



Published in final edited form as:

J Exp Child Psychol. 2012 July ; 112(3): 338–350. doi:10.1016/j.jecp.2011.12.002.

Preschool children and adults flexibly shift their preferences for auditory versus visual modalities, but do not exhibit auditory dominance

Nicholaus S. Noles and Susan A. Gelman

University of Michigan

Abstract

The goal of the present study is to evaluate the claim that young children display preferences for auditory stimuli over visual stimuli. This study is motivated by concerns that the visual stimuli employed in prior studies were considerably more complex and less distinctive than the competing auditory stimuli, resulting in an illusory preference for auditory cues. Across three experiments, preschool children and adults were trained to use paired audio-visual cues to predict the location of a target. At test, the cues were switched so that auditory cues indicated one location and visual cues indicated the opposite location. In contrast to prior studies, preschool age children did not exhibit auditory dominance. Instead, children and adults flexibly shifted their preferences as a function of the degree of contrast within each modality (with high contrast leading to greater use).

Keywords

Categories; Modality Preferences; Similarity-based Models; Perception; Labels; Development

Understanding and explaining children's induction and categorization behaviors are important goals for developmental psychologists. Two broad approaches – one focusing on associative and similarity-based processes and the other focusing on children's theories – are at the center of an ongoing debate (e.g., Gelman & Waxman, 2009; Sloutsky, 2009). Similarity-based and associative learning approaches suggest that basic cognitive and perceptual processes, such as the ability to detect statistical regularities (Rakison, 2004) and judge perceptual similarity (Sloutsky & Napolitano, 2004), are sufficient to guide children's categorization and induction. In contrast, theory-based approaches argue that children additionally make use of domain-specific knowledge, ontologies, causal relations, and non-obvious properties (Carey, 2009; Gelman, 2003; Gopnik & Sobel, 2000; Wellman & Gelman, 1998).

One key point of contention concerns the role of category labels. Hearing the same label for two items increases the likelihood that children (and adults) will generalize information from one item to the other (Balaban & Waxman, 1997; Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Sloutsky & Fisher, 2004; Xu, Cote, & Baker, 2005). Yet there are at

© 2011 Elsevier Inc. All rights reserved.

Address correspondence to Nicholaus Noles: Department of Psychology, University of Michigan, 530 Church Street, Ann Arbor, MI 48109-1043, nicholaus.noles@gmail.com, Phone : 517-505-5561.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

least two possible explanations for this result. One possibility is that conceptual similarity guides the effect. On this view, the label conveys information regarding category membership, and items from the same category are assumed to share many important features. Evidence in support of this interpretation is that the same effect holds when synonyms are used (Gelman & Markman, 1986, Study 2), or when conceptual similarity is detectable by subtle featural cues (Gelman & Markman, 1987, Picture-Only Condition). Furthermore, when children hear non-linguistic sounds (e.g., tones; emotional expressions) rather than linguistic labels, children do not make use of them to guide their categorization (Balaban & Waxman, 1997; Fulkerson & Waxman, 2007; Xu, 2002; Xu, Cote, & Baker, 2005).

In contrast, an alternative possibility is that the labeling effect is due to perceptual similarity (Sloutsky & Fisher, 2004; Sloutsky, Fisher, & Lo, 2001). In one prominent similarity model, the SINC (Similarity, Induction, Naming, and Categorization) model, labels are perceived as perceptual features of objects, and items with the same label are judged to be more similar to one another than are items that are unlabeled, or that receive different labels. Moreover, the model stipulates that young children find auditory cues to be more salient than visual cues (Napolitan & Sloutsky, 2004; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). Thus, labels affect children's inferences because they are auditory cues, and therefore have greater perceptual salience.

The present paper does not examine the SINC model in general, but examines the foundational claim that auditory cues are more salient than visual cues for young children. In prior research testing the auditory dominance theory, young children were presented with stimuli with both auditory and visual components, and the relative salience of the two cues was tested (Napolitan & Sloutsky, 2004; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). Specifically, a modified switch design was employed (e.g., see Werker, Lloyd, Cohen, Casasola, & Stager, 1998) that trained participants to use cross-modal cues to complete a task. Visual and auditory stimuli were paired (e.g., a visual display including a circle, pentagon, and triangle might be paired with a burst of white noise), and each pair reliably indicated that a target stimulus would appear at a certain location (e.g., one pair appeared on the right side of the computer screen; another pair appeared on the opposite side of the screen). At test, conflicting cues were presented so that one modality indicated that the target was in one location, whereas the other modality indicated the opposite location. Participants did not receive feedback, so they could only rely on a preferred modality to predict the target's location. In two sets of studies, 4-year-olds displayed an auditory preference, whereas adults showed a visual preference (Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003). Napolitano and Sloutsky (2004) expanded on these results by exploring the mechanisms that might underlie auditory dominance in children, concluding that children exhibit "a default auditory dominance" (p. 1869) when stimuli are unfamiliar, but that when the visual stimuli are familiar, the pattern can reverse.

Before concluding that children overall are biased to attend to auditory cues, however, it is important to demonstrate that the stimuli provide a fair test by including cues that are as evenly balanced between the two modalities as possible. One can readily imagine tests in which the two modalities are not evenly balanced. For example, if one were to compare an "easy" auditory distinction (e.g., a pleasant bell tone versus a loud undulating siren) with a "hard" visual distinction (e.g., two different paintings of waterlilies by Monet), it might not be surprising to find that people show an auditory preference. In contrast, if one were to compare a "hard" auditory distinction (e.g., two musical tones played on a synthesizer, less than 0.5 decibels apart in volume) with an "easy" visual distinction (e.g., a small black dot versus a large red star), people might then show a visual preference. In other words, intuition suggests that some comparisons may be more balanced than others.

Pretesting in prior experiments demonstrated that the stimulus sets used were individually discriminable (i.e., listeners could tell the difference between the two stimuli within a given modality), but this does not mean that the contrasts between stimuli within a given modality were of the same magnitude between stimuli across modalities. For example, both a small contrast in visual stimuli and a large contrast in auditory stimuli may be individually discriminable. However, when presented in concert, one modality may draw more attention than another, or the combination of the two cues may diminish the relative contrast between the different cues within the low-contrast modality.

