

Running head: COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

Incidental category learning and cognitive load in a multisensory environment across childhood

Broadbent, H.J., Osborne, T., Rea, M., Peng, A., Mareschal, D., and Kirkham, N.Z.

Centre for Brain and Cognitive Development, Birkbeck, University of London, London, UK

Corresponding author: Dr Hannah Broadbent, CBCD, Malet Street, Birkbeck, University of London,  
WC1E 7HX. Email: [h.broadbent@bbk.ac.uk](mailto:h.broadbent@bbk.ac.uk).

## Abstract

Multisensory information has been shown to facilitate learning (Bahrick & Lickliter, 2000; Broadbent, White, Mareschal, & Kirkham, 2017; Jordan & Baker, 2011; Shams & Seitz, 2008). However, although research has examined the modulating effect of unisensory and multisensory distractors on multisensory processing, the extent to which a concurrent unisensory or multisensory cognitive load task would interfere with or support multisensory learning remains unclear. This study examined the role of concurrent task modality on incidental category learning in 6- to 10-year-olds. Participants were engaged in a multisensory learning task whilst also performing either a unisensory (visual or auditory only) or multisensory (audiovisual) concurrent task (CT). We found that engaging in an auditory CT led to poorer performance on incidental category learning compared with an audiovisual or visual CT, across groups. In 6-year-olds, category test performance was at chance in the auditory-only CT condition, suggesting auditory concurrent tasks may interfere with learning in younger children, but the addition of visual information may serve to focus attention. These findings provide novel insight into the use of multisensory concurrent information on incidental learning. Implications for the deployment of multisensory learning tasks within education across development and developmental changes in modality dominance and ability to switch flexibly across modalities are discussed.

**Keywords:** Multisensory Integration; Cognitive Development; Incidental Learning; Cognitive Load

## **Introduction**

Successful integration of information from multiple sensory systems is imperative to constructing meaningful representations of the environment. This is particularly salient in formal learning environments, which typically demand attention from stimuli to different modalities. The ability to discriminate between relevant and extraneous environmental information is critical, particularly as learning tasks do not ordinarily occur in isolation. Developmental changes during the primary school years in ability to suppress irrelevant background information (Osman & Sullivan, 2014; Ridderinkhof, van der Molen, Band, & Bashore, 1997) have important implications for learning. However, age-related changes in ability to attend to and learn information on a primary task, whilst engaged in a concurrent unisensory or multisensory task has scarcely been examined.

The presentation of redundant multisensory cues can facilitate learning in adults (Fifer, Barutchu, Shivdasani, & Crewther, 2013; Lehmann & Murray, 2005; Seitz, Kim, & Shams, 2006; Shams & Seitz, 2008) and modulate attention in infants (Bahrick, Flom, & Lickliter, 2002; Bahrick & Lickliter, 2000; Bahrick, Lickliter, & Flom, 2004; Gogate & Bahrick, 1998; Lewkowicz, 2000; Lewkowicz & Kraebel, 2004; Richardson & Kirkham, 2004). These findings suggest that amodal properties presented unimodally are less salient than when experienced redundantly across two senses. However, despite our understanding of the development of multisensory integration, the efficacy of multisensory learning tools during childhood remains relatively unaddressed. In addition, multisensory cues in educational tools may only moderately resemble the stimuli used to address cross-modal sensitivity in laboratory studies.

In naturalistic learning tasks, the beneficial effects of complementary or redundant multisensory compared to unisensory information continue to mature between 5 and 10 years of age (Broadbent et al., 2017; Jordan & Baker, 2011). However, in the presence of competing unisensory and multisensory distractors, facilitation effects of multisensory information are diminished both in children and adults (Downing, Barutchu, & Crewther, 2014). Similarly, the processing of multisensory information is modulated in the presence of background auditory noise, as evidenced particularly by slower reaction times to targets (Steenken, Colonius, Diederich, & Rach, 2008).

## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

When unimodal visual stimulation is extraneous to the current task, such as in a vibrant classroom display, attention allocation is affected leading to poorer learning outcomes in 5-year-olds, compared with learning in sparse classroom environments (Fisher, Godwin, & Seltman, 2014). Likewise, in selective attention tasks, young children have more difficulty than older children and adults in avoiding processing peripheral stimuli (Enns & Akhtar, 1989) and more difficulty entirely filtering out irrelevant stimuli (Ridderinkhof et al., 1997). Also, 7-year-olds are more distracted by peripherally-presented letters than young adults (Huang-Pollock, Carr, & Nigg, 2002). These age-related improvements in maintaining focussed attention and reduced distractibility are attributed to the protracted development of working memory and inhibitory control (Diamond, 2006).

Although informative as to the role of same-modality or cross-modal background distractors or “noise” on multisensory processing, such findings tell us little about the role of unisensory and multisensory information when pertaining to concurrent tasks that increase attentional load. Learning exercises often require an individual to allocate their attention appropriately between two or more parallel tasks. For example, in a classroom environment, children are often expected to engage in learning tasks whilst attending to further instructions from a teacher or whiteboard. Currently, little is understood regarding the effect of engaging in a concurrent unisensory or multisensory task on learning in the primary task. That is to say, would the presentation of multisensory concurrent task information support attention to one or both tasks, or encumber the cognitive system and therefore serve as a greater distraction than unimodal information? Furthermore, does the role of concurrent task stimuli in different modalities differ across development, particularly given changes in sensory dominance during childhood (Nava & Pavani, 2013)?

The limited nature of attentional resources as proposed by Perceptual Load Theory (Lavie, 1995, 2005; Lavie & Tsai, 1994) determines the extent to which stimuli external to the primary task are processed. Due to limited attentional capacity, when perceptual load is high (there is a large amount of information involved in the processing of task stimuli), distractors and peripheral stimuli are less likely to be processed. In contrast, in conditions of low load, all available stimuli are processed, including distractors (Lavie, 2005).

## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

In adults, under conditions of high cognitive load, attentional resources have been shown to be shared across modalities. For instance, Macdonald and Lavie (2011) found that on a perceptually demanding visual task, conscious awareness of an auditory tone was reduced compared to a similar task with lower perceptual demands. Likewise, high visual load has been found to negatively affect recall for auditory information, further indicating a possible cross-modal perceptual load effect on memory (Murphy, Groeger, & Greene, 2016). Furthermore, such findings suggest that it is not just visual attention that is affected by visual load; the amount of input for a given sensory modality may be limited by attentional resources on supporting systems.

