Received: 19 January 2022 | Revised: 18 August 2022 | Accepted: 22 August 2022

DOI: 10.1111/desc.13321

Developmental Science WILEY

RESEARCH ARTICLE

Hearing water temperature: Characterizing the development of nuanced perception of sound sources

Tanushree Agrawal | **Adena Schachner**

Department of Psychology, University of California, San Diego, California, USA

Correspondence
Tanushree Agrawal, Department of Psychology, University of California, San Diego, CA, USA.
Email: t2agrawa@ucsd.edu

Funding information
Center for Academic Research and Training in Anthropogeny

Abstract

Without conscious thought, listeners link events in the world to sounds they hear. We study one surprising example: Adults can judge the temperature of water simply from hearing it being poured. We test the development of the ability to hear water temperature, with the goal of informing developmental theories regarding the origins and cognitive bases of nuanced sound source judgments. We first confirmed that adults accurately distinguished the sounds of hot and cold water (pre-registered Experiments. 1, 2; total N = 384), even though many were unaware or uncertain of this ability. By contrast, children showed protracted development of this skill over the course of middle childhood (Experiments. 2, 3; total N = 178). In spite of accurately identifying other sounds and hot/cold images, older children (7–11 years) but not younger children (3–6 years) reliably distinguished the sounds of hot and cold water. Accuracy increased with age; 11-year old's performance was similar to adults. Adults also showed individual differences in accuracy that were predicted by their amount of prior relevant experience (Experiment 1). Experience may similarly play a role in children's performance; differences in auditory sensitivity and multimodal integration may also contribute to young children's failures. The ability to hear water temperature develops slowly over childhood, such that nuanced auditory information that is easily and quickly accessible to adults is not available to guide young children's behavior.

KEYWORDS
auditory event perception, crossmodal correspondence, development, experience, individual differences, sound source judgments

Highlights

1. Adults can make nuanced judgments from sound, including accurately judging the temperature of water from the sound of it being poured.
2. Children showed protracted development of this skill over the course of middle childhood, such that 7–11-year-olds reliably succeeded while 3–6-year-olds performed at chance.
3. Developmental changes may be due to experience (adults with greater relevant experience showed higher accuracy) and the development of multimodal integration and auditory sensitivity.

4. Young children may not detect subtle auditory information that adults easily perceive.

1 | INTRODUCTION

Without conscious effort, people make nuanced inferences about events in the world from sounds they hear (Gaver, 1993a, 1993b; Gibson, 1977; see Giordano, 2003 for a review). Subtle acoustic features shape our everyday actions and decisions, even outside of awareness (e.g., we stop pouring when a bottle sounds full; Cabe & Pittenger, 2000; Perfecto et al., 2018). People make causal judgments from sound, weighing alternate causal possibilities of sound sources (Ballas, 1993; Bronkhorst & Houtgast, 1999; Lemaitre & Heller, 2012), and linking acoustic input to a rich representation of events in the world (Cusimano et al., 2018; Gerstenberg et al., 2018; Traer et al., 2021). By adulthood, acoustic properties provide nuanced information about objects' shape, size, and texture (Grassi, 2005; Kunkler-Peck & Turvey, 2000; Lederman, 1979).

One category of sounds that people make unexpectedly accurate sound source judgments from is sounds produced by liquids (Spence & Wang, 2015). For example, people can use sound alone to judge the freshness of beverages (Roque et al., 2018), can hear subtle differences in the levels of carbonation in sparkling water (Zampini & Spence, 2005), and can hear wetness, such that a sound synchronously presented with tactile stimulation impacts whether the touch was felt as wet or dry (Jousmäki & Hari, 1998). People can also estimate relative differences in viscosity from the sounds of different types of milk being poured, and use this perceptual input to make useful inferences, such as which milk will make them feel fuller (Pellegrino et al., 2019). Last, most relevant to the current work, people can hear water temperature: Adults can accurately judge the temperature of water simply from the sound of it being poured (Velasco et al., 2013; for a demo, see https://osf.io/brp2a/?view_only=7f49783ebbf646b29af32ca645246f57). In order to judge temperature from the sounds of liquids, adults appear to use their sensitivity to subtle temperature-related changes in the pouring sound, caused by differences in the liquids' viscosity, the rate of air bubble vibrations, and the splash pattern at different temperatures (Gutmann & Simmons, 1952; Guyot et al., 2017; Parthasarathy & Chhapgar, 1955).

Judgments of temperature, wetness, or freshness from sound form a specific kind of sound source identification (Lutfi, 2008). In making these nuanced judgments, people are mapping the auditory stimulus to a multimodal concept or category (Murphy, 2004), often one with salient tactile features (e.g., wet, hot vs. cold). How do such nuanced auditory perception abilities develop? To what extent are these cross-modal mappings present from early infancy or birth, versus a product of protracted maturation and learning over childhood? For auditory perception of liquid materials, previous work has repeatedly raised the question of development, and the relative contributions of innate factors, maturation, and learning (Merrick & Filingeri, 2019; Spence, 2020). However, developmental data have not previously been brought to bear on this question.

Here we ask whether nuanced auditory perception of water temperature emerges early in life, or whether this ability emerges later in childhood, as a process of more protracted developmental change. These developmental questions are important: If children fail to detect subtle auditory information, this would have strong potential to impact everyday decisions, including both mundane activities, like avoiding spills from the sound of filling a bottle (Cabe & Pittenger, 2000), and also high-stakes decisions. For example, young children (in contrast with adults) show difficulty detecting and localizing an approaching vehicle using subtle auditory cues, with implications for their ability to cross the street safely (Barton et al., 2013; Pfeffer & Barnecutt, 1996). If young children also fail to hear water temperature, this would suggest a more general difficulty perceiving nuanced auditory information early in life.

1.1 | Reasons to predict early emergence

The ability to hear water temperature could plausibly be present from extremely early in infancy. This prediction is motivated in part by evidence of infants' early abilities to perceive and learn crossmodal correspondences, and to integrate information across senses. The ability to hear water temperature involves linking sounds to thermal features, which are typically perceived through the tactile modality (thermoreception; Zhang, 2015). As such, the ability may involve crossmodal correspondence, matching a sensory feature in one modality with a feature in another sensory modality (Spence, 2011).

Several developmental experiments have suggested that infants comprehend complex crossmodal correspondences extremely early in infancy. For example, infants less than 1 month old who had only felt the texture of an object were later able to recognize it through sight alone (Gottfried et al., 1977; Meltzoff & Borton, 1979; but see Maurer et al., 1999). Similarly, newborns were found to visually recognize the shape of an object they had held, but not seen; though this effect appeared only for the right hand (Streri & Gentaz, 2003, 2004). If these accounts are correct, then minimal experience and maturation may be required in order to recognize a feature (such as temperature) via a new or unusual modality (such as sound).

The ability to hear water temperature may also depend on multisensory integration, the ability to combine inputs from two or more senses (Stein et al., 2014). Human perception of another liquid quality, wetness, similarly occurs through multisensory integration of information from thermal and tactile input together with information from sound and vision (Merrick & Filingeri, 2019). Many aspects of multisensory integration develop in infancy (Bahrick & Lickliter, 2000; Kopp, 2014; Lewkowicz, 1992), and experience with crossmodal input in the first months of life is needed for typical development (Putzar et al., 2007; Wallace et al., 2004). Auditory information is also dominant in infants' intermodal processing in some contexts (Robinson & Sloutsky, 2010). Even in infancy, stimuli presented in one modality interfere with or facilitate the processing of stimuli in other modalities, showing that the modalities are not processed entirely independently (Robinson & Sloutsky, 2010). By 4 months of age, infants expect to hear a sound when they see an impact, such as a bouncing ball (Spelke, 1976, 1979), match speech sounds to a face moving in ways needed to produce that sound (Kuhl & Meltzoff, 1982), and map speech sounds to visual shapes with specific properties, in the same manner as adults (the "bouba/kiki" effect; Ozturk et al., 2013). Infants also link sounds to locations in space: Newborn infants visually orient to the location of a sound (Morrongiello et al., 1991, 1994), and by 4 months infants use such perceptual information to meaningfully guide their actions, such as pushing backward in response to looming sounds as if to avoid an approaching object (Freiberg et al., 2001; Neuhoff, 2018). If the ability to hear water temperature depends on these foundational multisensory integration abilities, the ability would be expected to emerge early in life, and be present in early childhood.

There is also strong potential for natural selection for nuanced auditory perception and comprehension, due to the adaptive value of auditory information. Auditory information provides a uniquely valuable source of information about the world: Properties of sound allow it to provide information even in darkness, through water and ground, around corners, and even when the eyes are closed or directed elsewhere (Horowitz, 2012). This adaptive value may have resulted in natural selection for a suite of abilities to make subtle judgments about the world from sound, which could result in earlier development.

The perception of liquids may hold specific adaptive value: For example, the ability to perceive wetness is thought to have evolved repeatedly in multiple species, due to its value when selecting appropriate habitats and ensuring physical health (Merrick & Filingeri, 2019). As the accurate manipulation of liquids at particular temperatures is crucial to fundamental human activities (e.g., cooking, washing), accurate perception of other liquid features may have adaptive value as well. The sounds of pouring of hot and cold liquids into containers have been part of the acoustic environment of humans since at least 20,000 years ago, with the use of pottery to cook, store, and pour liquids (Wu et al., 2012), and potentially earlier through the use of natural containers (dried bottle gourds as cups; e.g., Nwokolo, 1996). This timescale is long enough to see biological change through natural selection in response to these cultural innovations (Cochran & Harpending, 2009; Henrich & McElreath, 2007; Wrangham, 2009). For example, the relevant crossmodal associations may have been prepared by natural selection, making them faster and easier to learn (prepared learning; Dunlap & Stephens, 2014; Garcia & Koelling, 1966). Similarly, recent literature on crossmodal associations has found that some are more easily learned than others, particularly those that appear to be based on relationships consistently found in nature (Shams & Seitz, 2008). While testing this account is beyond the scope of the current paper, these evolutionary considerations motivate the prediction that hearing water temperature may emerge early in development.

1.2 | Reasons to predict later emergence

While it is therefore possible that the ability to hear water temperature could be present from early in life, other factors motivate the prediction that this ability may emerge much later in childhood. Some other nuanced sound source judgements develop late: Children aged 5–9 years show extremely poor detection and localization of approaching vehicles using sound (Barton et al., 2013; Pfeffer & Barnecutt, 1996), with performance improving somewhat by age 11 years (Pfeffer & Barnecutt, 1996).

