Title:

Do cues from multiple modalities support quicker learning in primary school children?

Authors:

Natasha Z. Kirkham, Michaela Rea, Tamsin Osborne, Hayely White, & Denis Mareschal

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

Authors' note: We are grateful to all the children and schools who voluntarily participated in this study. We are particularly grateful to Daniel Richardson, Caspar Addyman, and Jennifer Aydelott for helping to develop the materials used in this work.

The work was funded by Economic and Social Research Council (ESRC), grant reference: ES/K005308/1. DM was also partially funded by a Royal Society –Wolfson research merit award.

Correspondence address:

Natasha Z. Kirkham
Centre for Brain and Cognitive Development
Birkbeck, University of London
Malet Street
London, WC1E 7HX
United Kingdom
n.kirkham@bbk.ac.uk

Word Count (main text): 8071

Abstract (249 words)

The current study investigates whether informative, mutually redundant audiovisual cues supports better performance in a category learning paradigm. Research suggests that, under some conditions, redundant multisensory cues supports better learning, when compared with unisensory cues. This was examined systematically across two experiments. In Experiment 1, children aged 5-, 7- and 10- years were allocated to one of the three 'modality' conditions (*audio informative only*, *visual informative only*, and *audiovisual informative*) and explicitly instructed to learn the category membership of individual exemplars, as determined by a threshold of correct responses. Unisensory or redundant multisensory cues determined category membership, depending on the learning condition. In addition to significant main effects of age group and condition, a significant interaction between age and sensory condition was found, with five-year-olds performing better when presented with redundant multisensory cues compared to unisensory cues. 10-year-olds performed better with auditory informative only cues, compared to visual informative only cues, or informative but redundant multisensory cues, with no significant difference between the latter two. In Experiment 2, the multisensory condition was presented to separate groups of 5-, 7-, and 10-year-olds, examining explicit learning outcomes in the *audiovisual informative* condition. Results showed that children who reached threshold during training were faster, made fewer errors, and performed better during test trials. Learning appeared to be based on the visual informative cues. Findings are discussed in the context of age-related selective attention, suggesting that the value of providing multisensory informative cues to support real-world learning depends on age and instructional context.

.

Keywords: multisensory; learning; development; education

Children's learning occurs in a dynamic multisensory environment that is complex and noisy. Throughout development infants and children learn to select and attend to the appropriate information, across different modalities and varying domains (e.g., Bahrick & Lickliter, 2014; Lercher, Evans, & Meis, 2003; Matusz, Broadbent, Ferrari, Forrest, Merkley, & Scerif, 2015; Tummeltshammer & Kirkham, 2013; Wu & Kirkham, 2010; Wu, Gopnik, Richardson, & Kirkham, 2011). Sometimes the cues in the different sensory streams are mutually supportive (i.e., the cues are consistent and informative, as in the instance of mouth movements during speech), but at other times they can be temporally offset therefore appear only partly correlated to a distal observer because of the difference in travelling velocity for light and sound (e.g., seeing a baseball player hit the ball before hearing the crack of the bat contacting the ball, while sitting in the baseball stands), or even irrelevant (e.g., the sound of outside traffic while reading a book). In the midst of all this, however, children learn which cues to attend to, which modalities to focus upon, and what information to ignore. This sensitivity to multiple cues, particularly those that are arbitrarily associated (e.g., letter shapes and sounds), has suggested to both researchers and educators that information received from multiple modalities could be 'supportive' of more complex learning in formal educational settings. Indeed, this idea has been the basis for many educational programs in literacy and numeracy, dealing with both typically and atypically developing children (Bullock, Pierce, & McClelland, 1989; Carbo, Dunn, & Dunn, 1986; Luchow & Sheppard, 1981; Mount & Cavet, 1995). In fact, many primary school educators have adopted a 'multimodal' approach to teaching (e.g., Jewitt, 2005; Kress, 2001; Moats & Farrell, 1999; Rowsell & Decoste, 2012). It is important to note that in educational settings, the term multimodal (in reference to classroom learning) has varied definitions, ranging from the combination of verbal and non-verbal representations of content (e.g., Jewitt, Kress, Ogborn, Tsatsarelis, 2001; Moreno & Mayer, 2007), to the usage of multimedia learning tools, such as web design (Doering, Beach, & O'Brien, 2007). In addition, this term does not include the surrounding, possibly distracting, multisensory information present in the classroom (Fisher, Godwin, & Seltman, 2014). Thus, stating that learning is supported by multisensory information becomes laden with issues of context.

The impact of multisensory cue combinations, and whether they support learning in a real-world setting, is thus an active and important topic of both educational and developmental research (e.g., Jewitt, 2005; Kress, 2001; Roswell & Decoste, 2012; Bremner et al, 2013).

Sensitivity to multisensory cue combination across childhood

The literature surrounding children's multisensory learning abilities is far from straightforward. Although the ability to consolidate inputs from different sensory inputs has been linked with increased IQ in children (Barutchu, Crewther, Fifer, Shivdasani, Innes-Brown, Toohey, et al, 2011), the actual experimental literature detailing multisensory integration in children is conflicting, with many researchers suggesting that mature integration does not occur until middle childhood, and that it is not until at least 8 years of age that children can use information from multiple modalities to reduce sensory uncertainty as well as adults can (e.g., Gori, Del Viva, Sandini, & Burr, 2008; Nardini, Bales, & Mareschal, 2015; Nardini, Bedford, & Mareschal, 2010 Nardini, Jones, Bedford, & Braddick, 2008). Indeed, Sloutsky and colleagues (e.g., Sloutsky & Napolitano, 2003) suggest that when 5-year-olds are presented with stimuli, they rely predominantly on auditory cues, relative to visual information, rather than combining both cues. They go on to posit that this "auditory dominance" shifts flexibly in young children, depending on the stimulus conditions. For example, 4-year-old children are likely to process stimuli only in the preferred modality when both auditory and visual information are of equal salience, suggesting that auditory and visual information compete for attentional resources early in childhood (Robinson & Sloutsky, 2004) – especially in tasks involving inductive learning. It is worth noting that these experiments are designed to present the cues equated for salience simultaneously. While this has provided important results on how children are combining cues, it does not always reflect how cues are presented in a more natural learning setting. In the real world it is unlikely that modalities occurring naturally in the sensory environment are always equated for salience. So, while the work of Sloutsky and colleagues provides valuable information regarding how young children are processing sensory input when all aspects are controlled for, it does not necessarily mean that this is how it works in a natural learning environment.

Importantly, a lot of multisensory integration requires combining cues that are arbitrarily associated. Being able to recognize informative associations is the cornerstone of much higher cognitive processing (e.g., reading, writing, causal connections). At 7 months of age, infants can learn associations between vowel sounds and objects (Gogate & Bahrick, 1998), and by 14 months, infants are capable of learning arbitrary associations between words and objects (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Indeed, newborns have been shown to learn arbitrary auditory-visual associations when the association is temporally redundant (Slater, Quinn, Brown, & Hayes, 2001). This suggests that there is an early developing ability to learn informative connections regardless of the content of the stimuli.

Establishing that children are able to *integrate* information across modalities, however, is not the same thing as establishing that they *use* this emerging ability to enhance learning in real world contexts such as the classroom. Recent work has shown that redundant audiovisual cues support greater incidental learning in a category learning task, both immediately after testing and 24 hours later (Broadbent, White, Mareschal, & Kirkham, 2017; Broadbent, Osborne, Mareschal, & Kirkham, 2019). In a continuous performance task, primary school-aged children were asked to detect a specific target among a temporal sequence of images. Unbeknownst to the participants, the targets were identified by visual, auditory, or audiovisual cues. A subsequent categorization task showed that children in the audiovisual cue condition outperformed those in the unisensory conditions in picking up on the incidental categories. This finding holds even when testing is delayed for 24 hours (Broadbent et al., 2019). This certainly suggests that perhaps the presence of more modalities does support learning, leaving open the question of whether or not this works across all types of learning, and across different age groups.