However, there remains a developmental question. That is, in young children for whom integration of bimodal information may not be optimal (e.g., Barutchu, Crewther, & Crewther, 2009; Gori, Del Viva, Sandini, & Burr, 2008; Nardini, Bales, & Mareschal, 2015; Nardini, Jones, Bedford, & Braddick, 2008), and thus presentation of bimodal information not be processed as readily, would cognitive load be effected differently? On the one hand, if the presentation of cross-modal stimuli were to actually *increase* cognitive load, it can be inferred that concurrent multisensory information (that which is external to the primary task but still attended to) would result in a higher level of interference than unisensory information, as these tasks demand the engagement of attentional resources in two modalities. As a result, one hypothesis is that multisensory information on a concurrent task may interfere with learning to a greater extent in younger than in older children, by requiring heightened selective attention, a skill that undergoes protracted development. On the other hand, however, in accordance with the Intersensory Redundancy Hypothesis (Bahrick et al., 2004), the cross-modal stimuli on a secondary task would provide redundant information, which could result in enhanced perception of the concurrent task stimuli. If this is the case, amodal cross-modal stimuli might be processed quickly without increasing cognitive load.

In one study, on a visual selective attention task, (Matusz et al., 2014) examined the role of perceptual load using peripheral unisensory and multisensory distractor stimuli. The extent to which an increase in visual perceptual load decreased the effects of peripheral distractors was dependent on both the modality of the distractor and the age of the participant. At low cognitive load, unimodal

visual distractors had a greater effect on performance in younger children, whereas higher visual load on the central task eliminated distractor effects at all ages. When distractors were audiovisual (multisensory), however, adults showed distraction effects both at high and low cognitive load conditions. A different pattern of performance was seen in children. At low load, 6- and 11-year-olds were significantly distracted by multisensory information, whereas a higher visual load decreased distraction effects of multisensory items, particularly in the youngest group. These findings go some way to demonstrating that processing of multisensory stimuli is more flexible than the literature on redundancy suggests, and highlights the importance of examining the differential effects of multimodal information on learning with development.

It is also currently unknown whether changes in sensory modality dominance across development, from auditory dominance in infants and children aged 4 to 6 years of age to visual dominance in adults (Colavita, 1974; Nava & Pavani, 2013; Robinson & Sloutsky, 2004), would influence the role of concurrent task modality on multisensory learning. According to the auditory dominance account, where visual and auditory information compete for attentional resources in children, greater attentional resources are delegated to auditory stimuli (given their dynamic and transient nature) at a cost to visual processing (Robinson & Sloutsky, 2010a, 2010b; Sloutsky & Napolitano, 2003); although the auditory processing is not invulnerable to interference from the concurrent visual information (Thomas, Nardini, & Mareschal, 2017). In light of this, an alternative hypothesis for the current study is that age-related differences in multisensory processing of information on two concurrent tasks may be modulated by a transition from auditory to visual modality dominance. In particular, concurrent auditory information may result in greater levels of distraction than visual in younger children, with a change to visual dominance with age.

In line with our previous research examining the role of multisensory information on learning across childhood, an incidental category learning task was used in the current study. In naturalistic environments, learning does not always arise from explicit instruction. Indeed, classroom-based teaching predominantly includes incidental learning objectives. For example, on a mathematical learning task that involves counting different objects, the learning of concepts such as categorical

information or other perceptual properties of the items may be incidental to the initial task, but are important comprehensively. For this reason, it is important to examine the extent to which unisensory and multisensory concurrent response tasks interfere with or support incidental learning across childhood.

The aim of the present study, therefore, was to investigate the extent to which a concurrent unisensory or multisensory cognitive load task would interfere with or support incidental learning of multisensory categories across the primary school years. Consistent with previous findings, it was predicted that performance would improve with age on both a primary learning task and a concurrent task, but that incidental category learning would be differentially effected across development depending on the modality of a concurrent task in which the individual is engaged. In line with typical developmental changes, we predicted age-related improvements in incidental learning, and differential effects of concurrent task condition on learning across all groups. In particular, in concurrent tasks comprising multisensory (i.e., audiovisual) stimuli, it was predicted that incidental category learning in a primary task would be modulated to a different extent than with unimodal concurrent tasks, but that this would be differentially effected between 6 and 10 years of age.

### **Method**

#### **Participants**

Data from 180 children were included in the study. Participants were selected from three separate school years (1, 3 and 5;  $n=60$  per group). This resulted in three age groups as follows: 6-year-olds [mean (sd) age = 6.21 (0.32) years, 32 males], 8-year-olds [mean (sd) age = 8.24 (0.39) years, 33 males], and 10-year-olds [mean (sd) age = 10.15 (0.45) years, 29 males]. Participants in each group were randomly allocated to one of three concurrent-task (CT) conditions [visual (V), auditory (A) or audiovisual (AV)], in a between-subjects design ( $N=20$  per condition in each age group). 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 primary schools and informed written

parental consent was obtained for each participant, in accordance with the Department of Psychological Sciences, Birkbeck University of London Ethics Committee, 131453, project title: ‘Learning from multisensory cues’. 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 was 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 lasted approximately 20 minutes.

### **Stimuli**

The ‘Frogs and Stars’ task is a modified version of the audiovisual condition of the Multisensory Attention Learning Task (MALT; Broadbent et al., 2017), with an integrated concurrent task (CT). Using a between-subjects design, the CT consisted of either a *visual (V)*, *auditory (A)* or *audiovisual (AV)* task in which participants were asked to count the number of stars (*V*), dings (*A*) or dinging stars (*AV*) that occurred during the game, depending on the participant’s assigned condition. CT stimuli appeared quasi-randomly (hard-coded, with the same permutation for all participants) during inter-stimulus intervals (ISIs) on the MALT, for 500ms durations (ISIs were 1500/2000ms durations, so CT stimuli occurred for the first 500ms of an ISI). Visual CT stimuli (a yellow star subtending an approximate 3° visual angle) appeared in the centre of the screen (the same location as the target and non-target images on MALT trials). Auditory CT stimuli were presented at 44kHz and around 70-75dB through dual-channel closed-back headphones. For audiovisual CT stimuli (dinging stars), the visual (star) and auditory (ding) elements of the cues were presented concurrently with the same onset and offset. Task stimuli were presented using the Psychophysics Toolbox extension for MATLAB (Brainard, 1997), and presented on a 15” laptop screen approximately 50cm in front of the participant.