Substantial relevant experience may be needed for more nuanced aspects of sound source identification. Theories of sound source identification have hypothesized that it develops through a protracted process of learning (e.g., Gaver, 1993a, 1993b). By these theories, people accumulate experience with various events (pouring, striking, splashing) involving various materials (liquids, metal, wood), and learn to associate the acoustic features of each event with other aspects of the relevant experiences. These learned associations may then form a foundation for rational inferences about the causes of sounds (e.g., Cusimano et al., 2018; Shams & Beierholm, 2010; Traer et al., 2021). There is substantial evidence that experience plays a role in sound source identification: Listeners can use auditory experience to build statistical summary representations of different sound sources (Dean et al., 2005; Młynarski & McDermott, 2018); and use these representations to categorize new sounds and infer the perceptual continuation of a sound that is interrupted by noise (McDermott et al., 2013; McWalter & McDermott, 2019). Adults are also able to form new associations across arbitrary crossmodal inputs with training (Ernst, 2007; Hidaka et al., 2015; Seitz et al., 2007; Spence, 2011). If this kind of associative learning of crossmodal correspondences is the basis for nuanced sound source judgements, then experience with the relevant stimuli, in the relevant sensory modalities, may be necessary. If so, children may not be able to detect water temperature from sounds until they have direct, multimodal experience of both hot and cold water pouring events.

Multimodal integration also continues to develop throughout childhood, in ways that could impact the ability to judge temperature from sound. In particular, prior to 8 years of age, children show notable failure to integrate multimodal input, across multiple tasks including shape discrimination, spatial localization, and detection of visual–auditory events (Barutchu et al., 2009; Gori et al., 2008; Nardini et al., 2008). For example, children below 8 years fail to integrate tactile information with information from other modalities (vision), with one or the other sense dominating decisions (Gori et al., 2008). Even within a single sensory modality (vision), children fail to integrate multiple cues, such that certain aspects of sensory integration (sensory fusion, optimal uncertainty reduction) are not present at 6 years of age and are not adultlike until approximately 12 years of age (Nardini et al., 2010). These data suggest that if hearing water temperature depends on having an integrated, multimodal representation of the pouring event, then children may fail until later childhood.

Last, it is possible that young children fail to hear water temperature because they cannot hear the relevant auditory features. Perception of water temperature from sound appears to depend on the relative loudness of frequencies in the 200 Hz and 5–6 kHz range: When the loudness of these frequencies was artificially manipulated, participants judged pouring sounds to involve warmer water when higher frequencies (5–6 kHz) were reduced, and lower frequencies (approx. 200 Hz) amplified. Similarly, participants judged pouring sounds to involve colder water when the 5–6 kHz range of frequencies were amplified, and the ~200 Hz range reduced (Velasco et al., 2013; see also Wang & Spence, 2017).

There is evidence of developmental change in children's auditory sensitivity over childhood, and particularly for perception of the frequency ranges relevant to hearing water temperature. Auditory processing is not fully adult-like until 10–12 years of age (Fior, 1972; Moore et al., 2011; Saffran et al., 2006; Stollman et al., 2004), and notable improvement in low frequency discrimination is seen from 4 to 6 years of age (Jensen & Neff, 1993; Maxon & Hochberg, 1982). Children's ability to perceive auditory frequencies may be particularly poor at the ranges necessary to hear water temperature: In a study of children aged 4–12 years, children's auditory sensitivity followed an inverted U-shaped curve, with poor sensitivity at low frequencies close to 200 Hz (250–500 Hz), highest sensitivity at middle frequencies, and lower sensitivity at high frequencies including the relevant 5–6 kHz range (4–8 kHz; Maxon & Hochberg, 1982). Children's auditory sensitivity also improved with age, with 12 year olds showing sensitivity similar to adults' (Maxon & Hochberg, 1982). If childrens' lower auditory sensitivity prevents them from hearing the relevant auditory information, children may not be able to hear water temperature until later in childhood, potentially between 6 and 12 years of age.

1.3 | The current studies

In this paper, we therefore test whether the ability to hear water temperature develops early in life, or has a protracted developmental time course. In doing so, we aim to constrain developmental theories regarding whether early-developing aspects of perception and cognition are sufficient to support nuanced auditory sound source judgements, versus whether extensive relevant experience and/or maturation is needed. We first replicate findings of adults' ability to hear water temperature, and test whether adults often lack awareness of this ability, as anecdotally reported. We ask whether adults' varied levels of experience with relevant events (hot and cold liquid pouring) predict their accuracy in the task (Experiment 1, pre-registered). Second, we characterize the development of this ability in childhood, testing children across a wide range of ages (3–11 years). This age range was selected to begin at an age at which children typically understand and produce words about temperature concepts, such as "hot" and "cold" (Luce & Callanan, 2020), and have many foundational abilities at multimodal integration and crossmodal correspondence (Lewkowicz, 1992; Meltzoff & Borton, 1979). The upper end of the age range extended to an age at which children have nearly adult-like auditory sensitivity (e.g., Maxon & Hochberg, 1982), improved multimodal integration (Gori et al., 2008), and improved performance at another nuanced sound source judgment (perception of moving vehicles; Pfeffer & Barnecutt, 1996). We replicate our developmental findings across two samples and methods (Experiments. 2, 3; Experiment 3 pre-registered).

If hearing water temperature requires only minimal experience and maturation, like more foundational aspects of sound source judgments, then children of all ages in this range should succeed at differentiating hot and cold pouring events from sound alone (comprehension checks ensured that children understood the task, and the concepts of hot and cold). If instead more protracted experience or maturation is needed, then young children may fail to distinguish the sound of hot versus cold water, succeeding only later in childhood. If so, this would inform developmental theories regarding the development of mature auditory event perception, and test the hypothesis that young children may not be sensitive to nuanced auditory information that adults easily detect.

Stimuli, methods, data, and supplemental methods and results for all experiments can be found in an OSF repository at the following link: https://osf.io/brp2a/?view_only=7f49783ebbf646b29af32ca645246f57.

2 | EXPERIMENT 1

2.1 | Method

2.1.1 | Participants

As preregistered (https://aspredicted.org/RJR_COQ), based on a power analysis (see OSF Supplement), N = 280 undergraduates at a large public university in Southern California participated in an online study in exchange for course credit (Mean age = 21.5 years; SD = 1.9 years; 212 female). Additional participants were tested but their data were excluded based on preregistered exclusion criteria: Having some form of hearing loss (n = 15); Self-reporting not using headphones to listen to the sounds, despite being instructed to do so (n = 17). The experiment received ethics approval from the Institutional Review Board at the university where the work was conducted (protocol 161172), and written consent was obtained from participants.

2.1.2 | Stimuli

All acoustic stimuli were professional recordings of water pouring events, as used in previous work (Velasco et al., 2013; provided by the first author). Each sound was a 5-second-long pouring event, of either hot or cold water being poured into a cup (Figure 1). The only variable manipulated was the temperature of the water, with the pitcher, cup, height/speed of pouring, and quantity of water kept the same. Hot water was 180–183°F, and cold water was 43–46°F. The recordings involved pouring the same amount of water (20 ml) from the same

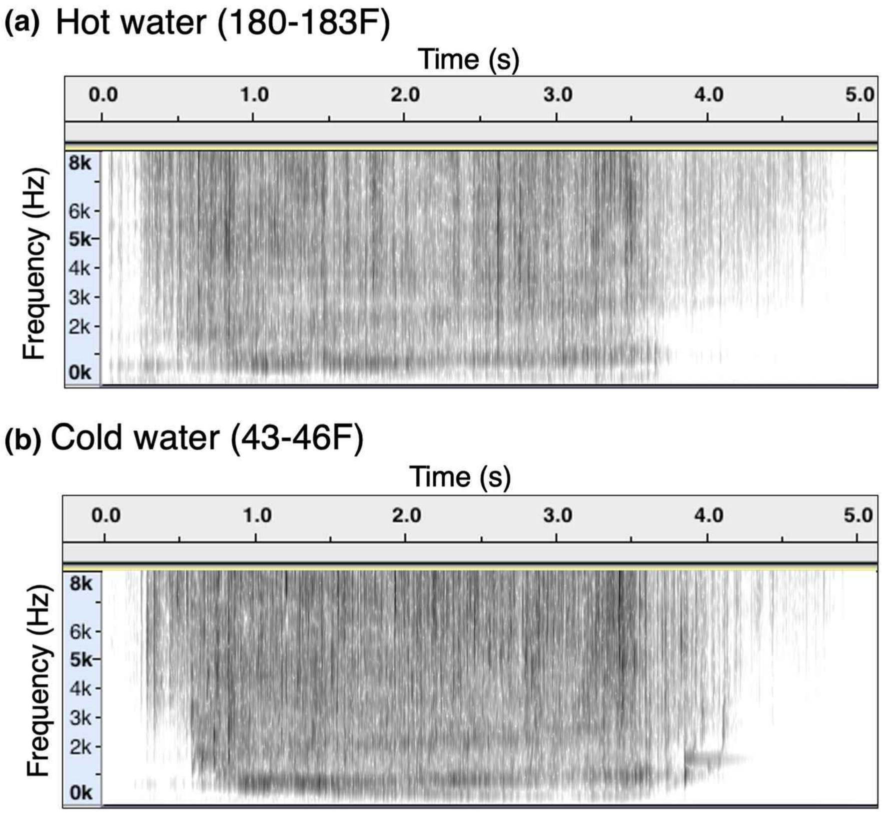

FIGURE 1 Sample spectrogram of acoustic stimuli. Spectrograms of hot (a) versus cold (b) water being poured from the same height at the same rate into the same glass cup. The x-axis represents time in seconds; y-axis is acoustic frequency in Hz. Darker parts of the spectrogram mean higher energy densities, lighter parts mean lower energy densities.

height (10 cm) at the same flow rate (40 ml/s). One pair of sounds (hot, cold) was created for four different cups, each of a different material (paper, plastic, porcelain, and glass). This resulted in a total of eight stimulus items, or four pairs, presented across eight unique trials.

2.1.3 | Design

As preregistered, participants each completed 32 trials (eight unique trials, four times each); this number of trials was selected based on previous work on adults' similar judgements (Velasco et al., 2013; Experiment 1). On each trial, participants heard one pouring sound (either hot or cold water being poured into one of four cups), and were asked to identify whether the sound was hot or cold water using a forced-choice radio button format. The order of each question (whether "hot" or "cold" was queried first) was counterbalanced across participants and the overall order of trials was randomized. Statistical analyses were conducted at an alpha level of 0.05.