When looking at educational research in which multisensory information is presented simultaneously, there have also been some promising outcomes. There are several applied educational theories that show a positive benefit for school children (across primary to secondary education: looking at literacy, numeracy and geometry) when information is simultaneously presented in two modalities (auditory and visual; Mayer & Anderson, 1991; 1992; Paivio, 1990; visual-haptic, Kalenine, Pinet, & Gentaz, 2011) rather than in successive presentations. However, given the variety of methodologies presented across these different studies and the performance differences across age groups, there remains a large gap between the well-controlled research on multisensory sensitivity and integration, and children's ability to exploit multisensory information in a more naturalistic educational environment in which cues are often temporally decorrelated, of unbalanced salience, and embedded within rich external noise.

In sum, even though the benefits of multimodal integration have been seized upon by educational training programs and clinical interventions, there is a dearth of systematic research on which to base this practice. The studies presented here have therefore been designed to investigate several questions: 1) Do informative multisensory cues support better performance, in comparison to unisensory cues, on a simple instructed learning task, 2) Does this performance differ across primary-school children aged 5 to 10 years, and 3) What exactly is being learned when multiple cues are provided?

In order to investigate this, a new paradigm, the *Multisensory Category Learning Task* (MCLT), was designed to test children's ability to use informative cues from different modalities in order to determine the category membership of exemplars. Category learning was selected as the dependent measure because it is central to cognition, cognitive development, and learning across multiple domains, as well as having a long history of robust research behind it (see Mareschal, Quinn & Lea, 2010 or Murphy, 2002). It was hypothesised that there would be age-related improvements in category learning, in line with previous developmental work. It was also hypothesised that, if informative multisensory information is indeed more supportive of learning, children at all ages would perform better during the multisensory learning condition.

Experiment 1

.

Method

Participants

One hundred and eighty primary school children (90 male, and 90 female) were separated into three age groups; 5-year-olds (31 male, 29 female; M= 67.1 months, SD= 3.6 months, Range= 58 to 73.4 months), 7-year-olds (29 male, 31 female; M= 90 months, SD=5 months, Range= 84.1 to 101.1 months) and 10-year-olds (30 male, 30 female; M=120.2 months, SD= 3.1 months, Range= 114.9 to 127.7 months). There were 60 children per age group who were randomly allocated to one of three sensory conditions of the MCLT (*visual informative only, audio informative only,* or *audiovisual informative*), resulting in 20 children per condition. All participants had normal hearing and normal or corrected-to-normal vision, and no known neurological or developmental disorder. Children were recruited from local primary schools and informed written parental consent was given for each participant. Ethical approval was obtained from Birkbeck, University of London ethics' committee ("Learning from multisensory cues," approval number: 131453). 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. Sample size was determined based on a power analysis that showed that a 2x2 ANOVA, with three levels in each factor, would require an N of 15 in each cell for a power of 0.8, at alpha=0.05, with a medium effect size of 0.3. To be conservative we ran 20 in each cell.

Apparatus and stimuli

In all conditions of the computerized MCLT, children were asked to learn the category membership of a series of individual exemplars (cartoon aliens). The game was presented on a 15.4 inch MacBrook Pro Laptop at 1280x800 pixels, using Macromedia Director MX. The display consisted of a black background, upon which two different spaceships would appear, one positioned in the top right corner of the screen, and one in the top left corner. The spaceships were similar in shape, but differed in color, with one colored orange and one colored silver. The category exemplars are brightly colored cartoon aliens that can vary on 3 different visual dimensions: Body length, stripe frequency, and number of legs (see Figure 1). The number of legs were 2, 3, 5, or 6. The stripe angle varied in 11 steps of 15º from -75º to 90º to the vertical. The stripe frequency varied in unit steps from 1 to 12 cycles per standard unit. Arm length, body width and body height all varied continuously, their exact values sampled from pre-specified statistical distributions (Addyman, 2009). Stripe frequency from each end of the distribution (i.e., thick or thin) or number of legs (many: 5 or 6, or few: 1 or 2) was used as category membership values, with exemplars' bodies varying within the two categories. The alien sounds varied across tone: breathy, fluttery, and pure (two of which were selected at random as the category distinguishing variables), with pitch varying within the two categories. Multiple versions of the syllable 'yow' (/jæw/) were created in Praat (Boersma, 2001) using KlattGrid acoustic synthesis. All stimuli were 750 msec in duration with a nominal amplitude of 90 dB, and consisted of an onglide (250 msec), steady state vowel (250 msec), and offglide (250 msec). Fifteen stimuli were generated for each voice quality 'family', with steady state F0 increasing in semitone (1.05946309436 Hz) steps from 100 Hz (Stimulus 1) to 224.4924097 Hz (Stimulus 15). The fundamental frequency contour of each stimulus was calculated based on the steady state F0 as follows: onset F0 = steady state – 50 Hz; offset F0 = steady state – 100 Hz. Formant frequencies for each portion of the stimulus were derived from the average F1, F2, and F3 values reported for men, women, and children in Peterson & Barney (1952) - onset of the onglide was based on values for /i/, steady state /æ/, offset of the offglide /u/. Voice quality 'families' were generated using the 'breathiness' and 'flutter' functions in Praat: for 'pure' stimuli, no flutter or breathiness was added; for 'breathy' stimuli, breathiness was added at an amplitude of 90 dB from stimulus onset; and for 'flutter' stimuli, flutter was set at a value of 1 from stimulus onset. Category membership values were counterbalanced across conditions and participants, as explained below.

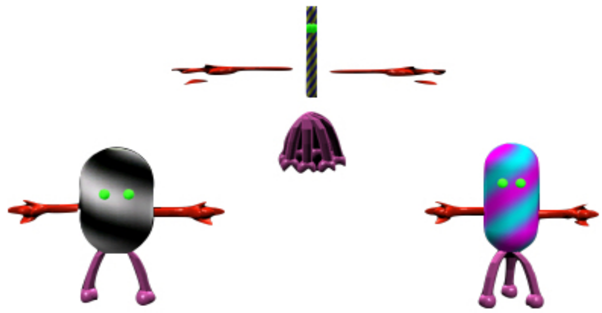
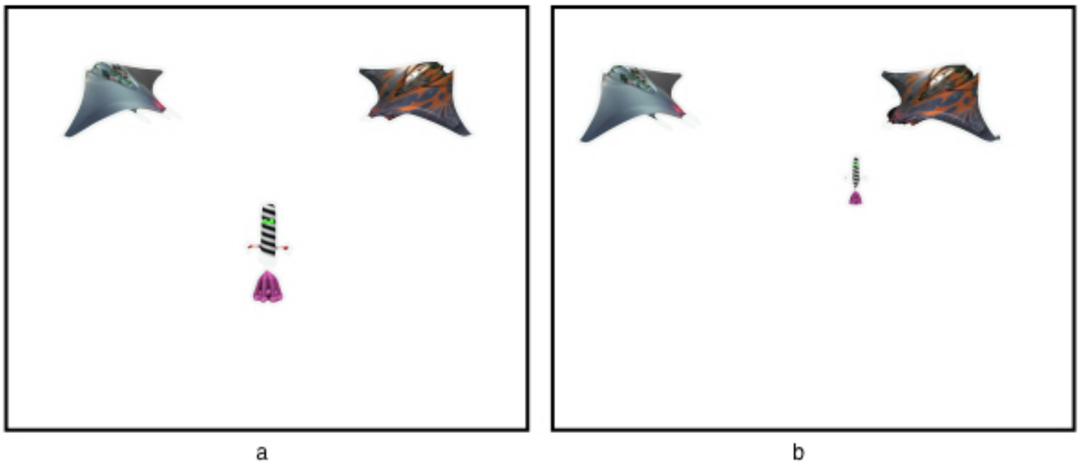


Figure 1. Examples of alien stimuli used in Experiment 1.

