Are two cues always better than one? The role of multiple intra-sensory cues compared to multi-cross-sensory cues in children's incidental category learning

Broadbent, H[1,2]., Osborne, T[2]., Mareschal, D[2]., and Kirkham, N[2].

[1]Royal Holloway, University of London

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

Address all correspondence to (present address): Dr Hannah Broadbent, Department of Psychology, Royal Holloway University of London, Egham, Surrey, TW20 0EX. Email: Hannah.broadbent@rhul.ac.uk

Abstract

Simultaneous presentation of multisensory cues has been found to facilitate children's learning to a greater extent than unisensory cues (e.g., Broadbent, White, Mareschal, & Kirkham, 2017). Current research into children's multisensory learning, however, does not address whether these findings are due to having multiple cross-sensory cues that enhance stimuli perception or a matter of having multiple cues, regardless of modality, that are informative to category membership. The current study examined the role of multiple cross-sensory cues (e.g., audio-visual) compared to multiple intra-sensory cues (e.g., two visual cues) on children's incidental category learning. On a computerized incidental category learning task, children aged six to ten years (N= 454) were allocated to either a visual-only (V: unisensory), auditory-only (A: unisensory), audio-visual (AV: multisensory), visual-visual (VV: multi-cue) or auditory-auditory (AA: multi-cue) condition. In children over eight years of age, the availability of two informative cues, regardless of whether they had been presented across two different modalities or within the same modality, was found to be more beneficial to incidental learning than with unisensory cues. In six-year-olds, however, the presence of multiple auditory cues (AA) did not facilitate learning to the same extent as multiple visual cues (VV) or when cues were presented across two different modalities (AV). The findings suggest that multiple sensory cues presented across or within modalities may have differential effects on children's incidental learning across middle childhood, depending on the sensory domain in which they are presented. Implications for the use of multi-cross-sensory and multiple-intra-sensory cues for children's learning across this age range are discussed.

Keywords: multisensory; cognitive development; incidental learning; intra-sensory cues; audiovisual

1. Introduction

Presenting information across multiple sensory modalities has been found to facilitate children's learning to a greater extent than when presenting information within a single sensory modality (Baker & Jordan, 2015; Broadbent et al., 2017). These findings have important potential implications for the use of multisensory learning tools within educational environments. Indeed, multisensory learning is frequently promoted in classrooms, particularly for the support of reading and mathematical development (Hulme, 1979; Jordan, Suanda, & Brannon, 2008; Ofman & Shaevitz, 1963; Thornton, Jones, & Toohey, 1983). Such multisensory learning tools rely on the assumption that a greater number of different sensory cues provides a more robust perceptual experience, and thus greater learning potential for the child. However, the existing research does not address the question of whether children's enhanced performance, following multimodal presentation of the learning stimuli, is due to the presence of a greater number of informative cues in the multisensory condition as compared to the unisensory condition. In other words, two cues may be better than one cue regardless of whether they are presented across two or more different sensory modalities, or within the same modality. Thus, it remains unclear whether it is the multisensory nature of the information that is beneficial to learning or simply the additional information associated with having two cues. The present study, therefore, examines whether the presentation of multiple informative cues within the same modality (e.g., two visual cues) are as effective at supporting children's learning as multiple cues presented across two different modalities (e.g., one audio and one visual cue).

On the one hand, it is possible that, in the context of multiple same-modality (or 'intra-sensory') cues, having additional information within the same modality may serve as a greater support to learning -and at an earlier developmental stage- than with multiple different-modality cues. This is because the presentation of two cues within the same modality would not necessitate the ability to successfully combine cues across different modalities. This is particularly relevant because the ability to integrate multisensory cues optimally undergoes protracted development across the primary school years (e.g., Ernst, 2008; Nardini, Bales, & Mareschal, 2015). However, several labs have also found that, even within a single modality, multiple sensory cues do not reach adult levels of integration until late in childhood or even early adolescence (Dekker et al., 2015; Mamassian, 2015; Nardini et al., 2010).

That said, it is important to note that the utilization of multiple sensory cues for learning does not necessarily require the *integration* of these sensory cues into a unified amodal representation. Instead, it is the ability to *combine* (i.e., to draw on the information present in) the two sensory cues to facilitate learning that is of significance here. These two processes (sensory integration versus cue combination) involve different underlying mechanisms and strategies (Ernst & Bülthoff, 2004). In tasks involving sensory integration (see Bahrick & Lickliter, 2012), sensory signals are typically redundant, presented simultaneously (appearing and disappearing together), and are limited to the same environmental unit (e.g., spatial localization or depth). In the case of sensory combination, however, a 'cue' refers to specific sensory information that is informative to category membership, rather than information that gives rise to a sensory estimate (Ernst & Bülthoff, 2004). In this latter case, sensory cues only need to be arbitrarily related across or within modalities. Furthermore, as is often the case with real-world multisensory learning environments, *simultaneous* presentation of sensory cues is not always necessary (e.g., Baines, 2008).

On the other hand, it is possible that information from multiple cues within the same modality (e.g., visual-visual) will not be as beneficial as information from multiple cues across different modalities (e.g., auditory-visual). For instance, cues in the visual modality are not always found to be easily averaged (Trommershauser, Kording, & Landy, 2011). This is particularly true when cues provide different types of visual information such as depth or distance cues (Landy, Maloney, Johnston, & Young, 1995). The extent to which two visual cues are combined for learning is also found to be dependent on the estimated reliability of each cue, which is in turn related to the level of cue ambiguity and whether the cue provides reliable versus unreliable information (Jacobs, 2002; Kirkham, 2010).

Multiple cues within the same modality may also compete for attention, and consequently not result in the same level of perceptual facilitation as with multisensory cues. This may be due to working memory constraints on domain-specific stores (Fougnie & Marois, 2011; Fougnie, Zughni, Godwin, & Marois, 2015). Studies examining multisensory working memory suggest that there is a performance benefit for remembering audiovisual stimuli compared to memory for modality-specific cues (for a review see Quak, London, & Talsma, 2015). Others have proposed that the auditory and visual cues are dealt with in separate stores that do not interfere with each other in the same way that cues from the same modality would (Fougnie et al., 2015). Despite this, cross-sensory interference can also occur; with the modality that engages attention faster, dominating later processing (Robinson & Sloutsky, 2010). However, there is some debate regarding the developmental trajectory of cross-modal interference; with some research finding that young children are less susceptible to cross-modal interference (Matusz et al., 2015), and others identifying stronger cross-modal interference early in childhood (Hirst, Kicks, Allen, & Cragg, 2019; Robinson, Hawthorn, & Rahman, 2018).