In the MALT, all target (frog) and non-target (cat, dog, elephant, goat, owl, pig) stimuli were presented with synchronous and complementary visual and auditory features. All visual images were



forward facing depicting a head and body with (front) legs for consistency and to maintain a level of similarity across stimuli. Auditory stimuli consisted of congruent animal sounds, consistent with the different visual animal stimuli.

Two categories of target (frog) stimuli could be differentiated by visual features [‘few spots (2-3)’ or ‘many spots (7-8)’] and auditory features [high and long-short “*rib-bit*” sound or a deep and short-long “*rib-bit*”]. Spots varied in colours and sizes across category members (10 within-category members). Similarly, within-category auditory features varied in pitch at 0.5 semitone intervals (10 within-category members). Auditory stimuli were created from digital animal sound (.wav) files and manipulated using ‘Audacity Digital Audio Editor Software’ into 300ms sound files. For counterbalancing, two combinations of audiovisual categorising features were used, i.e., either family 1: few spots with high ribbit and family 2: many spots with deep ribbit, or, family 1: few spots with deep ribbit and family 2: many spots with high ribbit. Participants were not made aware of these categories.

### **Procedure**

As a measure of auditory working memory, each participant initially completed the Digit Span Backwards (DSB) task from the British Ability Scales–II (Elliott, Smith, & McCulloch, 1996). This measure of auditory working memory was primarily included as a proxy for cognitive ability to assess whether age groups could be considered as performing as expected for age, in line with other previous studies (e.g., Broadbent et al., 2017).

**Stimuli familiarisation.** Before presentation of the ‘Frogs and Stars’ task, participants were first familiarised to the MALT stimuli to check that they were able to see and hear the exemplar of each animal in the task. Participants were shown one of each animal in turn and asked whether they were able to hear and see the exemplar. All participants answered affirmatively for each of the seven exemplars and so continued with the task.

**Frogs and Stars task.** Participants were instructed to press the space bar as quickly as possible whenever a frog (target animal) appeared (visually on the screen and auditorily through the headphones) whilst ignoring (inhibiting a response) to any other animal stimuli. Participants were not made aware that there would be two different families (categories) of target stimuli nor asked to learn their habitat locations. This was in order to examine *incidental* learning of these target categories.

The task consisted of 200 trials, separated into five blocks of 40 trials, with each block separated by a motivation screen to provide a rest break. 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 stimuli were presented randomly throughout the task. Participants were also asked to count the number of ‘stars’, ‘dings’ or ‘dinging stars’ throughout each block, depending on the CT condition to which they had been allocated (V, A, or AV, respectively). CT stimuli were presented during a number of inter-stimulus intervals (ISIs) following non-target trials in each block.

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. On each trial, an animal image (target or non-target) appeared individually in the centre of the screen for 300ms, simultaneously with the corresponding 300ms auditory cue (animal sound). If the space bar was (correctly) pressed after the presentation of a target stimulus, the same frog reappeared within a ‘net’ (see Figure 1, final screen). The frog then immediately travelled to the top left- or top right-hand corner of the screen to the correct frog habitat (e.g., 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). Travel time to habitat was 2000ms. The corresponding 300ms audio file for that frog was also played simultaneously three times until the frog reached the correct habitat. This was for consistency with exposure to the visual stimuli for incidental learning of categorical information. Following movement to the habitat, the target image was then paused for an additional 1000ms to avoid disorientation caused by an immediate appearance of the next stimulus. If the button was pressed incorrectly for a non-target animal, no feedback was given and the task continued to the next trial after either a 1500ms or 2000ms ISI.

As mentioned above, during some of the ISIs (following non-target MALT trials) in each block, a 500ms CT stimulus was played. During block 1, it was presented eight times, block 2 = 12 times, block 3 = 9 times, block 4 = 11 times, and block 5 = 10 times (Total = 50 CT stimuli). The CT stimuli, therefore, only appeared following 25 percent of the MALT trials. During rest breaks, participants were asked how many of each CT stimuli they had counted during that block. This was presented to participants both as a question on the screen and orally by the experimenter, to account for poorer reading and attentional skills in the youngest children. Another instruction screen was then displayed reminding them to start counting the CT stimuli again in the next block. As before, this instruction was also read aloud by the experimenter for all participants, to make sure that they remembered to begin counting from zero.

Across the whole 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.

FIGURE 1 HERE

**Category identification test.** To examine the extent of incidental category learning on the Frogs and Stars task, participants were then asked to complete a category identification task. Eight audiovisual exemplars from each category (16 total) were presented in a random order. Participants responded to whether the frog had lived at the lily pad or the log during the game. Participants were presented with each frog exemplar individually, and no feedback was given throughout the identification task. Total correct categorisation responses were recorded. Following the categorisation test, a measure of explicit categorisation knowledge was then given, where participants were asked, “Can you tell me how you decided where each frog lived? What made them belong to each family?”

## Results

### Auditory Working Memory

## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

Digit Span Backwards (DSB) raw ability scores were converted to standardised T-Scores and compared across groups using a one-way analysis of variance (ANOVA). No significant difference was found between groups; 6-years: Mean (SD) = 55.97 (8.76); 8-years = 53.90 (9.75); 10-years = 53.37 (8.53),  $F(2, 179) = 1.39, p = .252$ , showing participants in each group were performing at a cognitive level expected for their age.

### **Frogs and Stars Task**

**Accuracy score ( $d'$  prime).** To examine target-detection accuracy on the MALT, z-scores for Hit rates ( $H = \text{Correct hits} / 80 \text{ target trials}$ ) and FA rates ( $FA = \text{False alarms} / 120 \text{ non-target trials}$ ) were calculated. A  $d'$  prime [ $d' = z(H) - z(FA)$ ] measure of sensitivity was then calculated and mean values were analysed across groups (Figure 2). Results of a univariate ANOVA with two fixed factors of Age and CT Condition found a significant main effect of Age,  $F(2, 171) = 17.71, p < .001$ , partial  $\eta^2 = .17$ , but not of CT Condition ( $F < 1$ ); showing no effect of CT modality on accuracy performance across groups. No significant Age by Condition interaction was found,  $F(4, 171) = 1.47, p = .213$ . Bonferroni-corrected post hoc tests found that 6-year-olds were significantly less accurate than 8- and 10-year-olds ( $p = .013$ , and  $p < .001$ , respectively). In addition, 8-year-olds had significantly lower accuracy scores than 10-year-olds ( $p = .006$ ), demonstrating improvement in task accuracy with age.

