

The interplay between multisensory associative learning and IQ in children

Ayla Barutchu¹, Joanne M. Fifer², Mohit N. Shivdasani^{3,4}, Sheila G. Crewther² &
Antonio G. Paolini^{2,5}

¹Balliol College, University of Oxford, OX1 3BJ, UK

²School of Psychology and Public Health, La Trobe University, VIC 3086, Australia

³Graduate School of Biomedical Engineering, University of New South Wales,
Kensington, NSW 2033, Australia

⁴Bionics Institute, East Melbourne, VIC 3002, Australia

⁵ISN Psychology, Institute for Social Neuroscience, VIC 3084, Australia

Corresponding author:

Ayla Barutchu, PhD

Balliol College

University of Oxford

Broad Street, OX1 3BJ

Abstract

This study assessed the developmental profile of unisensory and multisensory processes and their contribution to children's intellectual abilities (8 and 11-year-olds, $N=38$, compared to adults, $N=19$) using a simple audiovisual detection task and three incidental associative learning tasks with different sensory signals: visual-verbal with pseudo-words, novel audiovisual, and visual-visual. The level of immaturity throughout childhood was dependent on both, the sensory signal type and the task. Associative learning was significantly enhanced with verbal sounds, compared to novel audio-visual and unisensory visual learning. Visual-verbal learning also remains the best predictor of children's general intellectual abilities. The results also demonstrate a separate developmental trajectory for visual and verbal multisensory processes and independent contributions to the development of cognitive ability throughout childhood.

Word count 120 (limit 120 words)

Children live and learn in a multisensory world. Within a typical school environment there is an abundance of facilitatory multisensory information (e.g., auditory and visual speech information, and written and spoken instruction), in addition to distracting stimuli (e.g., background noise and movement). Indeed multisensory processes have been shown to enhance the speed and accuracy of responses (Barutchu, Crewther, & Crewther, 2009; Barutchu et al., 2010; Jordan & Baker, 2011; Mahoney, Verghese, Dumas, Wang, & Holtzer, 2012; Miller, 1982), as well as improve learning and memory throughout the lifespan (Botta et al., 2011; Broadbent, White, Mareschal, & Kirkham, 2017; Brunetti, Indraccolo, Mastroberardino, Spence, & Santangelo, 2017; Fifer, Barutchu, Shivdasani, & Crewther, 2013; Heikkila & Tiippana, 2016; Lehmann & Murray, 2005; Matusz, Wallace, & Murray, 2017; Seitz, Kim, & Shams, 2006). Multisensory training has also been shown to enhance unisensory learning (e.g., Alais & Cass, 2010; Seitz et al., 2006). Thus, proficient processing of multisensory information within a typical school learning environment is likely to be related to the development of other important cognitive and general intellectual abilities. However, despite a multitude of studies demonstrating the efficacy of multisensory training on children's learning in the classroom (e.g., Guyer & Sabatino, 1989; Oakland, Black, Stanford, Nussbaum, & Balise, 1998; Tormanen, Takala, & Sajaniemi, 2008), few studies have systematically investigated the relationships between multisensory processes, associative learning and intellectual abilities in children.

Multisensory processes underlie many stages of information processing (Calvert, 2001; Driver & Noesselt, 2008). However, the nature of the neural mechanisms involved and their developmental trajectories are largely unknown, though it is apparent that the neural mechanisms depend on the sensory signal type

(learnt or novel) (Chen & Spence, 2013, 2017; Laine, Kwon, & Hamalainen, 2007; Molholm, Ritter, Javitt, & Foxe, 2004; Raij, Uutela, & Hari, 2000). Further, these mechanisms also depend on whether sensory signals are merged into a unified percept (i.e., multisensory integration), or whether specific sensory features are linked over time (e.g., transferring or matching information across the sensory systems, associative learning, etc.) (as reviewed in Calvert, 2001; Stein et al., 2010). These multisensory processes also appear to have different developmental trajectories. Many multisensory processes are apparent at birth and begin to modulate attention very soon thereafter, and continue to develop throughout the first year of life (e.g., Bahrack & Lickliter, 2000, 2004; Bremner, Lewkowicz, & Spence, 2012; Bremner, Mareschal, Lloyd-Fox, & Spence, 2008; Lewkowicz, 2000; Sai, 2005; Slater, Quinn, Brown, & Hayes, 1999). For example, both premature and newborn infants can orient their eyes and visual attention to the location of sounds at birth (e.g., Foreman & Fielder, 1989; Muir & Field, 1979). Infants can also associate and learn information across the senses within the first 6 months after birth (e.g., Bahrack & Lickliter, 2000). However, it is not until 9 months of age that optimal multisensory enhancements that cannot be predicted by unisensory processes start to emerge (Neil, Chee-Ruiter, Scheier, Lewkowicz, & Shimojo, 2006). These processes continue to mature throughout late childhood and adolescence (e.g., Barutchu et al., 2009; Brandwein et al., 2011; Dionne-Dostie, Paquette, Lassonde, & Gallagher, 2015; Downing, Barutchu, & Crewther, 2014; Gori, Del Viva, Sandini, & Burr, 2008; Nardini, Jones, Bedford, & Braddick, 2008).

In children, enhanced multisensory processing capabilities have long been associated with better general intellectual abilities (Barutchu et al., 2011; Birch & Belmont, 1965; Rose, Feldman, Futterweit, & Jankowski, 1998; Rose, Feldman, & Wallace, 1992); however, the nature of this relationship requires further investigation.

Firstly, it is necessary to confirm that the relationships with cognition are specific to multisensory processing and not to more generalised unisensory abilities (e.g., Birch & Belmont, 1965). Secondly, findings regarding the relationship between multisensory processing and various cognitive abilities have not always been equivocal. Using a simple detection task, Barutchu et al. (2009) failed to find significant relationships between the auditory-visual reaction time facilitation and non-verbal intellectual skills in late childhood. Yet, later sub-grouping of children based on their ability to gain from multisensory stimulation in different auditory noise backgrounds demonstrated that high verbal abilities were related to how well children integrate audiovisual information in auditory background noise (Barutchu et al. 2011). In addition, children who showed multisensory facilitation consistently in different levels of auditory background noise had significantly higher general intellectual abilities (Barutchu et al., 2011). These results suggest that the relationships between multisensory processing and various cognitive processes is complex, and varies depending on the intellectual skill assessed, the multisensory task employed, and testing conditions, such as background noise.

It is also important to note that aspects of multisensory processing related to learning and cognition in both children (Giannopulu, Cusin, Escolano, & Dellatolas, 2008) and adults (Laurienti, Kraft, Maldjian, Burdette, & Wallace, 2004) have also been linked to other language-related cognitive abilities, such as semantic processing. Previous tasks investigating visual-verbal associative learning have used a traditional paired-associate learning paradigm, in which two items are paired using a stimulus-response format and participants have to verbalise responses (e.g., Hulme, Goetz, Gooch, Adams, & Snowling, 2007; Messbauer & de Jong, 2003; Windfuhr & Snowling, 2001). However, there are limitations inherent in a protocol of this type

when the aim is to assess arbitrary associations using novel stimuli that are not readily verbalised. We have extended on these findings in adults and demonstrated a phonetic verbal-visual associative learning advantage over the learning of visual stimuli that were paired with sounds that could not be pronounced (Fifer et al., 2013). Indeed visual-verbal learning (even without semantics) represents an important variant of multisensory learning; verbal labelling of objects and events is evolutionarily an important function, frequently required, particularly in children, as new concepts and words are learnt (Gogate, Walker-Andrews, & Bahrick, 2001; Sloutsky & Napolitano, 2003). Therefore, it is important to establish whether the observed phonetic verbal-visual associative learning advantage in adults is present in school-aged children. Given that unisensory associative learning has previously been related to intellectual and cognitive functions (Siegler, 1988, 1991), it is also important to clarify whether this particular multisensory capability of linking incidental verbal information to novel symbols, is related to children's intellectual function and learning during primary school years, and whether it is to the same, or to a greater degree than other, within-modality learning capabilities.

The aim of the current study was to investigate multisensory processes and incidental unisensory and multisensory learning in primary school-aged children. To investigate how task specific children's multisensory processing is, we assessed the relationship between multisensory processes across two different tasks; a typical simple speeded detection task and a relatively complex incidental trial-and-error associative learning task. It was also of interest as to whether visual-verbal learning affords faster and more accurate learning performances in children, as it has been previously shown in adults (Fifer et al., 2013). Thus, the learning tasks required participants to learn the associations between unfamiliar symbols and brief unfamiliar

sounds (the verbal-AV task used pseudowords and the novel-AV task used novel non-speech sounds that could not be vocalised by either adults or children). The non-verbal sounds were designed to have the characteristics and discriminability akin to stimuli in the verbal condition (but without the material that could be verbalised). In addition, the relationships between multisensory processes and IQ were also investigated using the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV; Wechsler, 2003). Based on previous findings, it was predicted that children would show less multisensory facilitation than adults, as assessed using a simple speeded detection task, and that multisensory facilitation on this task would not be significantly related to children's intellectual abilities (Barutcu et al. 2009). In addition, it was predicted that incidental learning would be slower in children than adults, and that verbal-visual learning would result in faster and more accurate learning (Fifer et al., 2013). It was further hypothesised that children's incidental learning abilities would be related to their intellectual abilities (Siegler, 1988, 1991), and that this relationship would be stronger with the verbal-AV than the unisensory visual-VV and multisensory novel-AV incidental learning task.