With the use of either arbitrarily-related audiovisual cues or two unrelated cues within the same modality, although such cues may provide ancillary information that is task-relevant, these sensory cues would not be integrated into a single percept for storage, as the two units of sensory information would not be related to a shared perceptual dimension. However, it remains unclear whether there is an additional benefit to the use of cross-sensory audiovisual cues for children's *learning* compared to the use of two cues presented within one single modality. The current study, therefore, aimed to further examine the role of multiple cross-sensory versus multiple intra-sensory cues in children's learning. The presence of two informative cues, regardless of modality, was hypothesized in the current study to facilitate children's learning to a greater extent than with unimodal cues *(hypothesis 1)*. However, given that two cues within one modality may interfere with each other within the same sensory store (Fougnie et al., 2015), it was also predicted that multiple cues from two different modalities (i.e., visual and auditory combined) would be more beneficial to learning than with two cues from the same modality *(hypothesis 2)*. In view of the developmental changes in the ability to use multisensory cues for learning (e.g., Broadbent et al., 2017; 2018; 2019; Kirkham et al, 2019), we hypothesized age-related differences in the ability to combine sensory information (within or across modalities) to facilitate learning between five and ten years of age *(hypothesis 3)*.

We explore these questions in the context of incidental category learning. Incidental learning is a ubiquitous facet of children's learning (Reber, 1993; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997), and refers to the development of knowledge that emerges without explicit instruction. Previous research has found that the presentation of multisensory cues leads to superior immediate incidental learning (Broadbent et al., 2017) and delayed retention of incidental information (Broadbent, Osborne, Mareschal, & Kirkham, 2018) across childhood. In addition, we focus on the domain of category learning because category and concept learning underlie generalization and reasoning, and are core elements of primary school curricula (Mareschal, Quinn, & Lea, 2013; Murphy, 2002). The current study was therefore designed to examine the role of multiple sensory cues in these specific aspects of children's learning (incidental and categorical).

To examine the abovementioned points, we used an incidental category learning paradigm presented in Broadbent et al. (2017), but with the addition of a number of novel sensory learning conditions. In the study by Broadbent and colleagues, unimodal (either visual or auditory) cues that were intrinsic to category exemplars resulted in inferior incidental learning of category boundaries compared to intrinsic multisensory (audiovisual) cues. Perceptual features can be classified as intrinsic (referring to information that is within or an aspect of the stimulus itself, such as the colour of the stimulus) or extrinsic (referring to information that is surrounding or external to the stimulus, such as a patterned background) (Ecker, Maybery, & Zimmer, 2013). The current study compared data from Broadbent et al (2017) (with intrinsic cues) to performance in the following novel sensory learning conditions: (i) two informative visual cues (one extrinsic and one intrinsic to the category exemplars), (ii) two informative auditory cues (one extrinsic and one intrinsic to the category exemplars), and (iii) two extrinsic audiovisual (one audio and one visual) informative cues.

The use of both intrinsic and extrinsic cues in the two-cues conditions allowed for sensory cues to remain perceptually distinct, and for consistency across sensory conditions. It has been proposed that different binding mechanisms are used depending on whether perceptual features are intrinsic or extrinsic to the object (Ecker et al., 2013; Humphreys, 1998). For instance, intrinsic features, which are considered as incidental aspects of a stimulus itself (Troyer & Craik, 2000), are given more attention during encoding than extrinsic features, with the binding of extrinsic features found to be slower and less automatic than intrinsic feature integration (e.g., Ecker, Zimmer, & Groh-Bordin, 2007). Similarly, on a visual task in which participants made old/new judgements regarding the shape and/or color of the test stimuli, Ecker et al. (2013) found that intrinsic color effected shape recognition but extrinsic color did not. These findings reveal differences in attentional processes for intrinsic and extrinsic cues. However, little research has examined this in relation to the role of multisensory intrinsic and extrinsic features on incidental learning. As an additional aim of the current study, therefore, the role of intrinsic and extrinsic cues when presented within and between different sensory modalities was also examined, with poorer performance expected with extrinsic cues than intrinsic cues *(hypothesis 4)*.

2. Methods

2.1. Participants

Two hundred and seventy-three children (new to the testing paradigms) were included in the current study and compared to data from one hundred and eighty-one children previously reported in Broadbent et al (2017); total N= 454. Participants newly-recruited for the current study were selected from three separate age groups: six-year-olds (N=90, M_{age}= 5.59 years, SD = .32 years, Range = 5.1 to 6.2 years); eight-year-olds (N=91, M_{age}= 7.70 years, SD = .26 years, Range = 7.1 to 8.2 years); and ten-year-olds (N=92, M_{age}= 9.67 years, SD = .31 years, Range = 9.1 to 10.2 years). Participants in the previously-reported study were similarly aged: six-year-olds (N=60, M_{age}= 6.05 years, SD = .52 years, Range = 5.0 to 6.8 years); eight-year-olds (N=60, M_{age}= 8.26 years, SD = .31 years, Range = 7.6 to 8.8 years); and ten-year-olds (N=61, M_{age}= 10.20 years, SD = .41 years, Range = 9.0 to 10.8 years). Sample sizes for each group, per condition, were determined by power analysis for ANOVA with df = 1, f = 0.40.

Children participating in the current study were randomly allocated to one of five learning conditions using a between-subjects design; Visual-Visual (VV, multi-cue), Auditory-Auditory (AA, multi-cue), Extrinsic Auditory (EA, unisensory), Extrinsic Visual (EV, unisensory), or Extrinsic Audiovisual (EAV, multisensory). The data taken from Broadbent et al (2007) provided the complementary experimental conditions of Intrinsic Visual (IV, unisensory), Intrinsic Auditory (IA, unisensory) and Intrinsic Audiovisual (IAV, multisensory).

Children (in both studies) were recruited from local primary schools and by opportunity from the University research unit database. Informed written parental consent was obtained for each participant, in accordance with the University ethics committee guidelines. All participants had normal hearing (no known hearing impairments) 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, or in the University research unit. Children were thanked for participating with a certificate and stickers. The testing session for each participant lasted approximately 15 to 20 minutes.

2.2. Stimuli

The Multisensory Attention Learning Task (MALT) is a novel computerised incidental category-learning task, based on a modified version of a classic continuous performance task, and adapted for use with primary-school aged children. Visual stimuli consisted of seven different animal line drawings, subtending an approximate 3° visual angle, presented on a 15" laptop screen approximately 50cm in front of the participant. Animal stimuli consisted of one target animal ('frog') and six non-target animals ('owl', 'dog', 'goat', 'pig', 'elephant', and 'cat'). 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. Intrinsic auditory stimuli consisted of congruent 300ms animal sounds (e.g., croak, meow), consistent with the different visual animal stimuli. Extrinsic auditory stimuli consisted of 300ms sounds considered external to (and yet naturally associated with) the animal stimuli (e.g., water sound with frog, bell sound with goat). Auditory features (intrinsic and extrinsic) were all chosen to be natural (e.g., frog croaks and water sounds), but were arbitrarily paired with visual cues. For reference, visual and auditory cues that were included in each condition are presented in Table 1 and described in more detail in 2.2.1 – 2.2.8. Auditory stimuli were presented at 44 kHz and around 70-75dB through closed-back headphones. Stimuli were presented using the Psychophysics Toolbox extension for MATLAB (Brainard, 1997).

Table 1.