FIGURE 2 HERE

**Errors on concurrent task (CT).** Mean number of errors made on the concurrent task (counting stars (V), dings (A) or dinging stars (AV)) were calculated for each age group and across CT conditions (Figure 3). Given that counting was started anew for each block, error scores for each block were calculated as an absolute deviation score from correct number of stars, dings, or dinging stars (deviation either higher or lower). Total error score for each participant was the sum of deviation errors across blocks.

Results of a univariate ANOVA found a significant main effect of Age on number of errors

on concurrent task,  $F(2, 170) = 16.57, p < .001$ , partial  $\eta^2 = .16$ , but no effect of Condition,  $F(2, 170) = 2.25, p = .11$ . Bonferroni-corrected post hoc tests found significantly more errors in 6-year-olds than 8- and 10-year-olds ( $p < .001$  for both), but not between 8- and 10-year-olds ( $p = .11$ ).

With the inclusion of DSB score (as a proxy for working memory) as a covariate in this analysis, the pattern of results remained, with main effect of Age,  $F(2, 169) = 19.23, p < .001$ , partial  $\eta^2 = .16$ , and no effect of Condition,  $F(2, 169) = 1.85, p = .16$ . No significant Age by Condition interaction was found ( $F < 1$ ).

FIGURE 3 HERE

### **Incidental Category Learning Task**

**Total correct.** Mean number correct (out of 16 trials) on the incidental category learning task were calculated for each group (Figure 4). Results of a univariate ANOVA found a significant effect of Age,  $F(2, 171) = 6.38, p = .002$ , partial  $\eta^2 = .07$ , an effect driven by 6-year-olds scoring significantly fewer correct than 8- and 10-year-olds (Bonferroni-corrected:  $p = .01$  and  $.001$ , respectively). A significant effect of CT Condition was also identified,  $F(2, 171) = 4.19, p = .02$ , partial  $\eta^2 = .05$ , with fewer correct following the auditory-only (A) than the audiovisual (AV) concurrent task conditions ( $p = .004$ ), with no significant differences between any other conditions (Bonferroni-corrected:  $p > .05$  for all). Although no significant Age by CT Condition interaction was identified ( $F(4, 171) = 2.05, p = .09$ , partial  $\eta^2 = .05$ ), planned comparisons were conducted to examine the effects of condition within each group separately. Results found a significant effect of CT Condition only in the 6-year-old group,  $F(2, 60) = 6.43, p = .003$ , partial  $\eta^2 = .18$ . Bonferroni-corrected comparisons showed that this was due to poorer category test performance in the 6-year-olds who were given an auditory concurrent task compared to either a visual-only ( $p = .043$ ) or an audiovisual ( $p = .003$ ) concurrent task. No significant difference in category learning was identified between visual and audiovisual CT ( $p = 1$ ).

FIGURE 4 HERE

**Performance from chance.** Performance from chance on the category identification task was calculated for each age group and condition. Results of two-tailed t-tests with a test value of 8 showed that participants scored significantly above chance in each group and for all conditions ( $p < .001$  for all), except for 6-year-olds in the Auditory-only CT condition,  $t(21) = 1.45$ ,  $p = .16$ .

**Relationships between CT, incidental category learning and DSB.** In an investigation of whether performance on the CT task (number of errors on counting concurrent dings/stars/dinging stars) was related to incidental category learning (number correct on category test), Pearson correlations found no significant relationship in any CT condition ( $p > .05$  for all). In addition, no significant relationships were found between CT and DSB T-scores in any CT condition ( $p > .05$  for all). An examination of relationships between DSB and category learning performance also found no significant relationships in any CT condition ( $p > .05$  for all).

**Age and performance on incidental category learning task.** To examine changes in performance on the categorization task with age, we carried out Pearson's correlations, which revealed a significant positive relationship between Age and Total Correct in the incidental category learning task in the auditory-only condition ( $r = .43$ ,  $p = .001$ ), and a trend in the AV condition ( $r = .25$ ,  $p = .06$ ) but not in the Visual condition ( $p = .95$ ).

## **Discussion**

This study investigated the effect of concurrent task modality on incidental category learning, across middle childhood. Performance in the youngest children, aged 6 years, on an incidental category learning task was significantly poorer when engaged in an auditory CT than with a visual or audiovisual CT; a finding not seen in the older children. Our findings are, therefore, indicative of age-related changes in susceptibility to distraction when a concurrent task requires attention only to the auditory modality, in this case counting 'dings'. Paradoxically, this was not found to be the case

when a visual element was presented simultaneously with the concurrent auditory information in the bimodal (audiovisual) CT condition in any age group. Here, the concurrent visual information seemed to somehow compensate for the negative effect of auditory information described above. Similarly, in the youngest group, when CT stimuli consisted only of unimodal visual information, notably less detriment to incidental category learning was found than with an auditory CT.

These results are suggestive of cross-modal differences in the effects of cognitive load across childhood. In particular, engagement in an additional auditory task is particularly detrimental to performance on a multisensory learning task. In contrast, the findings indicate that the addition of visual information may serve to focus an individual's attention on the visual features of the learning task, resulting in better incidental learning of category information. This may have been by the automatic capturing of attention to the most salient or familiar stimuli (Christie & Klein, 1995), or by filtering out the auditory information (Macdonald & Lavie, 2011).

On the main MALT task, age-related differences in performance were found, in line with expected developmental changes in attention, inhibition, and performance on continuous performance tasks (Levy, 1980; Lin, Hsiao, & Chen, 1999) and concurrent load tasks (Irwin-Chase & Burns, 2000; Karatekin, 2004). Importantly, however, no differences in task elements such as MALT accuracy score or concurrent task performance were found across CT conditions, indicating that incidental category learning in the current study was not a reflection of differences in performance on the main task. Instead, our findings suggest that children's attentional resources are differentially affected depending on the modality of the concurrent task stimuli.

It was initially proposed that there might be heightened cognitive load from concurrent tasks that required attention to bimodal properties (in the present case, audiovisual information), and that this would result in poorer learning outcomes on a central task than with a unimodal concurrent task. This was anticipated particularly in younger children for whom the ability to integrate bimodal information into a unitary percept may not be fully mature (e.g., Gori et al., 2008; Nardini et al., 2015), thus resulting in a greater level of information to be processed in two sensory modalities. In contrast, however, when a concurrent task included bimodal information, this resulted in better

performance on an incidental learning task than a concurrent task that only included auditory information. This is in line with the Intersensory Redundancy Hypothesis (Bahrick et al., 2004), that would suggest amodal cross-modal stimuli might be processed more efficiently and without increasing cognitive load.

