THE DEVELOPMENT OF AUDIOTACTILE TEMPORAL PERCEPTION

THE DEVELOPMENT OF AUDIOTACTILE TEMPORAL PERCEPTION

By BRENDAN M. STANLEY, B.A. (hons.)

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements

for the Degree Doctor of Philosophy

McMaster University

McMaster University, DOCTOR OF PHILOSOPHY (2023)

Hamilton, Ontario (Department of Psychology, Neuroscience & Behaviour)

TITLE: The development of audiotactile temporal perception

AUTHOR: Brendan M. Stanley, H.B.A. (McMaster University)

SUPERVISOR: Dr. David I. Shore

NUMBER OF PAGES: xix, 196

LAY ABSTRACT

Perception relies on combining information from our senses. Multiple cues determine whether we integrate or segregate sensory information. Timing provides one crucial cue. Children's timing perception requires development to reach the same precision as adults. Most studies on the development of time perception between the senses have included vision. However, this thesis investigated the development of time perception between hearing and touch. The first two empirical chapters explored typical development using complementary tasks, while the third empirical chapter investigated the impact of congenital cataracts on timing perception. By studying children with cataracts who underwent early cataract removal, we can observe the effects of visual deprivation on these senses. These chapters shed light on the development of audiotactile temporal perception and propose that different combinations of senses may develop independently.

ABSTRACT

This thesis investigated developmental changes in temporal perception of hearing and touch (audiotactile). Three empirical chapters provide converging evidence on the unique characteristics of this modality pairing. In Chapter 2, a simultaneity judgment task assessed temporal perception. Three groups of children (aged 7-, 9-, and 11-years-old) were compared to a group of adults, examining measures such as the temporal simultaneity window and the point of subjective simultaneity. By age 11, mature temporal perception between hearing and touch was observed. Chapter 3 investigated developmental changes in temporal-based integration using the fission and fusion illusions. The study involved comparing three groups of children to adults (aged 9-, 11-, and 13-years-old). The measure of illusion strength combined with a signal detection analysis demonstrated that children did not exhibit adult-like integration until around age 13. Chapter 4 explored the potential impact of short-term congenital visual deprivation on hearing and touch temporal perception. An audiotactile simultaneity task was used to test a group of adults who received treatment for congenital bilateral cataracts. The results of this final experiment are considered preliminary because of limitations imposed by the COVID-19 pandemic; instead of the planned age- and gender-matched control participants, we utilized the adult data from Chapter 2 for comparison. The General Discussion provides a comprehensive account of how these findings relate to one another and how they situate in the broader literature. Additionally, a novel hypothetical theory is presented, incorporating the established causal inference framework, to offer insights into observed changes in multisensory perception across development.

ACKNOWLEDGMENTS

The innate fear of being forgotten is a fundamental human trait we all possess. Many spend their lives striving to build a lasting legacy, something that will continue to resonate through time. In my case, my legacy takes shape in the realm of science; even when I am no longer around, my work will continue to be part of the scientific conversation, subtly shaping the trajectory of future research and discoveries.

Yet, no endeavor is solely the result of one person's effort. As such, I hold immense gratitude for everyone involved in my journey—for their intellectual contributions, guidance, mentorship, and support.

First and foremost, I express my profound appreciation to my supervisor, mentor, and friend, David Shore. I know I would not be completing a PhD if it were not for your unwavering support, continual encouragement, and insightful mentorship. You provided me with countless opportunities to learn and grow, both as a researcher and as a person. Even when I was going through difficult times and considered giving up, you remained a pillar of support, continuing to believe in me even when I struggled to believe in myself.

Being mentored by my advisory team—Daphne Maurer, Terri Lewis, Yi-Chuan Chen, and Daniel Goldreich—has been nothing short of a privilege. I thank each of you for providing me with the wealth of knowledge, guidance, wisdom, and support throughout my graduate career. Daphne, thank you for always supporting and encouraging me, especially when I was struggling and needed it most. Terri, your wisdom has transformed my writing profoundly. Because of you, I will strive to never split the infinitive again! Yi-Chuan, you were the one who encouraged me to pursue graduate studies, and for that, I am forever grateful. Your enduring support throughout my journey has been invaluable. Dan, thank you for the advice over the years and never hesitating to provide opportunities to learn more about your field of research.

I am grateful to have shared this journey with my fellow graduate student colleagues, especially those from the Shore and Milliken labs. Lisa Lorentz, my partner-in-crime, always there to lift “Mahmood!” I will always cherish our spontaneous—but much needed—mid-day ‘rapid decalibrations’ at the Phoenix. Hanae Davis, the one with the face of a thousand expressions, thank you for sharing this journey with me and becoming my ‘pho-eva’ friend! *#NapkinForever*. My appreciation also extends to Andrew LoGiudice, my fellow entomology hobbyist and conference roomie, Irina Ghilic, a gem of a human who would give you her left kidney even if you didn’t need one, Kaian Unwalla, my TRG partner and the only person I know who eats biryani soup, Daniel Nienhuis, aka Danny-long-legs, Amy Pachai, the kindest no BS badass you will ever meet, and Michelle Cadieux, the emeritus monarch of the lab. I am so lucky to have met and developed friendships with each of you. To those I did not explicitly name, you are not forgotten. I have met too many wonderful people over the years to list them all, but please know that I appreciate each and every one of you.

To my family and friends, I could not have done this without your endless support and encouragement. To my partner Ivor, thank you for your patience even when the finish line never seemed to draw near. To my closest friends, Chris and Zack: thank you for always being there when I needed it most. To my sisters, Bree and Gab, thank you for simply existing—we make a great trio, each of us different in our own unique ways. I wouldn’t have it any other way.

Most importantly, I owe boundless gratitude to my parents, Cheryl and Steve, for their unwavering support over the years. Mom, while you always mention how proud you are of me, truth is, I am the one honoured to be your son. Despite the significant challenges you have faced in your life, you have always found the silver lining within the darkest of clouds. You have taught me patience, resilience, and persistence, and have always inspired and encouraged me to get back up when life knocked me down. Thank you for everything.

Finally, for me, being a Stanley is a source of immense pride. The title 'Dr. Stanley' reminds me of the lineage and legacy that has supported me. In honour of that legacy, I dedicate this journey to my grandparents, Granny Rita and Grandpa Edwin. Granny, you’ve always had this gentle way of motivating me to harness my full potential, always standing by my side and applauding my achievements. Growing up, I admired your intellect—you inspired me to cultivate that same level of intellect within myself. Grandpa, while growing up, I vividly remember you telling me repeatedly that I will be “the doctor in the family.” Well, you were right, I did it—and I know you would be so proud of me if you were still here with us today.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF ABBREVIATIONS AND SYMBOLS

CHAPTER 2: AUDIOTACTILE SIMULTANEITY DEVELOPMENT

PSS: Point of subjective simultaneity

SOA: Stimulus onset asynchrony

V1: Primary visual area

τ: Processing time difference

λA: Auditory processing variability

λT: Tactile processing variability

δ: Sensitivity

εAF: Auditory-leading response error

εS: Simultaneous response error

εTF: Tactile-leading response error

ANOVA: Analysis of variance

CHAPTER 3: AUDIOTACTILE INTEGRATION DEVELOPMENT

SOA: Stimulus onset asynchrony

#T#B: {number} Tap(s) {number} Beep(s) (e.g., "0T1B" = "zero taps one beep")

***d'*:** D-prime

***c*:** Criterion

MLE: Maximum likelihood estimation

ANOVA: Analysis of variance

CHAPTER 4: AUDIOTACTILE SIMULTANIETY IN CATARACT PATIENTS

PSS: Point of subjective simultaneity

SOA: Stimulus onset asynchrony

OD: Right eye

OS: Left eye

OU: Both eyes

ANOVA: Analysis of variance

MLE: Maximum likelihood estimation

DECLARATION OF ACADEMIC ACHIEVEMENT

The General Introduction of this thesis (Chapter 1) was written by Brendan M. Stanley and edited by Dr. David I. Shore, Dr. Daphne Maurer, and Dr. Terri L. Lewis.

This thesis contains three empirical chapters. The first (Chapter 2) was published in the *Journal of Experimental Child Psychology* and authored by Brendan M. Stanley, Dr. Yi-Chuan Chen, Dr. Terri L. Lewis, Dr. Daphne Maurer, and Dr. David I. Shore. The experiment design and methods were conceptualized by Dr. Yi-Chuan Chen, Dr. Terri L. Lewis, Dr. Daphne Maurer, and Dr. David I. Shore, with contributions to the methods by Brendan M. Stanley. The experiment was programmed (by modifying existing code authored originally by Dr. Yi-Chuan Chen) and conducted by Brendan M. Stanley. Formal analyses and visualizations were completed primarily by Brendan M. Stanley with the assistance of Dr. David I. Shore. The original draft of the manuscript was authored by Brendan M. Stanley and edited/revised by Dr. David I. Shore, Dr. Yi-Chuan Chen, Dr. Terri L. Lewis, and Dr. Daphne Maurer. Preparations and submissions for publication completed by Brendan M. Stanley.

The second data chapter (Chapter 3) was also published in the Journal of Experimental Child Psychology. All roles were identical to those described for Chapter 2.

The third and final data chapter (Chapter 4) was prepared as a manuscript for publication; however, because of significant limitations (see section *Circumstances that Affected Progress* below), the manuscript was not submitted to an academic journal. All roles were identical to those described for Chapter 2.

CIRCUMSTANCES THAT AFFECTED PROGRESS

My progress and completion of this degree were impacted significantly by two events beginning in March 2020.

The first event was the onset of the COVID-19 global pandemic, which led to the suspension of data collection on cataract patients at SickKids Hospital. Initially, the plan was to collect data from both bilateral and unilateral congenital cataract patients, as well as age- and sex-matched controls (with a minimum of two control participants per patient). By March 2020, data collection on the bilateral patients had been completed; however only a few unilateral patients, and no control participants, had been tested. Consequently, only the group of bilateral cataract reversal patients could be included in Chapter 4, and the control group used for comparison was the adults tested in Chapter 2; although the experiment parameters were similar, and the task was identical, this control group was not matched for age or sex of the patients. Aside from the significant limitations to the interpretations made in Chapter 4, the ongoing uncertainty about the timeline of the pandemic (i.e., how long it would be until data collection could be resumed) led to significant delays in my progress.

The second event, which took place in July 2020, involved the sudden and extended absence of my primary supervisor, Dr. David I. Shore. The specific details surrounding his absence are beyond the purview of this thesis. This unforeseen loss of my primary supervisor caused a substantial disruption to my progress due to various factors. The most noteworthy among them were the inability to communicate and collaborate with my supervisor, and the significant emotional impact it had on my overall well-being. It wasn't until approximately January 2023 that I managed to realign my efforts and successfully fulfill the requirements of this degree.

Chapter 1

General Introduction

Perception, from sensation to cognition, decodes the enigma that is our external world. The ambiguous signals originating from our sensory receptors require interpretation to become an internal representation of the external world. This internal representation is formed by combining signals from different senses to enhance perception—different senses typically provide converging and often redundant information, especially when the signals arrive in close temporal proximity to each other. These multisensory processes are constrained by the precision of the signals transduced from the sensory receptors. To add further complexity, the precision of each sensory system changes across the lifespan. Through prolonged developmental processes that rely on physical, neural, and cognitive maturation, all shaped by lived experience, multisensory perception eventually stabilizes and becomes adult-like. To chart the trajectory from immature to mature, we must rely on tasks that are assumed to provide proxies of multisensory capabilities. Given the importance of temporal coincidence for multisensory processing, this thesis explores the development of multisensory temporal perception. Given the constraints of a single thesis, I chose to focus on one modality pairing: hearing and touch. Worth noting is that touch is a multifaceted sensory modality encompassing mechanoreception (i.e., tactile perception), haptic perception, proprioception, and nociception. Unless otherwise specified, references to touch pertain to tactile perception, specifically pressure sensation resulting from a punctate stimulus. Using two unique measures of multisensory temporal perception, this thesis charts the development of audiotactile capabilities in middle-to-late childhood and examines one potential consequence of early *visual* deprivation.

Multisensory Perception

We undoubtedly live in a multisensory world. Perceiving this world requires combining information from all our senses, which presents a unique challenge for our perceptual system: when to integrate signals from different modalities and when to segregate them. To do this, our perceptual system relies on both low- and high-level cues (see Collignon et al., 2013; De Meo et al., 2015). Amodal cues such as space, time, and semantic labels, are considered low-level cues that contribute to this decision (Chen & Spence, 2017; Stein & Meredith, 1993). Interestingly, each sensory system codes these amodal cues uniquely: for space, vision has exquisite resolution whereas audition does not; for time, the auditory modality has the best resolution; and for semantic label, each object and context are going to provide unique contributions to resolution. Of these three, time provides the most crucial determinant of multisensory integration (Occelli et al., 2011; Vroomen & Keetels, 2010). Co-localized stimuli separated in time are perceived as segregated, but co-incident stimuli separated in space are integrated (see Spence, 2013). To understand our multisensory world, we must understand how temporal proximity develops and influences multisensory integration.