We hypothesized that the choice of stimuli employed in prior experiments may have significantly influenced modality preferences in several ways. First, in some experiments, there were more visual stimuli to encode and compare than auditory stimuli. For example, Robinson and Sloutsky (2004) pitted two sounds against six shapes (two sets of three each). Second, visual stimuli were more complex than auditory stimuli. Sloutsky and Napolitano (2003) contrasted simple computer-generated tones with detailed landscapes at different depths of view (e.g., close-up pictures of plant leaves and wide shots of forests). Third, the contrast between modalities was not well controlled. The auditory stimuli used in these studies varied integral stimulus dimensions (i.e., volume, pitch, and timbre), whereas the visual stimuli varied separable stimulus dimensions (e.g., shape) (see Robinson & Sloutsky, 2004). Individuals perceive separable and integral stimulus dimensions in different ways (Garner, 1974). For example, changes in integral dimensions yield categorical changes in perception. A small change in timbre categorically changes the nature of a sound (e.g., consider the difference between a note played on a piano and the same note played on a saxophone – both instruments produce the same note, but the two sounds are obviously different), but even large changes on separable dimensions do not necessarily influence the broader construal of perceptual objects (e.g., changing the color or size of a square doesn't make it any less square).

Napolitano and Sloutsky (2004) attempted to address some of the uncontrolled variability across dimensions mentioned above (e.g., single vs. multiple objects; shapes vs. scenes) by varying the complexity and familiarity of visual and auditory cues. Unfortunately, the experimental manipulations used in these studies introduced additional uncontrolled variability. For example, shapes (simple) were compared to landscapes, whole animals, and animal faces (complex). Likewise, computer-generated tones (simple) were compared to doorbells, barking dogs, and dial tones (complex). Furthermore, whole animals (familiar) were compared to animal faces (novel), but this comparison is not symmetrical because, in addition to the format change in the two stimuli sets, faces and whole animals may not be processed or categorized in the same way (e.g., even if children are familiar with a raccoon, they may not attend to a raccoon's specific and differentiating facial features). These contrasts cross a host of categorical and perceptual boundaries that are not captured by the simple dimension of "complexity" or "familiarity". In other words, participants were presented with layers of uncontrolled, structured variables that may guide their preferences and perhaps the deployment of their attention.

Finally, in prior tests, the specific auditory stimuli had additional attributes that may have increased their distinctiveness. Some of the sounds included a dynamic aspect such that they changed over time (e.g., a duck quacking, a dog barking, a laser sound; Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004), whereas the pictures were static, with no dynamic changes. Similarly, some of the sounds varied in their valence (e.g., a burst of white noise was included in Robinson & Sloutsky, 2004, which can be aversive), whereas the visual stimuli were more neutral in valence.

To summarize: Despite the wealth of prior evidence investigating the hypothesis that children show auditory dominance, this question has yet to be tested with stimuli for which auditory and perceptual contrasts are systematically manipulated, while controlling for item complexity and familiarity. We suggest that such a test provides a neutral test of modality processing biases.

Present Experiments

The current experiments were designed to test the one of the central assumptions of the SINC model: that the auditory modality is dominant in young children. We did so by making use of a modified switch design, as in prior research, but systematically varying the visual and auditory cues presented as stimuli. Experiment 1 was designed to employ stimuli that exhibit roughly equivalent levels of complexity and contrast across modalities. In particular, we aimed to include stimuli that were highly discriminable on both the auditory and visual dimensions in order to permit a test of auditory versus visual dominance when there are no barriers to perceiving or noticing the contrasts being displayed. In Experiments 2 and 3, we varied the nature of the contrasts presented, to examine flexibility in preference for auditory versus visual cues. Due to the importance of familiarity in prior work (Napolitano & Sloutsky, 2004), we kept the familiarity of the items constant throughout the experiments.

Napolitano and Sloutsky (2004) demonstrated that modality dominance is somewhat flexible, but they concluded that children exhibit auditory dominance and that preferences for visual stimuli were moderated by stimulus familiarity. In contrast, we will demonstrate that children do not exhibit auditory dominance and that modality preferences are driven by the relative “distance” between stimuli within each of the competing modalities. We also expand upon previous experiments examining modality flexibility by testing adult participants.

Experiment 1

Method

Participants—Although this investigation was primarily focused on children’s modality preferences, we tested both children and adults. Testing adults afforded us an opportunity to determine if child and adult preferences differed, as well as a means to fully characterize any developmental differences that we discovered. Sixteen 4-year-olds (6 female, 10 male, $M_{age} = 4.39$, $SD = .46$) and sixteen adult undergraduates (4 female, 12 male) participated. Children were recruited from [removed for blind review], representing a wide variety of social and economic backgrounds. Most participants were Caucasian-American. Adults were tested in a laboratory setting and received course credit for participation, whereas children were tested at their school and received a sticker for their efforts.

Materials—Stimulus design is a very important aspect of this investigation. Poor stimulus matching is the key confounding factor in prior investigations (see Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004), and our approach is based on presenting children with similar experiments employing stimuli that reveal their modality preferences. The ideal way to create stimuli that control within and between modality contrasts would be to individually calibrate the stimuli viewed by each participant. However, the task employed in both this study and previous investigations is very challenging for four-year-olds. Between 26% and 32% of the children who participated in Robinson and Sloutsky’s (2004) core experiments were unable to successfully complete the training task. Critically, the attrition rate for such a study might be massive, perhaps yielding a highly selected study sample that might not reflect the population-level predictions that are offered by the SINC model. Thus, our goal

in designing materials for the present study was to create systematically varied stimuli with large and reliably discriminable differences.