2.1.4 | Procedure

Participants completed the task online, in their own homes, on a computer web-browser (mobile devices were not allowed) using Qualtrics survey software. They were instructed to use headphones and do the study in a quiet room. Participants were told that they would be judging whether water being poured was hot or cold, from the sound of it being poured. Before the task, participants judged whether they expected to be able to hear water temperature, on a 5-point confidence scale from −2 to +2 ("Very confident that I *cannot* do this task" to "Very confident that I *can* do this task"), and completed a brief sound check.

Each trial was then presented on a separate page, and consisted of a prompt to listen to a sound, followed by a question: "Was the sound hot or cold water?" Participants could listen to the sound as many times as they liked before making their choice and clicking to move on to the next trial.

After completing all 32 trials, participants filled out a survey measuring their levels of experience with relevant sounds. Items were based on previously established measures of beverage intake (Hedrick et al., 2010), modified to ask specifically about acoustic experience. Items were: "Number of [hot/cold] drinks you hear being prepared/made per week"; "Number of [hot/cold] drinks you consume per week"; "Do you enjoy [hot/cold] beverages?"; and "Do you consider yourself to be someone who never/rarely drinks [hot/cold] drinks?" These were our main quantitative measures of adults' past experience with relevant pouring sounds.

Participants then answered two free-response questions about the listening strategies they used during the task, and about their past

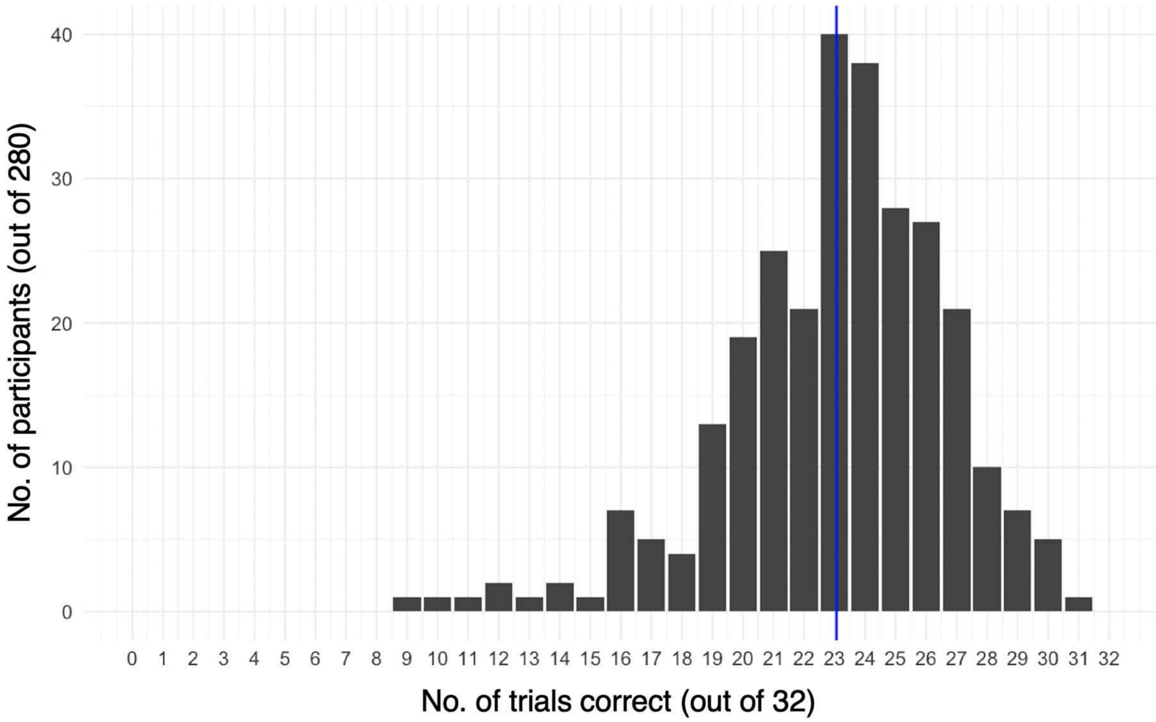

FIGURE 2 Histogram of adult accuracy in hearing water temperature (Experiment 1). Participants were significantly better than chance at judging the temperature of water from the sound of it being poured. The x-axis shows the number of correct responses, out of 32 trials per participant; chance is 16/32. Each trial required participants to identify whether a single sound stimulus was hot versus cold water. The blue vertical line shows the mean accuracy.

experiences with the sounds of hot/cold liquids being poured. Qualitative data from these questions were not formally analyzed, but were examined by the researcher to provide context for the quantitative measures above. Last, participants were asked if they used headphones, and were told that their answer would not affect their participation credit (to encourage honesty). The study concluded with a brief demographics section, followed by a debriefing if requested.

2.2 | Results

Accuracy on this task was high: Participants were correct on 72.10% of trials (23/32; SD = 3.66), and significantly above chance ($t(279) = 32.3$, $p < 0.001$, Cohen's $d = 1.93$, t-test vs. 50%; Figure 2), in line with previous findings (accuracy of 72.5%, Velasco et al., 2013). This high accuracy contrasted with participants' predictions before starting the task: 34.6% of participants were not confident in their abilities, and a further 11.4% were confident that they would *fail* at the task. These individuals still had high accuracy, statistically equivalent to the rest of the sample (Not confident: M = 70.6%, SD = 4.21, $t(157.79) = -1.60$, $p = 0.11$; Confident in failing: M = 70.7%, SD = 3.00, $t(46.60) = -1.35$, $p = 0.18$; two sample two-tailed *t* test vs. participants who were at least somewhat confident in their abilities).

As pre-registered, we modeled the data using a logistic regression. The full model predicted accuracy on each trial ("correct" vs. "incorrect") with the predictors of cup material (paper, plastic, porcelain, glass) as a fixed effect, and both trial number (1–32) and subject as random effects. Nested model comparisons revealed a significant effect of cup material, such that stimuli created with certain cup materials were more challenging than others ($\chi^2(3) = 347.0$, $p < 0.001$; Mean proportion correct: Plastic 59.8%, $SD = 17.4\%$; Paper 70.0%, $SD = 19.4\%$; Porcelain 74.9%, $SD = 20.2\%$; Glass 83.6%, $SD = 17.3\%$). There was no significant effect of trial number; that is, performance did not significantly change over the course of the session ($\chi^2(1) = 3.46$, $p = 0.063$). There were individual differences in performance ($\chi^2(1) = 109.7$, $p < 0.001$), such that some participants were more accurate than others.

In a post-hoc exploratory analysis, we asked whether performance on this task differed based on how much experience adult participants had with the sounds of hot and cold liquids being poured. We aimed to treat the amount of relevant experience as a continuous measure, rather than binning participants into groups (which would likely decrease power, see below). We found that both of our continuous measures of relevant experience significantly predicted accuracy at hearing water temperature, with a small effect size: Participant accuracy was positively correlated with the total number of hot and cold drinks participants consumed per week (Pearson $r = 0.163$, $p = 0.006$), as well as the number of times they heard hot or cold liquids being poured per week ($r = 0.220$, $p < 0.001$).

In our preregistration, we also mentioned a further exploratory analysis, which we therefore include here. For this analysis, we binned participants into low-experience and typical-experience groups based on a binary measure: Whether participants identified as rarely/never drinking hot or cold beverages; or not (see OSF Supplement for details). The low-experience participants performed marginally worse than the rest of the sample (Typical-experience group, n = 168: Mean trials correct = 23.41/32, SD = 3.53; Low-experience group, n = 112: Mean trials correct = 22.54/32, SD = 3.80; $t(226.1) = 1.921$, $p = 0.056$, Cohen's $d = 0.238$, two-tailed two sample *t*-test). However, even within each of the two bins, subjects substantially differed in the extent to which they had relevant experience. This was apparent in continuous measures: For both the typical- and low-experience groups, the number of times participants heard hot or cold liquids poured per week included the entire range of possible answers, and 17.9% of people in the low-experience group still heard hot/cold drinks being poured more than twice a day. Because this binned analysis does not account for these differences in experience within each group, it likely has lower statistical power than the analyses of continuous measures (above).

3 | EXPERIMENT 2

Data from Experiment 1 thus show that adults can hear water temperature, in spite of being often unaware or uncertain of this ability; and suggest a possible role of experience with relevant acoustic events on adult accuracy. While the effect of experience was relatively small, this may result from a ceiling effect: By adulthood, everyone has likely accumulated substantial experience with relevant pouring sounds. We next aimed to test children, a population with substantially lesser experience and maturation. To allow for detection of either early-developing or later-developing abilities, we tested children across a wide range of ages (3–11 years; see Introduction for predictions motivating this range). We replicated our findings across two samples and methods (Experiments 2, 3).

In order to make our task feasible for both younger and older children, we switched from an identification task with only one stimulus (as in Experiment 1) to an easier two-alternative forced-choice task (MacMillan & Creelman, 2005). To compare children's performance to adults' performance on the same task, we collected an additional adult sample using this new method, pre-registered at: http://aspredicted.org/blind.php?x=pp29ys. We also ensured that all children in our samples understood the concepts of "hot" and "cold" (i.e., accurately identified images of hot and cold stimuli), and understood the task itself (i.e., accurately identified other, non-pouring sounds using the same method).

3.1 | Method

3.1.1 | Participants

N = 113 children from the San Diego metro area and N = 104 undergraduate students from a large public university in Southern California participated. Children were recruited from a database of local families interested in research, and from local preschools. Children ranged from age 3–11 years (3 years, 0 months—11 years, 7 months; Mean age = 5.83 years; 46 female), and were tested by convenience within this range (see Figure 5 for age distribution). An additional 23 children were tested, but their data were excluded according to a priori exclusion criteria: Answering one or more pre-test questions incorrectly (n = 22; Mean age = 4.17 years); having participated previously (n = 1).

In our adult sample, N = 104 undergraduate students (Mean age = 21.30 years; SD = 2.32 years; 75 female) participated in exchange for course credit. Five additional participants were tested but their data were excluded based on pre-registered criteria: Answering before or without listening to the audio stimuli (n = 3); technical difficulties in presentation (n = 2).

The experiments received ethics approval from the Institutional Review Board (protocols 161173 and 161272). Written consent was obtained from adult participants and from children's parents/guardians, and verbal assent was obtained from children.

3.1.2 | Stimuli

Acoustic stimuli were the sounds of hot and cold water being poured, from Experiment 1. Adults heard all four pairs of sounds (the different cup materials). Children performed one trial, with one pair of stimuli: the sounds of hot/cold water poured into a glass cup. Additional pre-test trials with children consisted of two animal sounds (a cow mooing, a dog barking); and cartoon images of hot and cold scenes (a sunny desert; a snowy scene) and hot and cold drinks (iced lemonade; a steaming mug; See Figure 3).