Computing temporal proximity across modalities presents an additional challenge for perception. Although the signals that impact the sensory system originated from a single event, each sense receives, transduces, and transmits this signal idiosyncratically

(Stein & Meredith, 1993). Signal propagation (the amount of time it takes for a signal to arrive at the perceptual system from its origin) is nearly instantaneous for vision but limited by the medium and the speed of sound/vibration for audition and touch. Neural transduction (the amount of time it takes for the sensory receptor to convert the physical signal into a neural signal) is fast for audition and touch, but relatively slower for vision. Finally, neural transmission (the amount of time it takes for the neural signal to travel from the sensory receptor to the brain) varies depending on the location on the body for touch but is constant for vision and audition (see Harrar & Harris, 2008 for more detailed discussion). These complications force the perceptual system to tolerate some degree of temporal asynchrony when deciding whether to bind stimuli together; the temporal extent of this tolerance is called the *temporal binding window* (Dixon & Spitz, 1980; Keetels & Vroomen, 2012; Vroomen & Keetels, 2010). When signals arrive within this window, they will be integrated into a unified percept, whereas signals that arrive with a delay beyond the window will be segregated. The temporal binding window is the foundation of the temporal proximity rule of multisensory integration.

Sensory cue combination (see Ernst & Banks, 2002; Seilheimer et al., 2014) provides a computational approach to signal integration. Because all sensory signals are inherently noisy (Ernst, 2007), information from multiple modalities are combined to build our perception of the external world. For example, modifying the *auditory* frequency composition of rubbed sandpaper changes the *felt* roughness of the sandpaper (Guest et al., 2002). The final percept of roughness is changed because the auditory signal occurs at the same time and place as the tactile sensation: both amodal sensory cues—

temporal proximity and spatial location—drive the final percept (see Bahrick & Lickliter, 2012; Lewkowicz, 2011). When the signals are integrated, that integration often occurs optimally (Alais & Burr, 2004; Ernst & Bülthoff, 2004). Optimal, in the context of the model, means that each signal influences the final percept in proportion to its reliability, which is the inverse of the variance of the signal: noisy signals are weighted less. Although this model accounts for many multisensory phenomena, it assumes that integration occurs in all circumstances, and does not consider situations in which segregating the signals produces the appropriate percept (i.e., two separate events caused the stimulation; see Beierholm et al., 2009). This requires a model that can account for this limitation.

Causal inference provides a model within which optimal integration operates. This hierarchical model introduces a prior for the probability of integration. As such, the output of the model can account for both integration and segregation (see Noppeney, 2021; Shams & Beierholm, 2022 for reviews). This model builds on the optimal integration model by including a weighted prior of a common cause: an experience-dependent belief about whether the multiple signals originated from the same source (p-common = 1, or c = 1) or separate sources (p-common = 0, or c = 0). To highlight the importance of the prior for common cause, let’s consider an example from the sport of baseball (see Redden et al., 2017 for the same example in the context of prior entry). The first-base umpire’s task is to determine if the hitter is “safe” or “out” depending on whether the hitter contacts the plate before or after the first baseman catches the ball, respectively. In situations in which the visual cue of the hitter contacting the plate occurs very close in time with the sound of the ball contacting the first baseman's glove (i.e., a redundant temporal cue), the temporal order of these two events may be ambiguous, but regardless, the optimal strategy for the umpire would be to segregate these signals. The umpire's prior knowledge of separate causes (p-common = 0) should reduce the probability of integration, thus providing a more precise estimate of the temporal order in which these two events occurred. Although this is a hypothetical example, multisensory causal inference has stronger accounts for many multisensory phenomena than the model of optimal integration (see Shams & Beierholm, 2022). Achieving this hierarchical balance requires experience to fine tune both optimal integration within the c = 1 likelihood hypothesis and the prior likelihood of integration p(common). The experiments in the present thesis were not designed within the causal inference model, but we will make some speculations about how it might fit in the discussion of Chapter 3 and the General Discussion.

Development of Multisensory Perception

Newborns have primitive multisensory abilities that refine into childhood. At birth, the newborn experiences novel sensory input sourced from an unfamiliar world. Although their unisensory modalities are far from precise, newborns can combine sensory signals by relying on the amodal cues mentioned previously such as temporal proximity and spatial co-location (e.g., (Gibson, 1969; see Lewkowicz & Bremner, 2020). Reliance on these cues promotes the learning of meaningful relations while minimizing the formation of unhelpful arbitrary contingencies (Bahrick & Lickliter, 2000). This has been conceptualized as intersensory redundancy which posits that amodal cues promote intermodal learning by capturing attention and coding these meaningful relations (Bahrick & Lickliter, 2000, 2004). This process is thought to scaffold the development of later more complex intersensory relations, such as speech and language production and perception, social competencies, and multisensory object and event perception (for review, see Lewkowicz & Bremner, 2020)

Although infants continue to benefit from redundant amodal cues as they age, they also begin to demonstrate more advanced multisensory abilities. With neural circuitry maturing rapidly and a continuous accumulation of perceptual experiences (Ghazanfar & Schroeder, 2006; Stein & Meredith, 1993), infants transition to a reliance on higher-order experience-dependent and modality-specific features (Emberson, 2017; Lewkowicz, 2014; Lewkowicz et al., 2010). These advances in multisensory abilities come from perceptual narrowing: rarely experienced perceptual categories reduce in salience (Lewkowicz & Ghazanfar, 2009). For example, infants between 6 and 8 months of age can effectively match native (i.e., human) and non-native (i.e., monkey) faces with their corresponding voice calls by relying solely on amodal temporal cues (i.e., matched based on the synchronicity of the auditory and visual onset and offsets). By 8 to 10 months of age, infants were no longer able to perform this matching task for non-native (i.e., monkey) faces and voice calls. Presumably, this occurs because not only do infants transition to rely on more advanced intersensory cues (e.g., identity recognition), but also because infants are continuously exposed to native human faces and voices and consequently their ability to combine these multisensory cues is narrowed to filter out those of non-native species (Lewkowicz & Ghazanfar, 2009). This finding has been supported further with neurophysiological evidence using the same task showing that relative to monkey faces, the functional connectivity in the visual and frontal brain areas were more sensitive to human faces in older infants (Grossmann et al., 2012). Perceptual narrowing (in the behavioural sense) coincides with Hebbian-like neural strengthening and pruning that is also driven by experience with the environment (Maurer et al., 2013). These processes apply to a wide array of multisensory abilities; this thesis focuses on the development of temporal perception.

Temporal synchrony is one of the most basic cues necessary for scaffolding normal development (Bahrick et al., 2004). Although crude at birth, the ability to perceive cross-modal simultaneity supports the development of more advanced cognitive, social, language, and perceptual skills (see Chen et al., 2016). Several studies have demonstrated that the window of temporal simultaneity follows a protracted developmental trajectory, beginning wide and continuing to narrow into late-childhood and even adolescence before stabilizing to adult-like precision (e.g., Chen et al., 2016, 2018; Lewkowicz & Flom, 2014). Here, it is likely that the developing brain prioritizes temporal flexibility over precision, as maintaining a wider window of simultaneity lessens the probability of missing multisensory events (Chen et al., 2016). Similarly, a late and prolonged developmental trajectory is observed for temporal based cross-modal integration (Adams, 2016; Innes-Brown et al., 2011; O'Dowd et al., 2021). When young children combine multisensory signals, they tend to demonstrate sensory dominance (or modality switching). The modality that is more accurate for the task at hand typically dominates the percept (i.e., modality appropriateness; Welch & Warren, 1980). This dominance process is also proposed to calibrate the less accurate modality by the dominant modality (constant calibration hypothesis; Gori et al., 2012). By late childhood to early adolescence, after a vast array of perceptual experiences have been acquired and changes in body growth become more proportional (thus increasing reliability of neural transmission speeds), the precision of temporal perception begins to resemble that of adults. Studies charting the changes in the temporal simultaneity window for both audiovisual (Chen et al., 2016) and visuotactile (Chen et al., 2018) stimuli show that children reach adult like precision by age 7 and 9, respectively. The earlier maturation of audiovisual simultaneity perception is likely attributable to the ubiquitous nature of audiovisual contingencies relative to those involving vision and touch (Chen et al., 2018). This highlights the important role of perceptual experience in the typical development of multisensory perception.

Another approach to revealing the importance of normal sensory experience during development is to examine the outcomes of sensory deprivations. Seminal work conducted by Hubel and Wiesel demonstrates the importance of normal visual input in setting up the neural architecture for vision by showing deficits in kittens deprived of visual input only for a brief period following birth (Wiesel & Hubel, 1963, 1965). In typically developed kittens, most cells in the primary visual area (V1) respond to input from either eye (Hubel & Wiesel, 1962, 1963). However, if a newborn kitten is deprived of visual input to one eye for the first three months of life, almost none of the cells in V1 are responsive when visual stimulation is presented to the formerly deprived eye (Wiesel

& Hubel, 1963). Interestingly, when both eyes are deprived from birth, almost half of the cells in V1 were either abnormal or unresponsive, but surprisingly, over half of these cells responded normally (Wiesel & Hubel, 1965). In a follow-up study, Hubel and Wiesel (1970) found that the timing of deprivation is related to the ability to recover typical visual function. By depriving vision at different life-stages, they found that susceptibility to deprivation varies across the lifespan; kittens were most susceptible to visual deprivation (i.e., experienced the greatest irreversible damage to V1) between four and seven weeks of age and significant damage occurred with as little as three to four days of deprivation (Hubel & Wiesel, 1970). These studies combined revealed several key findings about visual development: 1) normal visual input is required to set up the neural architecture to support typical visual functions, 2) unilateral versus bilateral visual deprivation cause markedly different patterns of abnormal responses in V1, and 3) there are specific critical periods during development in which neural substrates are more susceptible to permanent damage caused by the deprivation of visual input.

In humans, we can take advantage of a rare population of newborns with congenital cataracts in one or both eyes. At birth, patterned visual input is blocked from reaching the retina because of a clouding within the lens, or lenses, of the eye(s).While several causes of congenital cataracts have been identified, most often they are a consequence of an inherited genetic mutation that causes aggregation of the proteins within the lens (Hejtmancik, 2008). Treatment of congenital cataracts involves surgically removing the cataractous lens(es), typically within the first year of life (Maurer, 2017). Corrective lenses are then prescribed to restore patterned vision (see Maurer et al., 2005).

Although vision is typically restored within the first year of life, these patients, when tested as adults, demonstrate lifelong deficits for specific spatial and temporal visual functions (Ellemberg et al., 2000, 2002). The severity of the deficits, however, differs depending on the function itself and whether the deprivation was in one or both eyes (for review, see Maurer, 2017). Deficits seen in unilateral patients, relative to bilateral patients, tend to be more pronounced for lower-level visual functions, likely resulting from abnormal neural connectivity resulting from unfair competition between the eyes (Maurer, 2017). As such, parents of infants born with a unilateral cataract are encouraged to patch the non-deprived eye for the first five years of life following treatment to minimize the effects of this competition (Lewis et al., 1995). Bilateral patients, on the other hand, tend to show more pronounced deficits in relatively higher-level visual functions (Ellemberg et al., 2002; Hadad et al., 2012). In this case, normal visual input in early infancy appears to be necessary to establish the neural architecture that supports higher-level visual functions, even though those functions do not typically emerge until later in development, a process coined the sleeper effect (Maurer et al., 2007). In this case, complete visual deprivation in early infancy, despite being corrected within the first year of life, has detrimental effects on these higher-level visual functions that emerge later in development (Lewis & Maurer, 2009). Just like the animal models discussed previously, both human patient populations demonstrate the existence of critical periods during development where typical sensory input is necessary for normal perceptual development.

More recently, the effects of congenital unilateral or bilateral cataracts have been explored for cross-modal temporal-based interactions. When tested for simultaneity perception, both patient groups, when compared to typically developed controls, demonstrated abnormal audiovisual simultaneity perception (Chen et al., 2017). As demonstrated with certain unimodal visual functions discussed previously, the abnormal pattern of results differed between unilateral and bilateral patients. Specifically, the simultaneity curve produced by unilateral patients was overall wider, resembling that of young developing children, whereas the curve produced by bilateral patients differed asymmetrically with a wider vision-leading side. This implies that the complete absence of vision at birth prevented the normal development of the neural architecture necessary to support audiovisual simultaneity perception. However, when vision was deprived only in one eye, visual input to the non-deprived eye was likely enough to support development of the neural architecture, albeit not to the extent allowing the precision of those with typical visual input since birth. This finding is in line with the differential effects discussed previously between unilateral and bilateral patients observed for unimodal vision. As for the results for visuotactile simultaneity perception, both unilateral and bilateral patients did not differ from typically developed adult controls. This result was surprising given the abnormal audiovisual simultaneity perception observed in both groups and given that both modality pairings involve the previously deprived visual modality, and the visual stimulus was the same in both experiments. This paradox is likely explained best by the cross-modal (or constant) calibration hypothesis (Burr & Gori, 2012; Gori, 2015) discussed previously, in that the more accurate signal calibrates the less accurate signal. During typical development, even though vision has poorer temporal resolution than audition, vision still provides a more reliable signal given that the propagation time for sound varies as a function of distance. Thus, for audiovisual temporal perception, vision tends to calibrate audition. Touch, like audition, also has a high temporal resolution, and given the body-centric nature of tactile stimulation (i.e., signal travel time is not a factor), the temporal processing times are relatively consistent, and only change slightly as a function of the distance from the stimulation on the body to the brain. As such, for visuotactile temporal perception, touch tends to calibrate vision (see Chen et al., 2017 for more details). Consequently, when vision is deprived by cataracts during early development, audiovisual temporal perception cannot be calibrated optimally, whereas visuotactile temporal perception is preserved because of the highly reliable influence of touch.

