, 95% CI [55.35, 59.82], (F (2, 180) = 12.88, $p< .001$), with 5-year-olds performing at a cognitive level significantly higher than the 7- and 9-year-old groups ($p< .001$ and $p= .004$, respectively). To confirm that the 5-year-olds were performing significantly below the older age groups in raw ability score, these data were also analysed. Results showed a significant effect of Age; 5 years: M = 8.27, 95% CI [7.33, 9.20]; 7 years: M = 11.59, 95% CI [10.67, 12.52]; 9 years: M = 15.33, 95% CI [14.40, 16.27], F(2, 172) = 55.89, $p< .001$, with significant differences between all groups (all $p< .001$).

DSB T-scores for each condition (collapsed across age groups) were also compared; Visual: M= 57.88, 95% CI [55.59, 60.18]; Auditory: M = 57.03, 95% CI [54.89, 59.17]; Audiovisual: M= 60.37, 95% CI [57.48, 63.25]. Results found no differences in working memory ability of children allocated to any sensory condition, F (2, 180) = 1.98, $p=.14$.

**3.1.2 Multisensory attention learning task (MALT).** To examine performance across groups on aspects of sustained attention on the learning element of the MALT, trials to criterion and accuracy (*d’prime*) scores were calculated. To examine differences across the two between-subjects factors; Age Group (3 levels : 5, 7, and 9 years) and Condition (3 levels: V, A and AV), data were analysed using mixed ANOVAs for each dependent variable (trials to criterion and d'prime), with Bonferroni-corrected pairwise comparisons for post-hoc analyses.

*3.1.2.1 Trials to criterion*

The mean number of learning trials on the MALT to reach the criterion of 50 correct target responses was calculated for each group; 5 years: M = 137.28, 95% CI [134.25, 140.32], 7 years: M = 139.89, 95% CI [136.88, 142.89], 9 years: M = 139.03, 95% CI [135.99, 142.07]. Results of the mixed ANOVA found no significant effects of Age or Condition, and no Age by Condition interaction was found (all Fs<1).

*3.1.2.2 Accuracy score (d'prime)*

To examine target-detection accuracy on the MALT, z-scores for Hit rates (Hr = total correct hits/ total target trials) and False alarm rates (FAr = Commission errors/ total non-target trials) were calculated. A *d'prime* [*d'*= z(Hr) - z(FAr)] measure of sensitivity was then calculated and mean values were analysed across groups. Results of a mixed ANOVA found a significant main effect of Age, F(2, 172) = 6.34, *p= .002*, partial $\eta^2$ = .07, driven by 5-year-olds (M= -.48) having significantly lower *d'prime* accuracy scores than 9-year-olds (M= .52), *p= .001*. No significant effect of Condition was found, F(2, 172) = 1.13, *p=.326*. In addition, no significant Age by Condition interaction was found, F (4, 172) = 2.31, *p= .06*. Analyses indicate a reliable improvement with age on the MALT, but with no significant differences across the sensory conditions on this task.

**3.2 Session 2**

**3.2.1 Retention of category learning.** As a measure of incidental category learning, mean correct on the category identification test was calculated for each age group and compared across learning conditions (Fig. 2a). Results of a mixed ANOVA with two between-subjects factors of Age

(5, 7, 9 years) and Condition (V, A, AV), revealed no significant effect of Age, F<1, but an effect of Condition, F (2, 172) = 9.34, *p< .001*, partial $\eta^2$ = .09, with scores following learning with AV cues significantly higher than with auditory-only (*p< .001*) or visual-only cues (*p= .013*) (Bonferroni-corrected). No significant Age by Condition interaction was found, F (4, 172) = 1.55, *p= .191*.