Rather than simply choosing shapes and sounds that were different, we created stimuli that systematically and symmetrically varied on matched stimulus dimensions. We focused on manipulating only integral stimulus dimensions of both visual and auditory stimuli, but hue and timbre were not manipulated because changing them often results in categorical changes within a stimulus dimension (e.g., playing different notes at different volumes does not make a saxophone sound like a different instrument, but playing the same note with both a piano and a saxophone results in categorically different sounds, even if the notes have the same volume and pitch). Also, rather than focusing on the specific values for different stimulus dimensions, our manipulations, and therefore our within and between modality contrasts, emphasized large magnitude differences (e.g., magnitude ratios ranged from 1:2 to 1:3 in all experiments) of the same degree and distance. One of the foundational elements of psychophysics is that humans can match the intensity of stimuli across modalities and compare different kinds of stimuli. So long as a person is presented with discriminably different stimuli, they can match the brightness of a light to the loudness of a tone, the intensity of an electric shock, the length of a line, the size of a shape, or a multitude of other sensations (e.g., Stevens' Power Law, for a review, see Stevens, 1975). By applying the same ratio manipulations (e.g., shape one is twice as saturated as a shape two, and the pitch of sound one is twice as high as sound two), we leveraged this aspect of perception to create equivalently spaced stimuli within each modality. This approach contrasts with the approach used by Sloutsky and Fisher (2004), who employed stimuli with different, unmeasured levels of contrast both within and across modalities. Their standard for a "good" contrast was one that children could discriminate when the stimuli were presented unimodally. Finally, because contrasts between novelty and familiarity were presented as important dimensions in previous investigations, we used familiar auditory and visual stimuli (piano notes and colored circles, respectively) and manipulated these stimuli on dimensions that are not easily nameable. For example, it is unlikely that most preschool age children have learned specific labels that differentiate between individual piano notes or specific saturation levels.

Stimuli consisted of two auditory and two visual stimuli. The auditory stimuli consisted of two computer-generated tones created by manipulating volume and pitch, and leaving timbre constant. One sound (AUD₁) was a low volume, low pitch note (G₂, 98 Hz); the other (AUD₂) was a high volume, high pitch note (D₄, 294 Hz). Children were tested in multiple locations, so precise estimations of volume are unavailable. However, the relative difference in volume between the two auditory stimuli was defined in the audio files and remained constant. The quiet sound (approximately 60 dB) was 34% of the volume of the loud sound (approximately 80dB). Auditory stimuli were created and manipulated using GarageBand. Visual stimuli consisted of colored circles 5.08 cm in diameter that were similarly manipulated. Visual stimuli consisted of three elements: hue, saturation, and brightness. As with the auditory stimuli, only two stimulus dimensions were manipulated. VIS₁ was a bright red circle (rgb values = 255, 0, 0); VIS₂ was a very dark red circle. We created VIS₂ by reducing the saturation and brightness values used in VIS₁ by approximately 66% using Adobe Photoshop (see Figure 1b). Cross-modal cues, consisting of paired visual and auditory stimuli were presented for 2000ms.

These stimuli were created to use large and roughly equivalent contrasts (i.e., 1:3 ratios) on two dimensions. In contrast to previous studies, we varied a restricted number of stimulus dimensions and we carefully avoided adding additional uncontrolled variables in the process of manipulating the variables of interest. Critically, the variability both between and, more importantly, within the competing modalities was dramatically cleaner than in previous

research. Additional stimuli included a digital photo of a reclining cat and a digital photo of the same cat wearing a novelty giraffe hat.

Procedure—The experimental session involved three phases: 1) warm-up, 2) training, and 3) testing. Each participant was presented with a short introduction to the task. The experimenter said, “Today we’re going to play a game called ‘Where’s Amy?’ Amy is my cat and she loves to play hide-and-go-seek. We play the game by pointing at the screen and guessing where Amy is hiding. She might be hiding here or here [experimenter indicated each side of a computer screen]. Amy might be hiding anywhere, so we’re going to give you hints to help you guess where she’s hiding.” The experimenter then began the warm-up phase, saying “Whenever this happens [experimenter played audio-visual pair 1], Amy is always hiding over here [e.g., on the left-hand side of the screen], and whenever this happens [experimenter played audio-visual pair 2] Amy is always hiding over here [e.g., on the right-hand side of the screen].” After each hint was presented, a picture of a cat appeared on the side of the screen indicated by the experimenter. The pairing of the auditory and visual stimuli and the side to which each pair refers was counterbalanced between subjects. After the experimenter presented each hint twice, the training phase began.

In the training phase, participants were presented with fourteen train trials with visual and verbal feedback. The experimenter controlled the pace of the experiment, waiting for the participant to respond to the stimuli before providing feedback and continuing on to the next trial. At the end of every training trial, visual feedback was provided. A picture of a cat would appear in the target location after children made their response. When a child provided a response, the experimenter would say, “Good job,” when the child guessed correctly, or “Oops! She’s over there” (pointing to the correct location) when the child was incorrect. If the child guessed incorrectly on one of the first six trials, then the experimenter replayed that trial, saying, “Remember, when this hint happens, Amy is always hiding here,” before continuing on to the next trial. After the first six training trials, the experimenter only provided verbal feedback (e.g., “good job” or “oops, she’s over there”) and did not replay trials. Each audio-visual pair perfectly predicted the cat’s location during the training phase. If participants correctly indicated the cat’s location on either four of the final six trials or on all of the final three trials, then they entered the test phase.

In the test phase, the experimenter provided the participant with additional instructions: “Now we’re going to keep playing the game, but Amy is going to stay hidden. You’re going to guess where she’s hiding, but she’s not going to appear. At the end, Amy is going to do a special trick for you.” Test trials were identical to training trials except that they lacked feedback, and the auditory/visual pairings were switched. For example, if the participants experienced VIS_1+AUD_1 and VIS_2+AUD_2 as training stimuli, then they received VIS_1+AUD_2 and VIS_2+AUD_1 during the test phase, so that each modality indicated a different, competing hiding place. After completing 12 test trials, participants were presented with a picture of the cat wearing a novelty giraffe hat and the experimenter said, “Look, there’s Amy in her Halloween costume!”