These same two groups of patients were tested for audiovisual temporal-based integration using the fission (and fusion) illusion (Chen et al., 2014 as cited in Maurer, 2017). Both the unilateral and bilateral patients demonstrated weaker fission illusions than typically developed adult controls, thus suggesting the auditory and visual signals were integrated to a lesser degree. Despite both groups showing abnormal temporal integration, the impact of visual deprivation was more pronounced in the bilateral group than in the unilateral group. Although both low- and high-level neural correlates have been shown to underly fission (see Hirst et al., 2020), the larger deficit observed in the bilateral patients compared to unilateral patients may be attributed to a sleeper effect in higher processing areas. Borrowing from the congenitally blind literature, Maurer (2017) proposes an intriguing hypothesis: perhaps the brief period of visual deprivation prevents the pruning of hyperconnected sensory cortical regions that is seen in normal development resulting in cortical takeover from competing modalities (Maurer et al., 2013). As such, it could be that the takeover process has begun for higher levels of the visual cortex during the period of visual deprivation, causing this sleeper effect in the bilateral patients. As for unilateral patients, it is possible that the normal visual input to one eye prevents this functional takeover of the higher areas of the visual cortex, thus minimizing the impact of visual deprivation on these higher-level functions. Along with the behavioural studies mentioned previously, fMRI results support this hypothesis, showing auditory-driven activation in these corresponding regions of the visual cortex in bilateral patients (Collignon et al., 2015). These patient populations have proven invaluable in not only understanding the effects of a sensory deficit on perception, but also better understanding processes related to normal development.

Audiotactile Perception

Interactions between audition and touch specifically appear to be neglected in the literature relative to cross-modal interactions involving vision (see Occelli et al., 2011; Soto-Faraco & Deco, 2009). This also appears to be the case in the developmental literature. Although audiotactile interactions may not be as ubiquitous as those that involve vision, we still rely on this sensory combination in situations when vision is not available (e.g., navigating a dark environment), or when vision is occupied (e.g., dribbling a ball while running toward a net, reaching for a vibrating and ringing phone while walking down the street, or swatting an insect buzzing around the head, etc.). Audition and touch also share some unique properties: both are highly sensitive to timing (Fujisaki & Nishida, 2009; Occelli et al., 2011), both transduce a physical signal to a neural signal via mechanoreceptors (thus sharing similar encoding mechanisms; von Békésy, 1959), and both share several properties that are not available to or as easily encoded by vision (such as pitch, loudness, volume, rhythm, etc.). This results in a more balanced mutual influence on one another than when either modality is paired with vision. During development, audition and touch are also unique in that they both become "online" before birth, whereas visual stimulation does not truly begin until after birth (Gottlieb, 1971; Lickliter & Bahrick, 2000). While in the womb, one of the first multisensory temporal contingencies the fetus is likely to experience is hearing and feeling the pulse/rhythm of the maternal heartbeat (Lickliter & Bahrick, 2000). Sensitivity to cross-modal interactions between audition and touch has even been measured in the human fetus; differential responses were measured when auditory and vibrotactile stimulation were combined compared to either signal in isolation (Kisilvesky & Muir, 1991), indicating that the fetus is indeed sensitive to temporal cues provided by combined auditory and vibrotactile stimulation in utero (see Lewkowicz, 2000). For this thesis, the audiotactile modality pairing was chosen for the reasons just presented, and because study of this modality pairing will help fill the apparent gap in the literature about cross-modal development of temporal perception and integration in both normal populations and those born with cataracts (see Scope of Thesis below).

Measuring Temporal Perception

A simultaneity judgment task provides one way to measure temporal perception (e.g., Chen et al., 2016, 2018; Stone et al., 2001; Zampini et al., 2005; see Vroomen & Keetels, 2010 for review). Two stimuli from different modalities are presented either at the same time or separated by one of a predetermined range of temporal delays known as stimulus onset asynchronies (or SOA). On each trial, the observer reports whether the two stimuli were perceived as simultaneous or not. Based on the proportion of trials labeled as simultaneous, two measures of temporal perception can be estimated: the point of subjective simultaneity (PSS) and the temporal simultaneity window. The PSS reflects the relative timing difference between the two modalities when the probability of simultaneity is most likely to be perceived. The temporal simultaneity window defines the maximum delay between the two modalities in which simultaneity is still likely to be perceived. Both the PSS and temporal simultaneity window are affected differentially by differences in signal propagation, neural transduction, and neural transmission times of the two modalities. While these measures provide strong proxies of multisensory temporal perception, they do not provide direct measures of multisensory temporal integration. In other words, temporal simultaneity does not *require* integration to determine if two stimuli are perceived as synchronous (see Chen et al., 2016 for discussion). This is not to say that these are not related; multisensory temporal perception is integral to perceiving multisensory temporal integration (e.g., Stevenson et al., 2018). For measuring multisensory temporal integration, a different task is required.

The fission and fusion illusions discussed previously provide a measure of sensory integration. Both illusions occur when an incongruent number of simple stimuli are presented to two modalities in quick succession. Fission, for example, occurs when two stimuli from a distractor modality are paired with a single stimulus from a target modality, resulting in the perception of two target stimuli (Shams et al., 2000, 2002). Fusion, on the other hand, occurs when one stimulus from a distractor modality is paired with two stimuli from a target modality, resulting in the perception of a single fused target stimulus (Andersen et al., 2004). These illusions provide an ideal measure of temporal based sensory integration because the width of the temporal simultaneity window has been shown to be correlated positively with the strength of the fission illusion in the audiovisual domain (Stevenson et al., 2018). Since their discovery, these illusions have been tested in all pairings of the physical senses. For each modality pairing, the modality that provides the more reliable signal is used as the distractor to thus bias the perception of the less reliable target signal (e.g., Bresciani et al., 2008). For audiovisual and visuotactile pairings, audition and touch are more temporally precise than vision (which is more precise in the spatial domain), and thus influence the number of visual stimuli perceived. The reverse illusion is often not observed—for example, two flashes rarely cause the presentation of one beep to be perceived as two beeps. As for the audiotactile pairing, both modalities are highly precise in the temporal domain, thus allowing for a stronger bidirectional influence (Occelli et al., 2011). Consequently, for the audiotactile pairing, both audition and touch have been shown to successfully induce the fission (Bresciani et al., 2005) and fusion (Bresciani & Ernst, 2007) illusions.

Scope of Present Thesis

This thesis explores the broad question of "how do audiotactile temporal interactions change across development?". The first data chapter (Chapter 2) measures the typical development of audiotactile temporal perception. To do this, three groups of children (aged 7-, 9-, and 11-years-old) and one group of adults were tested on a simultaneity judgment task. Results of this chapter are published in the Journal of Experimental Child Psychology, completing a triad of studies by our group that explored the temporal development of all three physical modality pairings (audiovisual: Chen et al., 2016; visuotactile: Chen et al., 2018; audiotactile: Chapter 2). The second data chapter (Chapter 3) charts the typical development of audiotactile temporal integration using the fission and fusion illusions. Results of this chapter are also published in the Journal of Experimental Child Psychology. The magnitude of both illusions was measured in three groups of children (aged 9-, 11-, and 13-years-old) and compared to a group of adults. This study determined not only the typical development of audiotactile integration, but also if temporal perception (as measured using the simultaneity judgment task in Chapter 2) must reach maturity prior to temporal integration, at least for audiotactile interactions. Finally, the third data chapter (Chapter 4) measured the potential impact of visual deprivation at birth on audiotactile temporal perception. To do so, we measured audiotactile simultaneity perception in a group of adults treated for congenital bilateral cataracts. By testing audiotactile temporal perception, we again complete a triad of physical modality pairings by our group (audiovisual and visuotactile: Chen et al., 2017;

audiotactile: Chapter 4). However, because of restrictions imposed on data collection by the COVID-19 pandemic, compromises on the data analyses had to be made (see Chapter 4 for details). The first two data chapters present a comprehensive perspective of the typical development of audiotactile temporal perception as it reaches adult-like maturity, whereas the third chapter explores possible disruptions to this typical development from early-life visual deprivation.

References

Adams, W. J. (2016). The development of audio-visual integration for temporal judgements. *PLoS Computational Biology*, *12*(4), 1–17. https://doi.org/10.1371/journal.pcbi.1004865

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*(3), 301–308. https://doi.org/10.1016/j.cogbrainres.2004.06.004

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, *36*(2), 190–201. https://doi.org/10.1037/0012-1649.36.2.190

Bahrick, L. E., & Lickliter, R. (2004). Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 137–147. https://doi.org/10.3758/CABN.4.2.137

Bahrick, L. E., & Lickliter, R. (2012). The role of intersensory redundancy in early perceptual, cognitive, and social development. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory Development* (pp. 183–206). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199586059.003.0008

Bahrick, L. E., Lickliter, R., & Flom, R. (2004). Intersensory redundancy guides the development of selective attention, perception, and cognition in Infancy. *Current Directions in Psychological Science*, *13*(3), 99–102. https://doi.org/10.1111/j.0963-7214.2004.00283.x

Beierholm, U. R., Quartz, S. R., & Shams, L. (2009). Bayesian priors are encoded independently from likelihoods in human multisensory perception. *Journal of Vision*, *9*(5), 1–9. https://doi.org/10.1167/9.5.23

Bresciani, J. P., & Ernst, M. O. (2007). Signal reliability modulates auditory-tactile integration for event counting. *NeuroReport*, *18*(11), 1157–1161. https://doi.org/10.1097/WNR.0b013e3281ace0ca

Bresciani, J. P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., & Kheddar, A. (2005). Feeling what you hear: Auditory signals can modulate tactile tap perception. *Experimental Brain Research*, *162*(2), 172–180. https://doi.org/10.1007/s00221-004-2128-2

Bresciani, J.-P., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin*, *75*, 753–760. https://doi.org/10.1016/j.brainresbull.2008.01.009

Burr, D., & Gori, M. (2012). Multisensory integration develops late in humans. In *The Neural Bases of Multisensory Processes* (pp. 345–362). CRC Press/Taylor & Francis. https://doi.org/10.1201/b11092-23

Chen, Y. C., Lewis, T. L., Shore, D. I., Spence, C., & Maurer, D. (2018). Developmental changes in the perception of visuotactile simultaneity. *Journal of Experimental Child Psychology*, *173*, 304–317. https://doi.org/10.1016/j.jecp.2018.04.014

Chen, Y. C., Shore, D. I., Lewis, T. L., & Maurer, D. (2016). The development of the perception of audiovisual simultaneity. *Journal of Experimental Child Psychology*, *146*, 17–33. https://doi.org/10.1016/j.jecp.2016.01.010

Chen, Y. C., Lewis, T. L., Shore, D. I., & Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology*, *27*(4), 583–589. https://doi.org/10.1016/j.cub.2017.01.009

Chen, Y. C., & Spence, C. (2017). Assessing the role of the 'Unity Assumption' on multisensory integration: A review. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00445

Collignon, O., Dormal, G., Albouy, G., Vandewalle, G., Voss, P., Phillips, C., & Lepore, F. (2013). Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. *Brain*, *136*(9), 2769–2783. https://doi.org/10.1093/brain/awt176

Collignon, O., Dormal, G., De Heering, A., Lepore, F., Lewis, T. L., & Maurer, D. (2015). Long-lasting crossmodal cortical reorganization triggered by brief postnatal visual deprivation. *Current Biology*, *25*(18), 2379–2383. https://doi.org/10.1016/j.cub.2015.07.036

De Meo, R., Murray, M. M., Clarke, S., & Matusz, P. J. (2015). Top-down control and early multisensory processes: Chicken vs. egg. *Frontiers in Integrative Neuroscience*, *9*. https://doi.org/10.3389/fnint.2015.00017

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, *9*(6), 719–721. https://doi.org/10.1068/p090719

Ellemberg, D., Lewis, T. L., Maurer, D., Brar, S., & Brent, H. P. (2002). Better perception of global motion after monocular than after binocular deprivation. *Vision Research*, *42*(2), 169–179. https://doi.org/10.1016/S0042-6989(01)00278-4

Ellemberg, D., Lewis, T. L., Maurer, D., & Brent, H. P. (2000). Influence of monocular deprivation during infancy on the later development of spatial and temporal vision. *Vision Research*, *40*(23), 3283–3295. https://doi.org/10.1016/S0042-6989(00)00165-6

Emberson, L. L. (2017). How does experience shape early development? Considering the role of top-down mechanisms. *Advances in Child Development and Behavior*, *52*, 1–41. https://doi.org/10.1016/bs.acdb.2016.10.001

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, *7*(5), 1–14. https://doi.org/10.1167/7.5.7

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

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

Fujisaki, W., & Nishida, S. (2009). Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception. *Experimental Brain Research*, *198*(2–3), 245–259. https://doi.org/10.1007/s00221-009-1870-x

Ghazanfar, A., & Schroeder, C. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences*, *10*(6), 278–285. https://doi.org/10.1016/j.tics.2006.04.008

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts.

Gori, M. (2015). Multisensory integration and calibration in children and adults with and without sensory and motor disabilities. *Multisensory Research*, *28*(1–2), 71–99. https://doi.org/10.1163/22134808-00002478

Gori, M., Giuliana, L., Sandini, G., & Burr, D. (2012). Visual size perception and haptic calibration during development. *Developmental Science*, *15*(6), 854–862. https://doi.org/10.1111/j.1467-7687.2012.01183.x

Gottlieb, G. (1971). Ontogenesis of sensory function in birds and mammals. *The Biopsychology of Development*, 67–128.