3.1.3 | Design

Each child completed one test trial. The task was shortened to a single trial to minimize inattention: This was the first developmental study with these stimuli, and we were uncertain how engaged children would be in a longer version of the task. The order of presentation of hot/cold sounds (and the right/left location of the corresponding icons on the screen) was counterbalanced across participants. During pre-test trials, the question order, order of sounds, and position of images on-screen was counterbalanced across participants.

Adult participants each completed 32 trials (four unique trials, eight times each). The order of trials (i.e., the cup materials) was randomized, except for the first four trials which occurred in a counterbalanced order (thus, the first four trials always included all four cup materials). The order of questions (i.e., whether "hot" or "cold" was queried first) was counterbalanced across participants. The order of the sounds (i.e., whether hot or cold was played first) was randomized.

For both adults and children, sounds were presented over headphones, and the experimenter was unaware of which sound was the correct answer.

3.1.4 | Procedure

Children were tested individually in the lab or at local preschools, in a quiet room seated at a child-size table across from an experimenter.

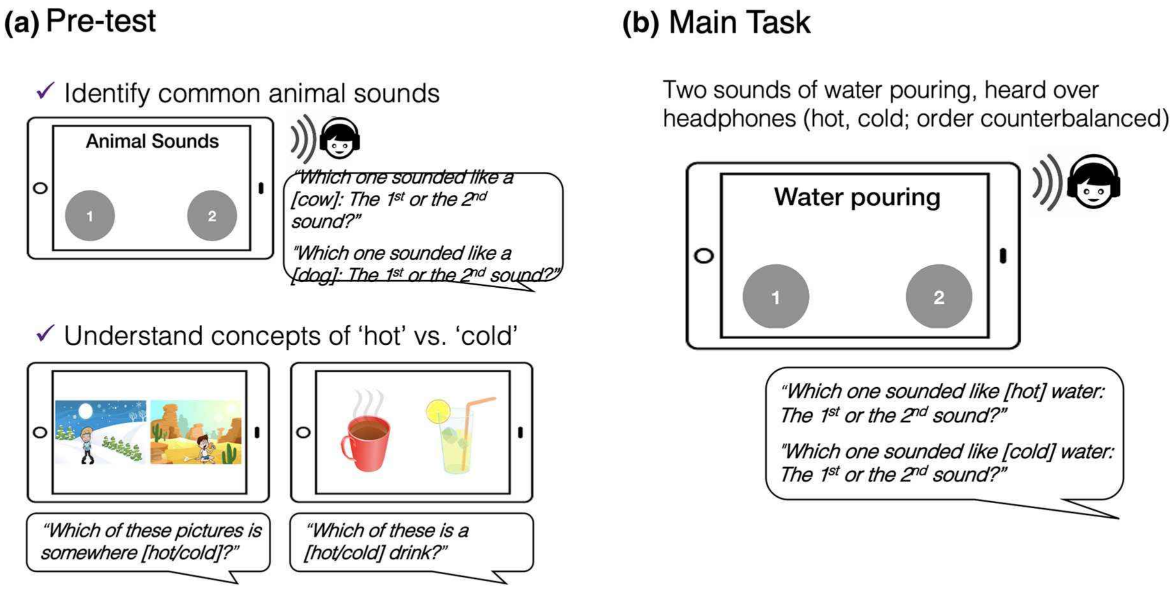

FIGURE 3 Overview of study procedure used with children (Experiment 2) (a) Participants first completed a pretest where they identified common animal sounds using the same iPad interface as the main task (to confirm that they could use this interface to identify sounds), and identified hot and cold images (to confirm that they understood the concepts of "hot" and "cold"). (b) Next, in the main task, they completed one trial involving hearing two sounds (one of hot water pouring and the other of cold water pouring; order counterbalanced) and then were asked to identify which sound was hot water and which was cold water, after both sounds had been presented (verbally or by pointing).

Sounds were presented via high quality headphones (MPOW 059 Bluetooth Headphones). The study was conducted on an iPad, with stimuli presented using Keynote, controlled by the experimenter via a remote control (with touchscreen disabled to prevent children from accidentally advancing the slides).

Three pre-test trials tested whether children understood the task, could identify other (non-pouring) sounds, and recognized concepts of hot and cold (see Figure 3a). Children were first asked to identify common animal sounds, using a similar procedure to the main task. The screen displayed two circles numbered 1 and 2; each corresponded to a sound by pulsing slightly as the sound played. Children listened to two animal sounds and were asked two questions: "Which one sounded like a [cow/dog], the first one or the second one?" Participants were then presented with two pictures of hot and cold scenes side by side, and asked: "Which of these pictures is somewhere [hot/cold]?" Last, participants were shown two pictures of hot and cold beverages, and asked "Which of these is a [cold/hot] drink?"

Children were then instructed that they would hear two sounds, one of hot water being poured and the other cold water, using the same format as the pre-test trials (see Figure 3b). After hearing both sounds, children were asked if they wanted to hear them again, and stimuli were replayed if requested. Children were then asked two questions: "Which one sounded like [hot/cold] water, the first sound or the second sound?"

Adult participants completed the task in the lab in a web-browser on an iMac desktop computer, using high-quality headphones (Bose QuietComfort 25) and Qualtrics survey software. Participants were instructed that they would be judging whether water being poured was hot or cold, from the sound of it being poured; and that every trial would include one hot and one cold sound. Each trial was presented on a separate page, and consisted of a prompt to listen to each of two sounds, followed by two questions: "Which of the sounds was hot water? (1, 2)" and "Which of the sounds was cold water? (1, 2)." Participants could listen to the clips as many times as they liked before making their choices, and clicking to move on to the next trial. All other aspects of the procedure with adults remained the same as in Experiment 1.

3.2 | Results

To test for developmental change in children's accuracy in hearing water temperature, we conducted a logistic regression, predicting accuracy based on age (Figure 4). Participants' age significantly predicted their accuracy in identifying the sound of hot and cold water (Nested logistic model comparisons, Likelihood ratio test, $\chi^2(1) = 12.91$, $p < 0.001$), such that the odds of succeeding on the task increased by 1.51 ($p = 0.0016$) for every year of age. Notably, children aged 5 and under performed at chance, in spite of our testing a substantial sample of 4- and 5-year-olds (79 children; 36/79 or 45.7% children correct, $p = 0.499$). In contrast, 85% of children age 6 and older answered correctly (29/34 children age 6+, $p < 0.001$, Cohen's $h = 0.784$; Figures 4 and 5). Crucially, younger children's failure was not due to lack of understanding of the task: To be included in the final sample, children had to successfully use a similar method to distinguish other sounds, and to categorize visual images of beverages and scenes as hot versus cold.

Adult participants were accurate on 93.0% of trials (29.8/32, SD = 2.78), performing significantly above chance, $t(103) = 107.4$, $p < 0.001$, Cohen's $d = 10.53$, t-test versus chance of 50%. For the

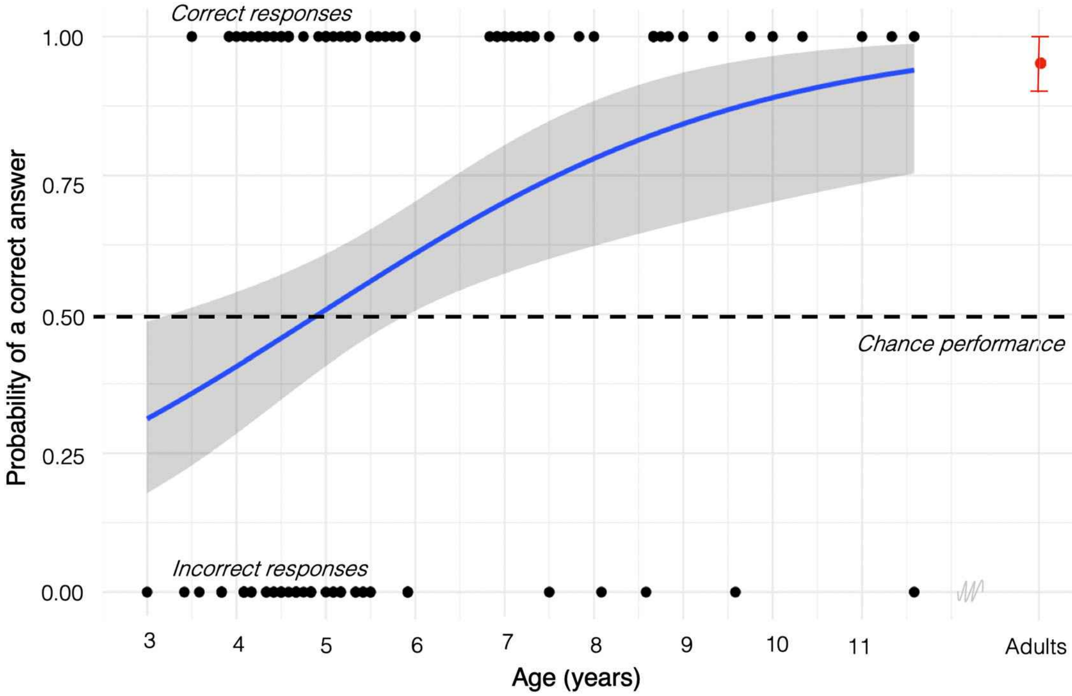

FIGURE 4 Developmental change in children's ability to hear water temperature (Experiment 2). Each dot represents one child's response, which was either correct (top of the figure) or incorrect (bottom of the figure). The blue line indicates the predicted probability of a correct response as a function of age (logistic regression; gray represents 95% confidence intervals). The red dot with its corresponding standard error bars represents adults' percent correct. Chance performance is represented by the dashed black line.