An exploration of the raw scores across groups revealed that a number of participants (six 5-year-olds; sixteen 7-year-olds; and fourteen 9-year-olds), particularly in the visual and auditory conditions, scored either 0, 1, or 2 on the delayed category learning test. Such low scores suggest that these participants were consistently allocating exemplars to the incorrect habitat. That is, they had retained the category boundaries, but not the appropriate category label (the habitat). Chi-squared analysis revealed a trend for older participants to consistently allocate category locations incorrectly more often than the youngest participants, $X^2(2)$ = 5.67, *p= .059*. When participant scores of 0, 1 or 2 were recoded as valid scores (i.e., original score of 0 = 16, a score of 1 = 15, and 2 = 14), the corresponding mixed ANOVA revealed a main effect of Age, F(2, 172)= 3.75, *p= .026*, partial $\eta^2$ = .04, with 5-year-olds performing significantly below 7-year-olds (*p= .037)*, but with no significant differences between any other groups (see Fig 2b). Importantly, the main effect of Condition remained, F(2, 172)= 8.41, *p< .001*, partial $\eta^2$ = .09, with audiovisual learners scoring significantly higher than auditory-only (*p= .001*) or visual-only (*p= .012*). No significant Age by Condition interaction was found (*p = .184*).

FIGURE 2 HERE

To examine whether incidental categorization performance differed from chance, data were analysed for each Age group and Condition separately, using one-sample t-tests with a test value of 8. Results showed that participants only scored significantly above chance following AV learning in all age groups (5 years, *p=.006*; 7 years, *p=.014*; 9 years, *p< .001*), with at-chance performance following auditory- or visual-only learning in all groups (*p>.05* for all). However, once scores were recoded to account for participants who remembered the category boundaries but not the habitat labels, then children scored significantly above chance in all conditions and at all ages (*p<.05* for all).

**3.2.2 Relationships between category score and working memory.** An examination of relationships between DSB score and category test performance in each learning condition found no significant correlations in any age group ($p > .05$ for all).

## 4. Discussion

A novel incidental category learning task, in which category membership was identifiable by either unisensory or multisensory cues, was used to examine the 24-hour retention of learning across childhood. When redundant visual and auditory cues were both informative of category membership and presented in synchrony, this resulted in superior retention of incidental category learning than with unisensory cues, across age groups.

Incidental learning is pervasive in formal educational environments, with learning in naturalistic environments not always arising from explicit instruction. The findings therefore convey meaningful insight as to the efficacy of multisensory learning tools across this age range for not just explicit learning tasks as previously examined (Jordan & Baker, 2011; Jordan et al., 2008; Thornton et al., 1983), but also for retention of incidental knowledge. Moreover, above-chance performance across age groups was found only when bimodal categorical information was provided, with at-chance performance in visual- and auditory-only conditions only found after recoding scores in instances of consistently-incorrect habitat choice. This suggests multisensory cues may have allowed for greater retention of learning of both the category information and the category label (habitat) over a 24-hour period. This finding is in line with previous research proposing a facilitatory effect of multisensory information on children's learning (Baker & Jordan, 2015; Broadbent et al., 2017; Jordan & Baker, 2011; Joshi et al., 2002). Our results also extend these findings, indicating greater retention of learning, which also consists of more detailed information following learning with multisensory cues compared to unisensory.

Research examining multisensory integration at the level of perception suggests that the ability to integrate bimodal cues has a protracted developmental course (Burr & Gori, 2012; Gori et al., 2008; Nardini et al., 2008), with cue integration not found to be optimal until 8-10 years of age. A notable difference between our research and those examining Bayes-optimal integration is that the current learning task does not require the sensory cues to be fused into a unitary percept or amodal representation, but may be used summatively. Although age-related changes were identified in the current study, our findings suggest that on educationally-relevant tasks, multisensory information can be used beneficially for learning from at least 5 years of age.

In relation to our expectations of developmental effects, although the youngest children performed more poorly than older children, the effect of age was not specific to the use of multimodal information on category learning. This lack of age x condition interaction is in contrast to the findings by Broadbent et al. (2017) on immediate learning of categories following unimodal and bimodal stimuli. In the abovementioned study, a positive correlation between age and category-learning performance was identified following learning with audiovisual cues and with a trend following auditory-only cues. These contrasting findings may be a direct result of differing levels of ability in the youngest groups across the two studies. In the current study, although there were differences in raw scores on a measure of working memory between groups, the 5-year-olds were significantly higher than 7- and 9-year-olds on the level expected for age. Our high-performing sample of 5-year-olds (in working memory) may therefore have influenced the ability to use bimodal information in this group.