Grossmann, T., Missana, M., Friederici, A. D., & Ghazanfar, A. A. (2012). Neural correlates of perceptual narrowing in cross-species face-voice matching: ERPs and face-voice matching. *Developmental Science*, *15*(6), 830–839. https://doi.org/10.1111/j.1467-7687.2012.01179.x

Guest, S., Catmur, C., Lloyd, D., & Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research*, *146*(2), 161–171. https://doi.org/10.1007/s00221-002-1164-z

Hadad, B. S., Maurer, D., & Lewis, T. L. (2012). Sparing of sensitivity to biological motion but not of global motion after early visual deprivation. *Developmental Science*, *15*(4), 474–481. https://doi.org/10.1111/j.1467-7687.2012.01145.x

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research*, *186*(4), 517–524. https://doi.org/10.1007/s00221-007-1253-0

Hejtmancik, J. F. (2008). Congenital cataracts and their molecular genetics. *Seminars in Cell & Developmental Biology*, *19*(2), 134–149. https://doi.org/10.1016/j.semcdb.2007.10.003

Hirst, R. J., McGovern, D. P., Setti, A., Shams, L., & Newell, F. N. (2020). What you see is what you hear: Twenty years of research using the Sound-Induced Flash Illusion. *Neuroscience and Biobehavioral Reviews*, *118*, 759–774. https://doi.org/10.1016/j.neubiorev.2020.09.006

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Hubel, D. H., & Wiesel, T. N. (1963). Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of Neurophysiology*, *26*(6), 994–1002. https://doi.org/10.1152/jn.1963.26.6.994

Hubel, D. H., & Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology*, *206*, 419–436.

Innes-Brown, H., Barutchu, A., Shivdasani, M. N., Crewther, D. P., Grayden, D. B., & Paolini, A. (2011). Susceptibility to the flash-beep illusion is increased in children compared to adults. *Developmental Science*, *14*(5), 1089–1099. https://doi.org/10.1111/j.1467-7687.2011.01059.x

Keetels, M., & Vroomen, J. (2012). Perception of synchrony between the senses. In *The neural bases of multisensory processes*.

Kisilvesky, B. S., & Muir, D. W. (1991). Human fetal and subsequent newborn responses to sound and vibration. *Infant Behavior and Development*, *14*(1), 1–26. https://doi.org/10.1016/0163-6383(91)90051-S

Lewis, T. L., & Maurer, D. (2009). Effects of early pattern deprivation on visual development. *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, *86*(6), 640–646. https://doi.org/10.1097/OPX.0b013e3181a7296b

Lewis, T. L., Maurer, D., & Brent, H. P. (1995). Development of grating acuity in children treated for unilateral or bilateral congenital cataract. *Investigative Ophthalmology and Visual Science*, *36*(10), 2080–2095.

Lewkowicz, D. J. (2000). The development of intersensory temporal perception: An epigenetic systems/limitations view. *Psychological Bulletin*, *126*(2), 281.

Lewkowicz, D. J. (2011). Development of multisensory temporal perception. In *The Neural Bases of Multisensory Processes* (pp. 325–343). https://doi.org/10.1201/b11092-22

Lewkowicz, D. J. (2014). Early experience and multisensory perceptual narrowing: Multisensory Perceptual Narrowing. *Developmental Psychobiology*, *56*(2), 292–315. https://doi.org/10.1002/dev.21197

Lewkowicz, D. J., & Bremner, A. J. (2020). The development of multisensory processes for perceiving the environment and the self. In *Multisensory Perception* (pp. 89–112). Elsevier. https://doi.org/10.1016/B978-0-12-812492-5.00004-8

Lewkowicz, D. J., & Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. *Child Development*, *85*(2), 685–694. https://doi.org/10.1111/cdev.12142

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences*, *13*(11), 470–478. https://doi.org/10.1016/j.tics.2009.08.004

Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy*, *15*(1), 46–60. https://doi.org/10.1111/j.1532-7078.2009.00005.x

Lickliter, R., & Bahrick, L. E. (2000). The development of infant intersensory perception: Advantages of a comparative convergent-operations approach. *Psychological Bulletin*, *126*(2), 260–280. https://doi.org/IO.I037//0033-2909.126.2.260

Maurer, D. (2017). Critical periods re-examined: Evidence from children treated for dense cataracts. *Cognitive Development*, *42*, 27–36. https://doi.org/10.1016/j.cogdev.2017.02.006

Maurer, D., Gibson, L. C., & Spector, F. (2013). *Synesthesia in infants and very young children*.

Maurer, D., Lewis, T. L., & Mondloch, C. J. (2005). Missing sights: Consequences for visual cognitive development. *Trends in Cognitive Sciences*, *9*(3 SPEC. ISS.), 144–151. https://doi.org/10.1016/j.tics.2005.01.006

Maurer, D., Mondloch, C. J., & Lewis, T. L. (2007). Effects of early visual deprivation on perceptual and cognitive development. *Progress in Brain Research*, *164*, 87–104.

Noppeney, U. (2021). Solving the causal inference problem. *Trends in Cognitive Sciences*, *25*(12), 1013–1014. https://doi.org/10.1016/j.tics.2021.09.004

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in temporal perception. *Psychonomic Bulletin & Review*, *18*(3), 429–454. https://doi.org/10.3758/s13423-011-0070-4

O'Dowd, A., Cooney, S. M., Sorgini, F., O' Rourke, E., Reilly, R. B., Newell, F. N., & Hirst, R. J. (2021). The development of visuotactile congruency effects for sequences of events. *Journal of Experimental Child Psychology*, *207*, 105094. https://doi.org/10.1016/j.jecp.2021.105094

Redden, R. S., d'Entremont, G., & Klein, R. M. (2017). Safe or out: Does the location of attention affect judgments at first base in baseball? *Canadian Journal of*

Experimental Psychology / Revue Canadienne de Psychologie Expérimentale, *71*(2), 146–159. https://doi.org/10.1037/cep0000118

Seilheimer, R. L., Rosenberg, A., & Angelaki, D. E. (2014). Models and processes of multisensory cue combination. *Current Opinion in Neurobiology*, *25*, 38–46. https://doi.org/10.1016/j.conb.2013.11.008

Shams, L., & Beierholm, U. (2022). Bayesian causal inference: A unifying neuroscience theory. *Neuroscience & Biobehavioral Reviews, 104619*.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152. https://doi.org/10.1016/S0926-6410(02)00069-1

Shams, L., Yukiyasu Kamitani, & Shinsuke Shimojo. (2000). What you see is what you hear. *Nature*, *408*(6814), 788–788.

Soto-Faraco, S., & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. *Behavioural Brain Research*, *196*, 145–154. https://doi.org/10.1016/j.bbr.2008.09.018

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule: Spatial coincidence and multisensory integration. *Annals of the New York Academy of Sciences*, *1296*(1), 31–49. https://doi.org/10.1111/nyas.12121

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. MIT Press.

Stevenson, R. A., Baum, S. H., Krueger, J., Newhouse, P. A., & Wallace, M. T. (2018). Links between temporal acuity and multisensory integration across life span.

Journal of Experimental Psychology: Human Perception and Performance, *44*(1), 106–116. https://doi.org/10.1037/xhp0000424

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., Port, M., & Porter, N. R. (2001). When is now? Perception of simultaneity. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1462), 31–38. https://doi.org/10.1098/rspb.2000.1326

von Békésy, G. (1959). Similarities between hearing and skin sensations. *Psychological Review.*, *66*(1).

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, *72*(4), 871–884. https://doi.org/10.3758/APP.72.4.871

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667. https://doi.org/10.1037/0033-2909.88.3.638

Wiesel, T. N., & Hubel, D. H. (1963). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology*, *26*(6), 1003–1017. https://doi.org/10.1152/jn.1963.26.6.1003

Wiesel, T. N., & Hubel, D. H. (1965). Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens. *Journal of Neurophysiology*, *28*(6), 1029–1040. https://doi.org/10.1152/jn.1965.28.6.1029

Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio–visual simultaneity judgments. *Perception & Psychophysics*, *67*(3), 531–544.

Chapter 2

Developmental Changes in the Perception of Audiotactile Simultaneity

Brendan M. Stanley[a], Yi-Chuan Chen[b], Terri L. Lewis[a],

Daphne Maurer[a], and David I. Shore[a]

Author Note

[a] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Canada

[b] Department of Medicine, Mackay Medical College, New Taipei City, Taiwan

This chapter has been published:

Stanley, B. M., Chen, Y. C., Lewis, T. L., Maurer, D., & Shore, D. I. (2019). Developmental changes in the perception of audiotactile simultaneity. *Journal of Experimental Child Psychology*, *183*, 208–221. https://doi.org/10.1016/j.jecp.2019.02.006

Correspondence concerning this article to be addressed to David I. Shore, Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, L8S 4K1, Canada. Email: dshore@mcmaster.ca

Keywords

Multisensory; Crossmodal; Development; Audition; Touch; Children; Simultaneity judgments; Point of subjective simultaneity

Acknowledgements

DIS was funded by a Discovery grant from NSERC and DM was funded by the James S. McDonnell Foundation and NSERC. We thank Sally Stafford, Miroslav Cika, and Lisa Lorentz for their administrative, technical, and editing contributions, respectively. We also thank Sehar Arfeen, Mathew Klein, Katherine Silang, Kaitlyn Gonsalves, Michelle Ho, Lesley Tsui, and Diana Chu for all of their help with recruiting and data collection.

Abstract

We charted the developmental trajectory of the perception of audiotactile simultaneity by testing three groups of children (7, 9, and 11 years old) and one group of adults. A white noise burst and a tap to the index finger were presented at one of 13 stimulus onset asynchronies (SOAs), and the participants were asked to report whether the two stimuli were simultaneous. Compared to adults, 7-year-olds made significantly more simultaneous responses at 9 of the 13 SOAs, whereas 9-year-olds differed from adults at only two SOAs. The fitted results indicated that the precision of simultaneity perception was lower, and response errors were higher, in younger children than in adults. Eleven-year-olds were adult-like on all measures, thus demonstrating that judgments about simultaneity for audiotactile stimuli mature by 11 years of age. This developmental pattern is similar to that for simultaneity perception for visuotactile stimuli, but later than that for audiovisual stimuli. The longer developmental trajectories of the perception of simultaneity between touch and vision and between touch and audition may arise from the need to coordinate and recalibrate between different reference frames and different neural transmission times in each sensory system during body growth; in addition, the ubiquity of audiovisual experience in everyday life may accelerate the development of that modality pairing.

Introduction

Temporal synchrony provides one of the fundamental cues governing multisensory perception (Stein & Meredith, 1993; Welch & Warren, 1980; see Vroomen & Keetels, 2010 for a review). When two or more stimuli are presented to different modalities in temporal proximity, they tend to be perceived as simultaneous, as if originating from a single event. Spatial proximity provides another cue. At birth, newborns are already sensitive to these parameters for auditory and visual signals (Lewkowicz et al., 2010; Morrongiello et al., 1998) and for visual and tactile signals (Filippetti et al., 2013). For audiotactile pairings, even the fetus is sensitive to temporal correspondence (Kisilvesky & Muir, 1991). Even though these early multisensory abilities establish a foundation, recent studies have demonstrated that the developmental trajectory for the perception of multisensory simultaneity is protracted (e.g., Chen et al., 2016, 2018; Lewkowicz & Flom, 2014; Röder et al., 2013).

Measuring simultaneity perception in older children and adults is typically accomplished with a simultaneity judgment task (e.g., Chen et al., 2016, 2018; Hillock et al., 2011; Hillock-Dunn & Wallace, 2012; Stevenson et al., 2018). In this paradigm, two stimuli, each delivered to a different sensory modality, are presented either simultaneously or separated by predetermined stimulus onset asynchronies (SOAs). Participants report whether the stimuli were perceived as synchronous. Based on the percentage of simultaneous responses as a function of the SOAs, two parameters about multisensory simultaneity perception can be estimated: (a) the point of subjective simultaneity (PSS), which reflects the estimated SOA that yields the highest probability that the stimuli were perceived as simultaneous, and (b) the width of the temporal window, which reflects the range of SOAs at which participants reliably perceived simultaneity above a certain criterion. These two measures of performance are independent of each other given that they tap theoretically different mechanisms; the PSS reflects the relative processing times of the signals in each modality, whereas the width of the window is a proxy for the variability of the arrival times for those signals (see García-Pérez & Alcalá-Quintana, 2012a, 2012b). How these two measures change across development can be used to assess their independence. Specifically, Chen et al. (2016) demonstrated that, for audiovisual pairings, 5- and 7-year-olds have a significantly wider temporal simultaneity window, which reaches adult-like width by 9 years of age[1]. The PSS of the audiovisual simultaneity window was reported on the vision-leading side at 5

[1] It is worth noting that other studies have demonstrated a later age of maturation for the perception of audiovisual simultaneity (Hillock et al., 2011; Hillock-Dunn & Wallace, 2012; Stevenson et al., 2018). However, methodological differences likely accounting for this discrepancy are listed as follows: First, the three studies reporting a later maturation age tested wider age ranges within groups (± 30 months in Hillock-Dunn & Wallace, ± 42 months in Stevenson et al., and unavailable in Hillock et al., versus ± 3 months for Chen et al., 2016 and 2018, and the current study). Second, Hillock et al., Hillock-Dunn & Wallace, and Stevenson et al. used general mathematical fitting methods to estimate the width of the window, whereas the model used by Chen et al., 2016, Chen et al., 2018, and the current study used the model designed specifically for simultaneity judgment tasks developed by García-Pérez & Alcalá-Quintana (2012a,b). This model is based on assumptions of human sensory processing and accounts for response errors and lapses in attention. Third, Hillock et al., Hillock-Dunn & Wallace, and Stevenson et al. do not report how or if they rejected outliers, which can be critical to ensure that the reported results represent only children who concentrated on the task. Any of these factors may contribute to poorer measures and estimates of changes across development, making direct comparisons across the literature impossible (see Chen et al. for a discussion of these factors). Fourth, different criteria were used to determine the width of the audiovisual simultaneity window: 75% simultaneity responses were used in Hillock et al.'s, Hillock-Dunn & Wallace's, and Stevenson et al.'s studies, while 50% simultaneity responses were used in Chen et al., (2016), Chen et al., (2018) and the current study. On the same simultaneity judgment curve, different criterion can yield different measures of the width of the temporal window (see Kaganovich, 2016). In the present study, we used the same methodological details and fitting function as the Chen et al. (2016, 2018) studies, making these the best points of comparison across modality pairings.

years of age (the youngest age group tested), suggesting that, from at least 5 years onward, the visual signal needs to be presented earlier in order for it to be perceived as simultaneous with the auditory signal. This is most likely caused by slower sensory transduction for vision than for audition (Chen et al., 2016). For visuotactile pairings (Chen et al., 2018), children did not reach an adult-like width of the window until 11 years of age, but the PSS was already located on the tactile-leading side by 7 years (the youngest age group tested). Given the rapid transduction of vibrotactile stimulation, this PSS shift toward the tactile-leading side probably results from the neural transmission time taking longer from the finger to the brain than the time it takes to process the visual signal.