Prima facie, it seems unclear from the current results whether allocating attention to a concurrent auditory task is *detrimental* to performance or whether, to the contrary, it is the exposure to a visual or audiovisual concurrent task that leads to an *improvement* in category learning per se. That is, the presentation of multisensory information may lead to more efficient processing of the concurrent task stimuli, thus reducing cognitive resources required to complete both tasks. However, this would also result in a difference in incidental learning between visual and audiovisual CT conditions, a difference not found in the current study, suggesting that this cannot fully explain the results.

To further address this point, performance in each group would have to be compared to performance on the primary learning task without a concurrent task. In the study by Broadbent et al. (2017), a comparable attention task with audiovisual stimuli resulted in performance in incidental category learning that was above chance in 6-year-olds; a finding of multisensory facilitation on learning that is also supported by other studies (e.g., Jordan & Baker, 2011). Although a direct comparison of data from the two studies was not possible due to task modifications, at-chance performance in the present study using an audiovisual task with an auditory CT suggests that attention to the concurrent auditory information *did* result in a detriment to learning in 6-year-olds.

The effects of multisensory distractors are reduced in conditions of high perceptual load in young children, but not in adults (Matusz et al., 2014). Although it is unclear the extent to which the MALT involves a low or high perceptual load, the current study differs from the task used by Matusz and colleagues in that the peripheral multisensory information was presented as an intermittent concurrent task to which participants were asked to attend, rather than serving as distractor items to be ignored. Our findings therefore provide a unique insight into the role of unisensory and multisensory



information that is secondary to a primary learning task when the stimuli from both are to be processed.

Van der Burg, Olivers, and Theeuwes (2012) suggest one mechanistic explanation for how attentional focus may narrow with an increase in task demands in children. According to these authors, multisensory information may shield children from distraction. This could only be corroborated in future studies with the addition of adult participants, to examine whether mature cognitive systems would show multisensory distractor effects.

In the study by Matusz and colleagues (2014), at levels of low cognitive load, 6-year-olds responded significantly faster in the presence of auditory distractors than with visual-only and audiovisual distractors, suggesting that unimodal auditory information is not processed as readily as visual information in this age group. This finding is somewhat in contrast with our results, which suggest additional auditory information may be processed to a greater extent, or at least impacts to a greater extent on incidental learning than visual-only or audiovisual concurrent stimuli. However, our findings may also reflect developmental changes in a modality processing dominance. Our results could, therefore, be considered as consistent with an auditory processing dominance in younger children (Sloutsky & Napolitano, 2003), with auditory information in our task seemingly more salient and therefore more likely to be processed than concurrent visual stimuli. Others have also reported age-related differences in auditory processing (e.g., Nava & Pavani, 2013). However, unlike in the current study, learning was not evaluated, nor were there any secondary or ‘distractor’ stimuli to attend to. In addition, in modality dominance studies, cross-modal cues are typically presented in synchrony and so it remains unclear whether the presentation of auditory information in a concurrent task that occurred after the visual input on a primary task would pull attention in a comparative way.

If modality dominance were a causal factor in our results, it might also be expected that participants engaged in an audiovisual concurrent task would perform at a similar level to those in the auditory-only condition, with the auditory information serving as an attention-grabber, particularly in 6-year-olds. In fact, engaging in a multisensory concurrent task resulted in high learning performance in all groups. Moreover, no correlation was identified between performance on the CT (ability to

count CT stimuli) and incidental learning in any condition, indicating that performance in category learning was not necessarily a reflection of greater processing or enumeration of auditory concurrent task stimuli, but a result of processing information in different modalities across two separate tasks.

The ability to switch flexibly across modalities is a skill undergoing development in the early years, with a gradual maturation of voluntary control over shifts of attentional set during childhood (Colombo, 2001; Garon, Bryson, & Smith, 2008; Jones, Rothbart, & Posner, 2003). This may go some way to explaining age-related differences when faced with a concurrent task modality that differs from the central task modality. On an audiovisual task such as the MALT, where the integration of multisensory information may not be optimal in younger children (e.g., Barutchu et al., 2009; Gori et al., 2008; Nardini et al., 2015), it may be that young children in particular were attending primarily to one element of the stimuli. For instance, focussing predominantly on the visual aspect of the MALT stimuli would have resulted in less distraction when faced with a concurrent visual or audiovisual task than with a concurrent auditory task, which would require the individual to flexibly shift across modalities. That said, this can only be assumed from the data as it is plausible that participants across conditions could rely exclusively on one modality for both the CT and MALT aspects of the task. Alternatively, it could be argued that the youngest children were actually focusing more on the auditory aspects of the MALT targets in order to acquire categorical information. In this situation, CT stimuli in the same modality might have heightened cognitive load and resulted in greater distraction. However, this is unlikely to be the case, given that when young children are only presented with unimodal auditory cues that are informative to categorical learning, they perform more poorly than with visual-only or bimodal audiovisual information (Broadbent et al., 2017). On this task, Broadbent and colleagues compared category learning performance (without an additional concurrent load task) in 5- to 10-year-olds when category membership was defined either by visual-only, auditory-only or multisensory information. Here, participants demonstrated superior performance following exposure to bimodal cues, and with the poorest performance seen in the youngest children who were presented with auditory-only category information.

Findings from dual-task studies suggest that focusing attention to stimuli in one modality can encumber performance in a different modality (e.g., see Strayer & Johnston, 2001). However, if it can be assumed that participants were switching of attention between modalities on the two tasks in the present study, this was not seen to negatively impact the detection of visual events in the MALT. That is, the concurrent task modality did not impede the ability to identify the target (frog), as shown on the accuracy measure on the MALT, but only the subsequent incidental learning of categorical features from multisensory targets. It could be suggested therefore that shifts of attentional set across modalities may be more deleterious to incidental learning of category information than to young children's ability to respond to a target. Alternatively, these findings may merely reflect the different effects of CT stimuli on a complex categorical learning task compared to a relatively simpler detection task. Of note is that although age-related differences in performance were seen following an auditory concurrent task, only a trend towards a significant age-by-condition interaction was found on the incidental learning task, and so planned comparisons to examine these developmental changes should be considered with some vigilance. This said, a strong positive correlation between age and incidental learning in the auditory CT condition group is further support of developmental changes in the effects of an auditory concurrent task on learning.