Above-chance performance in the current study may also be a reflection of high verbal working memory ability of the youngest participants. That said, our results indicate there was no relationship between age and working memory. A direct comparison of the performance of the two youngest groups across the two studies would help elucidate this further. Given that no significant developmental effects were found, extending the current study to examine performance in younger children would also be of interest to identify age-related changes in the ability to retain categorical information from multisensory cues compared to unisensory. However, task length and interface may have to be adapted to suit younger participants for whom performance may be limited by poorer attentional and motivational abilities.

The current findings also prompt other matters that warrant further investigation. In particular, these relate firstly, to the type of learning and retention that multisensory information affords, and secondly, whether it is imperative that the cues required for this type of learning be from multiple different modalities, or whether multiple cues in the same modality would suffice. To clarify the first point, it remains unclear whether multimodal information leads to enhanced stimuli perception and thus greater encoding of categorical features, or whether multimodal cues result in a stronger memory trace that is not subject to as great a level of decay as with unimodal information.

In relation to point two above, it is unclear from our findings whether the results are due to having multisensory information available or a matter of having multiple cues that are informative of category membership. The use of multiple cues compared to unitary information has been found to bolster learning in a number of studies (Kirkham, 2010); with the successful integration of multiple redundant cues able to bolster learning as young as 8 months of age (e.g., Kirkham, Slemmer, & Johnson, 2002; Kirkham, Slemmer, Richardson, & Johnson, 2007). For example, sequence learning was found to be supported using two visual cues of shape and color on a range of novelty preference paradigms (Kirkham et al., 2002). That said, multiple cues presented within the same modality are not easily averaged (Trommershauser, Kording, & Landy, 2011), and may compete for attention due to working memory constraints on domain-specific stores (Fougnie, Zughni, Godwin, & Marois, 2015). Two intra-modality cues may consequently not result in the same level of perceptual facilitation as with multisensory cues. Further research would, therefore, benefit from also including category exemplars with intra-modal informative cues to examine this underlying mechanism.

In conclusion, our findings indicate a facilitative effect of multisensory information on incidental category learning across middle childhood, with greater depth of learning than with unisensory information alone. Our findings, from the use of a novel incidental learning task with a 24-hour delay between task and test of learning, speaks to the efficacy of multimodal cues to support the retention of learning in children. This greater depth of knowledge retained through exposure to multimodal sources of information is an important finding for our understanding of the use of multisensory tools in the classroom.

## References

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology, 36*(2), 190.

Bahrick, L. E., & Lickliter, R. (2004). Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: a developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective, & Behavioral Neuroscience, 4*(2), 137-147.

Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory Development*. Oxford, UK: Oxford University Press.

Baker, J. M., & Jordan, K. E. (2015). The influence of multisensory cues on representation of quantity in children. In D. Geary, D. Berch, & K. M. Koepke (Eds.), *Math Cognition Vol 1: Evolutionary Origins and Early Development of Basic Number Processing* (Vol. 1): Elsevier.

Brandwein, A. B., Foxe, J. J., Russo, N. N., Altschuler, T. S., Gomes, H., & Molholm, S. (2011). The Development of Audiovisual Multisensory Integration Across Childhood and Early Adolescence: A High-Density Electrical Mapping Study. *Cerebral Cortex, 21*(5), 1042-1055. doi:10.1093/cercor/bhq170

Bremner, A. J., Lewkowicz, D. J., & Spence, C. (2012). *Multisensory Development*. Oxford, UK: Oxford University Press.

Broadbent, H. J., Osborne, T., Rea, M., Peng, A., Mareschal, D., & Kirkham, N. Z. (2018). Incidental category learning and cognitive load in a multisensory environment across childhood. *Developmental Psychology, 54*(6), 1020-1028. doi:10.1037/dev0000472

Broadbent, H. J., White, H., Mareschal, D., & Kirkham, N. Z. (2017). Incidental learning in a multisensory environment across childhood. *Developmental Science, 21*(2). doi:10.1111/desc.12554

Burr, D. C., & Gori, M. (2012). Multisensory integration develops late in humans. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton, FL: CRC Press.

Carter, M., & Stephenson, J. (2012). The use of multi-sensory environments in schools servicing children with severe disabilities. *Journal of Developmental and Physical Disabilities, 24*(1), 95-109.

Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British Ability Scales II*. Windsor: NFER-Nelson.