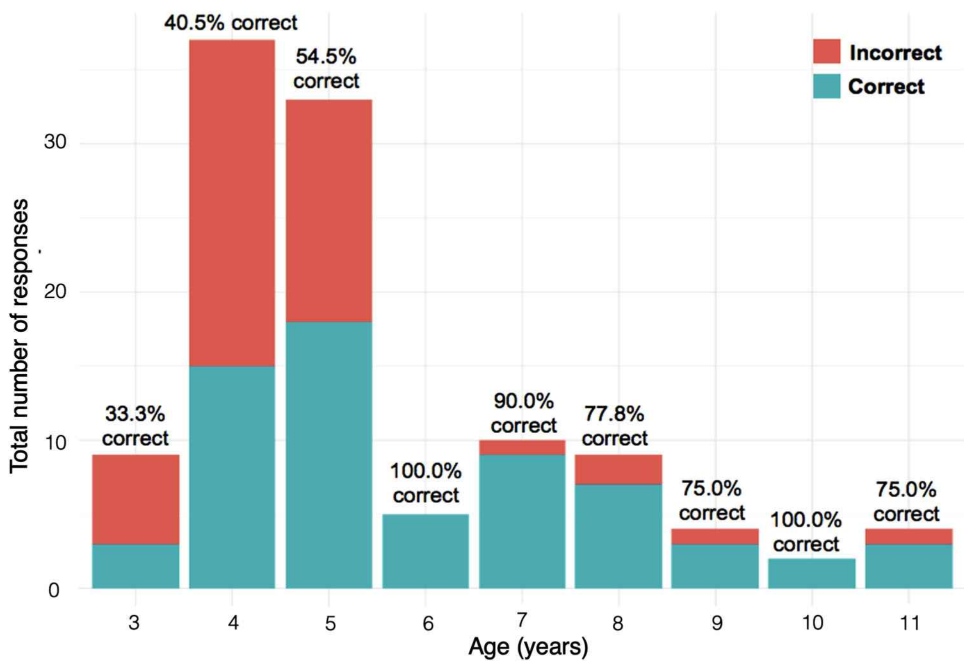

FIGURE 5 Accuracy and sample size by year of age (Experiment 2). Children under the age of 6 performed at chance, while older children succeeded. The overall height of each bar shows the number of children tested in each age group; 4 and 5 year-olds performed at chance in spite of a large sample of participants tested at these ages.

purposes of comparison to the child data, we calculated adults' performance on their first trial with the glass cup stimuli (since children completed one trial with this stimulus pair). 95.2% (99/104) of adult participants succeeded at this trial, an accuracy similar to the predicted accuracy (from the logistic fit) of the oldest children tested (11-year-olds, predicted accuracy 92.4%; see Figure 4). We note that adult accuracy on this task was higher than in Experiment 1. This is as expected, due the change to a two-alternative forced choice discrimination task (MacMillan & Creelman, 2005).

4 | EXPERIMENT 3

Data from Experiment 2 suggest that unlike many foundational aspects of auditory event perception and multimodal integration, the ability to judge the temperature of liquids from sound develops in a protracted manner, emerging in middle childhood. Here, we aimed to replicate and extend this finding in a pre-registered study. We observed that adults in Experiment 2 tended to mark a preliminary answer after hearing the first sound, potentially lowering working memory load. To test whether this task difference explained children's lower performance, here we similarly allowed children to mark preliminary answers. We conducted four trials per child, using four unique stimulus pairs (pouring into four different cup materials) in order to increase power and test whether children's abilities were due to any unique feature of the single stimulus pair used in Experiment 2. Last, we asked whether 6-year-old children would again be the first to succeed; in Experiment 2, the sample had few 6-year-old participants, leaving some uncertainty regarding the first age of robust success.[1]

Our methods and analyses were pre-registered at: https://aspredicted.org/blind.php?x=/MOB_XPW. As a result of University-wide and state-mandated COVID-19 protocols, we were forced to deviate from our pre-registered data collection plan, by stopping data collection with a partial sample of 4- to 6-year-old children. Our pre-registration specified that if 6-year-old children performed at chance, data from older children would be collected. During in-person data collection, a convenience sample of older children (7–11 years) had also been tested (see OSF Supplement). We therefore report the data from this new sample of older children, in addition to the pre-registered partial sample of 4- to 6-year-old children. We expected to observe developmental change (as in Experiment 2), such that older children but not younger children would succeed at hearing water temperature.

4.1 | Method

4.1.1 | Participants

N = 65 children from the San Diego metro area participated in the lab or at local preschools. Children ranged from age 4 to 11 years (Pre-registered sample of 4- to 6-year-olds: N = 23, Mean age = 5.71 years, Range = 4 years 5 months–6 years 8 months, 13 female; n = 4 4-year-olds, n = 11 5-year-olds, n = 8 6-year-olds; Additional older sample: N = 42, Mean age = 9.16 years, Range = 7 years 1 month–11 years 11 months, 22 female; n = 12 7-year-olds, n = 8 8-year-olds, n = 10 9-year-olds, n = 5 10-year-olds, n = 7 11-year-olds). An additional 14 participants were tested, but their data were excluded according to a priori exclusion criteria: Answering one or more pre-test questions incorrectly (n = 12, Mean age = 6.66 years); Technical difficulties (n = 1); Having some form of hearing loss (n = 1). The experiment received ethics approval from the Institutional Review Board (protocol 161173). Written consent was obtained from children's parents/guardians, and verbal assent was obtained from children.

4.1.2 | Stimuli

As pre-registered, acoustic stimuli were the sounds of hot and cold water being poured into glass, plastic, paper, and porcelain cups, as in Experiment 1. Stimuli for pre-test trials were the same as in Experiment 2, with one additional pre-test trial depicting cartoon images of hot and cold food (See Figure 6a) added to give children more practice with the method.

4.1.3 | Design

Each participant completed four test trials, with the order of presentation of hot/cold sounds counterbalanced across participants, and the order of presentation of answer choices in the question (hot/cold) also counterbalanced across participants. This number of trials allowed for one trial for each unique pair of acoustic stimuli, while still maintaining a short duration (allowing the task to be feasible for young children). Participants also completed four pre-test trials with other stimuli, during which the order of sounds and position of images was also counterbalanced across participants. The study session lasted approximately 4–8 min.

4.1.4 | Procedure

The procedure and materials used (Figure 6) were identical to those in Experiment 2, except for the addition of tokens, which were used to indicate answers (added to reduce memory load, by allowing children to mark an intended answer while listening; see OSF Supplement). Participants were given two tokens: a blue, "COLD" token with a cartoon picture of an icy cloud wrapped in a scarf, and a red, "HOT" token with a cartoon picture of the sun sweating in front of a fan. Participants were told that they would be deciding which of two things was hot and which was cold, and that they would use the tokens to answer. They were given a demonstration on how to answer, and told: "There's a red token that means Hot, and there's a blue token, which means Cold. So, if you think something is Hot, you put this Red one on top of it, like this; And if you think something is Cold, you put down this Blue one on top of it." Participants then used the tokens as part of the pre-test and test trials.

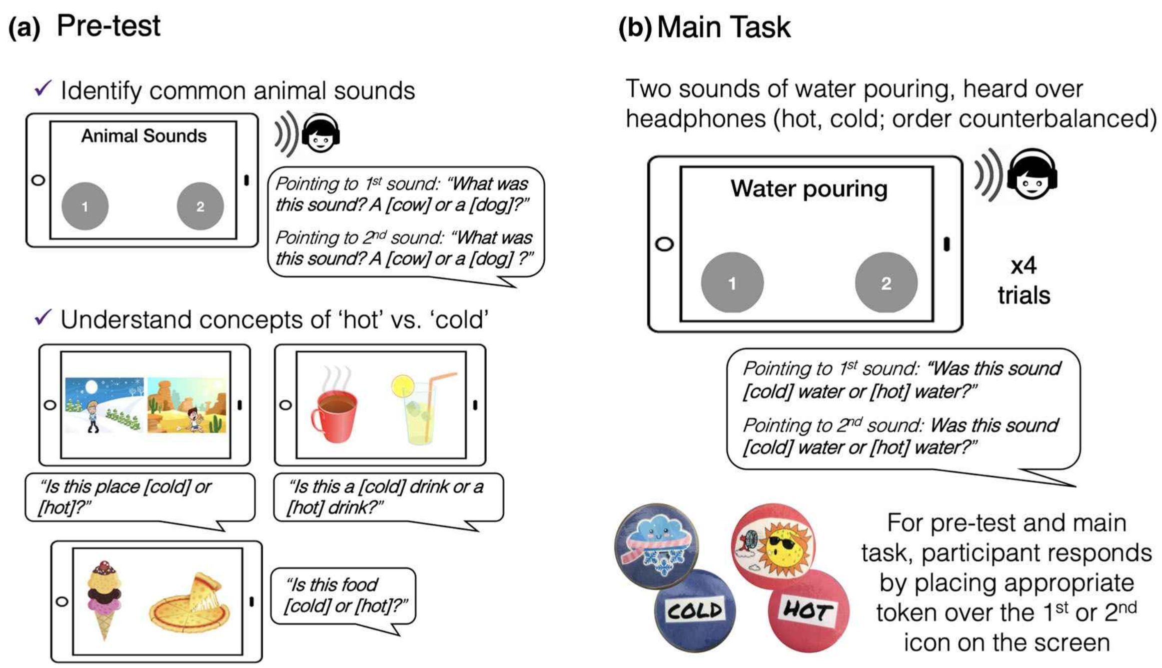

FIGURE 6 Overview of study procedure (Experiment 3). (a) As in Experiment 2, children first completed a series of pretest trials where they identified hot and cold images, and common animal sounds using the same iPad interface as the main task. (b) In the main task, they completed four trials, each involving a different pair of pouring sounds. Each trial involved hearing the two sounds (one of hot water pouring and the other of cold water pouring; order counterbalanced) and identifying which sound was hot versus cold water by placing the appropriate "hot" or "cold" token on top of the first or second sound icon on the screen. Participants were encouraged to do this at any time during stimuli presentation.

For each of the test trials, as in Experiment 2, children saw two circles corresponding to the sound of either hot or cold water being poured. They were asked to identify whether each sound was hot or cold water ("Was this sound [cold] water, or [hot] water?"), by placing the appropriate tokens on top of the circles at any time while listening to the two sounds. Verbal answers or pointing rarely occurred; in this case, children were given a prompt to use the tokens to indicate their answers instead.

4.2 | Results

As pre-registered, we first analyzed data from 4-, 5-, and 6-year-olds. A logistic regression, predicting children's accuracy based on age as a fixed effect, and cup material, trial order, and participant as random effects revealed no significant effects. This group of participants, overall, performed at chance levels (Mean score = 2.30/4 trials correct, SD = 1.11; $t(22) = 1.32$, $p = 0.20$; two-tailed t-test; 4-year-olds: M = 2.25/4; 5-year-olds: M = 2.36/4; 6-year-olds: M = 2.25/4).

Following the intention of our pre-registration to test older children in this case, we next analyzed data from children at all ages tested, including those tested due to convenience from 7–11 years of age (Figure 7). With this broader range of ages, participants' age significantly predicted their accuracy in identifying the sound of hot and cold water (Nested logistic model comparisons, Likelihood ratio test, $\chi^2(1) = 4.86$, $p = 0.028$), such that for every additional year of age, the odds of succeeding at the task increased by 1.233($p = 0.032$). There were no item effects of the four different stimuli pairs ($\chi^2(1) = 2.13$, $p = 0.145$) nor effects of trial order ($\chi^2(1) = 0$, $p = 1$). There were significant individual differences in performance across children even after accounting for effects of age ($\chi^2(1) = 7.92$, $p = 0.005$).