Presentation software was used to present stimuli on laptop computers with 14.1-inch screens. Children were tested in free spaces at their preschools and adults were tested in private rooms. Experimental sessions were identical for children and adults except that adults did not receive a sticker for their participation.

Results and Discussion

All participants tested in Experiment 1 mastered the training phase, predicting the location of the cat with impressive accuracy ($M = 96\%$ and 100% correct over the 14 training trials, for children and adults respectively). After completing the test trials, participants were

classified into different responder types using the same criteria employed by Sloutsky and colleagues. If participants relied on auditory cues on 75% or more of test trials, then they were designated as an Auditory Responder; if they relied on visual cues on 75% or more of test trials, then they were designated as a Visual Responder. Participants who did not meet either criterion were designated as Mixed Responders. Note that Mixed Responders did not significantly prefer either modality (i.e., they responded at chance as calculated by the binomial theorem).

Most participants, 93.75% of children and 87.5% of adults, were Visual Responders, relying on visual cues when presented with conflicting cross-modal cues. The remaining participants were Auditory Responders (2 adults) and a Mixed Responder (1 child) (see Table 1).

This basic response pattern was also reflected in participants' overall use of auditory cues (see Figure 2). Including the responses of all participants, children relied on auditory cues on only 3% of trials, and adults relied on auditory cues on only 13% of trials. Rather than exhibiting a preference for auditory stimuli, our participants overwhelmingly relied on visual cues when guessing the cat's location (97% of children's trials; 87% of adults' trials). Preference for the visual stimulus was significantly greater than chance for both children, $t(15) = 22.43, p < .001$, and adults, $t(15) = 4.35, p = .001$. An independent samples t-test found no statistically significant differences in auditory cue use between age groups, $t(30) = 1.19, p = .244$.

Experiment 2

The pattern of responses exhibited by participants in Experiment 1 did not result in the auditory preferences consistently reported in prior research (e.g., Robinson & Sloutsky, 2004). When exposed to cues that were matched across modalities in terms of complexity, stimulus dimensions, and relative degree of contrast, both children and adults demonstrated a strong preference for visual cues. These results would seem to undermine the argument that children display an auditory bias. However, the strong visual bias found in Experiment 1 may be explained in several ways. First, it is possible that participants possess an overarching preference for the visual modality, but that children's lack of visual dominance in prior research is attributable to the structure of the stimuli in those studies. Second, individuals may flexibly shift their modality preferences in response to the particular contrasts that they observe. If children's modality preferences are guided by the structure of test stimuli, then they will appear to be "auditory dominant" or "visual dominant" as a function of the stimuli selected for use, as reported by Napolitano and Sloutsky (2004). Third, it is possible that the auditory cues employed in Experiment 1 were indistinguishable and that participants could only distinguish between the visual cues.

In Experiment 2, we address these possibilities by retaining the same auditory contrast used in Experiment 1, but reducing the contrast between the visual stimuli. Thus, participants were presented with a 'difficult,' low-contrast visual pair and an 'easy,' high-contrast auditory pair. If participants possess an overarching preference for the visual modality, then they should continue to employ that modality, even when it is made less salient due to low contrast. Conversely, if participants flexibly shift modalities in response to changes in cue contrast, then this argues for flexibility rather than a dominant modality. Furthermore, such a finding would demonstrate that participants could distinguish the auditory stimuli employed in Experiment 1 but did not rely on them.

Participants

Fifteen 4-year-olds (7 female, 8 male, $M_{age} = 4.39$, $SD = .50$) and sixteen adult undergraduates (10 female, 6 male) participated in this experiment. Three additional children were tested but failed to meet the training criterion. Child participants were recruited from [removed for blind review]. The participants came from a variety of ethnic and racial backgrounds. Adults were tested in a laboratory setting and received course credit for participation; children were tested at their school and received a sticker for their efforts. None of the children or adults had participated in Study 1.

Materials

The materials included the same auditory cues as in Experiment 1 but different visual cues. Specifically, the contrast between visual cues was reduced in Experiment 2 compared to Experiment 1. In Experiment 2, VIS₁ was a bright blue circle (rgb values = 0, 0, 255) and VIS₂ was a pale blue circle. VIS₂ was created by reducing the saturation of VIS₁ by 50% (see Figure 1a). In contrast to the visual stimuli used in Experiment 1, these visual stimuli varied on one dimension (saturation) and the manipulation was smaller (a 50% reduction in saturation instead of a 66% reduction).

Procedure

The procedures employed in Experiment 2 were identical to those used in Experiment 1. Participants who did not meet this criterion were thanked for doing an excellent job and their session was halted.

Results and Discussion

In Experiment 2, adults exhibited perfect accuracy during training ($M = 100\%$), and children, though they were less accurate than in Experiment 1, were also quite accurate ($M = 74.0\%$). Table 1 reports the results during the testing phase. Applying the same criteria used in Experiment 1, 66.6% of children and 68.8% of adults were Auditory Responders. The rest of the children were Mixed Responders, but the rest of the adults – consisting of almost one-third of the adult sample – were Visual Responders. This pattern was also reflected in an overall preference for auditory cues exhibited by both age groups. Children used auditory cues in 78% of test trials and adults used them in 67% of test trials. Children's auditory preference was significantly above chance, $t(14) = 4.12$, $p = .001$, whereas adults' auditory preference did not significantly differ from chance, $t(15) = 1.46$, $p = .164$. However, there was no significant difference in performance between the two age groups, $t(29) = .77$, $p = .768$.

There was striking difference in performance between Experiments 1 and 2. The use of auditory cues increased significantly for both children, $t(29) = 10.94$, $p < .001$, and adults, $t(30) = 3.74$, $p = .001$. Considered together, the results of the first two experiments indicate that children prefer visual cues when both visual and auditory cues are salient and accessible, but that they also flexibly rely upon auditory cues when they are more discriminable than visual cues. Experiment 2 also demonstrates that children's visual preference in Experiment 1 was not due to an inability to discriminate the auditory stimuli in that experiment, as they were able to use the auditory cues successfully in Experiment 2. Finally, we demonstrated a shift in cue preference without any associated changes in "familiarity" within or between modalities.