Method

Participants

In total, 20 adults and 41 children were recruited and subdivided into three groups based on age: 8-year-olds (age range: 8:2 – 9:11), 11-year-olds (age range: 10:0 – 11:11) and adults (age range 18:2 – 34:0). Note that age is denoted in years:months. Children were recruited from local Catholic Schools, Melbourne, Australia. Further details of the groups are presented in Table 1. Pilot studies showed that a large proportion of children below the age of 8 had difficulty following instructions and completing all incidental learning tasks, therefore only children 8

years of age and above were assessed. All participants were right-handed with no reported neurological or psychological disorders. All participants spoke English as a first language and were assessed to have Full-Scale IQ (FSIQ) and index scores above the 10th percentile on the WISC-IV (WISC-IV; Wechsler, 2003). Three children were excluded from analyses as visual or hearing screening assessments revealed results that were outside normal limits (visual acuity less than 6/7.5 or hearing threshold above 20 dB on any frequency from 250 Hz to 8000 Hz moving up in octave steps assessed using a calibrated audiometer with sound attenuated headphones) on at least one of the measures (see screening and psychometric measures subsection for further details of screening tests used). In addition, one adult was excluded due to a history of probable learning difficulties leaving 38 children and 19 adults who participated in further testing. Informed consent was obtained from adults and from the parents of children who participated. All procedures were carried out according to prescribed ethical standards and protocols approved by the Human Research and Ethics Committees at La Trobe University and The Royal Victorian Eye and Ear Hospital. Approval to conduct this research was also obtained from the Catholic Education Office, Melbourne.

Screening and psychometric measures

Screening measures were carried out on all participants. Auditory detection thresholds were tested using a calibrated audiometer with sound attenuated headphones to ensure all participants could detect 20 dB sound pressure level at each octave step between the frequencies of 250 Hz to 8000 Hz. Visual screening measures included the assessment of near and distance vision (Patti Pics MassVat ® Logarithmic Visual Acuity Charts), binocular vision (Randot ® Stereotest) and colour vision

(Pseudoisochromatic Plates Ishihara Compatible – IPIC ®). Motor coordination and dexterity were assessed on the Purdue Pegboard (Tiffin, 1999).

For children, a prorated version of the WISC-IV (Wechsler, 2003) was used that consists of 10 subtests in total and provides a measurement of the FSIQ and four subscales: Verbal Comprehension Index (VCI, i.e., verbal IQ), Perceptual Reasoning Index (PRI, i.e., non-verbal IQ), Working Memory Index (WMI), and Processing Speed Index (PSI) was administered. The prorated version of the WISC-IV was employed due to testing time constraints so that the VCI and PRI could be derived from two subtests rather than the three subtests in a complete assessment.

Audiovisual detection task

Multisensory processing was assessed using a simple speeded auditory, visual and audiovisual detection task. The auditory stimulus was a 1500 Hz sinusoidal wave with a 5 ms rise and fall time, delivered via closed, sound attenuating headphones. The visual stimulus was a centrally positioned white disk, a diameter of three degrees of visual angle on a grey background, displayed on a CRT monitor positioned 1 meter in front of the participant. A button box was placed in front of participants and they were required to press a single button with their right index finger to all target stimuli. Target stimuli were presented briefly (102 ms), in one of three conditions: vision alone, audition alone, or vision and audition simultaneously. The onsets of the audiovisual stimuli were calibrated within 1 ms using an oscilloscope. A fourth ‘invalid’ condition, in which no stimuli were presented, was employed as a control to ensure that children were paying attention and not randomly pressing the button. All four stimuli types were presented in a random order with equal probability within each

testing block. The inter-stimulus interval between the stimuli was randomly varied between 1500 and 2500 ms (see Figure 1A).

Participants were required to press a button with their right index finger when a stimulus was detected. The task was explained to the participants in detail and practice trials were administered prior to testing. Testing consisted of two consecutive blocks of 80 stimuli (20 of each type), with a short break in between. Administration time for the detection task was about ten minutes. Accuracy and reaction times (RTs) were recorded for each trial.

Associative learning tasks

The associative learning tasks used to assess visual-verbal (verbal-AV) learning, novel-sound auditory-visual (novel-AV) learning, and visual-visual (shape-VV) learning have previously been published in Fifer et al. (2013).

Visual stimuli were presented on a 22-inch cathode-ray-tube (CRT) monitor, which was positioned at a distance of 1 meter from participants, at central fixation. The visual stimuli were approximately 53 mm in size in order to cover 3 degrees of the central visual field. Auditory stimuli (used in both the novel-AV and verbal-AV tasks) were presented through closed headphones, and verbal and non-verbal sounds were calibrated to a peak volume of 68 dB SPL. Auditory and visual stimuli were presented simultaneously. In all tasks, participants were required to learn an association between novel black symbols and either a novel auditory sound (novel-AV), a verbal sound (verbal-AV) or a red visual shape (shape-VV). For the novel-AV task, non-verbal auditory sounds that could not be physically pronounced were digitally created (sampling rate = 48.8 kHz, duration ~ 620 ms, 5 ms rise-fall time, peak intensity of 68 dB SPL), consisting of combinations of four amplitude-modulated tones of different

carrier and modulation frequencies (only frequencies within the speech range were used). In the verbal-AV task, four single syllable pseudowords with a consonant-vowel-consonant [CVC] structure were used (“jat”, “doot”, “chel” and “shoap”). The uni-modal shape-VV task employed a second visual stimulus, which was an irregular solid red shape, constructed by overlaying rectangular, triangular and circular shapes in different combinations. In this visual only task, the black symbol was presented superimposed over, and contained within, the red shape such that both were presented at the participants’ central fixation point. See Figure 1B and Fifer et al. (2013) to view all visual and auditory stimuli used, and the experimental design.

Instructions to children were built around a story to increase interest. Participants were told that they would be playing a game where they needed to learn about an alien language and writing system that had recently been discovered. They were to be shown alien symbols – which were like letters or numbers - and they would need to learn the alien sounds and names which matched the letters (the novel-AV and verbal-AV conditions respectively), and in the shape-VV condition, were instructed to learn which alien numbers were assigned to each red letterbox. The task procedure was explained and the practice task was administered.

For each condition, participants were required to learn the associations between the four stimulus pairs out of a possible 16 stimulus combinations via trial-and-error (note that the non-matching stimuli comprised the twelve other possible pairings of the stimuli). For each learning task, there were four blocks of 32 trials (128 total trials). A short break up to 5 minutes was provided between testing blocks to prevent fatigue and maximise attention across the experiment. Matching pairs were presented on 50% of the trials, and the remaining trials consisted of non-matching stimuli. The presentation of stimulus pairs within each block was pseudorandom, such that each block consisted

of 16 presentations of matching stimuli with each stimulus pair presented four times. Stimulus pairs were presented with simultaneous onset and offset times. Participants were instructed to make a button response on a special response box with their right index finger when a matching stimulus pair was detected. A “no response” was deemed to be an indication that the participant did not consider the stimuli to be a matched pair. Participants were allowed 3000 ms to respond, after which feedback was provided on every trial using an ascending tone pip (for a correct response) or descending tone burst (for an incorrect response), presented for a duration of 200 ms. The feedback was followed by an inter-stimulus interval, which randomly varied between 1000 to 2000 ms.

As an instructional phase was not included in the protocol (i.e., participants were not told what the matching pairs were), a practice task was employed prior to each experimental learning task. The practice task was analogous to the experimental task; however, there were only three matching stimulus pairs (as opposed to four), and the stimuli were not abstract (e.g., visual stimuli were images of a square, circle, and triangle rather than novel symbols). Participants were encouraged to ask questions about the practice task, and up to 60 trials were administered until the participant demonstrated an understanding of, and sufficient practice on, the task. Also, immediately before the experimental tasks began, participants were presented once with each of the eight individual stimuli that made up the four stimulus pairs. This familiarisation process ensured that participants were aware of the characteristics of each stimulus but did not receive any information regarding whether they were matching or not. Depending on the experimental task at hand, participants were instructed that they would be shown novel symbols, for which they would need to learn the associated names (verbal-AV task), sounds (novel-AV), or shapes (shape-

VV). Each of the four blocks within each task was of approximately 4 minutes duration, and participants were offered a short break of up to 1 minute between each block.

General procedure

Children were assessed in a quiet, unused schoolroom during class time. All testing was completed across three sessions held on separate days, with one learning task included in each session. Adults were assessed in a quiet laboratory room with all tests completed on the same day with a 10-minute rest period between each task. The order of learning task administration was counterbalanced across participants, however, screening tasks were always performed first, with the detection task and the associative learning tasks to follow. For children, all sessions were of 45 minutes to 1 hour duration. The total testing time per participant was approximately 2½ hours.

Data analysis

Differences between the two groups of children in standardised peg-board and IQ measures were assessed using a series of t-tests.

Audiovisual detection task

The RTs for correct responses were recorded and averaged for each participant. All RTs faster than 150 ms or slower than 3SD above the mean were excluded as probable late responses to the previous stimulus or a false alarm. Overall, less than 1% of total RTs were excluded on this basis.

Accuracy. For all participants, the percentage of errors for each stimulus type was calculated. Percentage error rates for all age groups were very low leading to violations of the assumption of normality. Therefore, group differences in accuracy

across the different stimuli types were assessed using the non-parametric Friedman's Test followed by Wilcoxon Signed Rank tests as pairwise post hoc tests. Differences in error rates across the three groups for each stimulus type were assessed using Independent-samples Kruskal-Wallis Tests followed by Mann-Whitney tests for pairwise post hoc analyses (Gravetter & Wallnau, 2017).