Ernst, M. O. (2008). Multisensory integration: a late bloomer. *Current Biology, 18*(12), R519-521. doi:10.1016/j.cub.2008.05.002

Flom, R., & Bahrick, L. E. (2010). The effects of intersensory redundancy on attention and memory: infants' long-term memory for orientation in audiovisual events. *Developmental Psychology, 46*(2), 428-436. doi:10.1037/a0018410

Fougnie, D., Zughni, S., Godwin, D., & Marois, R. (2015). Working memory storage is intrinsically domain specific. *Journal of Experimental Psychology: General, 144*(1), 30-47. doi:10.1037/a0038211

Gillingham, A., & Stillman, B. W. (1997). *The Gillingham Manual: Remedial Training for Children with Specific Disability in Reading, Spelling, and Penmanship*. Cambridge and Toronto: Educators Publishing Service.

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology, 18*(9), 694-698. doi:10.1016/j.cub.2008.04.036

Jordan, K. E., & Baker, J. (2011). Multisensory information boosts numerical matching abilities in young children. *Developmental Science, 14*(2), 205-213.

Jordan, K. E., Suanda, S. H., & Brannon, E. M. (2008). Intersensory Redundancy Accelerates Preverbal Numerical Competence. *Cognition, 108*(1), 210-221. doi:10.1016/j.cognition.2007.12.001

Joshi, R. M., Dahlgren, M., & Boulware-Gooden, R. (2002). Teaching reading in an inner city school through a multisensory teaching approach. *Annals of Dyslexia, 52*(1), 229-242. doi:10.1007/s11881-002-0014-9

Kirkham, N. Z. (2010). All together now: Learning through multiple sources. In S. Johnson (Ed.), *Neoconstructivism: The New Science of Cognitive Development* (pp. 32-44). USA: Oxford University Press.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition, 83*(2), B35-B42. doi:10.1016/S0010-0277(02)00004-5

Kirkham, N. Z., Slemmer, J. A., Richardson, D. C., & Johnson, S. P. (2007). Location, location, location: Development of spatiotemporal sequence learning in infancy. *Child Development, 78*(5), 1559-1571. doi:10.1111/j.1467-8624.2007.01083.x

Kirkham, N. Z., Wagner, J. B., Swan, K. A., & Johnson, S. P. (2012). Sound support: intermodal information facilitates infants' perception of an occluded trajectory. *Infant Behavior and Development, 35*(1), 174-178. doi:10.1016/j.infbeh.2011.09.001

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: an epigenetic systems/limitations view. *Psychological Bulletin, 126*(2), 281.

Mareschal, D., Quinn, P., & Lea, S. E. G. (2013). *The Making of Human Concepts*. Oxford, UK: Oxford University Press.

Montessori, M. (1912). *The Montessori Method*. New York: Frederick A. Stokes Company.

Moyer-Packenham, P., Baker, J. M., Westenskow, A., Anderson, K., Shumway, J., & Rodzon, K. (2013). A study comparing virtual manipulatives with other instructional treatments in third- and fourth-grade classrooms. *Journal of Education, 193*(2), 25-39.

Nardini, M., Bales, J., & Mareschal, D. (2015). Integration of audio-visual information for spatial decisions in children and adults. *Developmental Science*. doi:10.1111/desc.12327

Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology, 18*(9), 689-693. doi:10.1016/j.cub.2008.04.021

Ogden, S., Hindman, S., & Turner, S. D. (1989). Multisensory programs in the public schools: A brighter future for LD children. *Annals of Dyslexia, 39*(1), 247-267.

Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *Journal of Experimental Psychology: General, 133*(1), 46-62. doi:10.1037/0096-3445.133.1.46

Singleton, C. (2009). *Intervention for Dyslexia*. The Dyslexia-SpLD Trust.

Thomas, R. L., Nardini, M., & Mareschal, D. (2017). The impact of semantically congruent and incongruent visual information on auditory object recognition across development. *Journal of Experimental Child Psychology, 162*, 72-88.

Thornton, C. A., Jones, G. A., & Toohey, M. A. (1983). A multisensory approach to thinking strategies for remedial instruction in basic addition facts. *Journal for Research in Mathematics Education, 14*(3), 198-203. doi:10.2307/748382

Trommershauser, J., Kording, K., & Landy, M. S. (2011). *Sensory Cue Integration*. Oxford, UK: Oxford University Press.