Experiment 3

Before we can conclude that observers flexibly shift between modalities in response to characteristics of test stimuli, we must demonstrate that participants can distinguish between

the two low-contrast visual stimuli presented in Experiment 2. If the visual cues used in Experiment 2 were indiscriminable or functionally identical to children, then participants may have had no choice but to rely on auditory cues. In this situation, using the auditory modality would be a necessity and not a shift in preferences. However, if children could discriminate between VIS₁ and VIS₂ in Experiment 2, but attended to the auditory modality anyway, then our results suggest that children and adults flexibly attend to more high-contrast modalities when presented with competing cross-modal cues. The goals of Experiment 3 were two-fold. First, we intended to demonstrate that children can discriminate between the visual stimuli employed in Experiment 2. Second, we intended to replicate Experiment 1 with low-contrast auditory and visual stimuli.

Methods

Participants—Sixteen 4-year-olds (8 female, 8 male, $M_{age} = 4.37$, $SD = .39$) and 24 adult undergraduates (9 female, 15 male) participated in this experiment. Thirteen additional children were tested but failed to meet training criterion. An additional six children ($M_{age} = 4.78$, $SD = .13$) participated in pretesting of the materials (see Materials section below). Child participants were recruited from [removed for blind review]. The participants came from a variety of ethnic and racial backgrounds. Adults were tested in a laboratory setting and received course credit for participation; children were tested at their school and received a sticker for their efforts. None of the children or adults had participated in Experiments 1 or 2.

Materials—The visual stimuli employed in Experiment 3 were identical to those used in Experiment 2. New auditory stimuli that were roughly equivalent to the low-contrast visual stimuli were created for Experiment 2. AUD₁ was a high pitch piano note (C₄, 262 Hz) and AUD₂ was a lower pitch piano note (C₃, 131 Hz). Both notes were the same volume (approximately 70 dB) and timbre (computer-generated piano notes). Across dimensions, both auditory and visual stimuli varied by roughly the same amount on only one dimension (i.e., VIS₁ and AUD₁ were double the brightness and pitch of VIS₂ and AUD₂, respectively). Because the auditory stimuli used in Experiment 3 were created using a 1:2 ratio, placing them in the same pitch class, and because the difference between each auditory cue was smaller than the ratio used in Experiments 1 and 2, six additional children were tested in a separate discrimination task using only the auditory stimuli in training and at test. One child refused to complete the test session, but the remaining five children all used these auditory stimuli to accurately predict the cat's location ($M = 100\%$), indicating that children can discriminate and employ the auditory cues used in this experiment.

Procedure—The procedures employed in Experiment 3 were identical to those used in Experiments 1 and 2.

Results and Discussion

Experiment 3 was more difficult than Experiments 1 or 2. Thirteen children failed to meet training criteria in Experiment 3, in contrast to the earlier experiments (only 3 children failed to meet criteria in Experiment 2 and no children failed to meet the training criteria in Experiment 1). However, children who successfully completed training were as accurate ($M = 75.0\%$) as children tested in Experiment 2, and adults were very accurate ($M = 100\%$) during training.

The majority of adults (87.5%) were Visual Responders and the remaining adults (12.5%) were Auditory Responders. Most children, 50% of our sample, were Visual Responders, 43.75% of children were Mixed Responders, and 6.25% were Auditory Responders (see Table 1). On average, adults employed auditory cues on 13% of trials and visual cues on

87% of trials (above-chance in use of visual cues, $t(23) = 5.44, p < .001$). Similarly, on average, children used auditory cues on 37% of trials and visual cues on 63% of trials (above-chance in use of visual cues, $t(15) = 1.98, p = .067$).

In Experiment 3, unlike Experiments 1 and 2, children employed auditory cues significantly more often than adults (see Figure 2), $t(38) = 2.44, p = .019$. Although most children and adults relied on visual cues in Experiment 3, adults employed auditory cues to the same extent that they did in Experiment 1, $t(38) = .05, p = .962$, whereas children's use of auditory cues increased significantly, $t(14) = 4.12, p = .001$. Although this developmental difference is consistent with findings reported by Robinson and Sloutsky (2004), our analysis of individual response patterns indicates that only one child was an Auditory Responder. Thus, the increased use of auditory cues in this experiment does not reflect a strong preference for auditory cues, but rather reflects a dramatic increase in Mixed Responders. This response pattern could have several possible sources. It is possible that the additional difficulty of this task resulted in relatively more random or mixed responding. It is also possible that children either employed both kinds of cues at test or responded at random. Neither random responding nor shifts in cue preferences between individual trials represent a response pattern that reflects auditory dominance.

Whereas Robinson and Sloutsky reported an absolute difference in preference (i.e., children preferred auditory cues whereas adults preferred visual cues), we obtained a difference in degree. Children preferred visual cues overall, but their use of auditory cues was greater when contrasts within cross-modal stimuli were small. Critically, this effect of age was driven by an increase in the number of children who used both visual and auditory cues, and not by individual children who exhibited an overarching preference for auditory cues.

General Discussion

After equating visual and auditory cues more closely than in previous studies, we found that both preschool children and adults exhibited a strong preference for visual cues. This finding runs counter to predictions offered by the SINC model, and undermines the argument that children possess a consistent preference for auditory cues. Instead, our data suggest that children (like adults) tend toward a visual bias, though the extent of this bias shifts as a function of the test items. When one dimension presents lower contrast than another dimension, both children and adults are less likely to make use of that dimension on the modified switch design. We attribute the difference between our findings and those of prior reports as due to differences in stimuli. We deliberately created highly discriminable stimuli (i.e., magnitude ratios between 1:2 and 1:3) that would be easy to process, where auditory and visual dimensions were equated as much as possible, and where extraneous additional factors (such as aversive and/or dynamic auditory stimuli) were excluded.