Reaction times. A mixed 3(stimulus: auditory, visual, audiovisual) x 3(age groups: 8-year-olds, 11-year-olds, and adults) ANOVA was conducted to ascertain stimulus and age-related effects on RTs.

For RTs, audiovisual facilitation scores were also calculated for each participant. Multisensory facilitation was calculated by subtracting the mean audiovisual RT from the faster of the unisensory auditory or visual RT. This allowed for an estimate of the magnitude of facilitation in ms, irrespective of motor speed differences. We also calculated the percentage difference between the fastest mean RT on unisensory conditions (either auditory or visual) and the mean RT on the audiovisual condition to control for any effects that may be related to a fundamental difference in motor speed between the age groups. The two facilitation measures highly correlated with each other ($r > .95$, $p < .001$), therefore, only the percentage of facilitation was used in further data analyses. Age-related effects on multisensory facilitation were assessed using a one-way independent samples ANOVA, with Tukey tests as follow-up post hoc tests.

We also assessed for violations of the race model inequality. For each stimulus type, cumulative probabilities were calculated from .05 to .95 in increments of .1. We also calculated the bound probabilities of the unisensory conditions by summing the RT probabilities for the two unisensory conditions (refer to Miller, 1982 for a detailed

outline of the race-model test procedure). The race model assumes that if the fastest RTs of the multisensory RTs are faster than when the probabilities for the unisensory condition are bound, then race models cannot predict the level of facilitation under multisensory conditions. Initially, we used a three-way 3(age groups) x 2(stimulus type: audiovisual vs. bound auditory and visual condition) x 10(probabilities) mixed ANOVA to assess for significant differences between the cumulative density functions (CDFs) for the multisensory and the bound unisensory conditions. The significant three-way interaction was followed up with planned contrast assessing if the multisensory CDF was significantly faster than the bound unisensory CDF at each probability value.

Associative learning tasks

For each associative learning task, accuracy and RTs were recorded. Errors on target trials indicated errors of omission and involved the failure to correctly respond to a target pair. Only the first 120 trials completed by all participants were analysed. Due to technical errors, the data for the last 8 trials of one participant was not saved. For RTs, values less than 300 ms and greater than 3 SD above the mean were excluded, and the response recorded as an error. Less than 1% of responses were excluded on the basis of this criterion.

Accuracy. Learning trends across the three tasks were assessed using Signal Detection Theory with d-prime calculated as the measure of discriminability. The d-prime statistic, that minimises response bias to target and non-target stimuli (for a detailed explanation see Macmillan & Creelman, 2004), was calculated by finding the difference between Z-scores for hits and false alarms [i.e., $d\text{-prime} = Z(\text{hit rate}) - Z(\text{false alarm rate})$]. To visualise changes in discriminability with learning across

time, moving d-prime values were calculated for 10-trial moving windows in one step consisting of a total of 111 overlapping windows.

Measures of d-prime were also calculated for overall accuracy measures across all the 120 trials in order to ascertain the effects of age, task and stimulus relationship on overall performance levels, and analysed using a 3(age group: 8-year-old, 11-year-old and adult) x 3(task: novel-AV, verbal-AV and shape-VV) mixed ANOVA. To ascertain changes in discriminability across trials, d-prime values were calculated for consecutive non-overlapping blocks of 10 stimuli (Brown & Heathcote, 2003), and analysed using a 3(age groups: 8-year-old, 11-year-old and adults) x 3(task: novel-AV, verbal-AV and shape-VV) x 4(blocks: 1, 2, 3, 4) repeated measures ANOVA. Note that only the first 5 blocks were included in the analysis. This was because preliminary analyses showed a violation in the normality for block 5 and above, as most adults learnt the stimulus pairs within the first five blocks and showed a ceiling effect for accuracy.

Reaction times. Changes in RTs with learning were also assessed. Only correct responses to matching trials were analysed as responses to non-matching trials represented errors of commission (i.e., false alarms). Based on the moving d-prime curves, the first 40 trials were defined as the ‘learning’ phase and the last 40 trials were defined as the ‘learnt’ phase. Differences in RTs during these phases were assessed using a 3(age group: 8-year-old, 11-year-old, and adult) x 2(phase: learning and learnt) x 3(task: novel-AV, verbal-AV, and shape-VV) mixed ANOVA.

All significant ANOVA main and interaction effects were followed up with post hocs and simple effects analyses using Bonferroni corrections.

Multisensory processing and IQ in Children

For children, we performed exploratory analyses assessing the relationships between age, IQ measures, multisensory facilitation on the detection task, and the overall d-prime measures for each of the three learning tasks using Pearson's correlations. We also performed stepwise regression to assess whether FSIQ could be predicted by age, d-prime, RT measures on the three learning tasks and multisensory facilitation on the detection task as independent variables (Tabachnick & Fidell, 2007). Stepwise regression uses a combination of a forward and a backward selection criteria to derive the optimal model to explain the dependent variables (i.e, FSIQ in this case).

Results

Demographic and Psychometric Measures

Overall, standardised IQ and right-hand motor coordination and dexterity measures were very similar and did not significantly differ between the two groups of children (see Table 1).

Audiovisual detection task

As can be seen in Figure 2A and 2B, responses were significantly faster, $F(2, 108) = 76.01, p < .001, \eta^2 = .59$, with children making significantly fewer errors ($p < .05$), for audiovisual than the unisensory auditory and visual stimuli. Error rates in adults were very low and did not significantly differ across the four stimuli type. In addition, adults were significantly faster, $F(2, 54) = 9.99, p < .001, \eta^2 = .27$, and made significantly fewer errors than children for all stimuli except the multisensory audiovisual signals ($p < .05$) (Figure 2A and 2B). Although 11-year-old children were faster and more accurate than 8-year-olds, these differences did not reach significance (Figure 2A and 2B).

The level of facilitation in milliseconds and the percentage gain in RT from the faster of the unisensory signals (Figure 2C) showed a very similar pattern, suggesting that the slower mean RT of the younger children's group could not account for age-related differences in the observed multisensory facilitation levels (Figure 2C). The percentage of multisensory facilitation for adults was significantly higher than for both groups of children, $F(2, 54) = 11.08, p < .001, \eta^2 = .29$, suggesting that not only the motor responses but associated multisensory facilitation processes are still developing in children.

An analysis of the CDFs also showed an immature multisensory facilitation pattern in children. As can be seen in Figure 2D, the audiovisual CDF is faster than both the unisensory CDFs and the bound unisensory CDF in adults. In contrast, the differences between the bound unisensory and the multisensory CDF in the two child groups were relatively small. The three-way mixed ANOVA showed a significant interaction effect between the three age groups, stimuli and the probability values of the CDFs, $F(18,486) = 9.31, p < .001, \eta^2 = .26$. Follow-up pairwise planned comparisons showed that for the 8-year-old group, the multisensory CDF was not significantly faster than the bound unisensory CDF. For the 11-year-old group, the multisensory CDF was significantly faster than the bound CDF, but only for probability values .35 and .45. In contrast, the multisensory CDF in adults was significantly faster than the bound CDF for all RTs below .85 probability.

Associative Learning Tasks

Both adults and children showed improvements in performance suggesting learning throughout the tasks. The greatest level of improvement in discriminability (i.e., d-prime measures) was observed for the verbal-AV task for all age groups (see

Figure 3), suggesting that verbal information has a large enhancing effect on learning processes even when the verbal content is meaningless. In contrast, although some generalised audiovisual advantage was observed for adults, there was a relatively large overlap in performance between the novel-AV and the shape-VV task for both child groups (Figure 3A). Indeed, overall d-prime measures were significantly higher for the verbal-AV task than both the novel-AV and shape-VV learning task, $F(2,108) = 45.27$, $p < .001$, $\eta^2 = .46$, (Figure 3B). The d-prime measures were also higher for the novel-AV than the unisensory shape-VV condition ($p = .03$). However, this comparison failed to reach significance following a Bonferroni correction ($p = .07$ following the correction). For overall d-prime measures (Figure 3B), the main effect for age groups was also significant, $F(1,36) = 29.65$, $p < .001$, $\eta^2 = .52$, with d-prime measures gradually increasing across all age groups (see Figure 3B).

To assess for differences in learning patterns across the three learning conditions, block analyses were also employed. Adults learned the stimulus pairs relatively quickly, particularly for the verbal-AV condition leading to ceiling effects and large violations of normality, therefore only the first 4 blocks were included in the ANOVA analysis. A similar pattern was observed with block analyses of d-prime measures (see Figure 3C). D-prime measures were significantly higher for the verbal-AV condition than the novel-AV conditions, both of which significantly differed from the shape-VV condition, $F(2,324) = 35.18$, $p < .001$, $\eta^2 = .40$. Learning across blocks was assumed if d-prime measures significantly increased across blocks. Significant increases in d-prime measures were observed by block 3 which significantly differed from block 1, with d-prime measures in block 4 being significantly higher than all the preceding 3 blocks, $F(3,324) = 23.09$, $p < .001$, $\eta^2 = .30$. Overall, d-prime measures did not significantly differ between 11-year-olds and 8-year-olds across the first four

blocks, while adults scored significantly higher d-prime measures than either of the child groups, $F(2,54) = 20.43, p < .001, \eta^2 = .43$. The interactions between age group and block, $F(6,324) = 3.43, p = .003, \eta^2 = .11$, and stimulus type and block were also significant, $F(6,324) = 2.41, p = .03, \eta^2 = .04$. There were no significant differences between the three age groups in block 1. By block 2, d-prime measures were significantly higher in adults than 8-year-old children, and by block 3, adults were significantly outperforming both of the child groups. The verbal-AV advantage was observed from the beginning of the trial, with verbal-AV significantly differing from both novel-AV and shape-VV across all blocks, including block 1. For the shape-VV and the novel-AV conditions, significant increases in d-prime were observed only at block 4, which differed from the first two blocks and the first block, respectively.