The current study aimed to chart the developmental trajectory of audiotactile pairings in order to complete the comparison across all combinations of the three physical senses (i.e., those senses with a spatial representation such as vision, audition, and touch). Importantly, similar stimuli were used across the measurement of the three pairings (audiovisual, visuotactile, and audiotactile). This comparison will help to resolve the question about whether a common mechanism or separate mechanisms underlie simultaneity perception for different modality pairings (e.g., see Vroomen & Keetels, 2010). If the perception of audiotactile simultaneity develops similarly to audiovisual simultaneity, we could infer that audition is the determining factor driving its development, perhaps because of its high temporal resolution and/or processing speed. Alternatively, if the developmental trajectory is found to be similar to visuotactile simultaneity perception, we could infer that the determining factor would be touch, perhaps because it has the slowest processing speed and/or a later unimodal developmental trajectory (see Section Discussion). Finally, if the age of maturation is different from both audiovisual and visuotactile pairings, we would conclude that each modality pairing has its own temporal processing mechanism, and thus, they mature independently of one another.

Audiotactile interactions provide an interesting pairing because of the unique relation between these modalities (see Occelli et al., 2011). Unlike vision, both modalities are stimulated by mechanical displacement of a membrane by air/physical pressure (Soto-Faraco & Deco, 2009). Compared with vision, both modalities also have shorter transduction latencies (Barnett-Cowan & Harris, 2009). In adults, the temporal resolution for the pairing of sound and touch is higher than that for audiovisual or visuotactile pairings (Fujisaki & Nishida, 2009). Among the spatial senses, sound and touch experienced in close temporal proximity have a high likelihood of being causally related to self-generated actions (e.g., texture discrimination; Guest et al., 2002; Jousmäki & Hari, 1998). Taken together, these studies highlight the special nature of audiotactile interactions compared with those that involve vision.

The current study investigated the developmental trajectory for the perception of audiotactile simultaneity by comparing three groups of children (aged 7, 9, and 11 years) to adults. These age groups were chosen because we expected that audiotactile simultaneity perception matures during late childhood based on previous studies of the other modality pairings (Chen et al., 2016, 2018). The simultaneity judgment task required participants to judge whether a beep and tap were presented simultaneously. The stimuli (beeps and taps) were presented either simultaneously or at one of 12 SOAs, ranging from sound leading by 1200 ms to sound lagging by 1200 ms. The dependent variable of interest was the proportion of simultaneous responses at each SOA. These data were subjected to the bootstrap procedure developed by García-Pérez and Alcalá-Quintana (2012a, 2012b) to extract the parameters of PSS and the width of the simultaneity window while estimating parameters associated with peripheral sensory processing and response errors separately (see below). The experimental design, stimuli, and data analysis were similar to those in previous studies on the development of audiovisual and visuotactile simultaneity (Chen et al., 2016, 2018) except for the location of the auditory stimulus (headphones in the current study vs. free field in Chen et al., 2016) and the duration of the stimuli (10 ms in the current study vs. 17 ms in Chen et al., 2016, 2018). Hence, the developmental trajectory of audiovisual, visuotactile, and audiotactile simultaneity perception obtained in these three studies can be compared.

Method

Participants

Four age groups, each with 20 participants, were included in the analyses: 7-year-olds (mean age = 7.0 years, range = 6.8–7.3; 7 boys), 9-year-olds (mean age = 9.1 years, range = 8.9–9.3; 10 boys), 11-year-olds (mean age = 11.1 years, range = 10.8–11.3; 10 boys), and adults (mean age = 22.0 years, range = 17.9–37.1; 10 men). All participants were right-handed and had normal auditory and tactile acuity by self-report, and all passed the Randot test of stereoacuity (minimum of 40 s of arc achieved). The visual test was included to ensure that the participants were comparable to those tested in the previous studies of the perception of audiovisual and visuotactile simultaneity, all of whom had normal binocular vision, the visual function that is most susceptible to abnormal early visual experience. An additional 11 children were tested but excluded because of technical problems (three 11-yearolds) or because they did not pass the stereoacuity test (one 7-year-old and one 9-year-old), did not complete the experiment (two 7-year-olds), or did not pass criterion for the practice block (see below; two 7-year-olds and two 9-year-olds). Children were recruited from a database of parents who, at the time of their children’s birth, consented to be contacted at a later date about participation in developmental studies. The adults were students at McMaster University who participated in exchange for course credit or payment. Prior to the beginning of the experiment, verbal assent to participate was obtained from children in addition to written consent from their parents. Adult participants provided written consent. The study was cleared by the McMaster Research Ethics Board and conformed to the Tri-Council Statement on Ethical Conduct of Research Involving Humans (TCPS2; Canada).

Apparatus and Stimuli

The experiment was conducted in a dimly lit room. The auditory stimuli were 10-ms peak-to-peak white noise bursts with a flat amplitude envelope and 2-ms on- and off-ramping. Auditory stimuli were presented from closed-ear headphones (Sennheiser HDA-200) at 107 dB SPL (measured with a Brüel & Kjær artificial ear Type 4152 and a Brüel & Kjær sound level meter Type 2270). The tactile stimuli were taps with 10 ms duration and were delivered using one of two custom-made tap devices. The first machine consisted of a dull metal pin mounted on a solenoid. The second machine was introduced after the first machine became inoperable and used an electromagnetic solenoid driving a dull tactile stimulator from tactile stimulator (Dancer Design, http://dancerdesign.co.uk/products) mounted in a wooden plank. Both machines, when activated, indented the right index finger well above detection threshold, displacing the skin by approximately 3 mm. About two thirds (13/20) of the data from each group were collected with the first machine. Both the auditory and tactile stimuli were generated and controlled by MATLAB (MathWorks) and the Psychtoolbox-3 package (Brainard, 1997).

To reduce the possibility of the noise produced by the tap device influencing the perception of the auditory stimuli, free-field white noise was played continuously during the experiment (measuring 73 dB SPL at the ear while wearing headphones). In addition, the tactile devices were mounted inside a sound-attenuating box constructed from sound-dampening ceiling tile lined with shag carpet.

Design and Procedure

Two factors were manipulated: age (7-, 9-, and 11-year-old children and adults) and SOA (-1200, -800, -400, -300, -200, -100, 0, 100, 200, 300, 400, 800, and 1200). Negative values of SOA indicate that the auditory stimulus preceded the tactile stimulus, whereas positive values indicate that the auditory stimulus lagged the tactile stimulus. The onset timing and duration of auditory and tactile stimuli at all SOAs were verified using an oscilloscope. In the main experiment, each SOA was tested twice in each of 10 blocks, giving rise to a total of 260 trials.

Participants were instructed to sit facing forward and to keep their eyes closed throughout the experiment. The experimenter was present in the room and verified adherence to these instructions. Participants rested their right index finger over the tactile device, which was positioned approximately 40 cm away from their body along the axis of their midline. Participants were asked to say ‘‘yes’’ if they perceived that the tap and beep were presented at the same time or to say ‘‘no’’ if they perceived that the two stimuli were presented at different times. The experimenter keyed the responses into the computer manually and initiated the next trial.

Two practice sessions were completed prior to the main experiment. The first practice session consisted of 8 trials: 4 with large SOAs (-1200, -800, 800, and 1200) interspersed randomly with 4 0-ms SOAs. Participants were required to achieve 85%

accuracy (maximum of one error; three attempts allowed) to participate in the main experiment. The second practice session included one trial at each of the 13 SOAs used in the main experiment in order to familiarize participants with the entire set of stimuli. This second session, as well as the main experiment, had no accuracy requirement and no feedback except for general encouragement. Children were encouraged to take frequent breaks between blocks, whereas adult participants were encouraged to request a break if needed. Not including breaks, the experiment took approximately 40 min to complete for both children and adults.

Results

Proportion of simultaneous responses

The mean proportion of simultaneous responses was calculated for each participant at each SOA. The data were submitted to a three-way analysis of variance (ANOVA) with age (7-, 9-, 11-yearolds or adults) and machine (in-house built stimulator or Dancer Design tactor) as the between-participant factors, and SOA as the within-participant factor. The factor of machine produced no significant effect or any interaction (all *p*s > .05). As such, all subsequent analyses were based on a two-way ANOVA with the factors of age and SOA (see Figure 1). The Huynh–Feldt estimate of sphericity was used to adjust the *p*-values of this test because of the inclusion of the within-participant factor SOA. Both factors revealed significant main effects: age, $F(3, 76) = 8.53$, $p < .001$, $\eta_p^2 = 0.25$; SOA, $F(5, 362) = 347.92$, $p < .001$, $\eta_p^2 = 0.82$. Most important, the age by SOA interaction was significant, $F(15, 362) = 4.66$, $p < .001$, $\eta_p^2 = 0.16$. A total of 13 one-way ANOVAs—1 at each SOA—were conducted (see Table 1). The main effect of age was significant at 12 of 13 SOAs (all but the -100 ms SOA). Post hoc Dunnett tests (two-tailed) were used to compare each child age group with the adult group at the 12 SOAs (see Table 1). These tests showed that 7-year-olds differed significantly from adults (*p*s < .05) at all but 3 SOAs (-100, 100, and 200 ms), whereas 9-year-olds differed significantly from adults at only 2 SOAs (-1200 and 100 ms). There were no significant differences for 11-year-olds.

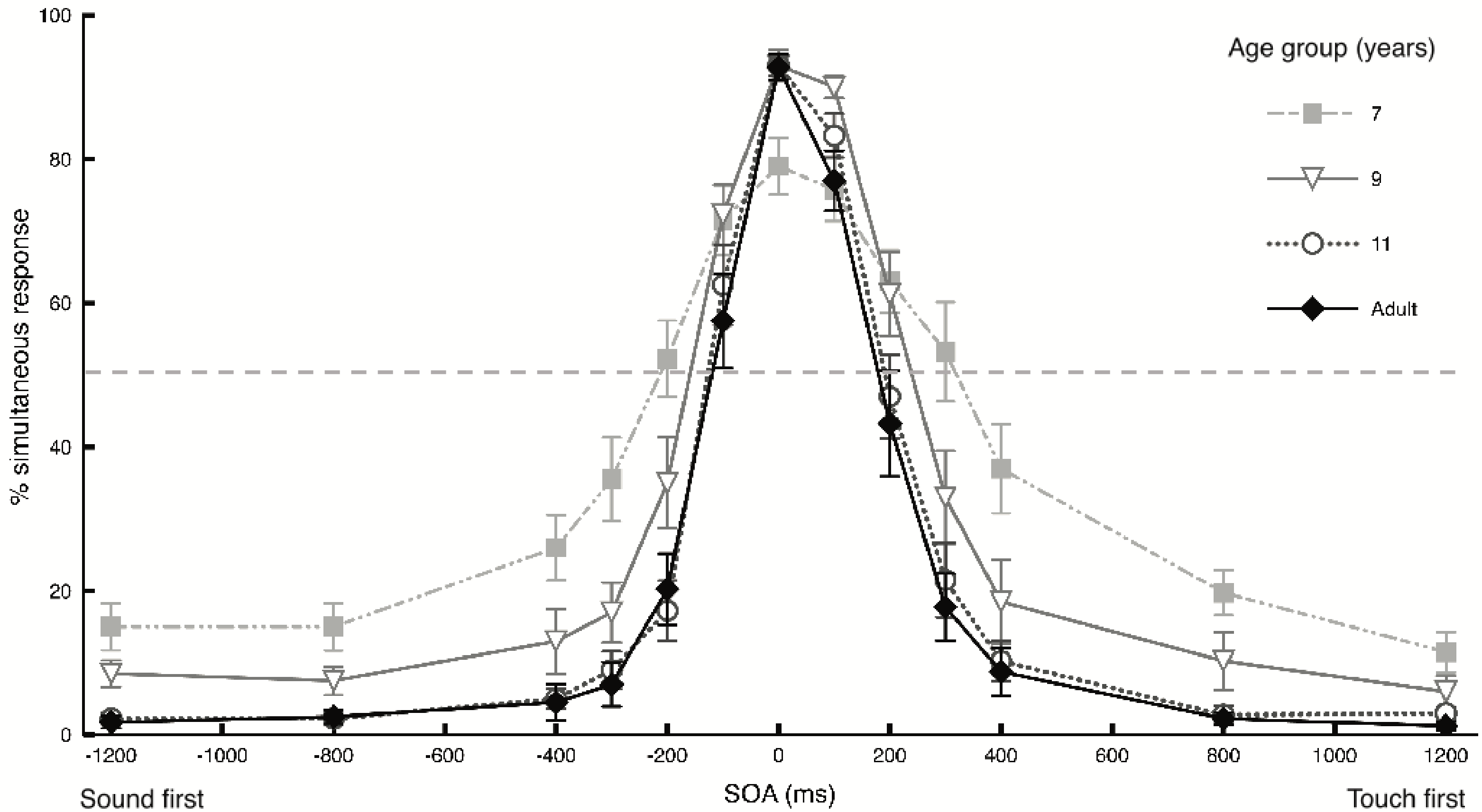

Figure 1. Mean percentage of simultaneous responses at each stimulus onset asynchrony (SOA) for each of the four age groups. Error bars are +/- 1 standard error of the mean.