In discussing their findings, Robinson and Sloutsky (2004) argue that their results were not driven by greater complexity of the visual stimuli, citing that children effectively encoded and employed non-preferred stimuli in calibration studies. However, although children can discriminate between complex unimodal cues when they observe them in isolation, they may fail to fully process these stimuli or ignore them altogether when they are presented with competing, more distinguishable cues. This result parallels findings presented by Napolitano and Sloutsky (2004), who also found flexibility in children's modality preferences. However, the procedures and stimulus manipulations used in prior research introduced a number of uncontrolled, competing variables, including qualitative and quantitative differences between experimental conditions and specific sets of stimuli. Critically, when these extraneous competing factors are removed, as they are in the present studies, children exhibit flexible modality preferences that are heavily weighted by an overarching preference

for visual stimuli. The relative distance between stimuli within stimulus dimensions appears to predict modality preferences in both children and adults.

Robinson and Sloutsky (2004) also argue that when children do make use of visual cues, it is because the stimuli are familiar and nameable. In order to fully reject any possible effects of familiarity, we tested an additional eight adults and twelve children ($M = 4.57$, $SD = .46$) on a condition that was similar to Experiment 1, except that the visual stimuli were replaced with two novel shapes (an hourglass-shaped hexagon and a “blob” of three overlapping circles, see Figure 1c), both of the identical blue hue. The auditory stimuli were unchanged from Experiment 1. After the experimental task, children were asked to name the shapes, and only one child was able to come up with a description (a child who reported that shape B looked like “a cloud”). Thus, we concluded that we were successful in creating visual stimuli that were unfamiliar and unnameable. In the experimental task, 100% of the adults and 92% of the children (11/12) were classified as using visual cues to guess the cat’s location -- and indeed did so on every trial. The remaining child responded at chance. Because this test manipulated shape, and not color, these data are not directly comparable to the data garnered in our other experiments. Nonetheless, these results demonstrate that neither nameability nor familiarity was required in order for children to prefer visual cues.

The present data support three conclusions. First, children and adults exhibit an overarching preference for visual cues. We found no evidence for auditory dominance in preschool age children. When presented with roughly equivalent large contrasts in each modality, both children and adults preferred visual cues, and only one child in both experiments preferred auditory cues. Although it may not be possible to perfectly match the salience and discriminability of the auditory and visual modalities, our study is far more controlled than previous studies in both the selection of manipulated dimensions and the contrasts both within and between modalities. When presented with carefully matched cross-modal stimuli, participants preferred visual cues and did not exhibit the auditory dominance that is required by the SINC model. Perhaps most importantly, our data align more closely with the general findings dating back to some of the earliest experiments in the modern perception literature indicating that vision is the dominant modality for most non-infant humans (for a review of early evidence of visual dominance, see Posner, Nissen, & Klein, 1976).

Second, when presented with cross-modal stimuli, both children and adult observers flexibly employ the modality that possesses the greatest “distance” (contrast) between cues. The dimensions of familiarity and complexity used in previous studies may also elicit shifts in modality preference, but at present, these dimensions have not been sufficiently isolated from other confounding factors. This conclusion is consistent with Napolitano and Sloutsky’s (2004) argument that modality preferences are flexible. It differs, however, in two major respects: (1) We find no evidence for a default auditory preference—and indeed, our data are more consistent with a default visual preference. (2) In the present experiments, both sets of results are influenced by the relative contrast between stimuli in each modality, even when controlling for item familiarity. All other factors being equal, both children and adults appear to prefer cues that are more contrastive. Thus, the flexibility we have obtained does not rest on familiarity.

Finally, children are somewhat more likely than adults to employ auditory cues. This last result aligns with prior claims by Sloutsky and colleagues, yet the size and consistency of the effect is much smaller in our experiments. Furthermore, as noted earlier, this developmental difference is one of degree, rather than a qualitative shift in preference. In the one experiment finding developmental differences (Experiment 3), only one child was an Auditory Responder and most children were Visual Responders.

One of the central features of the SINC model is its framing of linguistic and auditory input as perceptual features of objects. Much of the predictive and explanatory power of this model relies on the assumption that children exhibit auditory dominance. By manipulating within-modality contrasts in cross-modal stimuli, we have garnered results that both replicate and refute the claims regarding auditory dominance in children. Based on the outcomes of the three experiments reported here, we conclude that the data supporting prior claims of auditory dominance in preschool age children resulted from the structure of the cross-modal stimuli employed in previous research, and do not reflect a stable or default preference for auditory stimuli by young children.

Although the stimuli and procedures employed in this study revealed that secondary elements of prior studies yielded a false preference for auditory cues, these studies are not ideal for evaluating the broader claims presented by Sloutsky and colleagues. For example, Sloutsky and colleagues suggest that auditory dominance creates automatic pulls on attention that result in the overshadowing of visual cues in young children. If children do not prefer auditory cues, then what aspects of attention and perception influence children to respond to cross-modal stimuli in a different manner than adults? Also, what is the source of the predictive power of the SINC model if its assumption of auditory dominance is incorrect? Future studies must focus on the role of attention in modality encoding and preferences, and carefully consider how innocuous experimental design choices may lead to illusory effects.

Researchers have tried to explain children's competencies in terms of their knowledge or in terms of their perceptual and associative capabilities, but similarity-based and theory-based approaches to understanding children's inductive behaviors have historically been incomplete or inaccurate without the observations, and sometimes objections, offered by the opposing approach. Rather than interpreting our findings as a dismissal of similarity-based approaches, we see our findings as an indication that models relying only on the influences of attention and perception to explain inductive reasoning present an incomplete picture of children's reasoning. Only by combining both similarity-based and theory-based approaches will true progress be made toward understanding children's induction and categorization behaviors.

Highlights

- We challenge prior claims that preschool-age children are “auditory dominant.”
- We demonstrate that within and across modality contrasts, and not overarching modality dominance, guides children's modality preferences.
- When presented with conflicting cross-modal cues, children attend to the modality with the largest contrast between cues.
- This finding rejects one of the central assumptions made by the prominent Similarity, Induction, Naming and Categorization (SINC) model, calling into question the validity of the model and questioning the appropriateness of strictly similarity-based approaches to understanding children's induction and categorization behaviors.
- We conclude that models of categorization and induction must account for both perceptual structure and intuitive theories.