Design and procedure

Each child was shown a screenshot of the computer game, containing one alien and the two spaceships, while the experimenter explained the instructions of the game. Children were told that they would see an alien appear in the middle of the screen and that this alien either lived in the silver spaceship or the orange spaceship; their job was to help the alien get home by pressing either the button for the silver spaceship or the button for the orange spaceship. No information about how to determine category membership was given. After the instructions, the screen went black and the two spaceships appeared in the top left corner and the top right corner, and an alien appeared in the center of the screen (trial one; see Figure 2a). The child was then asked to choose either the silver spaceship by pressing the left arrow button or the orange spaceship by pressing the right arrow button. The left and right arrow keys had a silver and orange sticker affixed to them, respectively, with pictures of the spaceships below. After the keypress, the alien moves to its correct spaceship, regardless of arrow pressed (see Figure 2.b), and the experimenter provides feedback by saying either *"Well done! That alien did live in the silver/orange spaceship"* or, if incorrect, *"The alien lived in the silver/orange spaceship that time."* The experimenter then pressed the spacebar and the next alien appeared for trial 2. The game continued until the child had performed correctly on six consecutive trials (*threshold*) or until a total of 60 trials had passed.



Figures 2a and 2b. Example of one test trial in Experiment 1. (a) The alien exemplar is presented in the middle of the screen, and after the participant makes a choice, (b) the alien moves towards the correct spaceship (regardless of participant choice being correct or incorrect).

Each child was randomly assigned to one of three conditions: *visual informative only* (V), *audio informative only* (A) or *audiovisual informative* (AV). In the *visual informative only* condition, the aliens differed on a visual feature (e.g., thin vs. thick stripes), with an accompanying sound that was the same for every alien (a 'whooshing' sound as the alien appeared, identical in length to the alien sounds; 750 msec). In the *audio informative only* condition, the same randomly selected alien was presented on each trial. The alien was accompanied by a sound that differed in one feature (e.g., breathy vs. pure tone); this sound occurred as the alien appeared. In the *audiovisual informative* condition, the auditory and visual features both differed, but were redundant (e.g., the breathy tone went with thick stripes and the pure tone went with thin stripes). There were three possible auditory comparisons that the children could learn the categories from: Pure versus Breathy, Pure versus Fluttery or Breathy versus Pure, which were randomly presented within the game. There were two visual comparisons: thin stripes versus thick stripes or many legs versus few legs, which were counterbalanced by the experimenter across participants. Each child participated in only one condition. The category learning game outputted results for each child including their spaceship choice and the reaction time for each trial.

At the end of this task, every fourth child also completed both a visual and an auditory discrimination task. Testing every fourth child instead of the entire sample was entirely due to the length of the task. When testing children as young as 5 years of age there is a danger of attention fatigue. The child would see/hear two aliens and was asked, "Do you think these two aliens look/sound the same or look different?" In total, there were 30 comparisons; 15 visual and 15 auditory, which were blocked with half the children receiving the 15 auditory comparisons first and half the children receiving the visual comparisons first. The stimuli were taken from the same variations of aliens as used previously in the learning game and were chosen to include a variety of how similar/different they were from each other. Six pairs were completely different (chosen from different categories), two pairs were different but chosen from the same category with a minimum distance apart, one pair was different but chosen from the same category and a medium distance apart, one pair was different but chosen from the same category and maximum distance apart, and six pairs were identical. The sum of the total items correctly discriminated for both audio and visual were calculated.

Data Analysis

Overall performance on this task was measured by looking at: (1) number of trials to threshold (six correct categorizations in a row), (2) total number of errors committed, and (3) reaction times to category choice. Total number of errors should be highly correlated with number of trials to threshold. For example, a child who achieved threshold at the 10th trial could have a maximum of four incorrect trials; whereas, a child who achieved threshold at the 20th trial could have a maximum of 14 incorrect trials; thus, it is more likely to make more errors the longer it takes to get to threshold. This dependent measure was included to ensure that this was the case, and that there were not a significant number of children almost reaching threshold multiple times in a row (i.e., 5 correct choices followed by 1 or 2 errors, followed by 5 correct choices). Performance was compared across age and condition. All participants were included, regardless of whether they achieved the threshold of six correct categorizations in a row. We hypothesized that if redundant multisensory information was supportive of learning category membership, we should see fewer trials to threshold, fewer numbers of errors and shorter reaction times across all age groups in the *audiovisual informative* (AV) condition, compared to the *audio informative only* (AO) and *visual informative only* (VO) conditions. In addition, we hypothesized that overall performance would improve as age increased, with trials to threshold and total number of errors decreasing. Although reaction times are notoriously noisy in children (and indeed, we did not stress quick reactions to the participating children), we hypothesized that if reaction time differences were detected, slower reaction times would be found in the harder conditions (as determined by number of trials to threshold).

Results

Discrimination task

A 3 (age-group: 5-, 7-, and 10-year-olds) X 3 (condition: AO, VO, AV) two-way multivariate analysis of variance (ANOVA) was performed on visual and auditory discrimination trials. There was a significant effect of age, on both the visual discrimination trials ($F[2,45] = 6.47$, $p = .003$, $\eta_p^2=0.22$) and the auditory discrimination trials ($F[2,45] = 4.42$, $p = .018$, $\eta_p^2=0.16$). This was driven by the 5-year-olds performing significantly worse than either 7- or 10-year-olds on the visual discrimination task (*M* [5-year-olds] = 13.2; SD = 2.23; *M* [7-year-olds] = 14.7; SD = .69; *M* [10-year-olds] = 14.6; SD = .75) and the 5-year-olds performing worse than the 10-year-olds on the auditory discrimination task (*M* [5-year-olds] = 9.2; SD = 2.37; *M* [10-year-olds] = 11.1; SD = 1.62) but there was no main effect of Condition for either dependent measure. There were no other differences or effects. One-sample t-tests were performed to investigate whether performance on the discrimination task was different from chance: When looking at each age group individually, results showed that all age groups were better than chance on both the auditory [5-year-olds: $t(14) = 3.69$, $p = .002$; 7-year-olds: $t(16) = 5.00$, $p < .001$; 10-year-olds: $t(17) = 10.01$, $p < .001$], and the visual [5-year-olds: $t(14) = 7.7$, $p < .001$; 7-year-olds: $t(16) = 25.34$, $p < .001$; 10-year-olds: $t(17) = 38.19$, $p < .001$] discriminations.

Multisensory category learning task

Tables 1a and 1b show the *number of trials to threshold, number of errors,* and the *reaction time* for each age group across all three conditions. The three dependent variables were analyzed with a 3 (age group: 5-, 7-, and 10-year-olds) x 3 (sensory condition: AO, V, and AV) two-way multivariate analysis of variance (ANOVA).

Number of trials to threshold.

Results for *number of trials to threshold* revealed significant main effects of age group (F[2,171]=7.49, p=.001, η_p^2=0.08) and sensory condition (F[2,171]=5.82, p=.004, η_p^2=0.06), and a significant interaction between age and sensory condition (F[2,171]=7.62, p<.001, η_p^2=0.15). As hypothesized, mean *number of trials to threshold* decreased with age (M [5-year-olds] = 47.32, SD = 18.65; M [7-year-olds] = 39.85, SD = 19.99; M [10-year-olds] = 34.38. SD =21.73).

a	**Age group**	**Mean**	**95% Confidence Interval**	
			Lower Bound	*Upper Bound*
Number of trials to threshold	5-year-olds	47.317	42.634	51.999
	7-year-olds	39.850	35.167	44.533
	10-year-olds	34.383	29.701	39.066
Total errors	5-year-olds	23.533	20.835	26.232
	7-year-olds	18.517	15.818	21.215
	10-year-olds	14.917	12.218	17.615
Mean score of reaction times	5-year-olds	2877.741	2468.188	3287.295
	7-year-olds	2427.486	2017.933	2837.040
	10-year-olds	2681.189	2271.635	3090.742

b	**Sensory Condition**	**Mean**	**95% Confidence Interval**	
			Lower Bound	*Upper Bound*
Number of trials to threshold	Auditory Only	34.133	29.451	38.816
	Visual Only	45.183	40.501	49.866
	Audiovisual	42.233	37.551	46.916
Total errors	Auditory Only	15.600	12.901	18.299
	Visual Only	21.717	19.018	24.415
	Audiovisual	19.650	16.951	22.349
Mean score of reaction times	Auditory Only	2588.495	2178.942	2998.048
	Visual Only	2819.838	2410.284	3229.391
	Audiovisual	2578.084	2168.530	2987.637