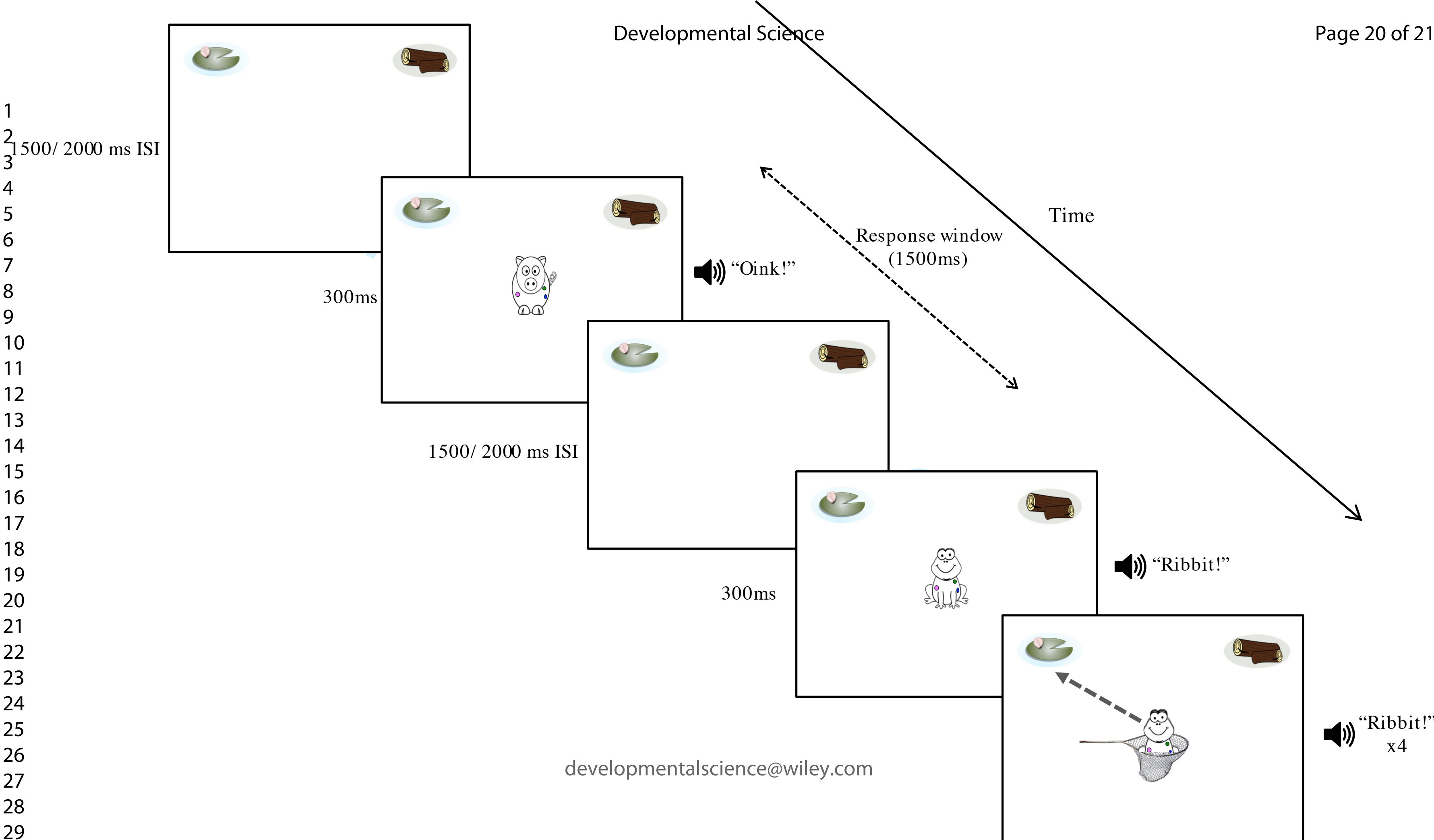
Wu, R., & Kirkham, N. Z. (2010). No two cues are alike: Depth of learning during infancy is dependent on what orients attention. *Journal of Experimental Child Psychology, 107*(2), 118-136. doi:10.1016/j.jecp.2010.04.014