Table 1.
Results of audiotactile simultaneity judgment task

SOA (ms)	Age group				$F(3,79)$	p	η_p^2	*Post-hoc* tests (*: $p < .05$; **: $p < .01$)
	7	9	11	Adults				
-1200	15.0	8.5	2.3	1.75	10.09	< .001	0.29	7 > Adults **; 9 > Adults *
-800	15.0	7.5	2.3	2.5	8.78	< .001	0.26	7 > Adults **
-400	26.0	13.0	5.0	4.5	8.17	< .001	0.24	7 > Adults **
-300	35.5	17.0	9.0	7.0	10.03	< .001	0.28	7 > Adults **
-200	52.3	35.0	17.3	20.3	9.46	< .001	0.27	7 > Adults **
-100	71.5	72.3	62.5	57.5	1.79	= .16	0.07	
0	79.0	93.0	93.0	92.8	7.56	< .001	0.23	7 > Adults **
100	75.8	90.0	83.3	77.0	3.57	< .05	0.12	9 > Adults *
200	63.0	61.3	47.0	43.3	2.83	< .05	0.10	
300	53.3	33.0	21.5	17.8	7.37	< .001	0.23	7 > Adults **
400	37.0	18.5	10.3	8.8	7.36	< .001	0.23	7 > Adults **
800	19.8	10.3	2.8	2.3	9.46	< .001	0.27	7 > Adults **
1200	11.5	6.0	3.0	1.3	5.90	< .005	0.19	7 > Adults **

Note. Mean percentage of simultaneous responses for each age group, the results of one-way ANOVAs, and post-hoc tests (Dunnett) for the proportion simultaneous at each SOA. Negative SOAs indicate that the sound was presented first, while positive SOAs indicate that the tap was presented first.

Estimated parameters of simultaneity judgments

To estimate the PSS and the width of the temporal simultaneity window, each individual's data were fitted in MATLAB using a bootstrap curve-fitting routine for the simultaneity judgment task (Alcalá-Quintana & García-Pérez, 2013). One strength of this model is that the parameters associated with sensory processing speeds, sensory processing variability, and response errors (e.g., lapses in attention and/or motor response errors) are isolated in order to produce the most accurate estimates of PSS and sensitivity. This approach, first proposed by García-Pérez and Alcalá-Quintana (2012a), was formulated using an independent-channels model (Sternberg & Knoll, 1973), which assumes that signals in each modality (in this case audition and touch) are processed independently before arriving at a central comparator. For each sensory signal, the peripheral processing time and variance from each sensory pathway are estimated using an exponential distribution. The arrival time difference between the two signals at the central comparator (i.e., perceived onset time differences) then forms a bilateral exponential distribution. The arrival time difference is compared against the observer's sensitivity to determine whether a simultaneous response will be made.

In this model, at each SOA a processing time difference between the auditory and tactile stimuli is estimated ($\tau = \tau A - \tau T$), whereas the processing variabilities of the auditory and tactile systems are estimated as λA and λT; respectively. The estimated sensitivity parameter (δ) is the criterion of simultaneity judgments; if the difference of the perceived onsets is smaller than δ, then the observer would make a simultaneous response. The estimated sensitivity parameter (δ) is defined as half the width of the simultaneity window when the simultaneous response criterion is set to 50% on both the auditory- and tactile-leading sides. A low sensitivity (large value of δ) describes a wide simultaneity window; conversely, a high sensitivity (small value of δ) describes a narrow simultaneity window such that the observer has the precision necessary to resolve small differences in stimulus arrival latencies. The PSS was computed as the midpoint of the full width of the temporal simultaneity window. Three response error parameters were also estimated: responding "simultaneous" to either auditory-leading trials (ϵAF) or tactile-leading trials (ϵTF) or responding "not simultaneous" at the 0-ms SOA (ϵS). By accounting for the response errors, the model can make more precise estimates of both the perceptual and sensory parameters.

A one-way ANOVA on the estimates of sensitivity with a between-participant factor of age (see Table 2) revealed a significant main effect, $F(3, 76) = 4.59$, $p < .01$, $\eta_p^2 = 0.05$. Each age group was compared with adults with a post hoc Dunnett test (one-tailed). The 7-year-olds had larger δ values ($p < .005$), indicating lower sensitivity. A similar trend was observed for the 9-year-olds, although this comparison was not significant ($p = .11$). The δ was similar between the 11-year-olds and adults ($p = .78$) (see Figure 2A).

Table 2.
Estimated parameters from fitting model

Parameter	Age group: 7	9	11	Adults	$F(3,76)$	p	η_p^2	*Post-hoc* tests (*: $p < .05$; **: $p < .01$)
δ	**271.7** (27.0)	**228.6** (25.7)	**172.3** (13.6)	**174.7** (20.0)	4.59	< .01	0.05	7 > Adults **
PSS	**57.4** (21.2)	**56.2** (12.6)	**38.5** (6.3)	**35.3** (8.0)	0.75	= .53	0.15	
λA	**0.17** (0.06)	**0.15** (0.07)	**0.34** (0.09)	**0.19** (0.07)	1.28	= .29	0.05	
λT	**0.15** (0.06)	**0.21** (0.06)	**0.11** (0.05)	**0.21** (0.07)	0.59	= .63	0.02	
τ	**-64.2** (31.8)	**-72.6** (21.8)	**-40.4** (15.5)	**-40.5** (14.4)	0.56	= .64	0.02	
εAF	**0.15** (0.03)	**0.08** (0.02)	**0.02** (0.01)	**0.02** (<0.01)	11.62	< .001	0.31	7 > Adults **; 9 > Adults *
εS	**0.11** (0.03)	**0.04** (0.01)	**0.02** (0.01)	**0.04** (0.02)	3.92	< .05	0.13	7 > Adults **
εTF	**0.14** (0.03)	**0.08** (0.02)	**0.03** (0.01)	**0.02** (0.01)	7.37	< .001	0.23	7 > Adults **; 9 > Adults *

Note. The mean (in bold) and SE (in parentheses) of each estimated parameter from the simultaneity judgment task. Results of one-way ANOVAs and *post-hoc* tests (Dunnett) were used. δ: resolution (threshold of simultaneity perception); PSS: point of subjective simultaneity; λA: processing variability of auditory stimulus; λT: processing variability of tactile stimulus; τ: processing time difference between auditory and tactile stimulus (τA-τT); εAF: response errors in the auditory-leading trials; εS: response errors in the simultaneous trials; εTF: response errors in the tactile-leading trials.

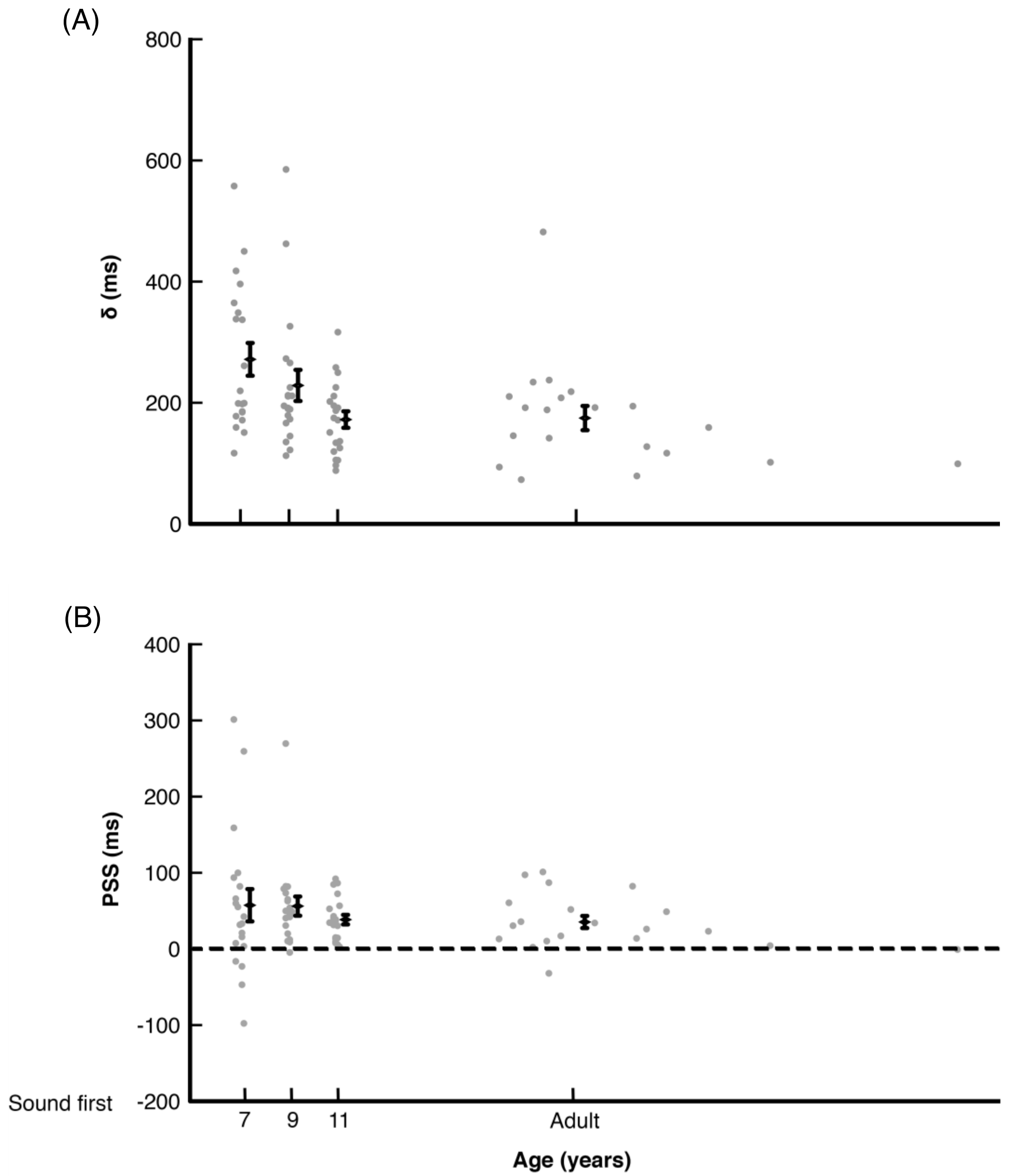

Figure 2. The threshold of simultaneity perception (δ) (A) and point of subjective simultaneity (PSS) (B) as a function of age. The mean PSS in each age group was positive, indicating the conditions where the tactile stimulus was presented first. Gray dots represent individual data, black dots represent the mean for each age group, and error bars indicate ±1 standard error of the mean.

The PSS, corresponding to the midpoint of the audiotactile simultaneity window, was on the tactile-leading side for all age groups: 7-year-olds, $t(19) = 2.70$, $p < .05$, Cohen's d = 0.60; 9-yearolds, t(19) = 4.45, p < .001, Cohen's d = 1.00; 11-year-olds, t(19) = 6.12, p < .001, Cohen's $d = 1.37$; adults, $t(19) = 4.43$, $p < .001$, Cohen's $d = 0.99$. A one-way ANOVA revealed no effect of age (p = .53). The PSS located at the tactile-leading side in all four age groups suggests that the simultaneity window became wider on the touch-leading side than on the auditory-leading side before 7 years, the youngest group tested (see Figure 2B). No age effect was found for the three sensory processing parameters: auditory processing variability (λA), tactile processing variability (λT), and processing arrival time difference (τ) (all ps > .25). Thus, sensory processing for these stimuli was adult-like by 7 years of age. The parameter of processing time difference ($\tau = \tau A - \tau T$) was negative and significantly different from zero in 9-year-olds, $t(19) = -3.33$, $p < .01$, Cohen's $d = 0.74$; 11-year-olds, $t(19) = -2.61$, $p < .05$, Cohen's $d = 0.58$; and adults, $t(19) = -2.81$, $p < .05$, Cohen's $d = 0.63$, with the same trend in 7-year-olds, $t(19) = -2.02$, $p = .06$, Cohen's $d = 0.45$. The negative τ in all age groups suggests that the auditory signal reaches the central comparator prior to the tactile signal when the onset of the two signals is at the same time.

Each of the response error parameters was submitted to a one-way ANOVA with the factor age; post hoc tests used a one-tailed Dunnett test. Auditory-leading (ϵAF), tactile-leading (ϵTF), and simultaneous (ϵS) response errors varied with age, $F(3, 76) = 11.62$, $p < .001$, $\eta_p^2 = 0.31$; $F(3, 76) = 7.37$, $p < .001$, $\eta_p^2 = 0.23$; and $F(3, 76) = 3.92$, $p = .01$, $\eta_p^2 = 0.13$, respectively. When the auditory or tactile stimulus was leading, 7-year-olds (*p*s < .001) and 9-year-olds (*p*s < .05) made more errors than adults, whereas 11-year-olds did not ($p = .35$ for auditory leading and $p = .29$ for tactile leading). When the stimuli were simultaneous (ϵS), only the 7-year-olds showed a higher error rate ($p < .01$); the 9year-olds ($p = .33$) and 11-year-olds ($p = .45$) were not different from adults. In sum, 7-year-olds made more errors in all three conditions, 9-year-olds made more errors only when the stimuli were non-simultaneous, and 11-year-olds made as few errors as adults under all three conditions.