Acknowledgments

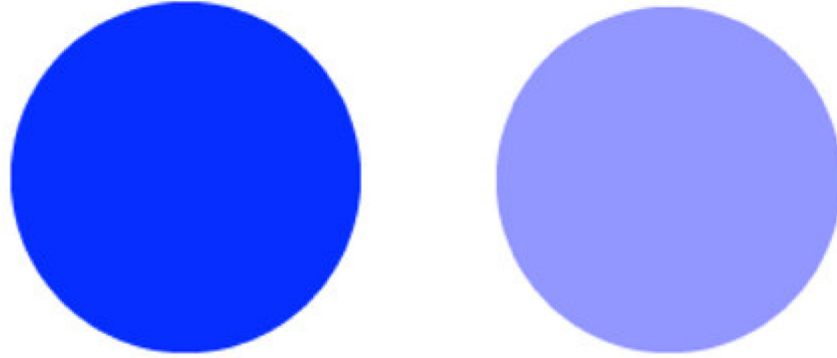
For assistance and helpful conversation and/or comments on earlier drafts, we thank Judith Danovitch, Margaret Evans, Meredith Myers, Liza Ware, Henry Wellman, and the members of the Language Lab at the University of Michigan. Thank you to the children, teachers, administrators, and staff of Annie's Children's Center, Childtime Children's Center, The Discovery Center, Little Folks Corner, Perry Nursery School, Play and Learn Children's Place, and Stony Creek Preschool Too. This work was supported by NICHD grant HD-36043 awarded to Gelman. We thank Nick Bauer-Levy, Gemma Bransdorfer, Liwen Chen, Taryn Konevich, Jessica McKeever, Meghan McMahon, Elana Mendelowitz, Miyuki Nishimura, Alex Pichurko, Kathryn Pytiak, Sarah Stilwell, and Alex Was, for research assistance.

References

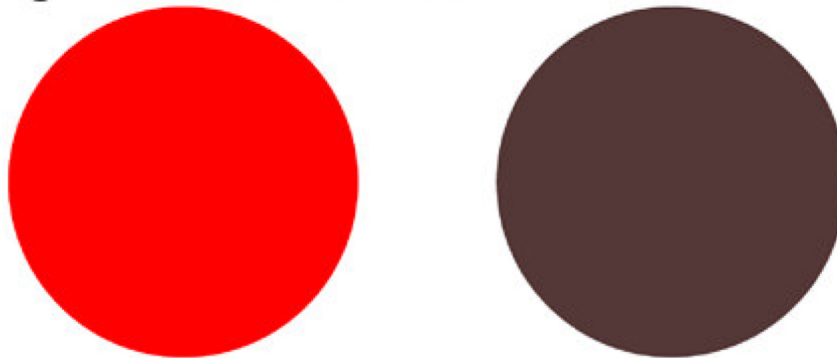
- Balaban MT, Waxman SR. Do words facilitate categorization in 9-month-old infants? *Journal of Experimental Child Psychology*. 1997; 64:3–26. [PubMed: 9126625]
- Carey, S. *The origin of concepts*. New York: Oxford University Press; 2009.
- Fulkerson AL, Waxman SR. Words (but not tones) facilitate object categorization: Evidence from 6- to 12-month-olds. *Cognition*. 2007; 105:218–228. [PubMed: 17064677]
- Garner, WR. *The processing of information and structure*. Maryland: Erlbaum; 1974.
- Gelman, SA. *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press; 2003.
- Gelman SA, Coley JD. The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*. 1990; 26:796–804.
- Gelman SA, Markman E. Categories and induction in young children. *Cognition*. 1986; 23:183–209. [PubMed: 3791915]
- Gelman SA, Markman E. Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*. 1987; 58:1532–1541. [PubMed: 3691200]
- Gelman SA, Waxman SR. Response to Sloutsky: Taking development seriously: Theories cannot emerge from associations alone. *Trends in Cognitive Sciences*. 2009; 13:332–333.
- Gelman SA, Wellman HM. Insides and essences: Early understanding of the obvious. *Cognition*. 1991; 38:213–244. [PubMed: 2060270]
- Goldstone RL. Effects of categorization on color perception. *Psychological Science*. 1995; 6:298–304.
- Gopnick A, Sobel DM. Detectingblickets: How young children use information about novel causal powers in categorization and induction. *Child Development*. 2000; 71:1205–1222. [PubMed: 11108092]
- Napolitano AC, Sloutsky VM. Is a picture worth a thousand words? The flexible nature of modality dominance in young children. *Child Development*. 2004; 75:1850–1870. [PubMed: 15566384]
- Posner MI, Nissen MJ, Klein RM. Visual dominance: An information-processing account of its origins and significance. *Psychological Review*. 1976; 83:157–171. [PubMed: 769017]
- Rakison DH. Infant's sensitivity to correlations between static and dynamic features in a category context. *Journal of Experimental Child Psychology*. 2004; 89:1–30. [PubMed: 15336916]
- Robinson CW, Sloutsky VM. Auditory dominance and its change in the course of development. *Child Development*. 2004; 75:1387–1401. [PubMed: 15369521]
- Sloutsky VM. Theories about 'theories': Where is the explanation? Comment on Waxman and Gelman. *Trends in Cognitive Sciences*. 2009; 13:331–332. [PubMed: 19646916]
- Sloutsky VM, Fisher AV. Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General*. 2004; 133:166–188. [PubMed: 15149249]
- Sloutsky VM, Fisher AV, Lo Y. How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development*. 2001; 72:1695–1709. [PubMed: 11768140]
- Sloutsky VM, Lo Y. How much does a shared label make things similar? Part 1. Linguistic labels and the development of similarity judgment. *Developmental psychology*. 1999; 35:1478–1492. [PubMed: 10563736]
- Sloutsky VM, Napolitano AC. Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development*. 2003; 74:822–833. [PubMed: 12795392]