Tables 1a and 1b: Table 1a shows the *number of trials to threshold, number of errors,* and the *reaction time* for each age group, and Table 1b shows the *number of trials to threshold, number of errors,* and the *reaction time* across all three conditions

Planned comparisons revealed that 5-year-olds needed significantly more trials to attain threshold than 10-year-olds (Tukey HSD; p<.001), and marginally more trials than 7-year-olds (Tukey HSD; p=.07), with no significant differences between the latter two age groups. In addition, 5-year-olds made significantly more errors during performance than 7-year-olds (Tukey HSD; p=.03), and 10-year-olds (Tukey HSD; p<.001). There were no differences between 7- and 10-year-olds. Planned comparisons performed on the sensory conditions show that the *audio informative only* condition (M = 34.13; SD = 22.24) required significantly fewer trials to threshold than *visual informative only* (M = 45.18; SD = 18.55; Tukeys HSD; p= .003) or *audiovisual informative* (M = 42.23; SD = 19.99; Tukeys HSD; p= .04). There was no significant difference between *visual informative only* and *audiovisual informative*. There were significantly fewer errors in the *audio informative only* condition (M = 15.60; SD = 12.8) than in the *visual informative only* condition (M = 21.72; SD = 11.18; Tukey HSD; p=.005), and marginally fewer errors than the *audiovisual informative* condition (M = 19.65; SD = 11.34; Tukey HSD; p=.09)

For *number of trials to threshold,* the significant interaction between sensory condition and age was unpacked using separate univariate ANOVAs for each age group (see Figure 3). For 5-year-olds, a significant main effect of condition was apparent (F[2,57]=4.90, p=.011, η_p^2=0.11. Post hoc analyses revealed that this age group was significantly better at category learning during the *audiovisual informative* condition (M = 37.35; SD = 23.13), compared to either the *visual informative only* condition (M = 53.2; SD = 13.56; Tukeys HSD; p=.016) or the *audio informative only* condition (M = 51.4; SD = 14.24; Tukeys HSD; p=.037). There were no differences between the *visual informative only* or the *audio informative only* condition. Analyses revealed no differences between the three sensory conditions for the 7-year-olds (M [VO] = 39.55; SD = 19.11; M [AO] = 35.05; SD = 22.86; M [AV] = 44.95; SD = 17.35). In the 10-year-old group, there was a main effect of sensory condition (F[2,57]=16.5, p<.001, η_p^2=0.37), with learning facilitated by *auditory informative only* (M = 15.95; SD

= 12.19) compared to *visual informative only* (*M* = 42.80; SD = 20.30; Tukeys HSD; p<.001) or *audiovisual informative* (*M* = 44.40; SD = 19.18; Tukeys HSD; p<.001), with no significant difference between the latter two.

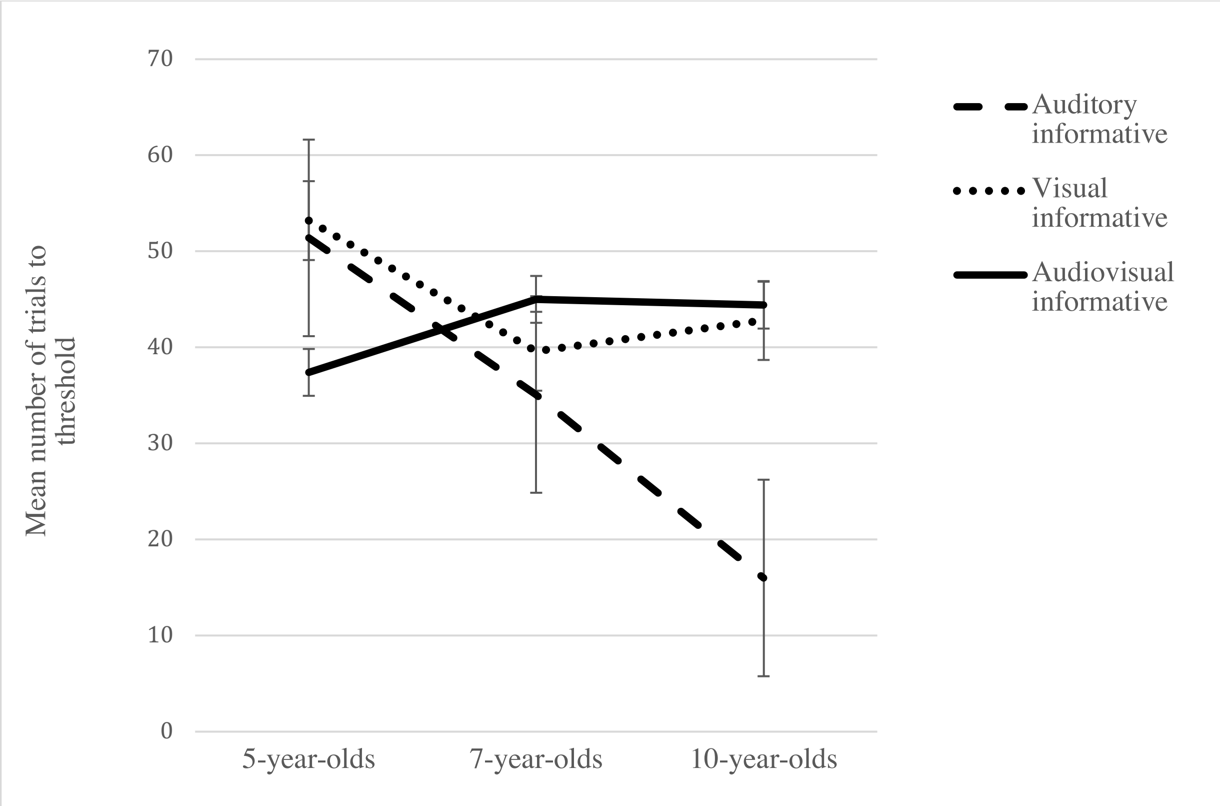


Figure 3. Experiment 1: Mean number of trials to threshold (six consecutive correct choices in a row), with standard errors, across age groups and sensory conditions.

Does children's performance on reaching the designated threshold of six correct in a row differ significantly from chance? Obtaining 6 successful trials in a row was selected as the criterion for success because the probability of getting 6 out of 6 binary choices correct is 1/64, which is 0.016. However, because there are 60 possible trials in a test session, there are, in fact, multiple independent 6-element windows in which the children could get 6 right answers correct by chance. When added up, the probability of getting 6 trials correct in a row out of a possible 60 is above the conventional p=0.05 limit. However, we can ask what the average number of trials to reach this threshold would be for a sample of 20 children selecting responses randomly. Monte Carlo simulations can be carried out to establish this chance-level group performance. To this end, we ran simulations (choosing by chance many times) to identify the expected value of the "trials to success" parameter, if 20 children chose by chance for a maximum of 60 trials each. The answer is 49.11 (estimated using 500000 samples of 20 simulated participants; Matlab simulation code can be supplied on request). We can then compare the performance of the samples of actual children in each of the test conditions to this chance level value to establish whether their performance differs from chance. As seen in the analyses, when children are assessed in this way only the 5-year-olds in *auditory alone* condition (trials to threshold = 51.4) and those in the *visual alone* condition (trials to threshold = 53.2 have behaviors that are not significantly different from chance. All other children reach threshold on significantly fewer trials than the 49.11 expected by chance. They are therefore all performing at better than chance level.

Numbers of errors.

Results for *number of errors* revealed significant main effects of age group (F[2,171]=10.02, p<.001, η_p^2=0.11) and sensory condition (F[2,171]=5.18, p=.007, η_p^2=0.06). There was a significant interaction between age and sensory condition (F[2,171]=7.17, p<.001, η_p^2=0.14). As hypothesized, mean *number of errors* decreased with age (M [5-year-olds] = 23.53; SD = 11.04; M [7-year-olds] = 18.52; SD = 11.44; M [10-year-olds] = 14.92; SD = 12.10).