In children aged 9 and 11 years, during the detection of auditory targets, auditory and audiovisual distractors were found to result in higher error rates than with visual distractors and with visual or audiovisual targets (Downing et al., 2014). This suggests that cross-modal distractors may have differential effects on the processing of multisensory targets. In contrast to our study, Downing et al. (2014) presented visual distractors in temporal synchrony with targets, but required a shift in spatial attention to the periphery. Visual concurrent task stimuli in the present study, however, were presented in the same spatial location as targets. CT stimuli were also not presented in temporal synchrony with the stimuli on the primary task and so may not have served as a cross-modal 'distractor' in the same way. This highlights an important distinction between the current study and previous work examining the role of unisensory and multisensory distractors on multisensory processing. In essence, the processing of multisensory stimuli for purposes of learning may be

differentially effected by engaging in cross-modal concurrent tasks compared to in the presence of extraneous and temporally synchronous distractors, particularly in younger children.

In summary, in children below 8 years of age, engagement in a concurrent auditory task modulated incidental learning performance to a greater extent than with visual or audiovisual stimuli. This suggests that intermittent auditory information on a concurrent task may have a different effect on performance than visual or bimodal stimuli. The results of the present study extend developmental research on cross-modal processing and provide novel insight into the differential role of unisensory and multisensory concurrent tasks on incidental category learning in children. Developmental improvements in multisensory integration and ability to switch flexibly between modalities are likely to underpin changes in the role of concurrent task modality on learning across the primary school years. Such findings have implications for designing educational environments and learning tasks and are indicative of changes with age in the ability to process multisensory information on two concurrent tasks, particularly when switching between modalities is required.

### **Acknowledgements**

This research was supported by funding from the Economic and Social Research Council (ESRC), grant reference: ES/K005308/1. DM was also partially funded by a Royal Society Wolfson Research Merit Award. We would like to thank all the children who participated in the study, and thanks to Carlton Primary, St John's CofE, Trafalgar Infants and Juniors, St Paul's CE Primary, and Sir John Lillie Primary, and SEYP youth project for taking part. Thanks also to Ruben Zamora and Anna Peng for help with task programming.

**References**

- Bahrick, L. E., Flom, R., & Lickliter, R. (2002). Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. *Dev Psychobiol*, *41*(4), 352-363. doi:10.1002/dev.10049
- 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., & Flom, R. (2004). Intersensory Redundancy Guides the Development of Selective Attention, Perception, and Cognition in Infancy. *Current Directions in Psychological Science*, *13*(3), 99-102. doi:10.2307/20182922
- Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: development of multisensory facilitation in children. *Dev Sci*, *12*(3), 464-473. doi:10.1111/j.1467-7687.2008.00782.x
- Broadbent, H. J., White, H., Mareschal, D., & Kirkham, N. Z. (2017). Incidental learning in a multisensory environment across childhood. *Dev Sci*.
- Christie, J., & Klein, R. (1995). Familiarity and attention: Does what we know affect what we notice? *Memory & Cognition*, *23*(5), 547-550. doi:10.3758/BF03197256
- Colavita, F. B. (1974). Human sensory dominance. *Attention, Perception, & Psychophysics*, *16*(2), 409-412.
- Colombo, J. (2001). The development of visual attention in infancy. *Annual Review of Psychology*, *52*(1), 337-367.

## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

Diamond, A. (2006). The early development of executive functions. In E. Bialystock & F. I. M. Craik (Eds.), *Lifespan Cognition: Mechanisms of Change* (pp. 70-95). New York: Oxford University Press.

Downing, H., Barutcu, A., & Crewther, S. G. (2014). Developmental trends in the facilitation of multisensory objects with distractors. *Frontiers in Psychology*, 5. doi:10.3389/fpsyg.2014.01559

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

Enns, J. T., & Akhtar, N. (1989). A developmental study of filtering in visual attention. *Child Development*, 1188-1199. doi:10.2307/1130792

Fifer, J. M., Barutcu, A., Shivdasani, M. N., & Crewther, S. G. (2013). Verbal and novel multisensory associative learning in adults. *F1000Res*, 2, 34. doi:10.12688/f1000research.2-34.v2

Fisher, A. V., Godwin, K. E., & Seltman, H. (2014). Visual Environment, Attention Allocation, and Learning in Young Children: When Too Much of a Good Thing May Be Bad. *Psychological Science*. doi:10.1177/0956797614533801

Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: a review using an integrative framework. *Psychological bulletin*, 134(1), 31.

Gogate, L. J., & Bahrick, L. E. (1998). Intersensory Redundancy Facilitates Learning of Arbitrary Relations between Vowel Sounds and Objects in Seven-Month-Old Infants. *Journal of experimental child psychology*, 69(2), 133-149. doi:10.1006/jecp.1998.2438

## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Curr Biol*, *18*(9), 694-698.

doi:10.1016/j.cub.2008.04.036

Huang-Pollock, C. L., Carr, T. H., & Nigg, J. T. (2002). Development of selective attention: perceptual load influences early versus late attentional selection in children and adults. *Dev Psychol*, *38*(3), 363-375.

Irwin-Chase, H., & Burns, B. (2000). Developmental changes in children's abilities to share and allocate attention in a dual task. *Journal of Experimental Child Psychology*, *77*(1), 61-85.

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science*, *6*, 498-504. doi:10.1111/1467-7687.00307

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

Karatekin, C. (2004). Development of attentional allocation in the dual task paradigm. *International Journal of Psychophysiology*, *52*(1), 7-21.

doi:10.1016/j.ijpsycho.2003.12.002

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, *9*(2), 75-82.

## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Attention, Perception, & Psychophysics*, *56*(2), 183-197.

Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Brain Res Cogn Brain Res*, *24*(2), 326-334.  
doi:10.1016/j.cogbrainres.2005.02.005

Levy, F. (1980). The development of sustained attention (vigilance) and inhibition in children: some normative data. *J Child Psychol Psychiatry*, *21*(1), 77-84.

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

Lewkowicz, D. J., & Kraebel, K. S. (2004). The value of multisensory redundancy in the development of intersensory perception. In G. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multi-sensory processes* (pp. 655-678). Cambridge, MA: MIT Press.

Lin, C. C., Hsiao, C. K., & Chen, W. J. (1999). Development of sustained attention assessed using the continuous performance test among children 6–15 years of age. *Journal of Abnormal Child Psychology*, *27*(5), 403-412.