- Sloutsky VM, Robinson CW. The role of words and sounds in infants' visual processing: From overshadowing to attentional tuning. *Cognitive Science*. 2008; 32:354–377.
- Stevens, SS. *Psychophysics: Introduction to its perceptual, neural, and social prospects*. New York: J. Wiley & Sons; 1975.
- Wellman, HM.; Gelman, SA. Knowledge acquisition in foundational domains. In: Damon, W.; Kuhn, D.; Siegler, RS., editors. *Handbook of Child Psychology: Cognition, Perception, and Language*. New York: J. Wiley & Sons; 1998.
- Werker JF, Cohen LB, Lloyd VL, Casasola M, Stager CL. Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*. 1998; 34:1289–1309. [PubMed: 9823513]
- Xu F. The role of language in acquiring kind concepts in infancy. *Cognition*. 2002; 85:223–250. [PubMed: 12169410]
- Xu F, Cote M, Baker A. Labeling guides object individuation in 12-month-old infants. *Psychological Science*. 2005; 16:372–377. [PubMed: 15869696]

A) Low Contrast Stimuli



B) High Contrast Stimuli



C) Novel Shapes



Figure 1.

A) Low contrast visual stimuli were manipulated on one dimension (saturation). B) High contrast visual stimuli were manipulated on two dimensions (brightness and saturation), and the size of the manipulation was larger. C) Novel, unnameable shapes.

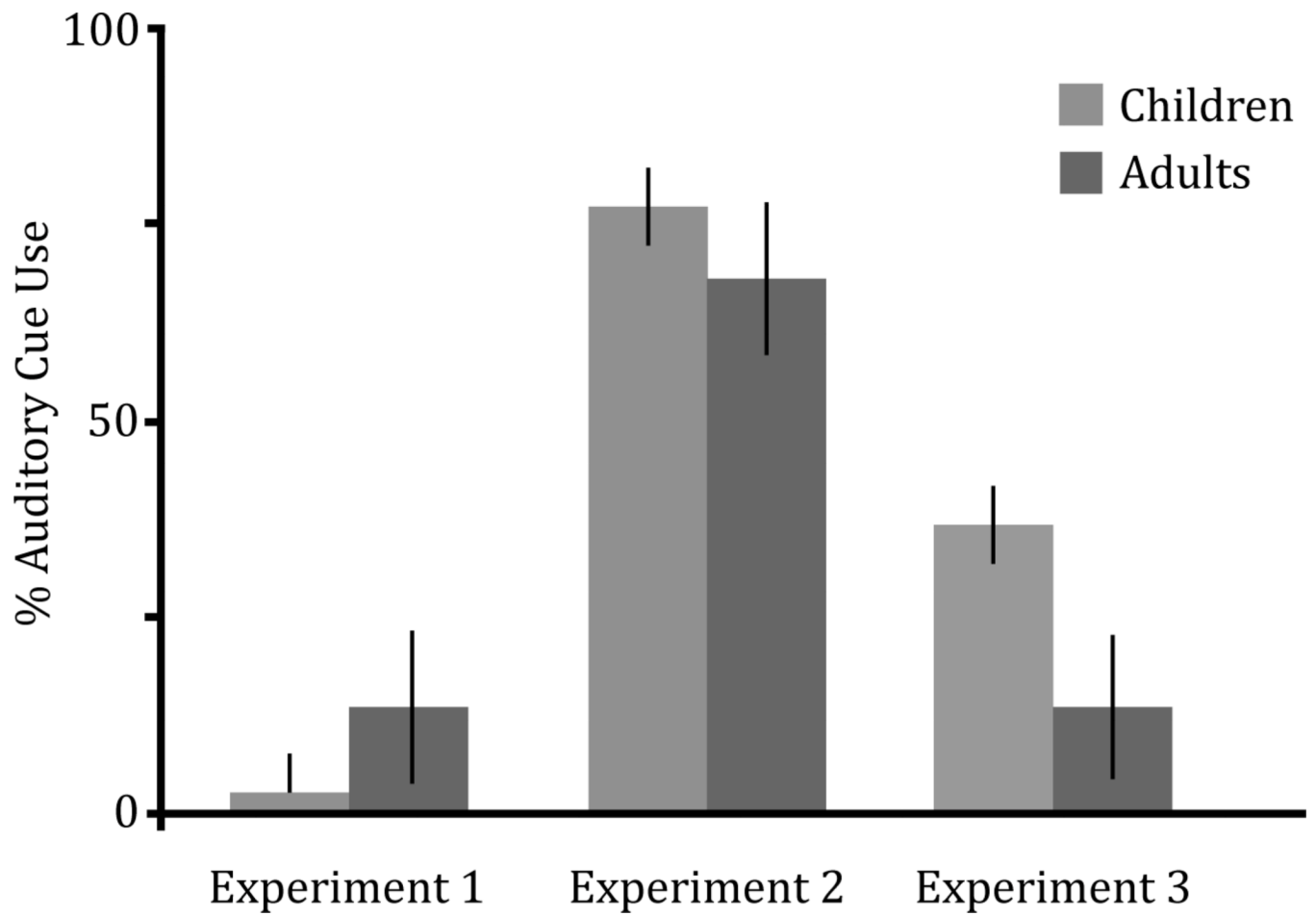


Figure 2. “% Auditory Cue Use” refers to the percentage of trials on which participants used auditory cues to predict target locations.

Table 1

Audio-Visual Dominance Data

Percentage of responder types as a function of experiment and age group. Participants who relied on auditory cues on 75% or more of test trials were designated as Auditory Responders; those who relied on visual cues on 75% or more of test trials were designated as Visual Responders. Participants who did not rely on a single modality were labeled Mixed Responders.

	CHILDREN			ADULTS		
	Auditory	Mixed	Visual	Auditory	Mixed	Visual
Experiment 1: High-Auditory High-Visual	0	6.25	93.75	12.50	0	87.50
Experiment 2: High-Auditory Low-Visual	66.66	33.34	0	68.75	0	31.25
Experiment 3: Low-Auditory Low-Visual	6.25	43.75	50.00	12.50	0	87.50