The RTs on the associative learning task also significantly differed across age groups and stimuli type (Figure 4). In general, RTs were significantly faster for the verbal-AV condition than for either the novel-AV or shape-VV conditions, $F(2,108) = 12.59, p < .001, \eta^2 = .19$. The RTs were also significantly faster at the learnt phase than the learning phase, $F(1,108) = 20.06, p < .001, \eta^2 = .27$. Consistent with the above reported d-prime measures, adults were significantly faster at responding than either child group, $F(2,54) = 4.59, p = .01, \eta^2 = .15$. The interaction between time and age-group was also significant, $F(2,108) = 4.68, p = .01, \eta^2 = .15$. Unlike the 8-year-old group of children, the RTs of both the 11-year-old and adult groups were significantly faster at the learnt than the learning phase. Furthermore, the differences in RTs between the three age groups during the initial learning phase of the associative tasks did not reach significance. However, adults were significantly faster at responding than either of the child groups during the learnt phase of the task. The interaction between stimulus type and time was also significant, $F(2,108) = 4.24, p = .02, \eta^2 = .07$. The RTs

for the learning and learnt phases of the task significantly differed for both of the multisensory novel-AV and verbal-AV conditions, but not for the unisensory shape-VV condition. In the initial learning phase of the task, only RTs for novel-AV stimuli significantly differed from verbal-AV. However, for the learnt phase of the verbal-AV task, RTs were significantly faster than those for the novel-AV and shape-VV tasks.

Correlation Analyses

Simple detection, associative learning, and age. For children, the relationships between d-prime and RT measures for the three associative learning tasks, RT measures on the detection task, age and IQ measures on the WISC were assessed (see Table 2 and Figure 5). D-prime measures on the shape-VV task showed a strong correlation with age suggesting that unisensory processes related to incidental learning are still developing (see Table 2 and Figure 5) during late childhood. D-prime measures on the novel-AV and verbal-AV tasks showed a moderate and a low correlation with age, respectively. In general, the relationships between measures on the simple speeded detection task (percentage facilitation) and the associative learning tasks (d-prime) were very weak. Significant correlations were only observed for the RTs; RTs for the verbal-AV and novel-AV tasks showed moderate and low correlations with RTs on the simple detection task, respectively (Table 2). In contrast, the correlations between the RTs on the Shape-VV task and the simple detection task approached zero. The facilitative effect of multisensory integration on the detection task also moderately correlated with the RTs on the verbal-AV task. Interestingly, there was a low correlation between d-prime on the verbal-AV task and multisensory facilitation on the simple detection task, which this failed to reach significance (see Figure 5).

Multisensory processing and IQ. D-prime measures on the three associative learning tasks moderately correlated with FSIQ (see Figure 5 for scatter plot) and PRI, but showed a low correlation with VCI and a relatively lower correlation that approached zero with WMI (see Table 2). This suggests that although performance on these associative learning tasks is partly dependent on children's non-verbal intellectual abilities, they are less likely to be dependent on children's working memory and verbal abilities. The Stepwise Regression with FSIQ as the dependent variable, and age, d-prime, RT measures on the three learning tasks, and multisensory facilitation as independent variables (note that the IQ subscales were not included due to violations of co-linearity with FSIQ), $R = .43$ $F(1,37) = 8.27$, $p = .007$. Only the d-prime measure on the verbal-AV task was identified as a significant predictor of FSIQ, accounting for 18% of the unique variance (part $r = 0.42$).

Discussion

This study shows developmental changes in multisensory processing and learning capabilities during late childhood. In general, adults outperformed both child groups, and the 11-year-old children achieved better overall learning accuracy and were faster to respond to learnt stimuli than the 8-year-old children suggesting a strong developmental trend in multisensory processing and incidental learning performance through childhood. Consistent with the findings of Fifer et al. (2013), children and adults performed best on the verbal-AV task, but unlike adults, children did not show a generalised multisensory advantage for novel multisensory stimuli. Although, all three associative learning tasks significantly correlated with IQ, only the d-prime measure on the verbal-visual associative task was identified as a unique predictor of FSIQ scores, suggesting that general intellectual abilities in children may be related to multisensory verbal associative learning abilities.

Multisensory processing: simple detection vs. associative learning

The results of this study provide novel comparisons of the behavioural performance on two different multisensory tasks in children; simple speeded multisensory detection and incidental multisensory associative learning. Consistent with a growing body of literature, multisensory processes were shown to be still immature throughout late childhood (e.g., Barutcu et al., 2009; Brandwein et al., 2011; Downing et al., 2014; Gori et al., 2008; Nardini, Bales, & Mareschal, 2016). On the simple detection task, speed, accuracy and the degree of multisensory facilitation improved with age. Consistent with the patterns observed on the detection task, age related improvements in RTs and d-prime measures, were also observed for all incidental learning tasks. D-prime measures were significantly higher for the multisensory learning tasks, gradually increasing with age, and only 11-year-olds and adults showed improvements in RTs at the learnt phase of the tasks. Further assessment of the relationship between the two types of tasks showed low to moderate correlations between the RTs on the detection task and the novel-AV and the verbal-AV condition of the learning task. In addition, a moderate correlation was observed between RTs on the verbal-AV learning task and the amount of multisensory facilitation in children. This relationship with multisensory facilitation and verbal audiovisual learning suggests that children who can derive more benefit from simple spatiotemporally linked stimuli are faster at learning associations between verbal and visual stimuli. It provides some support for the notion that low-level multisensory processes may facilitate more complex arbitrary word-object associations (Gogate et al., 2001).

The significant relationship between RT facilitation and RTs on the verbal-AV learning task did not generalise to d-prime. Although both the detection task and the

multisensory learning tasks involved the concurrent presentation of auditory and visual stimuli, the cognitive demands of each task varied greatly. Therefore, some variation is expected particularly across measures that are task specific, such as learning and discriminability of stimuli as assessed using d-prime. Thus, the significant relationships between RTs on the detection and RTs on the verbal-AV learning tasks may appear to represent an effect of enhanced motor speed, rather than generalised multisensory performance enhancement. However, it is important to note that the relationships between multisensory facilitation and RTs on the unisensory shape-VV task did not approach significance. This suggests that the significant relationship between RT facilitation on the detection task and the RTs on the verbal-AV learning task may represent a common multisensory factor. The lack of significant relationships with the RTs on the shape-VV task also suggests that a different processing mechanism might be involved in unisensory incidental learning, which impacted the speed of motor responses.

The findings of this study also indicated that the maturation of multisensory processes may not be gradual, and may be linked to the development of sensory systems and other task-related abilities (Barutchu et al., 2009; Hulme, Smart, Moran, & Raine, 1983). In children, the correlations between age, multisensory facilitation and d-prime measures on the verbal-AV task were low, unlike the observed significant moderate and high correlations between age and d-prime measures on the novel-AV and shape-VV learning tasks. Given that learning was faster and more accurate with verbal-phonetic information, these processes may reach maturity earlier, hence verbal measures showing a low correlation with age in the current study. The moderate to high correlations between age and d-prime measures on the non-verbal novel-AV and shape-VV tasks, respectively, suggests that younger children may have found these

tasks relatively harder, and that there is still relatively strong maturation taking place throughout late childhood for visual and novel multisensory processes. In adults, distinct neural processes are involved in novel and verbal multisensory processing (e.g., Calvert, 2001; Raij et al., 2000). The outcomes of this study are consistent with the idea that verbal and non-verbal multisensory processes may be dissociable, and that these may have independent maturational processes (Tremblay et al., 2007).

Multisensory Associative Learning

The verbal-AV learning task was consistently performed with greater accuracy and faster response times than both non-verbal tasks across all age groups. A verbal enhancement of information processing has been documented in infants and adults. Infants are unable to process dual auditory-visual information unless the auditory component is verbal in nature (Robinson & Sloutsky, 2010), and Hocking and Price (2009) found that verbal multisensory discrimination was performed faster than non-verbal multisensory discrimination in adults. In this study, this advantage may be partly related to children being able to rehearse the phonetic sounds, facilitating faster learning in the verbal-AV condition, unlike in the novel-AV learning condition. The results of this study further suggest that the verbal over the non-verbal auditory-visual advantage is not only present in 8-years-old children, but continues to develop throughout late childhood.