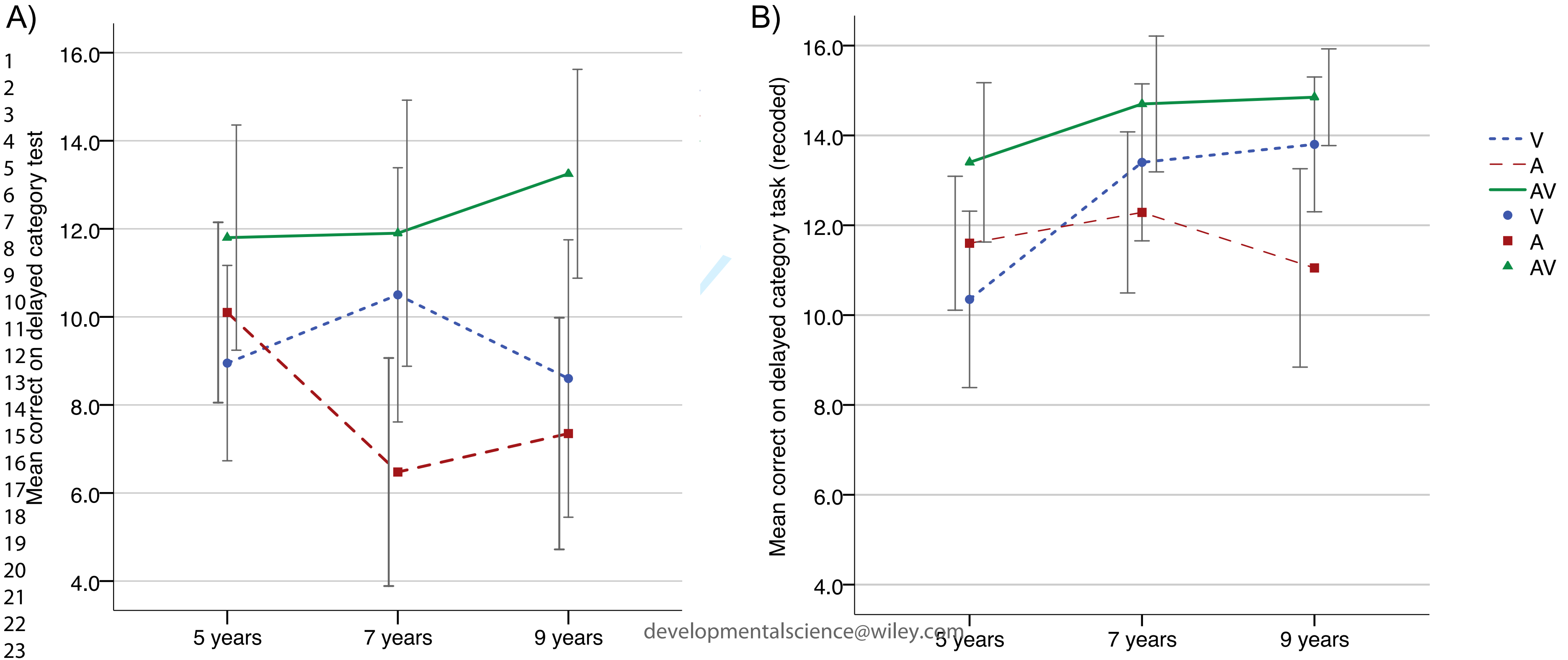


**FIGURE LEGENDS:**

Fig. 1. Schematic MALT presentation order. Final depicted screen appeared following a correct key-press response to the target stimulus, with dashed arrow indicating direction of movement to category habitat.

Fig. 1. A) Mean correct on category test after 24 hours delay across age groups and conditions. B) Mean correct on category test following recoding for habitat errors. Error bars represent 95% CIs.


1500/ 2000 ms ISI
300ms
“Oink!”
Response window
(1500ms)
Time
1500/ 2000 ms ISI
300ms
“Ribbit!”
“Ribbit!”
x4


A)
16.0
14.0
12.0
10.0
8.0
6.0
4.0
Mean correct on delayed category test
5 years
7 years
9 years
B)
16.0
14.0
12.0
10.0
8.0
6.0
4.0
Mean correct on delayed category task (recoded)
5 years
7 years
9 years
V
A
AV
V
A
AV