Informative sensory cues in each condition on the Multisensory Attention Learning Task (MALT)

		N cues	Modality condition	Visual cue	Auditory cue
Intrinsic	Visual (IV)	1	U	Spots	-
	Auditory (IA)	1	U	-	Croak
	Audio-visual (IAV)	2	CM	Spots	Croak
Extrinsic	Visual (EV)	1	U	Background	-

	Auditory (EA)	1	U	-	Water sound
	Audio-visual (EAV)	2	CM	Background	Water sound
Intrinsic + Extrinsic	Visual-Visual (VV)	2	WM	Spots + Background	-
	Auditory-Auditory (AA)	2	WM	-	Croak + Water

Note: U = unimodal; CM = cross-modality; WM = within-modality

2.2.1. Intrinsic Visual (IV, unisensory) condition (Broadbent et al., 2017)

In the IV learning 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. For exemplars of targets from the two IV families, see Figure 1a. Non-target animals were similarly marked with spots of varying colours, size and number, for consistency across stimuli. In the IV visual learning condition, auditory stimuli remained consistent across exemplars. That is, for target stimuli (frogs), only one of the two intrinsic auditory-cue ‘families’ (see below for further details) was used, counterbalanced across participants.

2.2.2. Intrinsic Auditory (IA, unisensory) condition (Broadbent et al., 2017)

In the IA condition, only unimodal intrinsic auditory features were used to differentiate family members. IA stimuli were presented for 300ms, consistent with visual presentation times. The visual ‘family’ for target stimuli remained consistent and was counterbalanced across participants. Target stimuli ‘families’ were distinguishable by two different frog croaks, each with a double-croak (*‘rib-bit’*) sound. Family 1 exemplars croaked with a ‘high and long-short’ sound, whilst family 2 exemplars croaked with a ‘deep and short-long’ croak (manipulated using ‘Audacity Digital Audio Editor Software’). Four different pitches of croak were used as a variant to denote different within-family members, varying in 0.5 semitone intervals. All other sound-file properties remained consistent across and within families.

2.2.3. Intrinsic Audiovisual (IAV, multi-sensory) condition (Broadbent et al., 2017)

In the IAV learning condition, both intrinsic visual (number of spots) and intrinsic auditory (croak type) features could be used to discriminate category membership. For example, family 1 members had few spots and a long-short croak, whilst family 2 members had many spots and a short-long croak. The two possible combinations of categorising audiovisual features were counterbalanced across participants.

2.2.4. Visual-Visual (multi-cue) condition

In the VV multi-cue condition, both intrinsic visual (spots) and extrinsic visual (background pattern) cues could be used to determine category membership. Intrinsic cues were as stated above (either few or many spots). Extrinsic visual cues consisted of a background square box surrounding the target cue either with diagonal (grey/white) line patterns (4 varying tones and directions) or zig-zag (grey/white) line patterns (4 varying tones and directions), for within-category variance. Background patterns resulted in visual stimuli subtending an approximate 7° visual angle. For exemplars of targets from the two VV families, see Figure 1b.

2.2.5. Auditory-Auditory (multi-cue) condition

In the AA multi-cue condition, both intrinsic (croaks) and extrinsic (water sound) cues could be used to determine category membership of target exemplars. Intrinsic cues were as stated above. Extrinsic auditory cues consisted of a ‘water’ noise presented for 300ms, with onset 100ms after onset of the intrinsic croak sound (to allow for 100ms of background sound to be separable from the intrinsic croak cue). Extrinsic sound features for Family 1 consisted of ‘light and effervescent’ water bubbles, and Family 2 of ‘slow and gloopy’ bubbles sound. Two different water sounds were chosen as extrinsic cues so that participants would have to distinguish between two categories of one nameable sound rather than two different and easily distinguishable sounds. Four different pitches of water sound in each category were used as a variant to denote different within-family members, varying in 0.5 semitone intervals. All other sound-file properties remained consistent across and within families. Extrinsic auditory cues were manipulated and combined with intrinsic croak cues of the corresponding family using Audacity software.

2.2.6. Extrinsic Visual (EV, unisensory) condition

In the EV condition, only the visual cues extrinsic to the target animal (background patterns, either stripes or zigzags) were used to denote category membership. All target images across families (as well as non-target stimuli) were plain with no dots. Auditory stimuli from only one intrinsic family (e.g., high-low ribbit sound) were used across all target exemplars, regardless of category membership.

2.2.7. Extrinsic Auditory (EA, unisensory) condition

In the EA condition, only the extrinsic auditory cues (two different water sounds) were used to denote category membership. Visual stimuli from only one intrinsic family (e.g., few dots) were used across all target exemplars, regardless of category membership.

2.2.8. Extrinsic Audiovisual (EAV, multi-sensory) condition

Extrinsic visual backgrounds and Extrinsic water sounds were both used in order to denote category boundaries in the EAV condition (e.g., zig-zag backgrounds with light water sound denoted family 1, and straight lines with gloopy water sound denoted family 2, counterbalanced across participants).

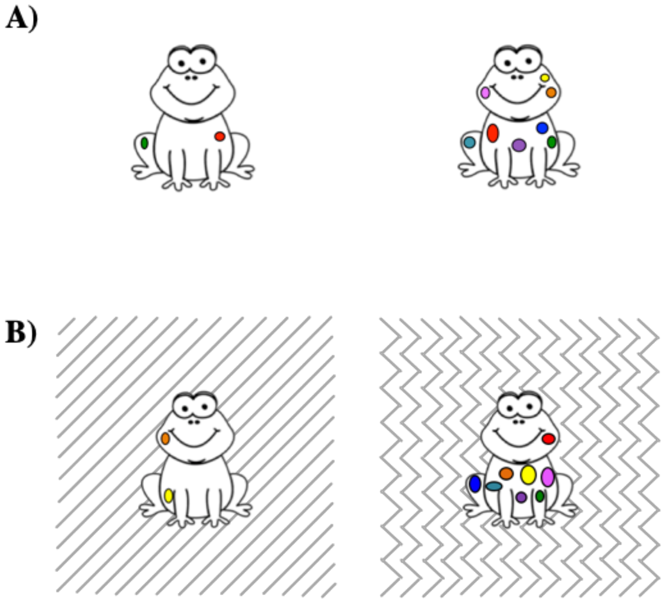

Fig 1. A) Exemplars of target stimuli from the two intrinsic visual families. B) Exemplars from the two different Visual-Visual families, denoting both intrinsic visual (dots) and extrinsic visual (background pattern) cues to category membership

2.3. Procedure

2.3.1. Auditory working memory

As a measure of auditory working memory, each participant completed the Digit Span Backwards (DSB) task from the British Ability Scales–II (BAS-II; Elliott, Smith, & McCulloch, 1996). This measure of auditory working memory was included as a proxy for cognitive ability to assess whether age groups could be considered as performing as expected for age, in line with previous studies (e.g., Broadbent et al., 2017). All participants received the DSB task before the following familiarization and MALT tasks in the current study, in line with presentation order in Broadbent et al (2017).

2.3.2. Familiarization Task

A short audio and visual task was used to familiarize participants with the stimuli before presentation of the MALT. Participants were shown one of each animal (target and non-targets) in turn and asked whether they were able to hear and see the exemplar. The version shown was matched with the participant's MALT condition, i.e. participants in the AA condition were familiarized with the relevant intrinsic and extrinsic auditory stimuli. All participants answered affirmatively for each of the seven familiarization exemplars and so continued with the task.