The generalised novel multisensory learning advantage (relative to unisensory incidental learning) found in adults by Fifer et al. (2013) was not evident in children. This finding is consistent with other studies that have shown immaturities in different types of multisensory learning, such as visual-haptic learning and incidental learning of auditory-visual object categories, throughout childhood (Broadbent et al., 2017;

Pirogovsky, Murphy, & Gilbert, 2009). The novel-AV learning advantage observed in adults may be explained by differences in multisensory integration and attention abilities. Previous studies have demonstrated that multisensory processes are not optimal in children and early adolescence (e.g., Downing et al., 2014). Using a simple speeded detection task, a similar pattern of results was found in the present study, suggesting that multisensory processes remain immature throughout childhood. Therefore, it is possible that an inferior ability to utilise multisensory information for facilitation at a lower stimulus processing level may have impacted on children's ability to exploit this for the benefit of cognitive operations, such as the learning of novel multisensory associations. Immature attention capacities in children may have also contributed to this result (e.g., Abundis-Gutierrez, Checa, Castellanos, & Rosario Rueda, 2014; Cromer, Schembri, Harel, & Maruff, 2015; Morrison, 1982; Pozuelos, Paz-Alonso, Castillo, Fuentes, & Rueda, 2014), as multisensory processes have been shown to be dependent on an individual's state of attention (e.g., Dean et al., 2017; Hillyard, Stormer, Feng, Martinez, & McDonald, 2016; Talsma, 2015; Talsma, Senkowski, Soto-Faraco, & Woldorff, 2010). Indeed, it is likely that the current task engaged significant attentional resources given the speeded nature of the task and the fact that an incidental trial-and-error learning paradigm was employed. Further research is required to establish the role of fundamental multisensory processes and attention in children's abilities to learn novel audiovisual associations.

Although our developmental findings may appear somewhat inconsistent with those of Tremblay et al. (2007), there are several important distinctions between the multisensory processes at play in the current study and those of Tremblay et al. (2007). Tremblay et al. (2007) investigated speech-based (i.e., McGurk Effect) and non-speech (i.e., flash-beep effect) multisensory illusory processes, that utilised lower-level ability

to integrate incongruent multisensory features of previously learnt associations, and as such represent a vastly different type of multisensory process. They concluded that non-verbal audiovisual integration matured relatively early compared to speech-based integration, which was found to continue developing throughout late childhood. At the level of incidental associative learning of new information investigated here, the familiarity of phonetic information likely assists with stimulus learning and consequently, associative learning, thus showing audiovisual incidental learning with novel unverbilizable sounds to be relatively immature.

Multisensory processing and IQ

Links between IQ and cross-modal temporal matching (Birch & Belmont, 1964), object transfer (Giannopulu et al., 2008; Rose, Feldman, Jankowski, & Futterweit, 1999), and detection (Barutchu et al., 2009) have previously been reported. To our knowledge, the relationships between incidental multisensory learning and IQ have not previously been explored in depth, and thus were addressed in the current study. All three learning tasks significantly correlated with FSIQ, which appeared to be largely driven by moderate correlations between the PRI and learning accuracy across the three tasks. The verbal IQ index consistently showed a low correlation with all three tasks, while the correlations with working memory abilities were much lower approaching zero. This suggests that performance on the associative learning tasks was not likely to be constrained by children's working memory abilities. Barutchu et al. (2011) also reported significantly reduced verbal IQ in children with reduced multisensory abilities, though this was only in children who were poor integrators in the presence of auditory background noise. As all three associative learning tasks here correlated with FSIQ and PRI scores, our results indicate that non-verbal or perceptual reasoning skills were required for successful performance on all the associative

learning tasks, including the verbal learning task. Furthermore, although performance on all the associative learning tasks yielded significant correlations with IQ measures, consistent with Barutcu et al. (2009), multisensory facilitation, as assessed using a simple detection task of lower cognitive demand, was not significantly associated with the IQ measures. Indeed the relationship between multisensory processing and IQ is complex, task-specific, and is not readily generalised to all multisensory abilities or IQ measures.

Despite the fact that the correlations between the VCI subscale and all three associative learning tasks were low, a regression analysis showed that d-prime measures on the verbal-visual associative learning task was the best predictor of FSIQ, uniquely accounting for 18% variability of this measure. This finding suggests that verbal-visual multisensory processes may independently contribute to general intellectual abilities. Indeed, infants are born with a predisposition to associate verbal and visual information, as evidenced in newborn babies who are reported to be able to form associations between faces and voices (Sai, 2005). This fundamental ability to attend to and associate verbal and visual information may also form a foundation for the development of other general intellectual abilities throughout childhood. For example, children who are better able to associate verbal information with visual objects may also be faster at processing other relations and instructions related to perceptual learning tasks, and thus enhancing their learning and general perceptual abilities. Indeed it is just as likely that children who have a predisposition for higher intellectual abilities are better at learning verbal associations. Either way, this finding suggests that the relationship between incidental verbal learning and children's general and perceptual reasoning abilities (i.e., non-verbal IQ) is independent of their general verbal abilities (i.e., VCI scores), which in hindsight, is not surprising. The Wechsler

VCI assessment is largely dependent on semantics, comprehension and verbal-semantic reasoning. In our verbal-AV learning task, in contrast, semantic information was not available; audio-visual associations were purely based on phonetic-visual processing, which appears to be more related to perceptual learning and reasoning as assessed by the PRI subscale of the Wechsler IQ test. Further research is needed to understand the relationship between multisensory verbal-phonetic abilities and non-verbal IQ, and if verbal multisensory abilities can be used to further enhance general intellectual abilities, most importantly, in children with generalised learning difficulties.

It is also important to note that despite the fact that the verbal-visual condition was the main predictor of FSIQ and PRI, none of the relationships between the learning tasks and IQ here were specific to multisensory conditions. Also, some observed low correlations may have failed to reach significance due to the low sample size, thus, possibly leading to a Type II error. This underscores the importance of unisensory comparison tasks (i.e., the visual-visual task) in evaluating the consequences and correlates of multisensory processes. In addition, due to perceptual limitations posed by participants' ability to identify temporally overlapping auditory signals, we were unable to include a unisensory auditory incidental learning task. Nevertheless, in the current study, it was evident that perceptual reasoning skills were related to how successfully children performed the learning task regardless of stimulus modalities employed. The difficult nature of the task, and the low sample size for the exploratory correlation analysis may, however, have obscured more subtle relationships with other skills. In other words, a higher sample size and a less difficult task may tap into more variability in results and uncover more subtle links with cognition and multisensory processing in future studies.

In conclusions, our data demonstrate that multisensory processes continue to develop throughout late childhood, and like adults, children as young as 8-year-olds show a large verbal-visual incidental associative learning advantage. Importantly, verbal multisensory associative skills are strongly linked to children's general and non-verbal intellectual abilities. The outcomes of this study have important implications for improving teaching strategies and children's learning environments to optimise learning outcomes. Further research is needed, however, to identify strategies that can be used to enhance multisensory verbal associative abilities, and in turn, intellectual abilities in children, and to investigate whether the outcomes can be generalised to children with perceptual and learning difficulties.

References

- Abundis-Gutierrez, A., Checa, P., Castellanos, C., & Rosario Rueda, M. (2014). Electrophysiological correlates of attention networks in childhood and early adulthood. *Neuropsychologia*, 57, 78-92.
doi:10.1016/j.neuropsychologia.2014.02.013
- Alais, D., & Cass, J. (2010). Multisensory perceptual learning of temporal order: audiovisual learning transfers to vision but not audition. *PLoS One*, 5, e11283.
doi:10.1371/journal.pone.0011283
- Bahrack, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Dev Psychol*, 36, 190-201.
- Bahrack, L. E., & Lickliter, R. (2004). Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: a developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective, & Behavioral Neuroscience*, 4, 137-147.
- Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: development of multisensory facilitation in children. *Developmental Science*, 12, 464-473. doi:10.1111/j.1467-7687.2008.00782.x
- 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, 877-885.
doi:10.1037/a0021903
- Barutchu, A., Danaher, J., Crewther, S. G., Innes-Brown, H., Shivdasani, M. N., & Paolini, A. G. (2010). Audiovisual integration in noise by children and adults. *J Exp Child Psychology*, 105, 38-50. doi:10.1016/j.jecp.2009.08.005
- Birch, H. G., & Belmont, L. (1964). Auditory-visual integration in normal and retarded readers. *American Journal of Orthopsychiatry*, 34, 852-861.
- Birch, H. G., & Belmont, L. (1965). Auditory-visual integration, intelligence and reading ability in school children. *Perceptual and Motor Skills*, 20, 295-305.
- Botta, F., Santangelo, V., Raffone, A., Sanabria, D., Lupianez, J., & Belardinelli, M. O. (2011). Multisensory integration affects visuo-spatial working memory. *Journal of Experimental Psychology: Human Perception and Performance*, 37, 1099-1109. doi:10.1037/a0023513
- Brandwein, A. B., Foxe, J. J., Russo, N. N., Altschuler, T. S., Gomes, H., & Molholm, S. (2011). The development of audiovisual multisensory integration across childhood and early adolescence: a high-density electrical mapping study. *Cerebral Cortex*, 21, 1042-1055. doi:10.1093/cercor/bhq170
- Bremner, A. J., Lewkowicz, D. J., & Spence, C. (2012). *Multisensory development*. Oxford: Oxford University Press.
- Bremner, A. J., Mareschal, D., Lloyd-Fox, S., & Spence, C. (2008). Spatial localization of touch in the first year of life: early influence of a visual spatial code and the development of remapping across changes in limb position. *Journal of Experimental Psychology: General*, 137, 149-162.
doi:10.1037/0096-3445.137.1.149
- Broadbent, H. J., White, H., Mareschal, D., & Kirkham, N. Z. (2017). Incidental learning in a multisensory environment across childhood. *Developmental Science*. doi:10.1111/desc.12554
- Brown, S., & Heathcote, A. (2003). Averaging learning curves across and within participants. *Behavior Research Methods, Instruments & Computers*, 35, 11-21.
- Brunetti, R., Indraco, A., Mastroberardino, S., Spence, C., & Santangelo, V. (2017). The impact of cross-modal correspondences on working memory performance. *Journal of Experimental Psychology: Human Perception and Performance*, 43, 819-831. doi:10.1037/xhp0000348

- Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral Cortex*, *11*, 1110-1123.
- Chen, Y. C., & Spence, C. (2013). The time-course of the cross-modal semantic modulation of visual picture processing by naturalistic sounds and spoken words. *Multisensory Research*, *26*, 371-386.
- Chen, Y. C., & Spence, C. (2017). Dissociating the time courses of the cross-modal semantic priming effects elicited by naturalistic sounds and spoken words. *Psychonomic Bulletin Review*. doi:10.3758/s13423-017-1324-6
- Cromer, J. A., Schembri, A. J., Harel, B. T., & Maruff, P. (2015). The nature and rate of cognitive maturation from late childhood to adulthood. *Frontiers in Psychology*, *6*, 704. doi:10.3389/fpsyg.2015.00704
- Dean, C. L., Eggleston, B. A., Gibney, K. D., Aligbe, E., Blackwell, M., & Kwakye, L. D. (2017). Auditory and visual distractors disrupt multisensory temporal acuity in the crossmodal temporal order judgment task. *PLoS One*, *12*, e0179564. doi:10.1371/journal.pone.0179564
- Dionne-Dostie, E., Paquette, N., Lassonde, M., & Gallagher, A. (2015). Multisensory integration and child neurodevelopment. *Brain Science*, *5*, 32-57. doi:10.3390/brainsci5010032
- Downing, H. C., Barutchu, A., & Crewther, S. G. (2014). Developmental trends in the facilitation of multisensory objects with distractors. *Frontiers in Psychology*, *5*, 1559. doi:10.3389/fpsyg.2014.01559
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron*, *57*, 11-23.
- Fifer, J. M., Barutchu, A., Shivdasani, M. N., & Crewther, S. G. (2013). Verbal and novel multisensory associative learning in adults. *F1000Res*, *2*, 34. doi:10.12688/f1000research.2-34.v2
- Foreman, N., & Fielder, A. (1989). Intermodal enhancement of stimulus localisation in infants born prematurely. *Perception and Motor Skills*, *69*, 43-50. doi:10.2466/pms.1989.69.1.43
- Giannopulu, I., Cusin, F., Escolano, S., & Dellatolas, G. (2008). Cognitive associations of bimanual haptico-visual recognition in preschoolers. *Child Neuropsychology*, *14*, 227-236. doi:10.1080/09297040701314881

- Gogate, L. J., Walker-Andrews, A. S., & Bahrick, L. E. (2001). Intersensory origins of word comprehension: An ecological-dynamic systems view. *Developmental Science*, 4, 451-467. doi:10.1111/1467-7687.00143
- Gori, M., Del Viva, M., Sandini, G., & Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology*, 6, 694-698.
- Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the Behavioral Sciences* (10th ed.). Boston, MA: Cengage Learning.
- Guyer, B. P., & Sabatino, D. (1989). The effectiveness of a multisensory alphabetic phonetic approach with college students who are learning disabled. *Journal of Learning Disabilities*, 22, 430-434.
- Heikkila, J., & Tiippana, K. (2016). School-aged children can benefit from audiovisual semantic congruency during memory encoding. *Experimental Brain Research*, 234, 1199-1207. doi:10.1007/s00221-015-4341-6
- Hillyard, S. A., Stormer, V. S., Feng, W., Martinez, A., & McDonald, J. J. (2016). Cross-modal orienting of visual attention. *Neuropsychologia*, 83, 170-178. doi:10.1016/j.neuropsychologia.2015.06.003
- Hocking, J., & Price, C. J. (2009). Dissociating verbal and nonverbal audiovisual object processing. *Brain and Language*, 108, 89-96. doi:10.1016/j.bandl.2008.10.005
- Hulme, C., Goetz, K., Gooch, D., Adams, J., & Snowling, M. J. (2007). Paired-associate learning, phoneme awareness, and learning to read. *Journal of Experimental Child Psychology*, 96, 150-166. doi:10.1016/j.jecp.2006.09.002
- Hulme, C., Smart, A., Moran, G., & Raine, A. (1983). Visual, kinaesthetic and cross-modal development: relationships to motor skill development. *Perception*, 12, 477-483.
- Jordan, K. E., & Baker, J. (2011). Multisensory information boosts numerical matching abilities in young children. *Developmental Science*, 14, 205-213.
- Laine, M., Kwon, M. S., & Hamalainen, H. (2007). Automatic auditory change detection in humans is influenced by visual-auditory associative learning. *Neuroreport*, 18, 1697-1701. doi:10.1097/WNR.0b013e3282f0d118
- Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research*, 158, 405-414. doi:10.1007/s00221-004-1913-2

- Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Brain Research Cognition Brain Research*, 24, 326-334. doi:10.1016/j.cogbrainres.2005.02.005
- Lewkowicz, D. J. (2000). The development of intersensory temporal perception: an epigenetic systems/limitations view. *Psychological Bulletin*, 126, 281-308.
- Lewkowicz, D. J., & Turkewitz, G. (1981). Intersensory interaction in newborns: modification of visual preferences following exposure to sound. *Child Development*, 52, 827-832.
- Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.
- Mahoney, J. R., Verghese, J., Dumas, K., Wang, C., & Holtzer, R. (2012). The effect of multisensory cues on attention in aging. *Brain Research*, 1472, 63-73. doi:10.1016/j.brainres.2012.07.014
- Matusz, P. J., Wallace, M. T., & Murray, M. M. (2017). A multisensory perspective on object memory. *Neuropsychologia*. doi:10.1016/j.neuropsychologia.2017.04.008
- Messbauer, V. C., & de Jong, P. F. (2003). Word, nonword, and visual paired associate learning in Dutch dyslexic children. *Journal of Experimental Child Psychology*, 84, 77-96. doi:10.1016/S0022-0965(02)00179-0
- Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247-279.
- Molholm, S., Ritter, W., Javitt, D. C., & Foxe, J. J. (2004). Multisensory visual-auditory object recognition in humans: a high-density electrical mapping study. *Cerebral Cortex*, 14, 452-465.
- Morrison, F. J. (1982). The development of alertness. *Journal of Experimental Child Psychology*, 34, 187-199.
- Muir, D., & Field, J. (1979). Newborn infants orient to sounds. *Child Development*, 50, 431-436.
- Nardini, M., Bales, J., & Mareschal, D. (2016). Integration of audio-visual information for spatial decisions in children and adults. *Developmental Science*, 19, 803-816. doi:10.1111/desc.12327
- Nardini, M., Jones, P., Bedford, R., & Braddick, O. (2008). Development of cue integration in human navigation. *Current Biology*, 18, 689-693. doi:10.1016/j.cub.2008.04.021

- Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science*, 9, 454-464.
- Oakland, T., Black, J. L., Stanford, G., Nussbaum, N. L., & Balise, R. R. (1998). An evaluation of the dyslexia training program: a multisensory method for promoting reading in students with reading disabilities. *Journal of Learning Disabilities*, 31, 140-147.
- Pirogovsky, E., Murphy, C., & Gilbert, P. E. (2009). Developmental differences in memory for cross-modal associations. *Developmental Science*, 12, 1054-1059. doi:10.1111/j.1467-7687.2009.00857.x
- Pozuelos, J. P., Paz-Alonso, P. M., Castillo, A., Fuentes, L. J., & Rueda, M. R. (2014). Development of attention networks and their interactions in childhood. *Developmental Psychology*, 50, 2405-2415. doi:10.1037/a0037469
- Raij, T., Uutela, K., & Hari, R. (2000). Audiovisual integration of letters in the human brain. *Neuron*, 28, 617-625.
- Robinson, C. W., & Sloutsky, V. M. (2010). Effects of multimodal presentation and stimulus familiarity on auditory and visual processing. *Journal of Experimental Child Psychology*, 107, 351-358. doi:10.1016/j.jecp.2010.04.006
- Rose, S. A., Feldman, J. F., Futterweit, L. R., & Jankowski, J. J. (1998). Continuity in tactual-visual cross-modal transfer: infancy to 11 years. *Developmental Psychology*, 34, 435-440.
- Rose, S. A., Feldman, J. F., Jankowski, J. J., & Futterweit, L. R. (1999). Visual and auditory temporal processing, cross-modal transfer, and reading. *Journal of Learning Disabilities*, 32, 256-266.
- Rose, S. A., Feldman, J. F., & Wallace, I. F. (1992). Infant information processing in relation to six-year cognitive outcomes. *Child Development*, 63, 1126-1141.
- Sai, F. Z. (2005). The role of the mother's voice in developing mother's face preference: Evidence for intermodal perception at birth. *Infant and Child Development*, 14, 29-50.
- Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current Biology*, 16, 1422-1427. doi:10.1016/j.cub.2006.05.048
- Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, 12(11), 411-417. doi: 10.1016/j.tics.2008.07.006

- Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development*, 59, 833-851.
doi:10.2307/1130252
- Siegler, R. S. (1991). *Children's thinking* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Slater, A., Quinn, P. C., Brown, E., & Hayes, R. (1999). Intermodal perception at birth: intersensory redundancy guides newborn infants' learning of arbitrary auditory-visual pairings. *Developmental Science*, 2, 333-338.
- Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*, 74, 822-833.
- Stein, B. E., Burr, D., Constantinidis, C., Laurienti, P. J., Alex Meredith, M., Perrault, T. J., Jr., . . . Lewkowicz, D. J. (2010). Semantic confusion regarding the development of multisensory integration: a practical solution. *European Journal of Neuroscience*, 31, 1713-1720. doi:10.1111/j.1460-9568.2010.07206.x
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.
- Talsma, D. (2015). Predictive coding and multisensory integration: an attentional account of the multisensory mind. *Frontiers in Integrative Neuroscience*, 9, 19.
doi:10.3389/fnint.2015.00019
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14, 400-410. doi:10.1016/j.tics.2010.06.008
- Tiffin, J. (1999). *Quick reference guide for the purdue pegboard #32020 - Test administrator's manual* (Rev. ed.): Lafayette Instruments.
- Tormanen, M. R. K., Takala, M., & Sajaniemi, N. (2008). Learning disabilities and the auditory and visual matching computer program. *Support for Learning*, 23, 80-88. doi:10.1111/j.1467-9604.2008.00374.x
- Tremblay, C., Champoux, F., Voss, P., Bacon, B. A., Lepore, F., & Theoret, H. (2007). Speech and non-speech audio-visual illusions: a developmental study. *PLoS One*, 2, e742. doi:10.1371/journal.pone.0000742
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children - Australian Standardisation* (4th ed.). Sydney: Harcourt Assessment.