For *number of errors,* the interaction between sensory condition and age was unpacked using separate univariate ANOVAs for each age group. For 5-year-olds, a significant main effect of condition was apparent (F[2,57]=5.11, p=.009, η_p^2=0.15. Post hoc analyses revealed that this age group made significantly fewer errors during the *audiovisual* condition (M = 17.50; SD = 12.92), compared to either the *visual informative only* condition (M = 26.70; SD = 8.71; Tukeys HSD; p=.007) or the *audio informative only* condition (M = 26.40; SD = 8.82; Tukeys HSD; p=.009). There were no differences between the *visual informative only* or the *audio informative only* condition. Analyses revealed no differences between the three sensory conditions for the 7- year-olds (M [AO] = 15.40; SD = 12.31; M [VO] = 19.25; SD = 11.74; M [AV] = 20.90; SD = 10.03). In the 10-year-old group, there was a main effect of sensory condition (F[2,57]=14.9, p<.001, η_p^2=0.34), with number of errors being significantly lower in the *audio informative only* condition (M = 5.00; SD = 6.27) compared to *visual informative only* (M = 19.20; SD = 11.62; Tukeys HSD; p<.001) or *audiovisual informative* (M = 20.55; SD = 11.16; Tukeys HSD; p<.001), with no significant difference between the latter two.

There were no significant effects for *reaction times*.

Discussion

The current study used a multisensory category learning task to examine the effects of multisensory redundant information on children's learning of category membership across primary school ages. Results from this experiment show some significant differences in the benefits of redundant multisensory information on a category learning task. As hypothesized, younger children required significantly more trials to achieve threshold than older children, showing that there is an increase in category learning ability across primary school years. In addition, number of errors on the task decreased with age, also in line with predictions. The effects of sensory modality are more complex. Informative multisensory information provided a learning benefit (in comparison to the unisensory informative conditions) only in the youngest age group. The 5-year-olds required significantly fewer trials to achieve threshold in the *audiovisual informative* condition than in either of the unisensory *informative* conditions, and showed performance significantly different from chance only in the *audiovisual informative* condition. Both the 7-year-olds' and the 10-year-olds' performance was significantly different from chance in all conditions. There were no differences between conditions for the middle group (7-year-olds); however, the 10-year-olds show significantly better performance on the *audio informative only* condition than on either of the other two conditions. Far from showing a blanket benefit for informative redundant multisensory cues, this study suggests that younger children benefit from the greater support provided by the presence of multiple informative but redundant sensory cues, but that by 10 years of age, children are capable singe cue (here auditory) learners. In fact, the addition of concurrent informative but redundant visual information at this age *reduces* performance as compared to that in the *audio informative only c*ondition. Two questions that emerge from this study are: 1) *What* is actually being learned during the *audiovisual informative* condition? Is it the case that the modalities are learned separately (or equally well), or have the modalities merged during the trials? And, 2) does reaching threshold in this task equate to learning of the categories?

Experiment 2

In this experiment, the *audiovisual informative* condition of the MCLT was presented to three new groups of 5-, 7-, and 10-year-olds, followed by a task that tested the informative visual and auditory cues separately.

Method

Participants

Sixty-one children (32 male, 29 female) were separated into three age groups; 5-year olds (10 male, 10 female; M= 75.2 months, SD= 3.5 months, Range= 70 to 80 months), 7-year olds (10 male, 11 female; M= 94 months, SD=5 months, Range= 81 to 101 months) and 10-year olds (12 male, 8 female; M=122.1 months, SD= 3.9 months, Range= 115 to 122 months). All 61 children were allocated to one condition of the MCLT (*audiovisual informative*). All participants had normal hearing and normal or corrected-to-normal vision, and no known neurological or developmental disorder. Children were recruited from local primary schools and informed written parental consent was given for each participant. Ethical approval was obtained from University of XXX ethics' committee ("Learning from multisensory cues," approval number: 131453). 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.

Apparatus and stimuli

The same computerized version of the MCLT, as was used for Experiment 1, was used in the current experiment. The only changes made was that all participants received the *audiovisual informative* condition only.

Design and procedure

As was the case in Experiment 1, each child was tested individually in a session lasting approximately 15 minutes. Testing took place either in a classroom or quiet room within the participating school. The procedure was identical to the *audiovisual informative* condition, with one exception: In this version, the game always continued until a total of 60 trials had passed (*training trials*), regardless of how many trials were correct. In other words, the task did not stop once the children had performed six correct choices in a row, as in Experiment 1. This was to ensure that individual children did not receive different numbers of trials depending on the speed with which they reached threshold (as happened in Experiment 1); each child received all 60 trials. At the end of the 60 training trials, each child was presented with a total of ten test trials: five audio alone trials and five visual alone trials, during which the sensory information determining category membership was 'pulled apart'. In the audio alone test trials, the screen was the same as during the previous 60 training trials except that when alien arrived with its accompanying sound, it was hidden completely behind a cloud containing a question mark (see Figure 4). In the visual alone test trials, the screen was the same as during the previous 60 training trials except the aliens appeared in silence, without their accompanying sounds. After the child pressed the button associated with either the silver or the orange spaceship the alien 'teleported' with a flash of green light to their spaceship (i.e., the spaceship did not fly home so there was no feedback regarding the veracity of the child's choice). The ten trials were presented in random order.

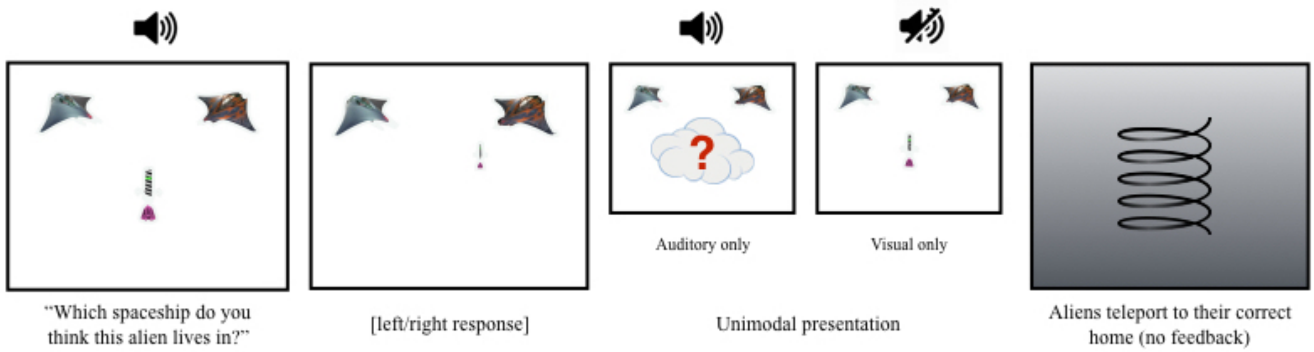


Figure 4. Test trials schematics in Experiment 2.

Data Analysis

As in Experiment 1, overall performance on this task was measured by looking at: (1) total number of errors committed, and (2) reaction times to category choice, during the 60 training trials. Whether the children reached the threshold value of 6 consecutive correct trials was also recorded. Performance on the ten test trials was coded as correct or incorrect, based on the choice of spaceship. Performance on test trials was compared across age and by whether 'threshold was reached.'

Results

Training trials

All participants.

A one-way (age group: 5-, 7-, and 10-year-olds) multivariate analysis of variance (ANOVA) was performed looking at the two dependent variables *number of errors* and *reaction times*. There was a significant main effect of age group for *number of errors* ($F[2,58]=3.45$, $p=.039$, $\eta_p^2=0.11$), with 7-year-olds making significantly more errors ($M = 23$, $SD = 8.46$; Tukey's HSD; $p = .039$) than 5-year-olds ($M = 15.5$; $SD = 9.19$) and *reaction times* ($F[2,58]=7.76$, $p=.001$, $\eta_p^2=0.21$), with 5-year-olds performing significantly slower that 10-year-olds (*M* [5-year-olds] = 3803.82msec; SD = 2074.01; *M* [10-year-olds] = 2027.88; SD = 982.33msec; $p<.001$), and 7-year-olds performing slower than 10-year-olds (*M* [7-year-olds] = 3009.39msec; SD = 956.79; $p=.03$) . There were no other main effects.