2.3.3. Multisensory attention learning task (MALT) (frog detection)

The MALT consisted of 200 trials, separated into four blocks by a motivation screen on which was written 'you are really good at this!', to allow for rest-breaks. Participants were instructed to press the space bar as quickly as possible whenever a frog (target animal) appeared on the screen, whilst ignoring (inhibiting a response to) any other animal (non-target) stimuli. The task screen consisted of a white background 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, a stimulus (target or non-target) appeared 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' (see Figure 2). The frog then immediately travelled to the top left- or top right-hand corner of the screen to the correct frog habitat (i.e., unbeknownst to the participants, 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 lasted 2000ms. Simultaneous with the travel to the habitat, the corresponding audio file for that frog was played on a loop until the frog reached the correct habitat. Thus, auditory stimuli exposure matched the continuous exposure to the visual stimuli during travel time.

On reaching 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 had been pressed incorrectly for a non-target animal, no feedback was given and the task continued to the next trial following either a 1500ms or 2000ms inter-stimulus interval (ISI); selected in line with findings that these timings are optimal for task performance when used with children (e.g., Chee, Logan, Schachar, Lindsay, & Wachsmuth, 1989; Okazaki et al., 2004). For a schematic of the MALT presentation sequence, see Figure 2.

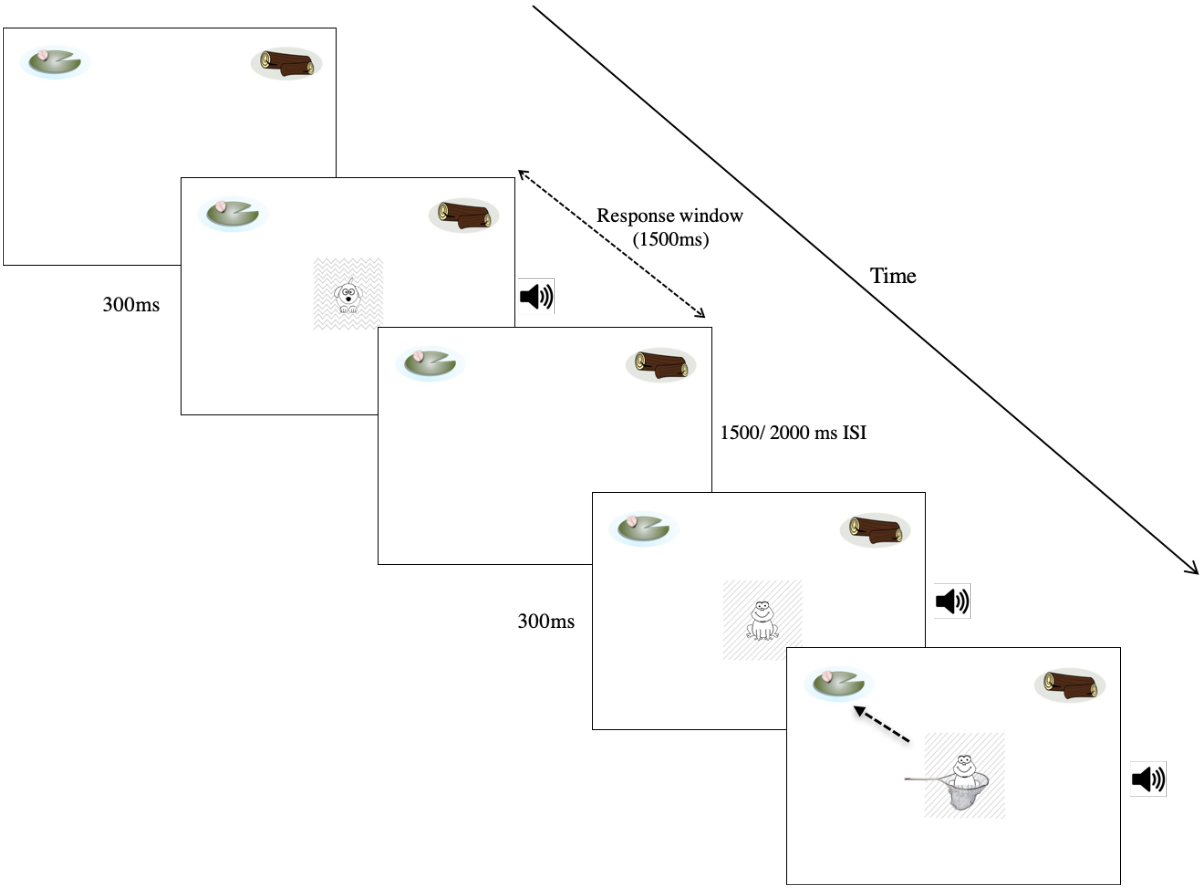

Fig 2. Multisensory Attention Learning Task (MALT) presentation order in Extrinsic Visual (EV) condition. Final screen depicts target animal with EV background caught in a net following correct keypress response. Dashed line indicates direction of travel to allocated habitat.

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 (calculated cumulatively across trials from task beginning), or until the maximum 200 trials were completed. Participants were therefore scored as having reached criterion or not. Data were analyzed only from those who met the 50-correct target responses criterion (n=1 participant did not meet criterion). As such, all participants included in the analyses had been exposed to the same number of category allocation trials (having observed 50 frogs travelling to their correct habitat).

2.3.4. Test of incidental category learning (primary outcome task)

The MALT task (2.3.3) was designed as a decoy task for participants to focus their attention on. However, the dependent variable of interest in this study was the degree of incidental learning that occurred during the MALT. At the beginning of the MALT, participants had not been informed that there were two different target categories that would travel to the two different habitats. Therefore, to examine the extent of incidental category learning at the end of the MALT game, participants were asked to complete a test of incidental category learning. Eight exemplars from each of the two target families (16 total) were presented in a random order. Participants responded verbally regarding whether they thought the frog had lived at the lily pad or the log during the game. The experimenter recorded the response by pressing one of two allocated laptop keys. Participants were presented with each frog exemplar individually, and no feedback was given throughout the identification task. Total number of correct categorisation responses were recorded. Following the categorisation test, a secondary 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?" This explicit knowledge task was used as a secondary measure to assess whether incidental learning had also resulted in explicit knowledge of categories.

2.3.5. Extrinsic exemplar discrimination task

To examine the discriminability of extrinsic cues, two discrimination tasks were used; one visual and one auditory, consisting of 12 pairs of frogs (6 pairs identical, 6 pairs different). In the extrinsic visual condition, participants were presented plain frogs in succession, with identical or different background patterns, and asked if the images *looked* the same or different. In the extrinsic auditory condition, participants were shown two plain frogs and played two consecutive water noises, and asked if they *sounded* the same or different. A number of participants from the EV, EA and EAV conditions in all three age groups were randomly selected to receive one discrimination task condition, relevant to their previous MALT task condition (6 years, Visual: N =15, Auditory: N =18; 8 years, Visual: N=11, Auditory: N= 11; 10 years, Visual: N= 11, Auditory: N=11). Note that discrimination performance for intrinsic cues is reported in Broadbent et al (2017).