Wertheimer, M. (1961). Psychomotor coordination of auditory and visual space at birth. *Science*, *134*, 1692.

Windfuhr, K. L., & Snowling, M. J. (2001). The relationship between paired associate learning and phonological skills in normally developing readers. *Journal of Experimental Child Psychology*, *80*, 160-173. doi:10.1006/jecp.2000.2625

Table 1

Mean (SD) of Demographic Variables and IQ measures for Children and Adults. Note, age ranges are provided in the ‘Participants’ section in Methods.

	8-year-olds	11-year-olds	Adults	<i>t</i> -value	<i>p</i> -value
<i>n total</i>	19	19	19		
<i>n males</i>	7	9	9		
Age (years:months)	8:11(0:7)	11:0(0:8)	24:10(3:6)		
WISC FSIQ	108(12.29)	105(11.67)	n/a	0.82	> .05
WISC VCI	108(14.19)	105(9.57)	n/a	0.62	> .05
WISC PRI	105(11.73)	105(12.41)	n/a	<0.01	> .05
WISC WMI	103(8.79)	100 (8.50)	n/a	0.20	> .05
WISC PSI	107(9.64)	103 (11.67)	n/a	0.31	> .05
Pegboard	13.58(1.58)	14.05 (1.58)	n/a	0.36	> .05

Table 2

Correlation Matrix Including Only the Children's Data (N = 38) Between Age, Overall Accuracy and Reaction Times on the Three Learning Tasks, the Audiovisual Detection Task (Det), and IQ Measures on the WISC and its Subscales.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
IQ																
1. Age	--	.03	.05	.09	-.19	.10	-.26	-.39*	-.37*	.27	.63**	.36*	.28	.002	-.26	-.22
2.FSIQ		--	.85**	.82**	.71**	.67**	.15	.11	.21	-.19	.40*	.33*	.43**	-.06	-.05	-.19
3.VCI			--	.48**	.57**	.47**	.20	.19	.26	-.19	.25	.26	.27	-.003	-.06	-.04
4.PRI				--	.53**	.45**	.16	.11	.17	-.02	.49**	.35*	.56**	-.05	<.001	-.34*
5.WMI					--	.22	.30	.28	.32*	-.20	.02	.09	.17	.10	.17	-.06
6.PSI						--	-.20	-.21	-.10	-.22	.38*	.22	.22	-.28	-.27	-.10
Detection task																
7.Det-A							--	.87**	.90**	-.35**	-.11	-.10	-.06	.09	.25	.33*
8.Det-V								--	.95**	-.36*	-.22	-.15	-.09	.02	.19	.28
9.Det-AV									--	-.55**	-.13	-.18	-.12	.05	.23	.37*
10.Det-f										--	-.01	.07	.20	.15	-.12	-.34*
Learning task																
11.Shape-VV											--	.47**	.58**	-.14	-.44**	-.49**
12.Novel-AV												--	.51**	-.14	-.11	-.49**
13.Verbal-AV													--	.16	-.07	-.44**
14.Shape-VV RT														--	.56**	.20
15.Novel-AV RT															--	.57**
16.Verbal-AV RT																--

Note: * = $p < .05$, ** = $p < .01$. WISC-IV: FSIQ = Full Scale Intelligence Quotient, VCI

= Verbal Comprehension Index, PRI = Perceptual Reasoning Index, WMI = Working

Memory Index, PSI = Processing Speed Index. Detection task: Det-A = RT for

audition, Det-V = RT for vision, Det-AV = RT for audiovisual, Det-f = percentage of

facilitation. Associative learning task: Shape-VV = d-prime for visual-visual learning,

Novel-AV = d-prime for novel audiovisual learning, Verbal-AV = d-prime for verbal

audiovisual learning, RT = reaction time measures.

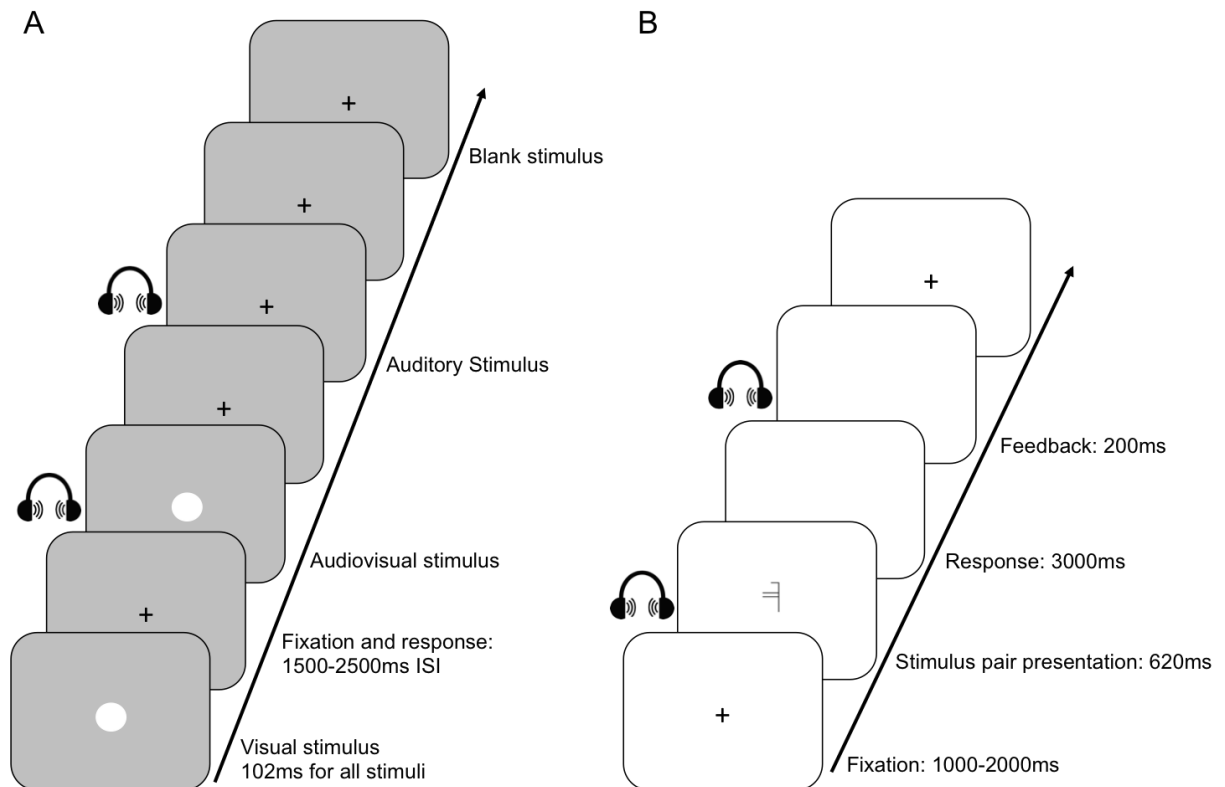


Figure 1. A. An illustration of a sequence of trials of the audiovisual detection task demonstrating each of the signal types. B. An illustration of a single trial of an audiovisual associative learning task.