Overall, children aged 6 and under performed at chance (n = 23, Mean score = 2.30/4, SD = 1.11, $t(22) = 1.32$, $p = 0.20$, compared to chance score of 2). Again, this failure was not due to lack of understanding of the task or of the concepts of "hot" and "cold"; all children in the sample accurately identified other hot and cold stimuli, and other sounds, using the same method. In contrast, children aged 7 and older performed significantly above chance (n = 42, Mean score = 2.93/4, SD = 1.16, $t(41) = 5.21$, $p < 0.001$, Cohen's $d = 0.803$).

5 | GENERAL DISCUSSION

For adults, acoustic features provide rich and nuanced information about objects and events in the world, including multiple properties of liquids (Giordano, 2003; Spence & Wang, 2015). Here we find that the ability to hear water temperature, which is robust in adults, is not present in young children. Instead, this ability appears only in middle-childhood, and develops in a protracted and gradual manner over the first decade of life. Across two samples, adults and older children (7–11 years) but not younger children (3–6 years) reliably distinguished hot and cold water from the sound of pouring alone. Accuracy increased with age, such that only the oldest children's accuracy matched that of adults. Children's abilities generalized across

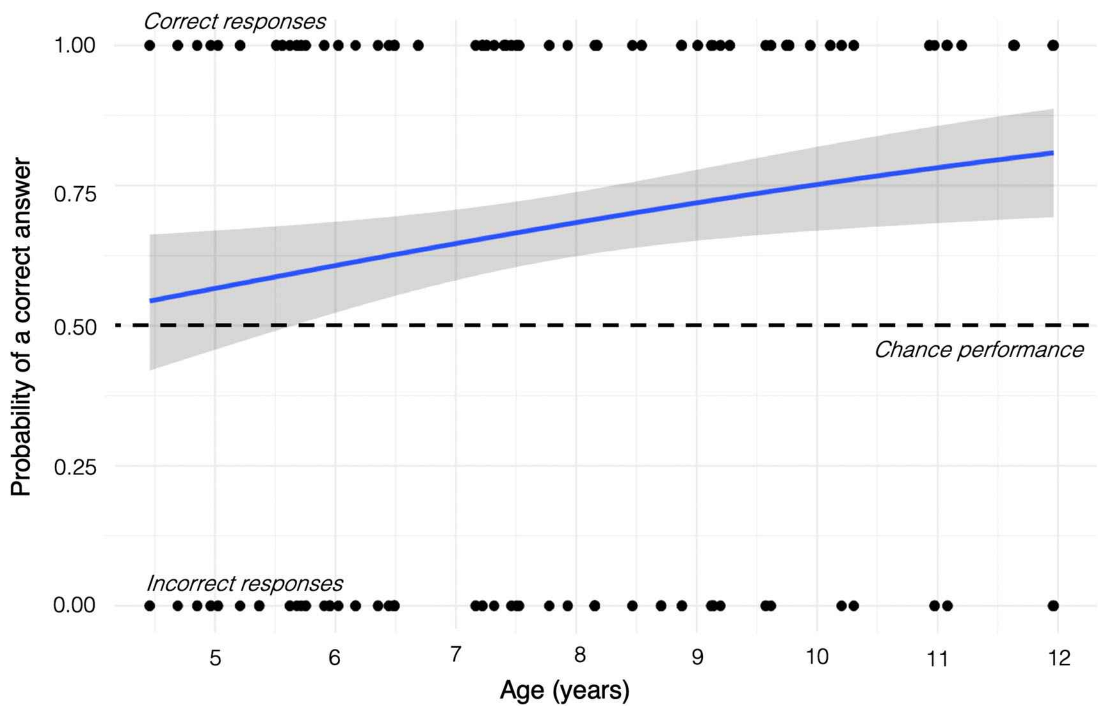

FIGURE 7 Developmental change in children's ability to hear water temperature (Experiment 3). Each dot represents an individual response (four responses per participant), which was either correct (on the top) or incorrect (on the bottom). The blue line indicates the probability of a correct response as a function of age (logistic regression; gray represents 95% confidence intervals). Chance performance is represented by the dashed black line.

multiple acoustic stimuli (pours into different material cups); and replicated across modifications of the method to minimize working memory load. Young children's failure was specific: These children accurately distinguished other sounds and hot/cold images in the same task. We also found evidence of individual differences; at age 6 in particular, some children succeeded (Experiment 2) but the ability was not yet robust (Experiment 3). Last, we found individual differences in adult abilities based on the amount of relevant experience (Experiment 1), suggesting that age-related differences in children may be partially driven by differing amounts of relevant experience.

Children are likely not consciously aware of limits to their perception of auditory events: We find that adults are often not aware of their ability to hear water temperature (Experiment 1). Combined with lack of awareness, children's failure to detect nuanced auditory information has potential to impact everyday decisions, as well as increase risks to children's safety. Overly hot beverages are among the most common sources of burns requiring emergency care in children (Battle et al., 2016; Kemp et al., 2014; Ramanathan et al., 1994). Inability to use the sound of a pouring event to judge a beverages' temperature may make it more challenging for children to avoid overly hot beverages, potentially contributing to the frequency of this common real-world injury in childhood.

Children's failures to hear water temperature may also be part of a more general difficulty detecting and comprehending subtle auditory information in early life. Young children of similar ages also show difficulty using sound to detect and localize an approaching vehicle (Barton et al., 2013; Pfeffer & Barnecutt, 1996). Children's ability to identify environmental sounds in challenging acoustic conditions (in the presence of noise) also improves from 7–12 years (Krishnan et al., 2013). In contrast, children easily succeed at basic identification of familiar environmental sounds under ideal listening conditions by 5 years of age and likely far earlier (e.g., the sounds of a bird, a baby, a vacuum; Finitzo-Hieber et al., 1980; Liu et al., 2013). 4- to 8-year-olds can also use sound to make rational causal inferences about the presence and numerosity of physical objects (e.g., shaking a box and listening to determine how many objects are inside; Siegel et al., 2021). Our findings, together with these data, suggest that children's sensitivity to, and their understanding of more nuanced aspects of auditory events may be impoverished in early life, in spite of their success at more coarse-grained sound source judgements.

What explains young children's inability to hear water temperature, and the protracted development of this skill over childhood? Multiple factors are likely to play a role, including the need for accumulation of relevant experience to support this skill, developmental changes in multisensory integration in middle-childhood, and changes to auditory sensitivity over childhood.

First, our data suggest that substantial relevant experience may be needed to hear water temperature. The experience-based individual differences seen in adults' accuracy (Experiment 1) suggests that experience with the relevant stimuli, in the relevant sensory modalities, may be necessary to hear water temperature. A lack of relevant experience may contribute to young children's failures as well: Children may require direct, multimodal experience of both hot and cold water pouring events. Sound source identification, and crossmodal correspondences more generally, may require protracted accumulation of experience with the relevant events, supporting learned associations between the acoustic features of each event and other aspects of the relevant experiences (Gaver, 1993a, 1993b; Hidaka et al., 2015; Spence, 2011; see also Piaget, 1954; Robinson & Sloutsky, 2010). Infants' early ability to use basic spatio-temporal information to perceive crossmodal correspondences (e.g., linking sound to an impact, or to a visual location; Spelke, 1979) may provide a foundation for further learning of more nuanced multimodal features of events over childhood.

In addition to the accumulation of experience with age, developmental changes in multimodal integration abilities in middle childhood may also contribute to the slow developmental trajectory of hearing water temperature. From 6 to 12 years of age, children slowly become able to integrate multiple sources of sensory information in adult-like ways (Nardini et al., 2010). Children show notable failures at multimodal integration prior to 8 years of age, including failing to integrate tactile input with other modalities (Barutchu et al., 2009; Gori et al., 2008; Nardini et al., 2008). We find that children begin to succeed at hearing water temperature by age 7 years, with accuracy continuing to improve until age 11, roughly aligning with these major changes in multimodal integration abilities. This suggests that an integrated, multimodal representation of the pouring event may not be present in a fully adult-like form until middle-childhood, potentially limiting younger children's ability to connect a typically tactile feature (temperature) to sound.

Developmental changes in auditory sensitivity may also play a role: Older children may be more able to hear the auditory frequencies that hold information about water temperature versus younger children. We find that children fail to hear water temperature before age 6–7 years. Hearing water temperature appears to depend on information present in the 200 Hz and 5–6 kHz ranges of frequencies (Velasco et al., 2013), and children's auditory sensitivity to frequencies in this range dramatically improves from age 4 to 6 (Jensen & Neff, 1993; Maxon & Hochberg, 1982), reaching adult-like levels only by 10–12 years (Saffran et al., 2006)—the age at which children in our samples perform as accurately as adults. Large individual differences have also been observed in the auditory processing abilities of 4–6 year olds, with some young children showing far higher ability than others (Jensen & Neff, 1993; Moore et al., 2011). This could help explain variation in performance in our dataset: Some children appeared to succeed at ~6 years of age, while many others failed.

Last, children's conceptual development may play some role, although our tasks were designed to fit well even with younger children's levels of conceptual development. Over childhood, children's physical understanding of the nature of temperature, heat, and related concepts becomes increasingly nuanced and scientifically-accurate, shifting from viewing heat as a substance or property of objects, to viewing heat as a process of transferring energy (Appleton, 1985; Luce & Callanan, 2020; Shayer & Wylam, 1981). Our tasks did not require children to think of heat as a process, but only to label stimuli as hot or cold. This task is consistent with children's early-developing theories of heat as a property, and even preschool-age children commonly hear and use the words "hot" and "cold" in parent-child conversation to describe objects and substances (Luce & Callanan, 2020). Our comprehension questions confirmed that all children included in our sample were able to accurately label visual stimuli as hot or cold. Thus, even at preschool age, children's conceptual understanding of temperature appears sufficient to support categorization of stimuli as hot or cold. However, it is possible that older children's greater mechanistic understanding of the impact of heat on liquids may guide their attention to relevant temperature-related differences in the sounds, such as the sounds of steam or bubbles.

Overall, we find evidence that experience plays a role in adults' ability to hear water temperature, suggesting it is likely to play some role in children's performance as well. The protracted development of hearing water temperature is likely to be the product of multiple factors, including a lack of sufficient relevant experience, developmental change to multimodal integration, and lower auditory sensitivity to the relevant frequencies in childhood.