2.4 Variables of Interest

Independent variables in this study consisted of Age (3 levels: 6-, 8-, and 10-years) and Condition (8 levels: IV, IE, IAV, VV, AA, EV, EA, and EAV). The primary outcome measure (main DV) was the mean number of correct categorisations (out of 16 trials) on the test of incidental category learning (2.3.4). As a secondary variable of interest, mean explicit categorization scores were analysed across Age and Conditions. A third outcome measure of 'on-task performance' (MALT accuracy) was also used to ensure that participants in each condition had performed at a comparable level on the MALT decoy task.

3. Results

3.1. Auditory Working Memory

Digit Span Backwards (DSB) raw scores (from both studies) were converted to standardized T-Scores and compared across groups using a one-way analysis of variance (ANOVA). A significant difference was found between groups; six-year-olds: N=150, M= 58.47, 95% CI [57.00, 59.95]; eight-year-olds: N=151, M= 53.61, 95% CI [52.31, 54.91]; 10-year-olds: N=153, M= 55.06, 95% CI [53.64, 56.48], F (2, 453) = 12.36, $p< .001$, with participants in the youngest group performing at a cognitive level significantly higher than the eight- and 10-year-old groups ($p< .001$ and $p<.002$ respectively). However, to confirm that the six-year-olds were performing significantly below the older age groups in raw ability score, these data were also analyzed. Results showed a significant effect of Age; six-year-olds: M= 76.03, 95% CI [72.23, 79.82]; eight-year-olds: M= 103.85, 95% CI [100.65, 107.05]; 10-year-olds: M= 117.93, 95% CI [114.72, 121.14], F(2, 453) = 152.71, $p< .001$, with significant differences between all groups (all $p< .001$).

3.2. MALT Task Performance (accuracy d' prime score)

To examine target detection accuracy (on-task performance) on the MALT decoy task, hit-rate (H = correct hits (50)/ number of target trials) and false-alarm rate (FA = commission errors/ number of non-target trials) were calculated for each participant and converted to z-scores. A *d'*prime [d' = z(H)

- z(FA)] measure of sensitivity was then calculated and mean values were analyzed across groups and conditions. Results of a two-way ANOVA with two fixed factors of Age (3 levels) and Condition (8 levels) found a significant main effect of Age, F(2, 430)= 22.59, $p< .001$, $\eta p^2 = .09$; with 6-year-olds < 8- and 10-year-olds ($p< .001$ for both). No significant effect of Condition, F (7, 430) = 1.93, $p= .06$, $\eta p^2 = .03$, and no Age by Condition interaction, F (14, 430) = 1.49, $p= .109$, $\eta p^2 = .05$ was found; showing that the effect of sensory learning condition on MALT accuracy performance across groups did not reach significance. Given that no significant differences in performance were found between any intrinsic and extrinsic conditions, data were collapsed across intrinsic and extrinsic conditions for clarity (Figure 3). A re-run of the ANOVA with collapsed conditions revealed comparable findings, with a significant main effect of Age, F(2, 439)= 24.71, $p< .001$, $\eta p^2 = .10$; with 6-year-olds < 8- and 10-year-olds ($p< .001$ for both). No significant effect of Condition, F (4, 439) = 1.77, $p= .133$, $\eta p^2 = .02$, and no Age*Condition interaction, F (8, 439) = 1.92, $p= .06$, $\eta p^2 = .03$.

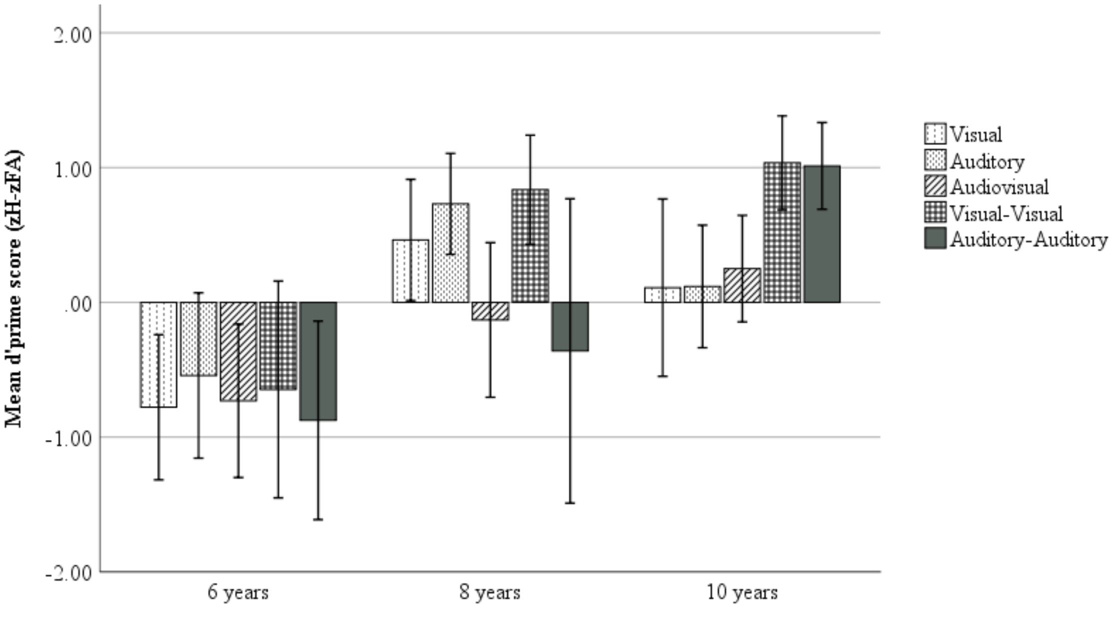

Fig 3. Mean accuracy score (*d*'prime) on MALT (zHit rate - zFalse Alarm rate) in each age group across conditions. Error bars 95% CI.

3.3. Test of Incidental Category Learning (primary outcome measure)

A two-way ANOVA with two between-subjects' factors of Age (3 levels: 6-, 8-, and 10-years) and Condition (8 levels: IV, IE, IAV, VV, AA, EV, EA, and EAV) was conducted to examine the mean number of correct scores (out of 16 trials) on the test of incidental category learning, with Bonferroni-corrected posthoc pairwise comparisons (Figure 4). No significant Age by Condition interaction was found (F<1). A significant main effect of Age, F(2, 426) = 20.87, $p< .001$, $\eta p^2 = .09$ was found, with 6-year-olds significantly below 8-year-olds and 10-year-olds ($p<. 001$ for both). A significant main effect of Condition was also found, F (2, 426) = 10.84, $p <.001$, $\eta p^2 = .15$. This was due to IV < IAV and EAV ($p< .001$ and $p= .008$, respectively); IA < VV ($p= .004$), IAV and EAV ($p= .001$ for both); VV > EV ($p< .001$) and EA ($p= .005$); EV < IAV, VV, EAV ($p<.001$ for all); EA < IAV and EAV ($p< .001$ for both). No significant difference was found between AA and any other condition, ($p>.05$ for all). As with the MALT accuracy score, no significant differences in performance were found between any intrinsic and extrinsic conditions within the same sensory category. For clarity, data were therefore collapsed across intrinsic and extrinsic conditions and re-analyzed using a 2-way ANOVA with Age (3 levels) and Condition (5 levels). The analyses indicate a main effect of Age, F(2, 435) = 11.81, $p<.001$, $\eta p^2 = .05$, with 6-year-olds significantly below 8-year-olds and 10-year-olds ($p<. 001$ for both). A main effect of Condition was also found, F(4, 435) = 8.95, $p<.001$, $\eta p^2 = .08$, with V< AV and VV ($p=.002$ and $p=.011$, respectively), and with A < AV, VV and AA ($p<.001$, $p<.001$, and $p=.044$, respectively).