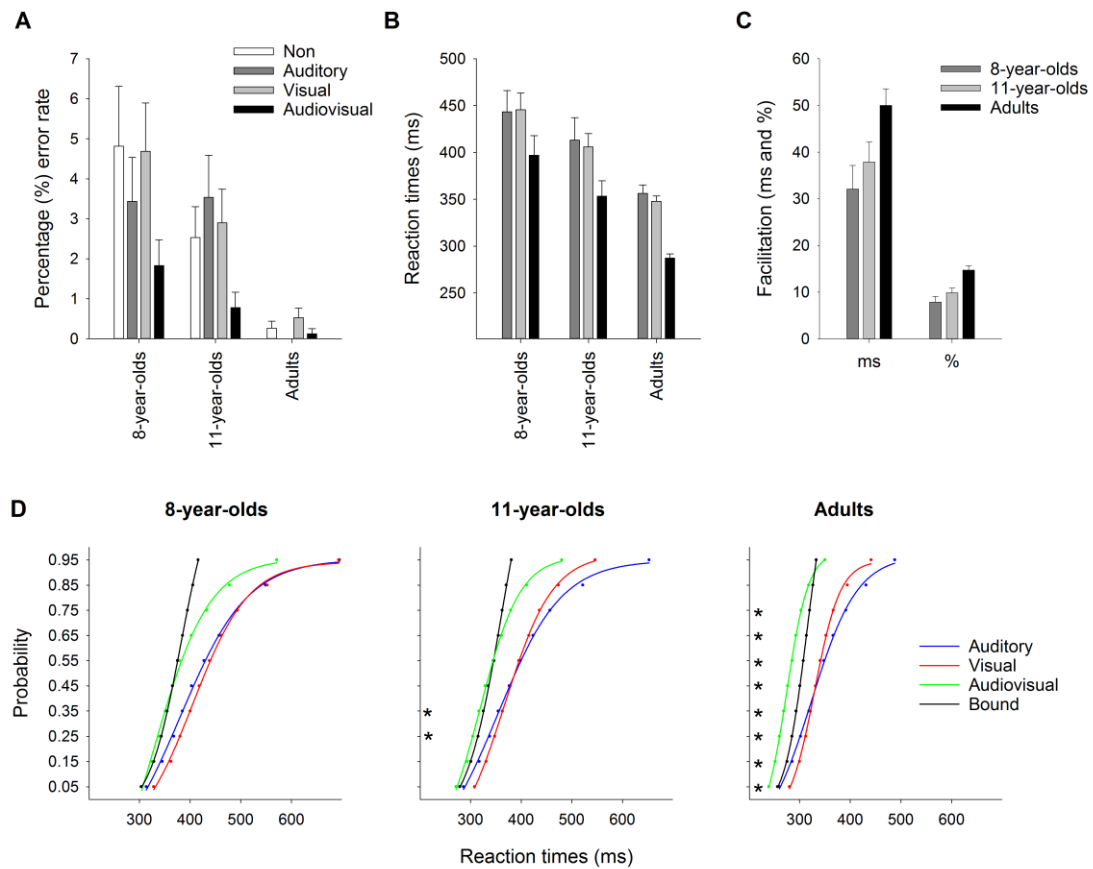


Figure 2. **A.** Mean percentage of error rate (+SD), **B.** reaction time (+SEM) and **C.** multisensory facilitation (+SEM) shown in milliseconds (ms; left of Figure 1C) and percentage (%; right of Figure 1C) on the audiovisual speeded detection task. **D.** Cumulative density functions (CDFs) for auditory, visual, audiovisual stimuli and for the bound auditory and visual conditions. Note that * = $p < .05$.

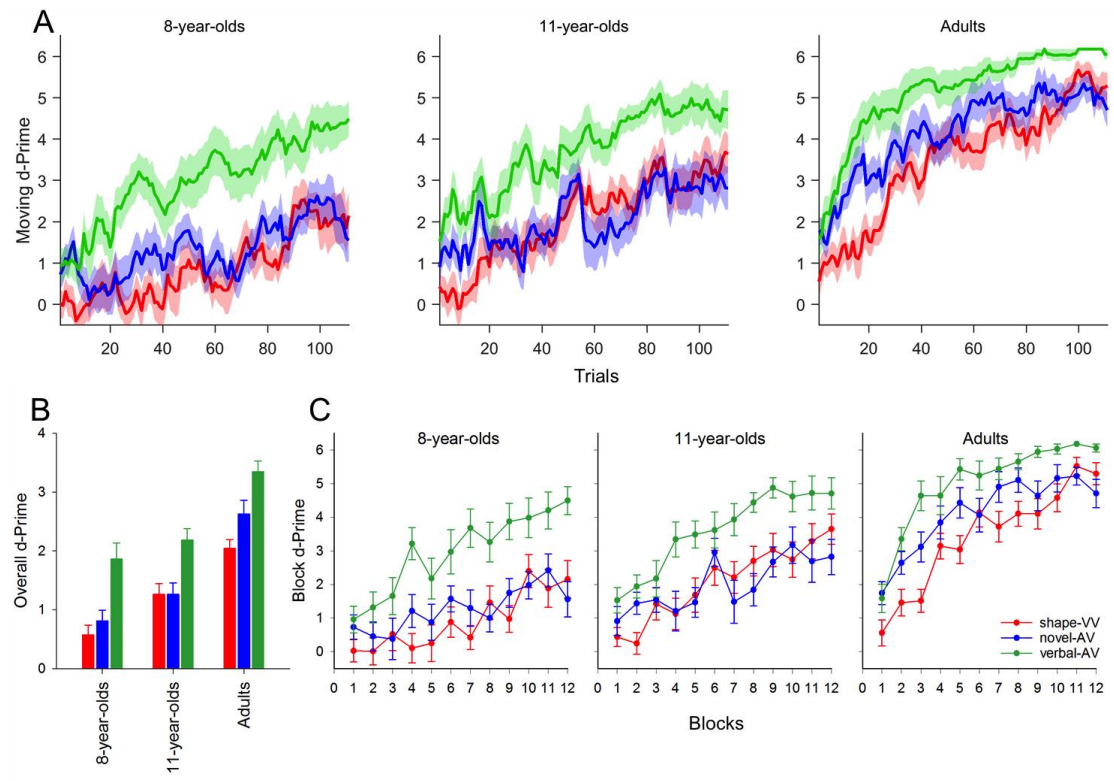


Figure 3. A. Moving average of mean d-prime measures for 8-year-olds, 11-year-olds and adults, and visual-visual (shape-VV- red line) tasks, auditory-visual (novel-AV – blue line), and visual-verbal (verbal-AV – green line). Shaded areas depict SEM along the moving average. **B.** Mean (\pm SEM) overall d-prime measures for each age group on the novel-sound auditory-visual (novel-AV), visual-verbal (verbal-AV), and visual-visual (shape-VV) learning tasks. **C.** Mean d-prime (\pm SEM) for the 12 blocks of trials for visual-novel sound (novel-AV), visual-verbal (verbal-AV), and visual-visual (shape-VV) learning tasks for children and adults.

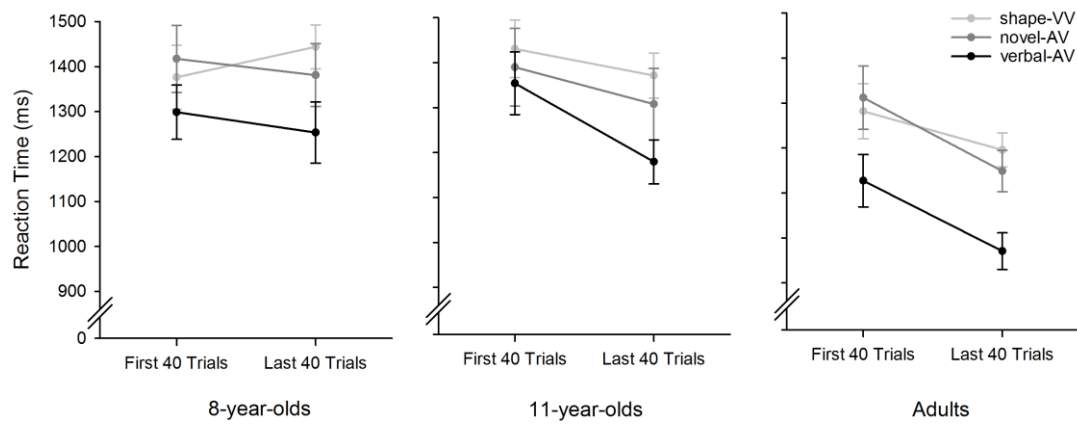


Figure 4. Mean RTs (\pm SEM) for the learning phase (trials 1 – 40) vs. the learnt phase (trials 80-120) in the novel-sound auditory-visual (novel-AV), visual-verbal (verbal-AV), and visual-visual (shape-VV) learning tasks for each age group.

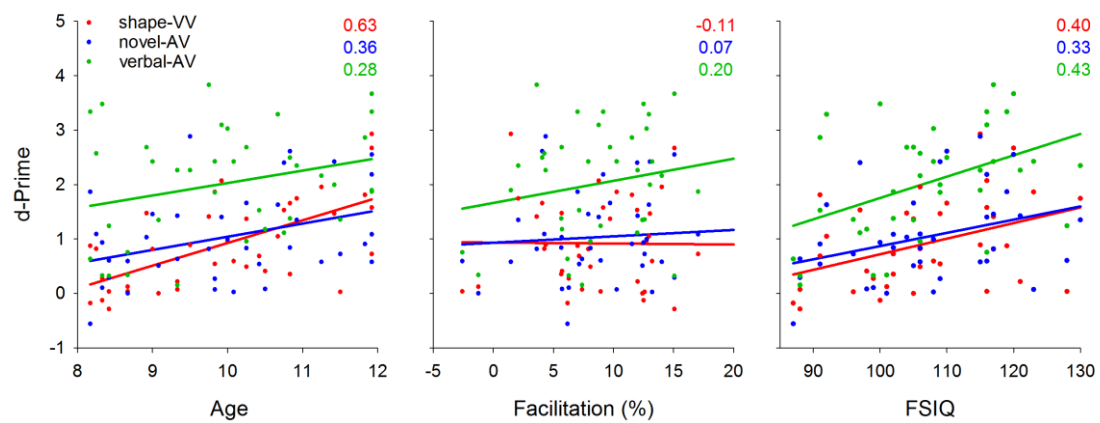


Figure 5. Scatter plots showing the relationship (including r values) between d-prime measures on the three learning tasks and age, percentage of multisensory facilitation and general intellectual abilities (FSIQ).