Discussion

The current study charted the developmental trajectory of the perception of audiotactile simultaneity by testing three groups of children (7-, 9-, and 11-year-olds) and adults. The PSS was tactile leading in all age groups, even at the youngest age tested. However, children differed from adults on many other measures. Children aged 7 years made more simultaneous responses across a wide range of SOAs, whereas children aged 9 years did so only at the -1200- and 100-ms SOAs. Based on an instantiation of the independent-channels model for stimulus temporal judgments (García-Pérez & Alcalá-Quintana, 2012a), 7-year-olds had a significantly wider simultaneity window (i.e., larger δ) and a significantly higher rate of response errors in the sound-leading, tap-leading, and simultaneous conditions. Children aged 9 years were still not adult-like on some of the parameters measured; compared with adults, 9-year-olds made more response errors when either the sound or tap led. Taken together, children aged 11 years have reached adult-level performance across all measures obtained from the simultaneity judgment task for audiotactile stimuli.

Children aged 7 years have not yet reached adult-like precision for the perception of audiotactile simultaneity. Specifically, the temporal simultaneity window was significantly wider (i.e., larger δ) for 7-year-olds compared with adults. In other words, these children likely perceive temporally close stimuli as originating from a single event, whereas adults may attribute the same pair of stimuli to different sources. The audiotactile simultaneity window was not statistically wider for 9-year-olds than for adults; however, qualitative observation of the data suggests that some children at this age have much wider simultaneity windows, highlighting the importance of considering individual differences during development (see Figure 2A). Thus, it appears that 9-year-olds are in a transitional stage toward reaching perceptual maturity.

Children aged 7 and 9 years also made more response errors than adults. In the parameter estimation model used here (García-Pérez & Alcalá-Quintana, 2012a), response errors represent participants' lapses in attention and mistakes in motor responses. Thus, the higher response errors in 7- and 9-year-olds suggest an immature system of executive and attentional control. The reduction in response errors to adult levels by 11 years of age is consistent with the literature on the maturation of these systems during late childhood (see Ridderinkhof et al., 1997; Rueda et al., 2004; Shore et al., 2006).

For all ages tested in the current study, the PSS was shifted toward the tactile-leading side, suggesting that the tap and beep were most likely to be perceived as simultaneous when the tap was presented slightly before the beep. The shifts in the PSS may result from at least three sources of time differences: physical signal propagation time in the environment, sensory transduction latency at the receptor, and/or neural transmission time to a central comparator (Stein & Meredith, 1993; Stone et al., 2001). In the current study, the propagation times for the auditory and tactile signals are negligible and essentially the same as each other (sound presented via headphones directly to the ear and tap administered directly to the skin of the fingertip). Sensory transduction latency, in which physical signals are transduced into neural impulses, is also similar for the auditory and tactile modalities given the mechanical nature of the receptor mechanism. Thus, the observed shift of the PSS to the tactile-leading side most likely originates in the differential neural transmission time from the receptors (hair cells in the tympanic membrane vs. mechanoreceptors in the fingertip) to the brain (Barnett-Cowan & Harris, 2009; Stein & Meredith, 1993). To be clear, the tactile stimulus must be earlier than the auditory stimulus because the neural signal has farther to travel (see also Fujisaki & Nishida, 2009; Zampini et al., 2005). The similarity in PSS across the ages tested implies one of two things: either that changes in arm length do not significantly change the time of neural transmission (perhaps because increased myelination speeds up the neural transmission time with age) or that the perceptual system continuously recalibrates to temporal changes as the arms grow (see Ernst, 2008). Critically, the constancy of PSS across development in the face of the aforementioned developmental trajectory for the temporal simultaneity window implies that the specific mechanisms underlying these two measures (PSS and sensitivity) are independent.

In summary, the overall flatter and wider simultaneity judgment curve for 7-year-olds than for adults can be attributed to the young children's immature sensitivity to audiotactile simultaneity, poor response execution, and poor attentional control. Children aged 9 years appear to be in transition between the immature 7-year-olds and the mature 11-year-olds. The 9-year-olds showed greater individual variability in their sensitivity to audiotactile simultaneity and made more errors than adults. In contrast, children aged 11 years performed like adults on all measures and estimated parameters examined.

Comparison across modality pairings

The perception of audiovisual simultaneity develops earlier than that of visuotactile and audiotactile simultaneity (Chen et al., 2016, 2018; current study; see Figure 3). This may derive from the ubiquitous use of audiovisual cues in everyday life, beginning during infancy and continuing throughout life. Specifically, infants demonstrate a rudimentary ability to discriminate the synchrony of auditory and visual stimuli at birth (Lewkowicz et al., 2010). Subsequently, the bulk of language acquisition occurs during early childhood (Bornstein et al., 2004; Vihman, 1993), and speech comprehension requires a combination of auditory and visual signals (Bristow et al., 2008; Chuen & Schutz, 2016; McGurk & Macdonald, 1976; Van Wassenhove et al., 2007; Vatakis & Spence, 2007, 2008; Weatherhead & White, 2017). In contrast, beyond infancy, interactions between touch and vision or touch and audition may not be as pertinent for everyday functions.

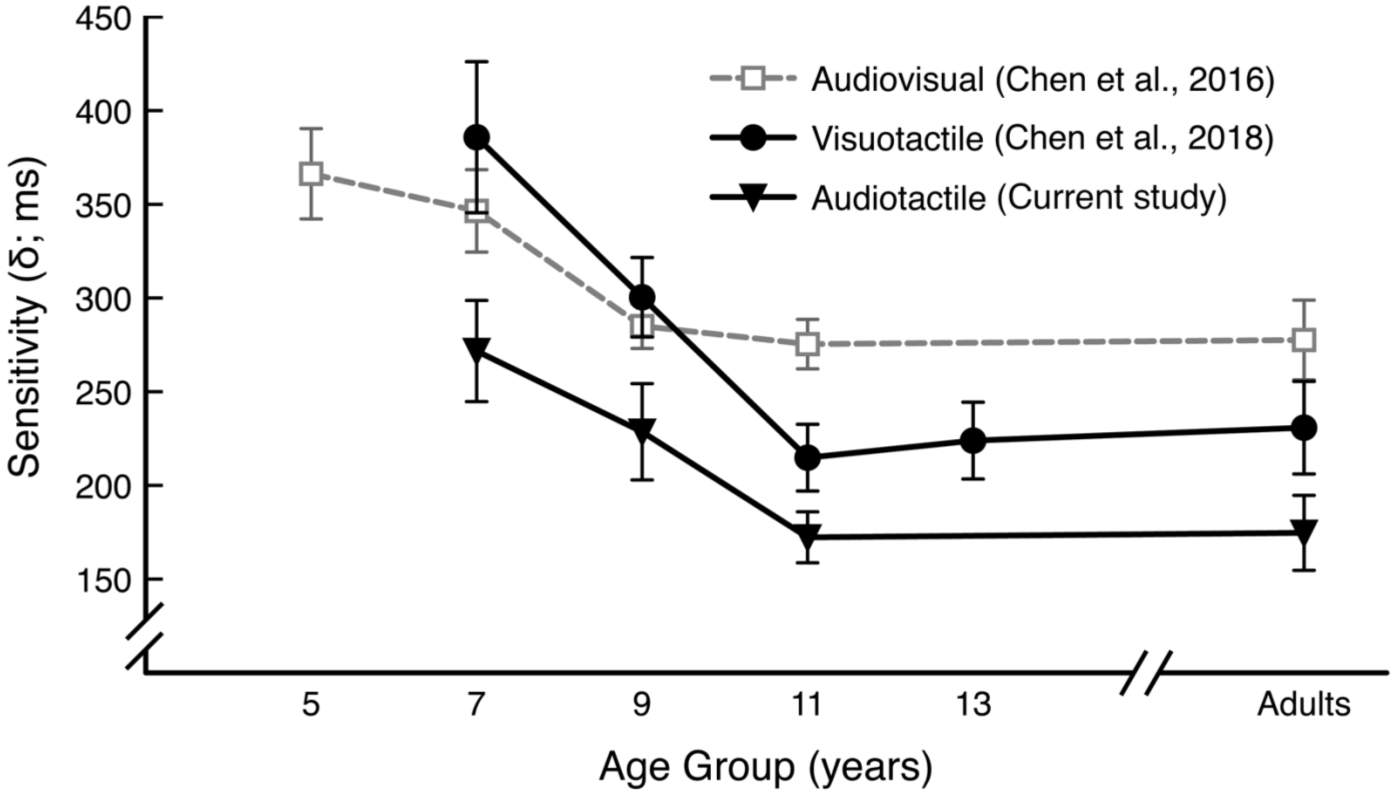

Figure 3. Mean sensitivity (δ) corresponding to half of the width of the temporal simultaneity window for the current audiotactile data compared against audiovisual and visuotactile simultaneity perception. Error bars are +/- 1 standard error of the mean.

Alternatively, the earlier maturation of adult-like precision in perceiving audiovisual simultaneity than of those pairings involving touch may derive from the unique body-based nature of tactile processing. Tactile signals are first encoded in a somatotopic frame of reference and are then translated into an environmental frame of reference (Azañón & Soto-Faraco, 2008; Heed & Azañón, 2014; Shore et al., 2002), whereas visual and auditory signals are represented in an environmental frame of reference from the start of their processing. Such coordination between somatosensory and environmental frames of reference continues throughout development but differs between childhood and puberty. Specifically, during childhood the size of the body and limbs change nonlinearly while the size of the head remains stable, and then during puberty the body, limbs, and head grow proportionally (see Bremner et al., 2012, for a review). Significant cognitive improvement in the coordination between the internal (somatotopic) and external (environmental) frames of reference, specifically in terms of visual perspective taking and body ownership, also occurs around the same ages during late childhood (see Pearson et al., 2016). Hence, the protracted development of the perception of visuotactile and audiotactile simultaneity may be caused in part by the need to coordinate changing frames of reference until the end of childhood. In contrast, the coordination between vision and audition may be completed earlier because they use the same frame of reference.

A third possibility arises from the travel distances of neural signals from sensory organs to the brain and how they change with body growth. Specifically, auditory and visual signals need to travel a short distance (i.e., cochlea to the temporal cortex and retina to the occipital cortex), which does not change significantly across development. In contrast, tactile neural signals need to travel from the receptors on the skin of the finger to the brain; this distance is longer and changes considerably as the body grows. Although this had no effect on the PSS, the required recalibration may prevent a fixed and stable temporal simultaneity window from developing until relative growth rates become proportional. Consistent with this extended flexibility, we note that unisensory temporal resolution for touch matures later (perhaps by 6 years of age; see Pagel et al., 2009) than for vision (by 4 years of age; Ellemberg et al., 1999) or audition (by 4 years of age; Wightman et al., 1989). Thus, we propose that the perceptual system constantly recalibrates the perceived onset of tactile stimuli but does not crystalize the temporal simultaneity window for this modality until the body grows proportionately.

Note that the three explanations are not mutually exclusive but rather more likely to contribute conjointly to the development of the perception of multisensory simultaneity. Other non-physiological factors such as the development of executive functioning, cognitive complexity, and decision making may also influence the developmental trajectories of simultaneity perception. Regardless, audiovisual interactions appear to be privileged in terms of perceptual maturation, likely because of the extensive daily occurrences of audiovisual events. In contrast, the challenge of coordination and recalibration of spatial and temporal representation or processing across sensory modalities during body growth may lead to the longer developmental trajectories for audiotactile and visuotactile than for audiovisual simultaneity perception.

Single or multiple mechanisms for multisensory simultaneity perception

The different developmental trajectories described above suggest unique mechanisms for each modality pairing rather than a unitary mechanism for all pairings (see also (Harrar & Harris, 2005, 2008; Van der Burg et al., 2013); but see (Hanson et al., 2008; Machulla et al., 2016). Studies of temporal recalibration (see (Van der Burg et al., 2013) provide another example supporting unique mechanisms. Audiovisual recalibration can occur on a trial-by-trial basis, whereas perception of visuotactile and audiotactile pairings recalibrates at a slower timescale (Van der Burg, Alais, et al., 2015; Van der Burg, Orchard-Mills, et al., 2015). This difference may arise because audiovisual signals occur external to the body and temporal recalibration must account for the constant changes in the distance of the signals (Engel & Dougherty, 1971; Sugita & Suzuki, 2003; see Vroomen & Keetels, 2010, for a review). Audiotactile and visuotactile signals, however, originate on the body surface, and so their propagation time differences (from stimulus origin to sensory receptors) are relatively constant, and recalibration might not need to occur on a rapid timescale. A third piece of evidence comes from studies of temporal perception in a population of adults treated for congenital cataracts (Chen et al., 2017). These individuals were born fully deprived of patterned vision in either one eye or both eyes because of a dense cataract or cataracts. In developed countries, full vision is typically restored within the first 6 months of life through the removal of the cataractous lens followed by the use of corrective lenses. When tested as adults, these patients show deficits in the perception of audiovisual simultaneity, whereas the perception of visuotactile simultaneity is spared (Chen et al., 2017). These findings further support the claim that different mechanisms are involved in the perception of simultaneity for different modality pairings.

That said, it remains unclear the extent to which the perception of visuotactile simultaneity and that of audiotactile simultaneity share the same process or mechanism. Current evidence shows that they mature at the same age (Chen et al., 2018; current study). Nevertheless, testing groups in finer age categories may reveal subtle developmental differences. In addition, examining the influence of transient early visual deprivation on audiotactile temporal perception in cataract-reversal patients should provide additional evidence on the extent to which the perceptions of audiovisual, visuotactile, and audiotactile simultaneity are mediated by overlapping or separate underlying mechanisms.