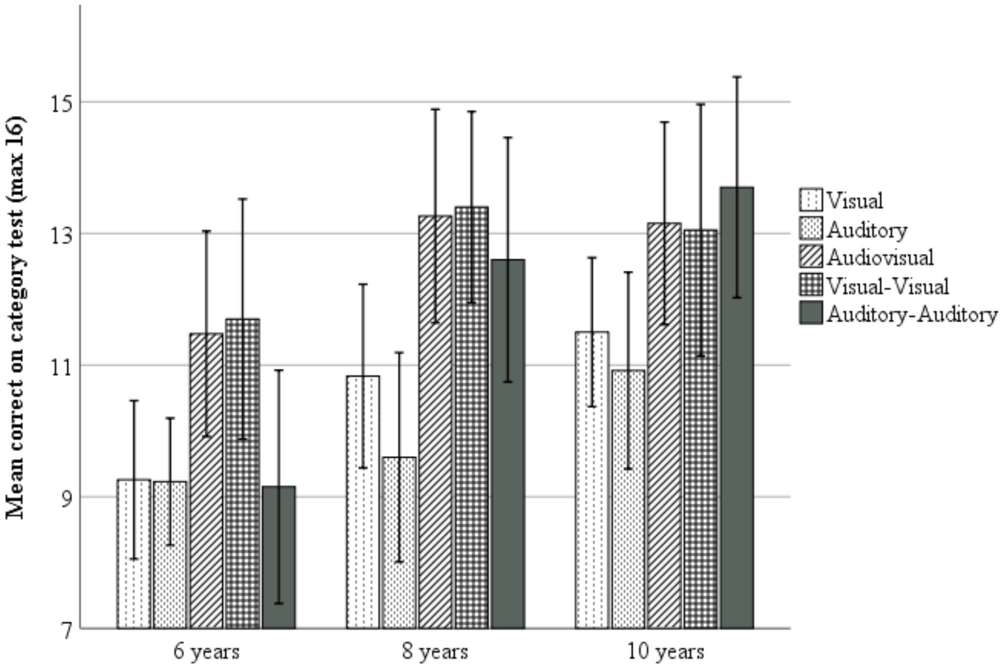

Fig 4. Mean correct on category test for each age group across conditions. Error bars 95% CI.

Given the effect of age and our a priori hypothesis *(hypothesis 3)* of differences in the ability to use different sensory information across development, we examined whether incidental categorization performance differed from chance in each age group and condition separately, using one-sample t-tests with a test value of eight. Results found that all groups performed significantly above chance (8) in all conditions (with *p= .041* for 6-year-olds V-only; *p=.014* for 6-year-olds A-only; *p=.049* for 8-year-olds A-only; all others *p(s) <.001,* except 6-year-olds in the AA condition for whom performance was not significantly different from chance, *t*(19)= 1.36, *p=.190.*

3.3.1. Relationship between MALT (decoy task) performance and incidental category learning

To ensure that performance on the incidental categorisation task was not related to accuracy on the initial MALT task, the relationship (controlling for age) between incidental categorisation scores and MALT accuracy (*d'prime*) scores, was examined. A partial correlation found that incidental category learning performance was not significantly related to accuracy score (*d'prime*) on the MALT, r = .041, *p = .383*.

3.4. Explicit Categorisation Knowledge

At the end of the category learning task, participants were asked to describe verbally how they knew where each frog lived, as a measure of explicit knowledge of category boundaries. Answers were coded to determine the particular sensory cue reported as having been perceived by each participant. Verbal responses were scored as follows; 0 points = don't know/no reason given; 1 point = related categorical description given but inaccurate (e.g., "they had different colored spots" in the IV condition); 2 points = partially correct family description (citing 1 feature but not both in IAV or EAV condition, e.g., different background patterns, but no mention of auditory features); 3 points = fully correct family description (e.g., "different number of spots and different croak sounds" in IAV condition or "croaks to log were deeper than croaks to lily pad" for IA condition). A mean explicit categorization score was calculated for each group and condition (Table 2). Results of a two-way ANOVA with two between-subjects factors of Age and Condition for explicit knowledge data found no Age by Condition interaction (F<1), but a significant main effect of Age, F(2, 430) = 23.76, $p < .001$, $\eta p^2 = .12$, (6 years < 8 and 10 years, $p < .001$) and a main effect of Condition, F(7, 430) = 4.06, $p < .001$, $\eta p^2 = .07$. Bonferroni-corrected posthoc tests revealed that significantly fewer points were scored in the Extrinsic Visual condition than in the Intrinsic Visual ($p = .002$) and in both the Intrinsic Multisensory (IAV) and Extrinsic Multisensory (EAV) conditions ($p < .001$ and $p = .009$, respectively). Points scored in the IAV condition were also significantly higher than in the two intra-modal conditions; AA ($p = .033$) and VV ($p = .046$).

Table 2.

Mean (SD) explicit categorization score for each age group across conditions (higher scores indicate stronger category learning)

	6 years (n= 150)	8 years (n=151)	10 years (n=153)
Intrinsic Visual	1.30 (1.26)	2.55 (1.09)	2.25 (0.91)
Intrinsic Auditory	1.55 (1.28)	1.80 (1.24)	2.05 (1.09)

Intrinsic Audiovisual	1.85 (1.04)	2.40 (0.99)	2.38 (0.59)
Extrinsic Visual	0.60 (0.51)	1.75 (1.18)	1.56 (1.15)
Extrinsic Auditory	1.33 (0.98)	2.07 (1.03)	2.19 (0.91)
Extrinsic Audiovisual	1.50 (0.95)	2.25 (0.44)	2.15 (0.49)
Visual-Visual	1.35 (0.75)	1.80 (0.69)	1.90 (0.64)
Auditory-Auditory	1.35 (0.86)	1.70 (0.57)	1.95 (0.61)

3.6. Discrimination Task

Mean number correct on the two (visual and auditory) extrinsic cues discrimination tasks conducted in six-year-olds (Visual: N=15, M_{corr}=10.87, SD_{corr}=1.46; Auditory: N=18, M=11.78, SD=.43); eight-year-olds (Visual: N= 11, M=11.91, SD=.30; Auditory: N= 11, M=11.64, SD=.67) and 10-year-olds (Visual: N=11, M=11.82, SD=.41; Auditory: N=11, M=11.91, SD=.30) were analyzed using a two-way ANOVA with two between-subjects variables of Age (six-, eight- and 10-year-olds) and Discrimination Condition (EV and EA). Analyses revealed a significant Age by Discrimination Condition interaction, F(2, 71) = 4.41, *p= .016.* Results also found a significant main effect of Age, F(2, 71) = 4.09, *p= .021*, but not of Discrimination Condition, F(1, 71) = 1.90, *p= .172*. Six-year-olds scored significantly below ten-year-olds (*p=.035*), with no other differences between groups (following Bonferroni correction). Further examination of the Age by Discrimination Condition interaction found that in six-year-olds, Visual discrimination scores were significantly below Auditory, t(31) = -2.53, *p= .017*. No other significant within-group differences were found (*p> .05*).