Participants who achieved threshold

A one-way (age group: 5-, 7-, and 10-year-olds) multivariate analysis of variance (ANOVA) was performed looking at the two dependent variables *number of errors* and *reaction times* on only those children who reached threshold. There was a significant main effect of age group for *number of errors* ($F[2,49]=4.14$, $p=.022$, $\eta_p^2=0.14$), with 7-year-olds making significantly more errors ($M = 21.27$; SD = 7.81) than 10-year-olds ($M = 13.47$; SD = 9.28; Tukey's HSD; $p=.03$), and a trend towards significance with 5-year-olds ($M = 14.74$; SD = 8.67; Tukey's HSD; $p=.06$). There was also a significant main effect for *reaction times* ($F[2,49]=8.12$, $p=.001$, $\eta_p^2=0.25$), with 5-year-olds performing significantly slower (M = 3753.05; SD = 2118.03) than 10-year-olds (M = 1780.21; SD = 684.02; Tukey's HSD; $p=.001$). There were no other main effects.

In order to test the hypothesis that reaching threshold would enhance performance on both the 60 training trials and the test trials, analyses were performed on all participants with the variable *Threshold Reached* introduced as a between-subject factor. A 3 (age group: 5-, 7-, and 10-year-olds) x

2 (threshold reached: yes, no) two-way between-subject multivariate analysis of variance (ANOVA) was performed looking at *number of errors* and *reaction time*.

Reaction time.

There was a significant main effect of *Age Group* [F(2,55) = 4.28, p = .019] with 5-year-olds performing significantly slower (M = 3803.782 msec; SD = 2074.01) than 10-year-olds (M = 2027.88; SD = 982.33 msec; Tukey's HSD; p = .001). There was no difference between 7-year-olds and either of the other two age groups. There was also a significant main effect of *Threshold Reached* [F(1,55) = 17.02, p < .001) with children who had reached threshold making fewer errors (M= 16.63; SD = 9.08) than those who did not reach threshold (M = 30.78; SD = 3.67). In addition, there was a trend for a significant effect of *threshold reached* on reaction time [F(1,55) = 3.59, p = .063], with those who reached threshold performing faster (M = 2853.84 msec; SD = 1605.24) than those who did not (M = 3492.40 msec; SD = 1391.38). There were no other significant effects (see Figure 5).

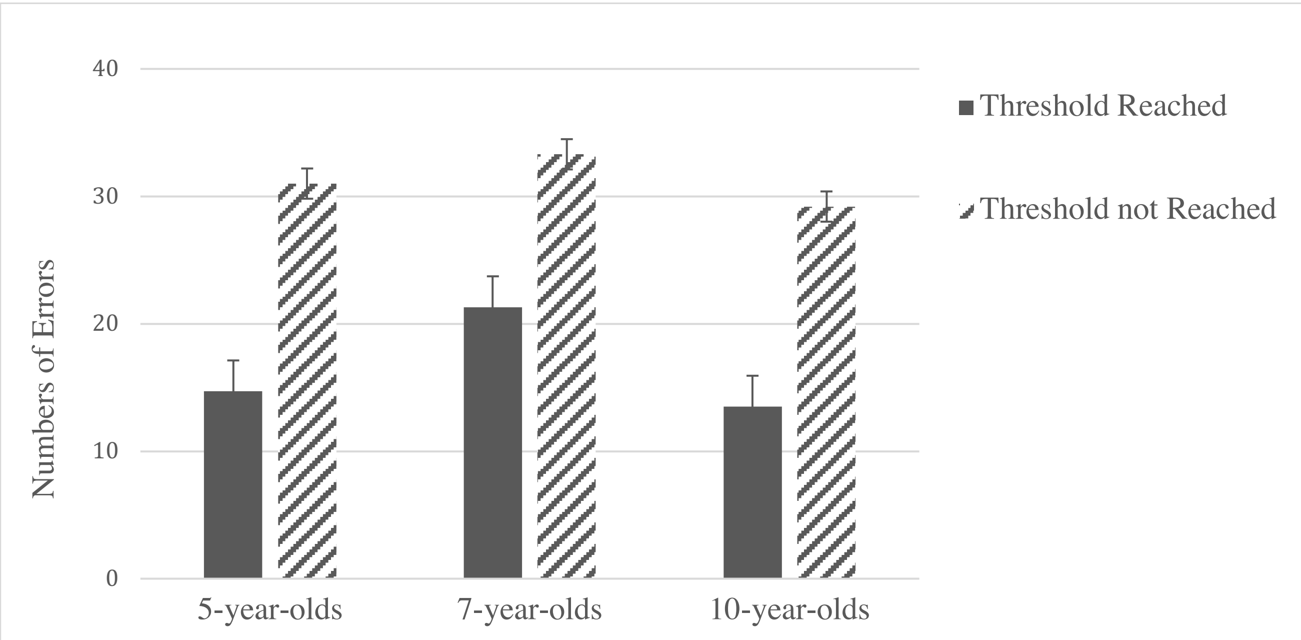


Figure 5. Mean number of errors, with standard errors, by age group, and whether threshold was reached in Experiment 2.

Test trials

In order to measure whether reaching threshold enhanced performance on the test trials, the variable *threshold reached* was again used as a between-subjects' variable when assessing performance on test trials. A 3 (age group: 5-, 7-, and 10-year-olds) x 2 (threshold reached: yes, no) two-way analysis of variance (ANOVA) was performed, with modality of test trial as a within-subjects' variable. There was a significant effect of threshold [$F(1,55)=4.01$, $p=.012$, $\eta_p^2= .11$], with children who reached threshold on the training portion of the task, performing better on the unimodal test (*threshold reached:* Audio M = 2.58; SD = 1.23; Visual M = 3.60; SD = 1.46; *threshold not reached*: Audio M = 2.0; SD = 1.22; ;Visual M = 1.89; SD = 1.27; see Figure 6). There were no other effects. Given the original question of which sensory modality was being relied on to learn the category, planned paired samples t-tests were conducted to compare audio test trials and visual test trials, at each age group. Both 5-year-olds [$t(19)= -3.74$, $p = .001$] and 10-year-olds [$t(19) = -2.99$, $p = .008$] showed significantly better performance on the visual test trials than on the audio test trials. In addition, while both age groups showed better than chance performance on the visual trials, neither age groups showed difference from chance on the audio trials: 5-year-olds *visual* [$t(19)= 2.9$, $p = .009$]; 5-year-olds *audio* [$t(19)= -.827$, $p = .42$]; 10-year-olds *visual* [$t(19) = 3.52$, $p = .002$]; 10-year-olds *audio* [$t(19) = -.21$, $p = .83$]. The 7-year-olds showed no difference from chance on either audio or visual test trials. A correlation looking at the two types of test trials (auditory and visual) for the 7-year-olds showed a significant relation between the two ($r=.52$, $p=.02$), which suggests that overall children were not divided between modalities (i.e., half attending to the auditory cues and half attending to the visual cues)