Macdonald, J. S., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, *73*(6), 1780-1789. doi:10.3758/s13414-011-0144-4

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkle, R., & Scerif, G. (2014). Multi-modal distraction: Insights from children's limited attention. *Cognition*, *136c*, 156-165. doi:10.1016/j.cognition.2014.11.031



## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

- Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next? *Psychonomic Bulletin & Review*, *23*(5), 1316-1340. doi:10.3758/s13423-015-0982-5
- Nardini, M., Bales, J., & Mareschal, D. (2015). Integration of audio - visual information for spatial decisions in children and adults. *Dev Sci*. doi:10.1111/desc.12327
- Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Curr Biol*, *18*(9), 689-693. doi:10.1016/j.cub.2008.04.021
- Nava, E., & Pavani, F. (2013). Changes in sensory dominance during childhood: converging evidence from the colavita effect and the sound-induced flash illusion. *Child Dev*, *84*(2), 604-616. doi:10.1111/j.1467-8624.2012.01856.x
- Osman, H., & Sullivan, J. R. (2014). Children's auditory working memory performance in degraded listening conditions. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1503-1511. doi:10.1044/2014\_JSLHR-H-13-0286
- Richardson, D. C., & Kirkham, N. Z. (2004). Multimodal events and moving locations: eye movements of adults and 6-month-olds reveal dynamic spatial indexing. *J Exp Psychol Gen*, *133*(1), 46-62. doi:10.1037/0096-3445.133.1.46
- Ridderinkhof, K. R., van der Molen, M. W., Band, G. P. H., & Bashore, T. R. (1997). Sources of Interference from Irrelevant Information: A Developmental Study. *Journal of Experimental Child Psychology*, *65*(3), 315-341. doi:10.1006/jecp.1997.2367
- Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Dev*, *75*(5), 1387-1401. doi:10.1111/j.1467-8624.2004.00747.x

## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

Robinson, C. W., & Sloutsky, V. M. (2010a). Development of cross - modal processing.

*Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 135-141.

doi:10.1002/wcs.12

Robinson, C. W., & Sloutsky, V. M. (2010b). Effects of multimodal presentation and stimulus familiarity on auditory and visual processing. *Journal of Experimental Child Psychology*, *107*(3), 351-358. doi:10.1016/j.jecp.2010.04.006

Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Curr Biol*,

*16*(14), 1422-1427. doi:10.1016/j.cub.2006.05.048

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends Cogn Sci*, *12*(11),

411-417. doi:10.1016/j.tics.2008.07.006

Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Dev*, *74*(3), 822-833.

Steenken, R., Colonius, H., Diederich, A., & Rach, S. (2008). Visual–auditory interaction in saccadic reaction time: Effects of auditory masker level. *Brain Research*, *1220*, 150-

156. doi:10.1016/j.brainres.2007.08.034

Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*,

*12*(6), 462-466. doi:10.1111/1467-9280.00386

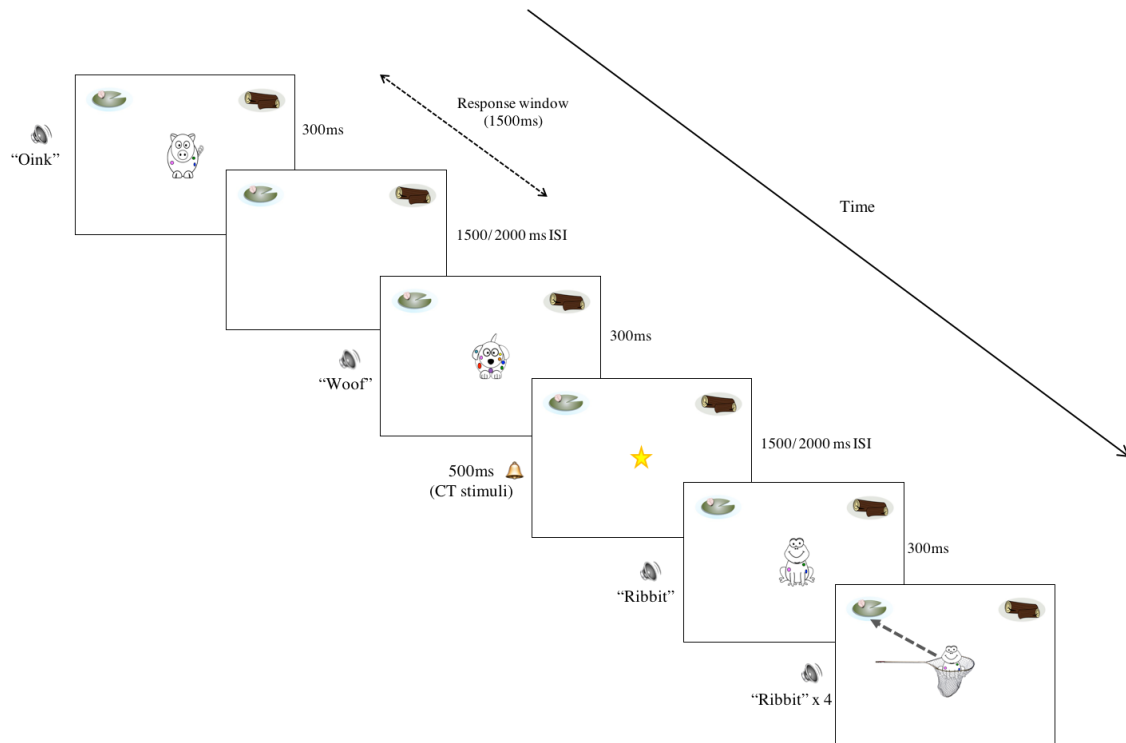
Thomas, R. L., Nardini, M., & Mareschal, D. (2017). The impact of semantically congruent and incongruent visual information on auditory object recognition across

development. *Journal of experimental child psychology*, *162*, 72-88.

Van der Burg, E., Olivers, C. N., & Theeuwes, J. (2012). The attentional window modulates capture by audiovisual events. *PLoS One*, 7(7), e39137.