Discrimination between intrinsic cues are reported in Broadbent et al (2017), with comparable findings in the youngest group (6-year-olds) of visual discrimination scores significantly below auditory discrimination scores (*p= .045*).

4. Discussion

The role of unimodal and multimodal cues presented either within or across modalities in children's incidental category learning was examined. In support of *hypothesis 1*, when two informative cues were available, regardless of whether the two cues were within the same modality (e.g., visual-visual), or across two different modalities (i.e., audiovisual), this resulted in a marked facilitative effect on children's incidental category learning as compared to learning with a single unimodal cue. Presenting children with multiple cues, whether this is within the same sensory modality or across two modalities, is beneficial for incidental learning, and particularly for older children.

Even though category learning was enhanced with multimodal as compared to unimodal stimuli, the beneficial effect of having two informative cues was not the same in conditions in which the cues were intra-sensory and those in which the cues were multisensory across all age groups. For instance, the youngest children (6-year-olds) did not score significantly higher than would be expected by chance when exposed to two informative auditory cues (AA) on the incidental category learning task. These findings provide partial support for *hypothesis 2* predicting improved performance in the multiple intra-modality conditions as compared to multiple cross-modality conditions, as well as *hypothesis 3* predicting that these benefits would change across development. These findings indicate that, for younger children, learning with multiple auditory cues may be relatively difficult. Although this conclusion may seem to stand in contrast to findings of auditory (over visual) dominance in younger children (Hirst, Cragg & Allen, 2018; Robinson & Sloutsky, 2004; Zupan & Sussman, 2009), this conclusion is in line with other reports demonstrating greater difficulty in learning of category boundaries using auditory cues than visual cues (e.g., Broadbent, Osborne, Rea, et al., 2018; Broadbent et al., 2017; Noles & Gelman, 2012). In particular, the current study suggests that the presence of multiple auditory cues may not facilitate learning in 6-year-olds to the same extent as with older children, or to the same extent as multiple cues within a different modality (VV), or to the same extent as with cues presented across two different modalities (AV). Of note, however, is that performance in the youngest group for *unimodal* auditory cues in the current study was reliably above chance.

Differences in the weighting of auditory compared to visual stimuli in relation to the informative nature of the cues (Ernst & Bülthoff, 2004) may also go some way to explaining the current findings. Specifically, a single item of auditory information places a larger load on memory than one item of visual information (Fougnie & Marois, 2011). Therefore, the greater retention load of auditory cues may have resulted in poorer auditory learning across groups. Notably, however, despite relatively poorer learning with auditory cues in the younger age group (6-year-olds), this difficulty was not reflected in children's ability to discriminate between category exemplars on the basis of auditory information alone, since visual patterns that were extrinsic to the target stimuli were found to be more difficult to discriminate than extrinsic auditory cues. An interesting avenue for future research, particularly given the large confidence intervals in the current study, would be to examine factors related to individual differences in young children's ability to learn from specific modality cues at higher and lower cognitive loads.

Although no significant differences were found in performance levels between multisensory (AV) and intra-sensory (VV or AA) conditions, findings of particularly poor AA learning in 6-year-olds indicate that there may be something additionally beneficial to learning when informative cues are spread across two separate sensory modalities or within the visual modality in younger children. Fougnie et al. (2015) suggest that auditory and visual information from multisensory cues are processed separately in modality-specific stores. Such a processing strategy would result in greater chances of recalling correct categorical information than when informative cues are located within the same store and compete for attention (Fougnie & Marois, 2011; Quak et al., 2015). In younger children, therefore, it may be that multiple auditory cues are either processed, encoded or recalled differently to multiple visual cues. Further examination of why the processing of multiple unimodal auditory cues is more problematic than processing multiple unimodal visual cues in younger children, and at which point auditory cues may compete for attention, is an interesting avenue for future research.

Results from the explicit knowledge task revealed that the most difficult sensory cues to label were extrinsic visual cues. Thus, learning category boundaries using just extrinsic visual cues may be particularly difficult for children across this age range. Poor scores on the explicit knowledge test were also found in the two intra-modal conditions in which the recollection of both an extrinsic and intrinsic informative cue was required. However, the ability to describe explicitly the two distinct informative cues in the two AV conditions (intrinsic and extrinsic) was found to be higher. This suggests that when two cues are presented within the same sensory modality, competition for attention to the cues results in poorer explicit recall performance than with cues from two different sensory modalities that do not result in interference during either encoding or recall (Fougnie & Marois, 2011; Fougnie et al., 2015; Quak et al., 2015). That said, cross-modal interference on audiovisual tasks can also be found with sufficiently high memory load (Morey & Cowan, 2004, 2005), as well as in younger children when the multisensory cues are conflicting or irrelevant (Matusz et al., 2015). This provides further partial support for *hypothesis 2* but only in relation to explicit knowledge rather than the incidental learning. Furthermore, it is indicative of better retention of featural information following the presentation of multisensory cues than following the presentation of intra-modal cues that may compete for attention or working memory capacity constraints on modality-specific stores (Fougnie & Marois, 2011; Fougnie et al., 2015).

One aim of the current study was to examine differences in performance between intrinsic and extrinsic informative cues in unimodal and bimodal learning conditions. Intrinsic cues were predicted *(hypothesis 4)* to be more beneficial to children's learning than extrinsic cues given differences in the binding automaticity of extrinsic and intrinsic features (Ecker et al., 2007). However, extrinsic cues were not found to result in reliably different levels of performance on incidental learning than unimodal intrinsic cues. Interestingly, one of the highest levels of performance across age groups occurred when two extrinsic cues were presented together (the EAV condition), despite relatively poor performance unimodally in both cases. These findings, alongside children's high level of performance with intrinsic AV cues, add to previous findings of a facilitative effect on learning of multisensory as compared to unimodal stimuli, even when cues are not integrated into a unitary percept (Baker & Jordan, 2015; Broadbent et al., 2017; Jordan & Baker, 2011). The current results extend these previous findings by suggesting that informative cues within a single modality can also enhance children's ability to learn category boundaries.