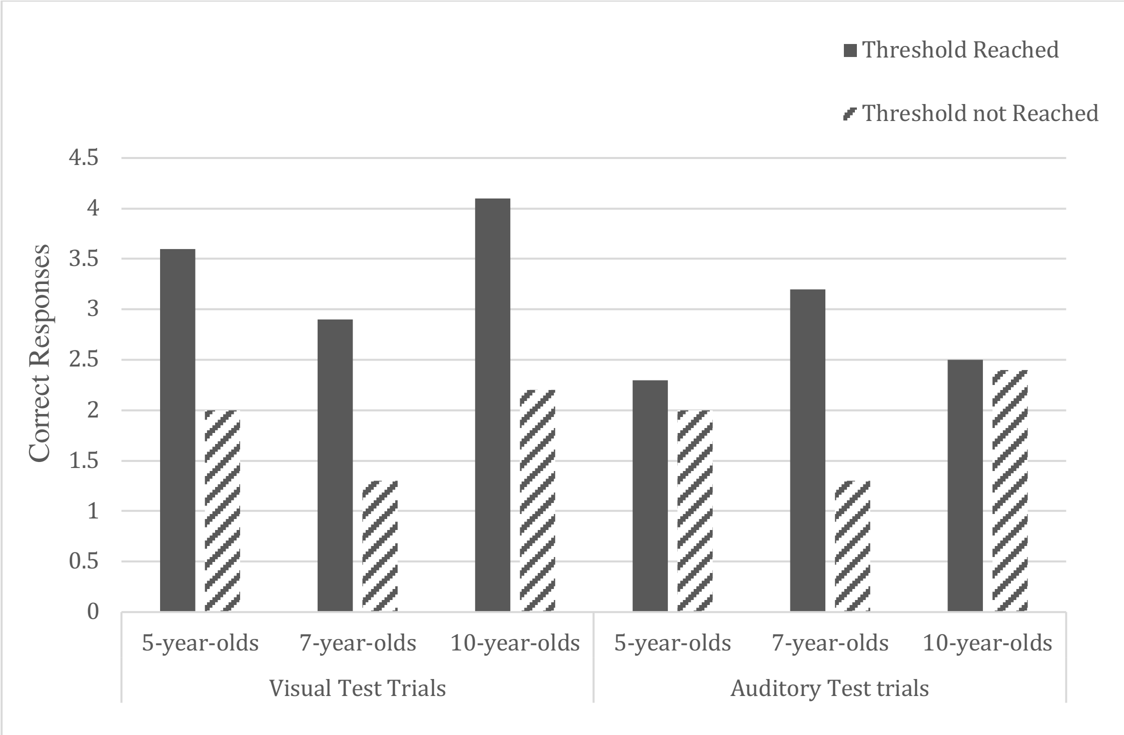


Figure 6. Mean correct responses, with standard errors, to the unimodal test trials, by whether threshold was reached in Experiment 2.

Discussion

Experiment 2 was designed to identify whether children in the *audiovisual informative* condition of the MCLT are learning the visual cues, the auditory cues, both cues, or neither cue. In addition, Experiment 2 was run to provide converging evidence for the threshold of six correct consecutive trials indexing learning on the MCLT. Results from the experiment showed that children who reached threshold (six consecutive correct category choices) performed significantly better than those who did not, with more correct responses during training and faster responses to category choice. Looking at the test trials, children who reached threshold during training performed better on the test trials. This supports the idea that reaching threshold indexes learning the categories. Results from the test trials suggest that it is the visual cues that are being learned, in comparison to the auditory trials. Both 5-year-olds and 10-year-olds showed evidence of learning only on the visual test trials, with 7-year-olds performing at chance on both visual and auditory test trials (or possibly they learned an audio-visual compound, which they were not able to demonstrate in the unimodal test phase).

As discussed in Experiment 1, it is possible that some children reached a threshold value of 6 consecutive correct responses in a row by chance. However, including such children in the "reached threshold" group would simply strengthen our conclusions since their presence in this group would, if anything, have diluted the positive significant effects reported in Experiment 2.

General Discussion

The two current studies employed a multisensory category learning paradigm designed to investigate whether children benefit from arbitrary redundant information from two modalities, in comparison to unisensory information, in a task in which children were explicitly instructed to learn. Importantly, this paradigm was created to look more closely at how redundant information may be presented in a formal learning environment. A second goal of this work was to determine to which sensory modality attention is deployed during the *audiovisual informative* condition. Are children learning only a unimodal cue during this condition, or are they learning both equally well? Based on the cited literature, which shows that multisensory information supports attention and learning, and the educational research suggesting that learning with multiple cues are superior to unisensory information, these experiments set out to systematically look at changes across primary school age in the effects of multisensory cues on an explicit learning task.

In Experiment 1, only the youngest age group showed any benefit from the informative redundant audiovisual cues, requiring fewer trials to reach learning threshold in the multisensory informative condition than in either of the two unisensory conditions. By 10 years of age, children were showing their best performance in the *auditory informative* unisensory condition. This suggests a developmental change in attention to multisensory information, with 10-year-olds relying more on auditory cues. This developmental trend has also been seen in infant work, with multiple cues supporting better attention to stimuli at younger age groups (e.g., Kirkham et al., 2012). In addition, this finding could reflect a longer tenure in formal education for the 10-year-olds, and thus more exposure to auditory information in an explicit instruction context.

Another interesting finding is that it appears that the children go from benefiting from auditory and visual information (5-year-olds) to benefiting from the presence of a single informative auditory cue (10-year-olds). Though not statistically significant, the pattern of 7-year-olds' data is consistent with these children being in a transition between the performance of younger and older children. In Experiment 2, both the 5-year-olds and the 10-year-olds focus on a single dimension in the audiovisual condition, but not the 7-year-olds. Again, this is consistent with the 7-year-olds being in a phase of strategy change, in which they either alternate between different strategies (Siegler, 1999) or adopt a locally more adaptive but globally less effective strategy (Karmiloff-Smith, 1978).

In Experiment 2, in the *audiovisual informative* task, children who reached the threshold of six correct responses in a row showed better performance on the unisensory test trials than those who did not reach threshold. This supports the hypothesis that the children who reached threshold were indeed learning the category distinction. Interestingly, results showed that in this experiment, visual information was relied upon more than auditory information when both visual and auditory information were equally predictive of category membership (although this was only found consistently in the youngest and oldest age groups).

Interestingly, it might seem that because of the difference in length of presentation time of the two modalities (i.e., the auditory cue stopped after 750 msec, whereas the visual cue remained on the screen until the child made a decision), children would unanimously rely on the visual cue. But this was clearly not the case. Although it is beyond the scope of our data to establish the reasons for this, it is possible that, as children were told to react quickly, the initial decision was made quite quickly. This something that will be explored in further work.

In Experiment 1 there were no differences between age groups or conditions on reaction times; whereas, in Experiment 2 younger children were consistently slower than older children, and those who reached threshold performed faster category choices than those who did not. The differences between experiments could be due to the increase in number of trials in Experiment 2. It is possible that the younger children slowed down in the later trials, which could have led to an overall increase in reaction time observed in the second experiment. In Experiment 2, 5-year-olds showed learning on only the visual test trials. This matches well with results from Experiment 1, in which 5-year-olds who reached threshold showed no difference in number of trials to threshold between the *audiovisual* condition and the *visual* condition.

At first, there may appears to be a discrepancy between the two experiments with 10-year-olds in Experiment 1 performing their best during the auditory-only informative condition; whereas, on test in Experiment 2, they appear to show selective learning for only the visual information. However, it is important to understand that the better performance using audio information in Experiment 1 was apparent only in the audio-only informative conditions. In the multisensory informative condition, the 10-year-olds performed like the other age groups. The finding that in the audiovisual condition they selectively focus on visual information, like the other age groups, is consistent across experiments.

Broadbent and colleagues have shown that multisensory information does support learning in these age groups, when the learning is incidental to the target task, even when testing has been delayed for 24 hours (Broadbent al., 2017; Broadbent et al., 2019). Those finding are very different to the current results, and suggest that there may be something about explicit instruction (as opposed to incidental learning) that is shifting performance. In the current experiments, children are told right at the beginning of the task the goal is to figure out the categories (although their attention is never drawn to the importance of the sensory features). There are many differences between incidental and intentional learning, in performance outcomes, learning retention, and neural underpinnings (e.g., Eagle & Leiter, 1964; Gabay, Dick, Zevin, & Holt, 2015; Rüsseler, Hennighausen, Münte, & Rösler, 2003), but they are both considered important teaching goals in current primary school classrooms.

So, why the apparent reduced sensitivity to multisensory information in the current studies? In adults, category distinctions that are easily verbalizable through a simple rule often result in explicit category learning, whereas category distinctions that are difficult to verbalize in a simple rule often lead to gradual implicit category learning (Ashby, Alfonse-Reese, Turkon & Waldron, 1998; Mareschal et al., 2010). Perhaps framing the current task as an explicitly instructed task led the children to form explicit hypotheses, testing simple unisensory categorization rules; thereby unintentionally reducing their sensitivity to the redundant (and supportive) multisensory information. In an incidental learning context, however, no such hypotheses would be generated because the children are focusing on the primary task, leading to greater sensitivity to the redundant multisensory information. This reasoning is supported by two lines of evidence. The first is the finding that children explore the possible hypothesis space less when attempting to solve a problem when explicitly instructed to do so by an adult, than when they are free to explore the problem by themselves (Bonawitz et al., 2011). The second is children's developing use of selective attention strategies (Miller, 1990). Children over the primary school years are gradually better at selectively tuning their attention to the relevant dimensions of a task when faced with stimuli made up of separable cues. Younger children focus their attention less well, and are therefore worse at learning about the primary central cues, but consequently better at revealing learning about the secondary cues. Indeed, 4-year-olds are better than adults at encoding and remembering incidental information in a category learning task precisely because they fail to focus their attention on to the relevant dimensions (Plebanek & Sloutsky, 2017). Multisensory stimuli with arbitrary audiovisual pairings, such as the ones we use, fall into this category of stimuli with separable dimensions.