In order to tease apart the relative contributions of these factors, we suggest testing the impact of training as a productive avenue for future study. A training study could provide an informative means of testing whether relevant experience is sufficient to hear water temperature, or whether other factors are at play—and how this may differ across developmental time. For example, using a pre- and post- test design, children of various ages could be tested on their ability to hear water temperature, as well as their auditory sensitivity to different frequencies, and multimodal integration abilities. Children who initially fail to hear water temperature could then undergo a training protocol, observing several hot and cold water pouring events. A post-test hearing water temperature task could then determine if children's performance improved as a result of training (as compared to a control group of children who did not get the relevant training). This proposed study could be informative in many potential ways. For example, it is possible that only children above some threshold of age or auditory sensitivity are able to benefit from training. In this case, these data would support the conclusion that experience as well as these other factors plays a causal role in the development of nuanced auditory sound source judgements. Alternatively, it is possible that even for young children, experience with the relevant events is sufficient to allow them to hear water temperature. This would suggest that children's limitations of auditory sensitivity and multimodal integration may not contribute to early failures.

ACKNOWLEDGMENTS

This work was supported by a fellowship to TA from the Center for Academic Research and Training in Anthropogeny (CARTA) at the University of California, San Diego. We thank Viola Stormer and Tim Brady for conversations inspiring this line of work, as well as Amanda Calcetas, Danielle Clarke, Finiva Ednalino, Maxine Covello, Michelle Lee, and Naomi Lin for their help with data collection. We note that data from Experiment 2 of this manuscript was presented at the 2020 Cognitive Science Society meeting and published in the conference proceedings (see Agrawal et al., 2020).

CONFLICT OF INTEREST

We have no known conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in an OSF repository https://osf.io/brp2a/?view_only=7f49783ebbf646b29af32ca645246f57.

ORCID

Tanushree Agrawal https://orcid.org/0000-0001-5814-3109
Adena Schachner https://orcid.org/0000-0001-6886-7098

ENDNOTE

[1] In this experiment, we also collected preliminary data on children's prior exposure to hot/cold beverages by asking parents to complete a beverage survey similar to the one completed by adults in Experiment 1. This was exploratory work, and was not preregistered. We ultimately had low confidence in the accuracy of parents' estimates. Incomplete data and ceiling/floor effects also limited our power. These data and a summary of findings are available online in our OSF Supplement.

REFERENCES

Agrawal, T., Lee, M., Calcetas, A., Clarke, D., Lin, N., & Schachner, A. (2020). Hearing water temperature: Characterizing the development of nuanced perception of auditory events. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings for the 42nd annual meeting of the Cognitive Science Society* (pp. 494–499). Cognitive Science Society.

Appleton, K. (1985). Children's ideas about temperature. In R. P. Tisher (Ed.), *Selections of refereed papers from the Sixteenth Annual Conference of the Australian Science Education Research Association, Capricornia Institute, Rockhampton, Queensland* (pp. 122–126). Springer. https://doi.org/10.3316/aeipt.176448

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, *36*(2), 190. https://doi.org/10.1037/0012-1649.36.2.190

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(2), 250. https://doi.org/10.1037/0096-1523.19.2.250

Barton, B. K., Lew, R., Kovesdi, C., Cottrell, N. D., & Ulrich, T. (2013). Developmental differences in auditory detection and localization of approaching vehicles. *Accident Analysis & Prevention*, *53*, 1–8. https://doi.org/10.1016/j.aap.2012.12.040

Barutchu, A., Crewther, D. P., & Crewther, S. G. (2009). The race that precedes coactivation: Development of multisensory facilitation in children. *Developmental Science*, *12*(3), 464–473. https://doi.org/10.1111/j.1467-7687.2008.00782.x

Battle, C. E., Evans, V., James, K., Guy, K., Whitley, J., & Evans, P. A. (2016). Epidemiology of burns and scalds in children presenting to the emergency department of a regional burns unit: A 7-year retrospective study. *Burns & Trauma*, *4*, s410380160047-7 . https://doi.org/10.1186/s41038-016-0047-7

Bronkhorst, A. W., & Houtgast, T. (1999). Auditory distance perception in rooms. *Nature*, *397*(6719), 517–520. https://doi.org/10.1038/17374

Cabe, P. A., & Pittenger, J. B. (2000). Human sensitivity to acoustic information from vessel filling. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 313. https://doi.org/10.1037/0096-1523.26.1.313

Cochran, G., & Harpending, H. (2009). *The 10,000 year explosion: How civilization accelerated human evolution*. Basic Books.

Cusimano, M., Hewitt, L. B., Tenenbaum, J., & McDermott, J. H. (2018). Auditory scene analysis as bayesian inference in sound source models. In C. Kalish, M. Rau, J. Zhu, & T. Rogers (Eds.), *Proceedings of the 40th annual meeting of the Cognitive Science Society (pp.* 510–516). Cognitive Science Society.

Dean, I., Harper, N. S., & McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience*, *8*(12), 1684–1689. https://doi.org/10.1038/nn1541

Dunlap, A. S., & Stephens, D. W. (2014). Experimental evolution of prepared learning. *Proceedings of the National Academy of Sciences*, *111*(32), 11750–11755. https://doi.org/10.1073/pnas.1404176111

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, *7*(5), 7. https://doi.org/10.1167/7.5.7

Finitzo-Hieber, T., Gerling, I. J., Matkin, N. D., & Cherow-Skalka, E. (1980). A sound effects recognition test for the pediatric audiological evaluation. *Ear and Hearing*, *1*(5), 271–276. https://doi.org/10.1097/00003446-198009000-00007

Fior, R. (1972). Physiological maturation of auditory function between 3 and 13 years of age. *Audiology*, *11*(5–6), 317–321. https://doi.org/10.3109/00206097209072598

Freiberg, K., Tually, K., & Crassini, B. (2001). Use of an auditory looming task to test infants' sensitivity to sound pressure level as an auditory distance cue. *British Journal of Developmental Psychology*, *19*(1), 1–10. https://doi.org/10.1348/026151001165903

Garcia, J., & Koelling, R. A. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science*, *4*(1), 123–124. https://doi.org/10.3758/BF03342209

Gaver, W. (1993a). How do we hear in the world? Explorations in ecological acoustics. *Ecological Psychology*, *5*(4), 285–313. https://doi.org/10.1207/s15326969eco0504_2

Gaver, W. W. (1993b). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology*, *5*(1), 1–29. https://doi.org/10.1207/s15326969eco0501_1

Gerstenberg, T., Siegel, M. H., & Tenenbaum, J. (2018). *What happened? Reconstructing the past through vision and sound*. PsyArXiv. https://doi.org/10.31234/osf.io/tfjdk

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67–82). Lawrence Erlbaum Associates.

Giordano, B. L. (2003). Everyday listening: An annotated bibliography. In D. Rocchesso & F. Fontana (Eds.), *The Sounding Object* (pp. 1–16). Mondo estremo.

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

Gottfried, A. W., Rose, S. A., & Bridger, W. H. (1977). Cross-modal transfer in human infants. *Child Development*, *48*(1), 118–123. https://doi.org/10.2307/1128889

Grassi, M. (2005). Do we hear size or sound? Balls dropped on plates. *Perception & Psychophysics*, *67*(2), 274–284. https://doi.org/10.3758/BF03206491

Gutmann, F., & Simmons, L. M. (1952). The temperature dependence of the viscosity of liquids. *Journal of Applied Physics*, *23*(9), 977–978. https://doi.org/10.1063/1.1702361

Guyot, P., Houix, O., Misdariis, N., Susini, P., Pinquier, J., & André-Obrecht, R. (2017). Identification of categories of liquid sounds. *The Journal of the Acoustical Society of America*, *142*(2), 878–889. https://doi.org/10.1121/1.4996124

Hedrick, V. E., Comber, D. L., Estabrooks, P. A., Savla, J., & Davy, B. M. (2010). The beverage intake questionnaire: Determining initial validity and reliability. *Journal of the American Dietetic Association*, *110*(8), 1227–1232. https://doi.org/10.1016/j.jada.2010.05.005

Henrich, J., & McElreath, R. (2007). Dual inheritance theory: The evolution of human cultural capacities and cultural evolution. In L. Barrett

& R. Dunbar (Eds.), *Oxford Handbook of Evolutionary Psychology* (pp. 555–570). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198568308.013.0038

Hidaka, S., Teramoto, W., & Sugita, Y. (2015). Spatiotemporal processing in crossmodal interactions for perception of the external world: A review. *Frontiers in Integrative Neuroscience*, *9*, 62. https://doi.org/10.3389/fnint.2015.00062

Horowitz, S. (2012). *The universal sense: How hearing shapes the mind*, Bloomsbury USA.

Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science*, *4*(2), 104–107. https://doi.org/10.1111/j.1467-9280.1993.tb00469.x

Jousmäki, V., & Hari, R. (1998). Parchment-skin illusion: Sound-biased touch. *Current Biology*, *8*(6), R190–R191. https://doi.org/10.1016/s0960-9822(98)70120-4

Kemp, A. M., Jones, S., Lawson, Z., & Maguire, S. A. (2014). Patterns of burns and scalds in children. *Archives of Disease in Childhood*, *99*(4), 316–321. https://doi.org/10.1136/archdischild-2013-304991

Kopp, F. (2014). Audiovisual temporal fusion in 6-month-old infants. *Developmental Cognitive Neuroscience*, *9*, 56–67. https://doi.org/10.1016/j.dcn.2014.01.001

Krishnan, S., Leech, R., Aydelott, J., & Dick, F. (2013). School-age children's environmental object identification in natural auditory scenes: Effects of masking and contextual congruence. *Hearing Research*, *300*, 46–55. https://doi.org/10.1016/j.heares.2013.03.003

Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*(4577), 1138–1141. https://doi.org/10.1126/science.7146899

Kunkler-Peck, A. J., & Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(1), 279. https://doi.org/10.1037/0096-1523.26.1.279

Lederman, S. J. (1979). Auditory texture perception. *Perception*, *8*(1), 93–103. https://doi.org/10.1068/p080093

Lemaitre, G., & Heller, L. M. (2012). Auditory perception of material is fragile while action is strikingly robust. *The Journal of the Acoustical Society of America*, *131*(2), 1337–1348. https://doi.org/10.1121/1.3675946

Lewkowicz, D. J. (1992). Infants' responsiveness to the auditory and visual attributes of a sounding/moving stimulus. *Perception & Psychophysics*, *52*(5), 519–528. https://doi.org/10.3758/BF03206713

Liu, S.-Y., Liu, T.-C., Teng, Y.-L., Lee, L.-A., Lai, T.-J., & Wu, C.-M. (2013). Environmental sounds recognition in children with cochlear implants. *Plos One*, *8*(6), e66100. https://doi.org/10.1371/journal.pone.0066100

Luce, M. R., & Callanan, M. A. (2020). Family conversations about heat and temperature: implications for children's learning. *Frontiers in Psychology*, *11*, 1718. https://www.frontiersin.org/article/10.3389/fpsyg.2020.01718

Lutfi, R. A. (2008). Human sound source identification. In (W. A. Yost, A. N. Popper, & R. R. Fay Eds.), *Auditory perception of sound sources* (pp. 13–42). Springer.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Lawrence Erlbaum Associates Publishers.