In summary, in line with previous findings, having two informative cues presented in different sensory modalities was more beneficial to children's incidental learning than a single informative cue presented in a single sensory modality. In addition, two informative cues within the visual modality were better than a single unimodal cue (auditory or visual) across all ages, and were better than two informative auditory cues in children over 8 years of age. Learning from auditory cues was found to be markedly more difficult than from visual cues, possibly because of modality dominance factors or different weighting of cue reliability. This is despite findings of greater difficulty in discriminating between visual exemplars than auditory exemplars in the current study. In addition, two cues extrinsic to the target stimuli were no poorer at facilitating children's learning than two cues that are intrinsic to the stimuli. These findings contribute to research examining the role of multisensory cues in children's learning, highlighting in particular that two informative cues within the visual modality can also facilitate learning in primary school children to a greater extent than unimodal cues, albeit not necessarily to the same extent as simultaneously presenting two cues across different modalities in relation to explicit knowledge of the specific informative cues available.

Acknowledgements

This research was supported by funding from the Economic and Social Research Council (ESRC), grant reference: ES/K005308/1. We would like to thank all the children who participated in the study, and thanks to St Stephen's CE Primary School Twickenham, St Stephen's CE Primary School Shepherd's Bush, Orleans Primary School, Carlton Primary School, and St Paul's CE Primary School for taking part. Thanks also to Hayley White for assistance with collecting data for the original study conditions.

References

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.

Baines, L. (2008). *A teacher's guide to multisensory learning: Improving literacy by engaging the senses*. USA: Association for Supervision and Curriculum development.

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.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*, 433-436.

Broadbent, H. J., Osborne, T., Mareschal, D., & Kirkham, N. Z. (2018). Withstanding the test of time: multisensory cues improve the delayed retention of incidental learning. *Developmental Science*.

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

Chee, P., Logan, G., Schachar, R., Lindsay, P., & Wachsmuth, R. (1989). Effects of event rate and display time on sustained attention in hyperactive, normal, and control children. *Journal of Abnormal Child Psychology, 17*(4), 371-391.

Dekker, T. M., Ban, H., van der Velde, B., Sereno, M. I., Welchman, A. E., & Nardini, M. (2015). Late development of cue integration is linked to sensory fusion in cortex. *Current Biology, 25*(21), 2856-2861. doi:10.1016/j.cub.2015.09.043

Ecker, U. K., Maybery, M., & Zimmer, H. D. (2013). Binding of intrinsic and extrinsic features in working memory. *Journal of Experimental Psychology General, 142*(1), 218-234. doi:10.1037/a0028732

Ecker, U. K., Zimmer, H. D., & Groh-Bordin, C. (2007). The influence of object and background color manipulations on the electrophysiological indices of recognition memory. *Brain Research, 1185*, 221-230. doi:10.1016/j.brainres.2007.09.047

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

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*(4), 162-169. doi:10.1016/j.tics.2004.02.002

Fougnie, D., & Marois, R. (2011). What limits working memory capacity? Evidence for modality-specific sources to the simultaneous storage of visual and auditory arrays. *Journal of Expimental*

Psychology: Learning, Memory and Cognition, 37(6), 1329-1341. doi:10.1037/a0024834

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

Hirst, R. J., Cragg, L., & Allen, H. A. (2018). Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. *Neuroscience & Biobehavioral Reviews*, *94*, 286-301.

Hirst, R. J., Kicks, E., Allen, H. A., & Cragg, L. (2019). Cross-modal interference-control is reduced in childhood but maintained in aging: A lifespan study of stimulus- and response-interference in cross-modal and unimodal Stroop tasks. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(5), 553.

Hulme, C. (1979). *Reading Retardation and Multisensory Teaching*: Londres: Routledge & Kegan Paul.

Humphreys, G. W. (1998). Neural representation of objects in space: a dual coding account. *Philosophical Transactions of the Royal Society B: Biological Sciences, 353*(1373), 1341-1351.

Jacobs, R. A. (2002). What determines visual cue reliability? *Trends in Cognitive Sciences, 6*(8), 345-350. doi:10.1016/S1364-6613(02)01948-4

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

Kirkham, N. Z. (2010). Altogether now: Learning through multiple sources. In S. P. Johnson (Ed.), *Neoconstructivism: The New Science of Cognitive Development*. New York: Oxford University Press.

Kirkham, N. Z., Rea, M., Osborne, T., White, H. & Mareschal, D. (2019) Do cues from multiple modalities support quicker learning in primary school children? *Developmental Psychology,55,* 2048-2059.

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research, 35*(3), 389-412.

Mamassian, P. (2015). Sensory Development: Late Integration of Multiple Cues. *Current Biology, 25*(21), R1044-R1046. doi:10.1016/j.cub.2015.09.048

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

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, R., & Scerif, G. (2015). Multi-modal distraction: Insights from children's limited attention. *Cognition*, *136*, 156–165.

Morey, C. C., & Cowan, N. (2004). When visual and verbal memories compete: Evidence of cross-domain limits in working memory. *Psychonomic Bulletin & Review, 11*(2), 296-301. doi:10.3758/bf03196573

Morey, C. C., & Cowan, N. (2005). When do visual and verbal memories conflict? The importance of working-memory load and retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31*(4), 703-713. doi:10.1037/0278-7393.31.4.703

Murphy, G.L. (2002) The Big Book of Concepts. Cambridge, MA, US: MIT Press

Nardini, M., Bedford, R., & Mareschal, D. (2010). Fusion of visual cues is not mandatory in children. *Proceedings of the National Academy of Science U. S. A. 107(39)*, 17041-17046

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

Noles, N. S., & Gelman, S. A. (2012). Preschool-age children and adults flexibly shift their preferences for auditory versus visual modalities but do not exhibit auditory dominance. *Journal of Experimental Child Psychology, 112*(3), 338-350. doi:10.1016/j.jecp.2011.12.002

Ofman, W., & Shaevitz, M. (1963). The kinesthetic method in remedial reading. *Journal of Experimental Education, 3*, 317-320.

Okazaki, S., Hosokawa, M., Kawakubo, Y., Ozaki, H., Maekawa, H., & Futakami, S. (2004). Developmental change of neurocognitive motor behavior in a continuous performance test with different interstimulus intervals. *Clinical Neurophysiology, 115*(5), 1104-1113.

Quak, M., London, R. E., & Talsma, D. (2015). A multisensory perspective of working memory. *Frontiers in Human Neuroscience, 9*, 197. doi:10.3389/fnhum.2015.00197

Reber, A. S. (1993). *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*: Oxford University Press.

Robinson, C. W., Hawthorn, A. M., & Rahman, A. N. (2018). Developmental Differences in Filtering Auditory and Visual Distractors During Visual Selective Attention. *Frontiers in Psychology*, *9.*

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development, 75*(5), 1387-1401. doi:10.1111/j.1467-8624.2004.00747.x

Robinson, C. W., & Sloutsky, V. M. (2010). Development of cross-modal processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1)

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental Language Learning: Listening (And Learning) out of the Corner of Your Ear. *Psychological Science, 8*(2), 101-105.

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.

Troyer, A. K., & Craik, F. I. (2000). The effect of divided attention on memory for items and their context. *Canadian Journal of Experimental Psychology, 54*(3), 161-171.

Zupan, B., & Sussman, J. E. (2009). Auditory preferences of young children with and without hearing loss for meaningful auditory–visual compound stimuli. *Journal of Communication Disorders, 42*(6), 381-396. doi:10.1016/j.jcomdis.2009.04.002