There are certainly many reasons to believe that redundant information across different sensory modalities would facilitate learning across development (e.g., the more cues there are to support a representation, the more robust that representation might be; e.g., Kirkham et al., 2012). However, there are also many reasons to suspect that multiple cues might interfere with learning (e.g., by requiring heightened selective attention, strong inhibitory control and the capacity to switch flexibly across modalities, none of which are strongly represented in early childhood; cf. Peng, Kirkham & Mareschal, 2018; Thomas, Nardini, & Mareschal 2017). Finally, the impact of receiving information across different sensory modalities on learning may be dependent in whether learning takes place in an explicitly instructed context or an incidental learning context.

References

Addyman, C (2009) *Rules and Similarity in the Development of Category Learning in Children and Infants*. (Unpublished doctoral dissertation). Birbeck, University of London, London, UK.

Ashby, F. G., Alfonso-Reese, L. A., Turken A. U. & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442-481.

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. (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.

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, S. G., Fifer, J., Shivdasani, M. N., Innes-Brown, H., Toohey, S., ... & Paolini, A. G. (2011). The relationship between multisensory integration and IQ in children. *Developmental Psychology*, *47*(3), 877.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, 5(9/10), 341-345

Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L (2011). The double-edged sword of pedagogy: Instruction limits spontaneous exploration and discovery. *Cognition, 120*, 322–330.

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

Broadbent, H., Osborne, T., Mareschal, D., & Kirkham, N. Z. (submitted). *Multisensory cues improve the delayed retention of incidental learning in children.*

Bullock, J., Pierce, S., & McClelland, L. (1989). *TouchMath*. Colorado Springs, CO: Innovative Learning Concepts, Inc.

Carbo, M., Dunn, R., & Dunn, K. (1986). *Teaching students to read through their individual learning styles*, Englewood Cliffs, NJ: Prentice-Hall.

Doering, A., Beach, R., & O'Brien, D. (2007). Infusing multimodal tools and digital literacies into an English education program. *English Education*, *40*(1), 41-60.

Eagle, M., & Leiter, E. (1964). Recall and recognition in intentional and incidental learning. *Journal of Experimental Psychology*, *68*(1), 58.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429.

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, 25*(7), 1362-1370. doi:10.1177/0956797614533801

Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology*, *43*(1), 238.

Gabay, Y., Dick, F. K., Zevin, J. D., & Holt, L. L. (2015). Incidental auditory category learning. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4), 1124.

Gingras, G., Rowland, B. A., & Stein, B. E. (2009). The differing impact of multisensory and unisensory integration on behavior. *Journal of Neuroscience*, *29*(15), 4897-4902.

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

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

Jewitt, C. (2005). Multimodality,"reading", and "writing" for the 21st century. *Discourse: studies in the cultural politics of education*, *26*(3), 315-331.

Jewitt, C., Kress, G., Ogborn, J., & Tsatsarelis, C. (2001). Exploring learning through visual, actional and linguistic communication: The multimodal environment of a science classroom. *Educational Review*, *53*(1), 5-18.

Jordan, K. E., & Baker, J. (2011). Multisensory information boosts numerical matching abilities in young children. *Developmental Science, 14*(2), 205-213. doi:10.1111/j.1467-7687.2010.00966.x

Kalenine, S., Pinet, L, & Gentaz, E. (2011) The visual and visuo-haptic exploration of geometric shapes increases their recognition in preschoolers. *International Journal of Behavioral Development, 35*, 18-26.

Kirkham, N. Z., Richardson, D. C., Wu, R., & Johnson, S. P. (2012). The importance of "what": Infants use featural information to index events. *Journal of Experimental Child Psychology*, *113*(3), 430-439.

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

Klatte, M., Bergström, K., & Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Frontiers in Psychology*, *4*, 578.

Kress, G. (Ed.). (2001). *Multimodal teaching and learning: The rhetorics of the science classroom*. A&C Black.

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

Lercher, P., Evans, G. W., & Meis, M. (2003). Ambient noise and cognitive processes among primary schoolchildren. *Environment and Behavior*, *35*(6), 725-735.

Levy, F. (1980). The development of sustained attention (vigilance) and inhibition in children: some normative data. *Journal of Child Psychology and Psychiatry, 21*(1), 77-84. doi:10.1111/j.1469-7610.1980.tb00018.x

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.

Luchow, J. P., & Shepherd, M. J. (1981). Effects of multisensory training in perceptual learning. *Learning Disability Quarterly*, *4*(1), 38-43.

Mareschal, D., Quinn, P. C., & Lea, S. G. W. (2010). The making of human concepts. Oxford, UK: Oxford University Press.

Matusz, P. J., Broadbent, H., Ferrari, J., Forrest, B., Merkley, 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

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*(4), 484.

Miller, P. H. (1990) The development of strategies of selective attention. In D. F. Bjorklund (Eds.) Children’s strategies Contemporary views on cognitive development (pp. 157-84). Hillsdale, NJ: Lawrence-Eralbaum.

Moats, L. C., & Farrell, M. L. (1999). Multisensory instruction. In J. R. Birsh (Ed.), *Multisensory teaching of basic language skills* (pp. 1-18). Baltimore, MD: Paul H. Brookes.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, *19*(3), 309-326.

Morrongiello, B. A., Fenwick, K. D., & Nutley, T. (1998). Developmental changes in associations between auditory-visual events. *Infant Behavior and Development*, *21*(4), 613-626.

Mount, H., & Cavet, J. (1995). Multi-sensory environments: an exploration of their potential for young people with profound and multiple learning difficulties. *British Journal of Special Education*, *22*(2), 52-55.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press

Nardini, M., Bales, J., & Mareschal, D. (2015). Integration of audio-visual information for spatial decisions in children and adults. *Developmental Science, 19*(5), 803-816. doi:10.1111/desc.12327

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

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

Nava, E., & Pavani, F. (2013). Changes in sensory dominance during childhood: converging evidence from the colavita effect and the sound-induced flash illusion. *Child Development, 84*(2), 604-616. doi:10.1111/j.1467-8624.2012.01856.x

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford University Press.

Peng, A., Kirkham, N. Z. & Mareschal, D. (2018) Information processes of task-switching and modality-shifting across development. *PLoS One, 13(6)*, e0198870.

Peterson, G.E., & Barney, H.L. (1952). Control methods used in a study of the vowels. Journal of the Acoustical Society of America, 24, 175–184.

Plebanek, D. J. & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science, 28(6),* 723-732.

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

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

Rowsell, J., & Decoste, E. (2012). (Re)designing writing in English class: a multimodal approach to teaching writing. *Pedagogies: An International Journal, 7*(3), 246-260.

Rüsseler, J., Hennighausen, E., Münte, T. F., & Rösler, F. (2003). Differences in incidental and intentional learning of sensorimotor sequences as revealed by event-related brain potentials. *Cognitive Brain Research*, *15*(2), 116-126.

Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current Biology, 16*(14), 1422-1427. doi:10.1016/j.cub.2006.05.048

Siegler, R. S. (1999). Strategic development. *Trends in Cognitive Science, 3*(11), 430-435.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Science, 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 Development, 74*(3), 822-833. doi:10.1111/1467-8624.00570

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America*, *26*(2), 212-215.

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.

Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: probabilistic variation and noise guide infants' eye movements. *Developmental Science*, *16*(5), 760-771.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34*(6), 1289-1309.

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

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology*, *47*(5), 1220.