Maurer, D., Stager, C. L., & Mondloch, C. J. (1999). Cross-modal transfer of shape is difficult to demonstrate in one-month-olds. *Child Development*, *70*(5), 1047–1057. https://doi.org/10.1111/1467-8624.00077

Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: Sensitivity and discrimination. *Ear and Hearing*, *3*(6), 301–308. https://doi.org/10.1097/00003446-198211000-00003

McDermott, J. H., Schemitsch, M., & Simoncelli, E. P. (2013). Summary statistics in auditory perception. *Nature Neuroscience*, *16*(4), 493–498. https://doi.org/10.1038/nn.3347

McWalter, R., & McDermott, J. H. (2019). Illusory sound texture reveals multi-second statistical completion in auditory scene analysis. *Nature Communications*, *10*(1), 5096. https://doi.org/10.1038/s41467-019-12893-0

Meltzoff, A. N., & Borton, R. W. (1979). Intermodal matching by human neonates. *Nature*, *282*(5737), 403–404. https://doi.org/10.1038/282403a0

Merrick, C., & Filingeri, D. (2019). The evolution of wetness perception: A comparison of arachnid, insect and human models. *Journal of Thermal Biology*, *85*, 102412. https://doi.org/10.1016/j.jtherbio.2019.102412

Młynarski, W., & McDermott, J. H. (2018). Learning midlevel auditory codes from natural sound statistics. *Neural Computation*, *30*(3), 631–669. https://doi.org/10.1162/neco_a_01048

Moore, D. R., Cowan, J. A., Riley, A., Edmondson-Jones, A. M., & Ferguson, M. A. (2011). Development of auditory processing in 6-to 11-yr-old children. *Ear and Hearing*, *32*(3), 269–285. https://doi.org/10.1097/AUD.0b013e318201c468

Morrongiello, B. A., Fenwick, K. D., Hillier, L., & Chance, G. (1994). Sound localization in newborn human infants. *Developmental Psychobiology*, *27*(8), 519–538. https://doi.org/10.1002/dev.420270805

Morrongiello, B. A., Hewitt, K. L., & Gotowiec, A. (1991). Infants' discrimination of relative distance in the auditory modality: Approaching versus receding sound sources. *Infant Behavior and Development*, *14*(2), 187–208. https://doi.org/10.1016/0163-6383(91)90005-D

Murphy, G. (2004). *The big book of concepts*. MIT press.

Nardini, M., Bedford, R., & Mareschal, D. (2010). Fusion of visual cues is not mandatory in children. *Proceedings of the National Academy of Sciences*, *107*(39), 17041–17046. https://doi.org/10.1073/pnas.1001699107

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

Neuhoff, J. G. (2018). Adaptive biases in visual and auditory looming perception. In T. L. Hubbard (Ed.), *Spatial biases in perception and cognition* (pp. 180–191). Cambridge University Press.

Nwokolo, E. (1996). Bottle gourd, buffalo gourd and other gourds. In E. Nwokolo & J. Smartt (Eds.), *Food and feed from legumes and oilseeds* (pp. 290–297). Springer.

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy: Evidence for sound–shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology*, *114*(2), 173–186. https://doi.org/10.1016/j.jecp.2012.05.004

Parthasarathy, S., & Chhapgar, A. F. (1955). Sound absorption in liquids in relation to their physical properties viscosity and specific heats. *Annalen der Physik*, *451*(5-8), 297–303. https://doi.org/10.1002/andp.19554510509

Pellegrino, R., Jones, J. D., Shupe, G. E., & Luckett, C. R. (2019). Sensitivity to viscosity changes and subsequent estimates of satiety across different senses. *Appetite*, *133*, 101–106. https://doi.org/10.1016/j.appet.2018.10.028

Perfecto, H., Donnelly, K., & Critcher, C. R. (2018). Volume estimation through mental simulation. *Psychological Science*, *30*(1), 80–91. https://doi.org/10.1177/0956797618813319

Pfeffer, K., & Barnecutt, P. (1996). Children's auditory perception of movement of traffic sounds. *Child: Care, Health and Development*, *22*(2), 129–137. https://doi.org/10.1111/j.1365-2214.1996.tb00780.x

Piaget, J. (1954). *The construction of reality in the child* (M. Cook, Trans.). Basic Books. https://doi.org/10.1037/11168-000

Putzar, L., Goerendt, I., Lange, K., Rösler, F., & Röder, B. (2007). Early visual deprivation impairs multisensory interactions in humans. *Nature Neuroscience*, *10*(10), 1243–1245. https://doi.org/10.1038/nn1978

Ramanathan, C., Ekpenyong, L., & Stevenson, J. H. (1994). Scald burns in children caused by hot drinks—The importance of the type of cup. *Burns*, *20*(2), 111–114. https://doi.org/10.1016/S0305-4179(06)80005-7

Robinson, C. W., & Sloutsky, V. M. (2010). Development of cross-modal processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 135–141. https://doi.org/10.1002/wcs.12

Roque, J., Lafraire, J., Spence, C., & Auvray, M. (2018). The influence of audiovisual stimuli cuing temperature, carbonation, and color on the categorization of freshness in beverages. *Journal of Sensory Studies*, *33*(6), e12469. https://doi.org/10.1111/joss.12469

Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. In D. Kuhn, R. S. Siegler, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Cognition, perception, and language* (pp. 58–108). John Wiley & Sons Inc.

Seitz, A. R., Kim, R., van Wassenhove, V., & Shams, L. (2007). Simultaneous and independent acquisition of multisensory and unisensory associations. *Perception*, *36*(10), 1445–1453. https://doi.org/10.1068/p5843

Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. *Trends in Cognitive Sciences*, *14*(9), 425–432. https://doi.org/10.1016/j.tics.2010.07.001

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences*, *12*(11), 411–417. https://doi.org/10.1016/j.tics.2008.07.006

Shayer, M., & Wylam, H. (1981). The development of the concepts of heat and temperature in 10–13 year-olds. *Journal of Research in Science and Teaching*, *18*(4), 419–434. https://doi.org/10.1002/tea.3660180506

Siegel, M. H., Magid, R. W., Pelz, M., Tenenbaum, J. B., & Schulz, L. E. (2021). Children's exploratory play tracks the discriminability of hypotheses. *Nature Communications*, *12*(1), 1–9. https://doi.org/10.1038/s41467-021-23431-2

Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, *8*(4), 553–560. https://doi.org/10.1016/0010-0285(76)90018-9

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology*, *15*(6), 626–636. https://doi.org/10.1037/0012-1649.15.6.626

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995. https://doi.org/10.3758/s13414-010-0073-7

Spence, C. (2020). Temperature-based crossmodal correspondences: Causes and consequences. *Multisensory Research*, *33*(6), 645–682. https://doi.org/10.1163/22134808-20191494

Spence, C., & Wang, Q. J. (2015). Sensory expectations elicited by the sounds of opening the packaging and pouring a beverage. *flavouring*, *4*(1), 1–11. https://doi.org/10.1186/s13411-015-0044-y

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews. Neuroscience*, *15*(8), 520–535. https://doi.org/10.1038/nrn3742

Stollman, M. H. P., van Velzen, E. C. W., Simkens, H. M. F., Snik, A. F. M., & van den Broek, P. (2004). Development of auditory processing in 6–12-year-old children: A longitudinal study. *International Journal of Audiology*, *43*(1), 34–44. https://doi.org/10.1080/14992020400050006

Streri, A., & Gentaz, E. (2003). Cross-modal recognition of shape from hand to eyes in human newborns. *Somatosensory & Motor Research*, *20*(1), 13–18. https://doi.org/10.1080/0899022031000083799

Streri, A., & Gentaz, E. (2004). Cross-modal recognition of shape from hand to eyes and handedness in human newborns. *Neuropsychologia*, *42*(10), 1365–1369. https://doi.org/10.1016/j.neuropsychologia.2004.02.012

Traer, J., Norman-Haignere, S. V., & McDermott, J. H. (2021). Causal inference in environmental sound recognition. *Cognition*, *214*, 104627. https://doi.org/10.1016/j.cognition.2021.104627

Velasco, C., Jones, R., King, S., & Spence, C. (2013). The sound of temperature: What information do pouring sounds convey concerning the temperature of a beverage. *Journal of Sensory Studies*, *28*(5), 335–345. https://doi.org/10.1111/joss.12052

Wallace, M. T., Perrault, T. J., Hairston, W. D., & Stein, B. E. (2004). Visual experience is necessary for the development of multisensory integration. *Journal of Neuroscience*, *24*(43), 9580–9584. https://doi.org/10.1523/JNEUROSCI.2535-04.2004

Wang, Q. J., & Spence, C. (2017). The role of pitch and tempo in sound-temperature crossmodal correspondences. *Multisensory Research*, *30*(3–5), 307–320. https://doi.org/10.1163/22134808-00002564

Wrangham, R. (2009). *Catching fire: How cooking made us human*. Basic Books.

Wu, X., Zhang, C., Goldberg, P., Cohen, D., Pan, Y., Arpin, T., & Bar-Yosef, O. (2012). Early pottery at 20,000 years ago in Xianrendong Cave, China. *Science*, *336*(6089), 1696–1700. https://doi.org/10.1126/science.1218643

Zampini, M., & Spence, C. (2005). Modifying the multisensory perception of a carbonated beverage using auditory cues. *Food Quality and Preference*, *16*(7), 632–641. https://doi.org/10.1016/j.foodqual.2004.11.004

Zhang, X. (2015). Molecular sensors and modulators of thermoreception. *Channels*, *9*(2), 73–81. https://doi.org/10.1080/19336950.2015.1025186

How to cite this article: Agrawal, T., & Schachner, A. (2023). Hearing water temperature: Characterizing the development of nuanced perception of sound sources. *Developmental Science*, *26*, e13321. https://doi.org/10.1111/desc.13321