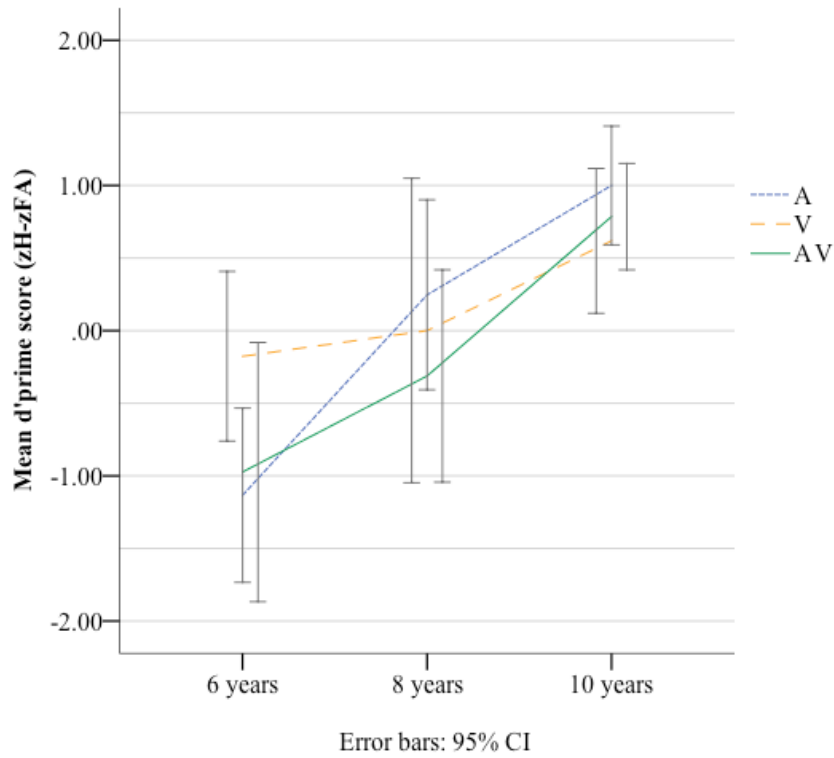
doi:10.1371/journal.pone.0039137

# COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN

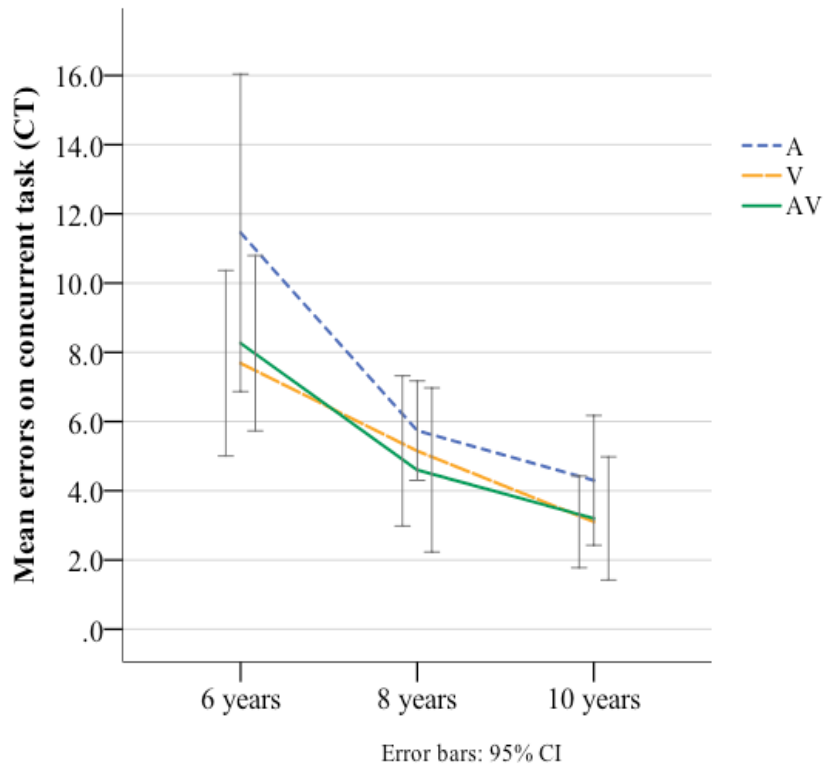


**Figure 1** Task presentation order. Screen with star indicates bimodal concurrent task stimulus (AV). Final depicted screen appeared following a correct keypress response to target stimulus, with dashed arrow indicating direction of movement to correct category habitat

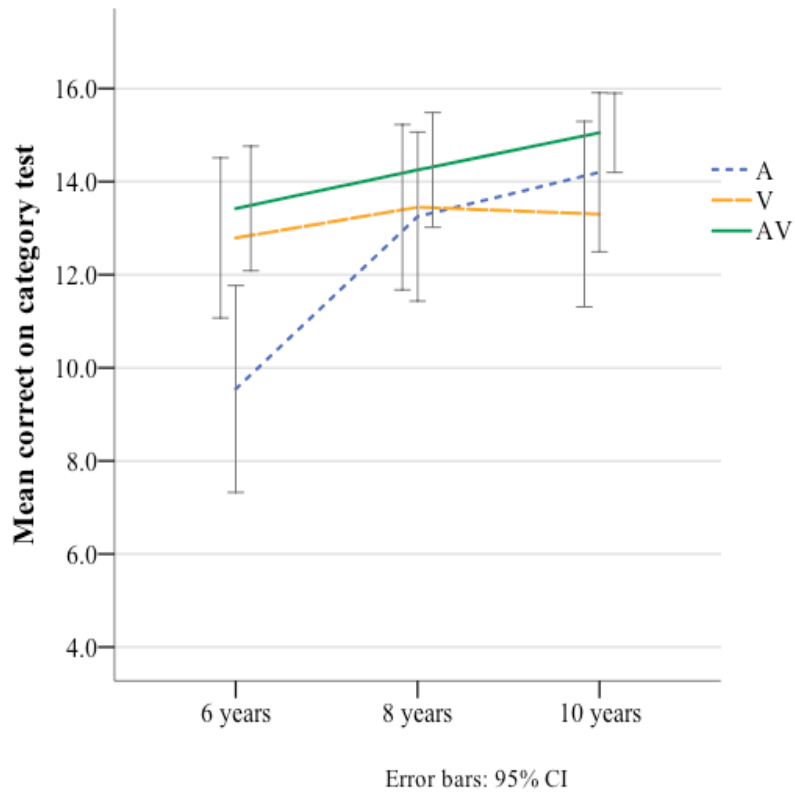
## COGNITIVE LOAD AND MULTISENSORY LEARNING IN CHILDREN



**Figure 2** Mean accuracy score (d'prime) on MALT (zHit rate - zFalse Alarm rate) in each age group across CT conditions.



**Figure 3** Mean number of errors on concurrent task (CT) in each age group across CT conditions



**Figure 4** Mean correct on category task in each age group across CT conditions