Relation between simultaneity perception and multisensory integration

Two distinct processes contribute to our coherent perception of multisensory events. One determines whether two signals presented to different modalities originate from the same event, known as the unity assumption (cf. Chen & Spence, 2017) or the unity prior in the process of causal inference (e.g., Odegaard et al., 2017; Odegaard & Shams, 2016). In this process, temporal coincidence provides a critical cue that two signals should be integrated into a single percept (Stein & Meredith, 1993; see Vroomen & Keetels, 2010, for a review). The other process follows the rules of multisensory optimal integration by weighing each sensory signal in terms of its reliability (e.g., Alais & Burr, 2004; Ernst & Banks, 2002; Ernst & Bülthoff, 2004). A comparison of the developmental trajectories of temporal perception and multisensory integration provides insight into the relation between these two processes. To the best of our knowledge, the only studies to examine the development of audiotactile integration demonstrated that 11-year-olds did not integrate audiotactile cues associated with object size (e.g., bigger object makes louder sound when hitting the floor) as optimally as adults (Petrini et al., 2014), and children aged 13–15 years are still in a transitional stage (Scheller et al., 2018). In the context of the current findings of adult-like perception of audiotactile simultaneity by 11 years of age, we might propose that the maturation of simultaneity perception is a prerequisite for optimal integration, at least for the audiotactile pairing. In other words, one may assume that unity must be established before accurate, or optimal, multisensory integration can occur (e.g., Welch & Warren, 1980).

Examinations of other modality pairings, however, present a different and more complex picture. The development of audiovisual optimal integration appears to be adult-like by 8 years of age in the temporal domain (the youngest age tested; see Adams, 2016, for evidence of optimal integration emerging at 10 years of age), whereas adult-like integration does not emerge until after 12 years of age in the spatial domain (the oldest age tested; Gori et al., 2012). In contrast, the maturation of the perception of audiovisual simultaneity occurs between these two ages, by 9 years (Chen et al., 2016). The optimal integration for visuotactile shape and orientation perception matures at around 8–10 years of age, similar to the maturation of visuotactile simultaneity perception (by 11 years; Chen et al., 2018). Thus, the order in which simultaneity perception and optimal integration reach maturity differs depending on the sensory pairing and the stimulus domain(s) under examination.

The above findings seem to suggest that the developmental trajectories of multisensory simultaneity perception and optimal integration might not be sequential. However, the nature of the tasks used to measure simultaneity perception and optimal integration does not support direct comparisons to test whether temporal perception and multisensory integration are sequential or interdependent. The only way to properly compare the maturation of temporal perception and optimal integration across modality pairings would require using the same stimuli to measure both abilities using tasks with similar demands. Consequently, the only statement that can be made with any degree of certainty is that the development of multisensory simultaneity perception and optimal integration are prolonged into late childhood or early adolescence for all modality pairings.

Conclusion

This is the first study to measure the developmental trajectory of the perception of audiotactile simultaneity—a modality combination that has rarely been studied to date. Children younger than 11 years had wider windows of simultaneity perception and were likely to make response errors. The PSS, on the other hand, was shifted to the tactile-leading side by 7 years, the youngest age tested. The later developmental trajectories for the perception of audiotactile and visuotactile simultaneity than for audiovisual simultaneity, in combination with evidence from cataract-reversal patients, suggest different underlying processes of simultaneity perception for different modality pairings.

References

Adams, W. J. (2016). The development of audio-visual integration for temporal judgements. *PLoS Computational Biology*, *12*(4), 1–17. https://doi.org/10.1371/journal.pcbi.1004865

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods*, *45*(4), 972–998. https://doi.org/10.3758/s13428-013-0325-2

Azañón, E., & Soto-Faraco, S. (2008). Changing reference frames during the encoding of tactile events. *Current Biology*, *18*(14), 1044–1049. https://doi.org/10.1016/j.cub.2008.06.045

Barnett-Cowan, M., & Harris, L. R. (2009). Perceived timing of vestibular stimulation relative to touch, light and sound. *Experimental Brain Research*, *198*(2–3), 221–231. https://doi.org/10.1007/s00221-009-1779-4

Bornstein, M. H., Hahn, C.-S., & Haynes, O. M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language*, *24*(3), 267–304. https://doi.org/10.1177/0142723704045681

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Bremner, A. J., Holmes, N. P., & Spence, C. (2012). *The development of multisensory representations of the body and of the space around the body*.

Bristow, D., Dehaene-Lambertz, G., Mattout, J., Soares, C., Gliga, T., Baillet, S., & Mangin, J.-F. (2008). Hearing Faces: How the Infant Brain Matches the Face It Sees with the Speech It Hears. *Journal of Cognitive Neuroscience*, *21*(5), 905–921. https://doi.org/10.1162/jocn.2009.21076

Chen, Y. C., Lewis, T. L., Shore, D. I., Spence, C., & Maurer, D. (2018). Developmental changes in the perception of visuotactile simultaneity. *Journal of Experimental Child Psychology*, *173*, 304–317. https://doi.org/10.1016/j.jecp.2018.04.014

Chen, Y. C., Shore, D. I., Lewis, T. L., & Maurer, D. (2016). The development of the perception of audiovisual simultaneity. *Journal of Experimental Child Psychology*, *146*, 17–33. https://doi.org/10.1016/j.jecp.2016.01.010

Chen, Y. C., Lewis, T. L., Shore, D. I., & Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology*, *27*(4), 583–589. https://doi.org/10.1016/j.cub.2017.01.009

Chen, Y. C., & Spence, C. (2017). Assessing the role of the ‘Unity Assumption’ on multisensory integration: A review. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00445

Chuen, L., & Schutz, M. (2016). The unity assumption facilitates cross-modal binding of musical, non-speech stimuli: The role of spectral and amplitude envelope cues.

Attention, Perception, & Psychophysics, *78*(5), 1512–1528. https://doi.org/10.3758/s13414-016-1088-5

Ellemberg, D., Lewis, T. L., Hong Liu, C., & Maurer, D. (1999). Development of spatial and temporal vision during childhood. *Vision Research*, *39*(14), 2325–2333. https://doi.org/10.1016/S0042-6989(98)00280-6

Engel, G. R., & Dougherty, W. G. (1971). Visual-auditory distance constancy. *Nature*, *234*(5327), 308–308.

Ernst, M. O. (2008). Multisensory Integration: A Late Bloomer. *Current Biology*, *18*(12), 519–521. https://doi.org/10.1016/j.cub.2008.05.003

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

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

Filippetti, M. L., Johnson, M. H., Lloyd-Fox, S., Dragovic, D., & Farroni, T. (2013). Body perception in newborns. *Current Biology*, *23*(23), 2413–2416. https://doi.org/10.1016/j.cub.2013.10.017

Fujisaki, W., & Nishida, S. (2009). Audio-tactile superiority over visuo-tactile and audio-visual combinations in the temporal resolution of synchrony perception. *Experimental Brain Research*, *198*(2–3), 245–259. https://doi.org/10.1007/s00221-009-1870-x

García-Pérez, M. A., & Alcalá-Quintana, R. (2012a). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model.

Psychonomic Bulletin & Review, *19*, 820–846. https://doi.org/10.3758/s13423-012-0278-y

García-Pérez, M. A., & Alcalá-Quintana, R. (2012b). Response errors explain the failure of independent-channels models of perception of temporal order. *Frontiers in Psychology*, *3*(APR), 1–13. https://doi.org/10.3389/fpsyg.2012.00094

Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in Integrative Neuroscience*, *6*(SEPTEMBER), 1–8. https://doi.org/10.3389/fnint.2012.00077

Guest, S., Catmur, C., Lloyd, D., & Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research*, *146*(2), 161–171. https://doi.org/10.1007/s00221-002-1164-z

Hanson, J. V. M., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Experimental Brain Research*, *185*(2), 347–352. https://doi.org/10.1007/s00221-008-1282-3

Harrar, V., & Harris, L. R. (2005). Simultaneity constancy: Detecting events with touch and vision. *Experimental Brain Research*, *166*(3–4), 465–473. https://doi.org/10.1007/s00221-005-2386-7

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research*, *186*(4), 517–524. https://doi.org/10.1007/s00221-007-1253-0

Heed, T., & Azañón, E. (2014). Using time to investigate space: A review of tactile temporal order judgments as a window onto spatial processing in touch. *Frontiers in Psychology*, *5,* 76–76. https://doi.org/10.3389/fpsyg.2014.00076

Hillock, A. R., Powers, A. R., & Wallace, M. T. (2011). Binding of sights and sounds: Age-related changes in multisensory temporal processing. *Neuropsychologia*, *49*(3), 461–467. https://doi.org/10.1016/j.neuropsychologia.2010.11.041

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science*, *15*(5), 688–696. https://doi.org/10.1111/j.1467-7687.2012.01171.x

Jousmäki, V., & Hari, R. (1998). Parchment-skin illusion: Sound-biased touch. *Current Biology : CB*, *8*(6), 190–191. https://doi.org/10.1016/s0960-9822(98)70120-4

Kaganovich, N. (2016). Development of sensitivity to audiovisual temporal asynchrony during midchildhood. *Developmental Psychology*, *52*(2), 232–241. https://doi.org/10.1037/dev0000073

Kisilvesky, B. S., & Muir, D. W. (1991). Human fetal and subsequent newborn responses to sound and vibration. *Infant Behavior and Development*, *14*(1), 1–26. https://doi.org/10.1016/0163-6383(91)90051-S

Lewkowicz, D. J., & Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. *Child Development*, *85*(2), 685–694. https://doi.org/10.1111/cdev.12142

Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy*, *15*(1), 46–60. https://doi.org/10.1111/j.1532-7078.2009.00005.x

Machulla, T.-K., Di Luca, M., & Ernst, M. O. (2016). The consistency of crossmodal synchrony perception across the visual, auditory, and tactile Senses. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 1026–1038. https://doi.org/10.1037/xhp0000191

McGurk, H., & Macdonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748.

Morrongiello, B. A., Fenwick, K. D., & Chance, G. (1998). Crossmodal learning in newborn infants: Inferences about properties of auditory-visual events. *Infant Behavior and Development*, *21*(4), 543–553. https://doi.org/10.1016/s0163-6383(98)90028-5

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in temporal perception. *Psychonomic Bulletin & Review*, *18*(3), 429–454. https://doi.org/10.3758/s13423-011-0070-4

Odegaard, B., & Shams, L. (2016). The Brain's Tendency to Bind Audiovisual Signals Is Stable but Not General. *Psychological Science*, *27*(4), 583–591. https://doi.org/10.1177/0956797616628860

Odegaard, B., Wozny, D. R., & Shams, L. (2017). A simple and efficient method to enhance audiovisual binding tendencies. *PeerJ*, *5*, e3143. https://doi.org/10.7717/peerj.3143

Pagel, B., Heed, T., & Röder, B. (2009). Change of reference frame for tactile localization during child development: Reference frame for tactile localization in children. *Developmental Science*, *12*(6), 929–937. https://doi.org/10.1111/j.1467-7687.2009.00845.x

Pearson, A., Marsh, L., Ropar, D., & Hamilton, A. (2016). Cognitive Mechanisms underlying visual perspective taking in typical and ASC children: Cognitive mechanisms of perspective taking. *Autism Research*, *9*(1), 121–130. https://doi.org/10.1002/aur.1501

Petrini, K., Remark, A., Smith, L., & Nardini, M. (2014). When vision is not an option: Children's integration of auditory and haptic information is suboptimal. *Developmental Science*, *17*(3), 376–387. https://doi.org/10.1111/desc.12127

Ridderinkhof, K. R., Van Der Molen, M. W., Band, G. P. H., & Bashore, T. R. (1997). Sources of Interference from Irrelevant Information: A Developmental Study. *Journal of Experimental Child Psychology*, *65*(3), 315–341. https://doi.org/10.1006/jecp.1997.2367

Röder, B., Pagel, B., & Heed, T. (2013). The implicit use of spatial information develops later for crossmodal than for intramodal temporal processing. *Cognition*, *126*(2), 301–306. https://doi.org/10.1016/j.cognition.2012.09.009

Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*(8), 1029–1040. https://doi.org/10.1016/j.neuropsychologia.2003.12.012

Scheller, M., Proulx, M., & Petrini, K. (2018). *Audio-haptic cue integration across the lifespan*. 19th International Multisensory Research Forum.

Shore, D. I., Burack, J. A., Miller, D., Joseph, S., & Enns, J. T. (2006). The development of change detection. *Developmental Science*, *9*(5), 490–497. https://doi.org/10.1111/j.1467-7687.2006.00516.x

Shore, D. I., Spry, E., & Spence, C. (2002). Confusing the mind by crossing the hands. *Cognitive Brain Research*, *14*(1), 153–163. https://doi.org/10.1016/S0926-6410(02)00070-8

Soto-Faraco, S., & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. *Behavioural Brain Research*, *196*, 145–154. https://doi.org/10.1016/j.bbr.2008.09.018

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. MIT Press.

Sternberg, S., & Knoll, R. L. (1973). The perception of temporal order: Fundamental issues and a general model. *Attention and Performance*, *IV*, 629–685.

Stevenson, R. A., Baum, S. H., Krueger, J., Newhouse, P. A., & Wallace, M. T. (2018). Links between temporal acuity and multisensory integration across life span. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 106–116. https://doi.org/10.1037/xhp0000424

Sugita, Y., & Suzuki, Y. (2003). Implicit estimation of sound-arrival time. *Nature*, *421*(6926), 911. https://doi.org/10.1038/421911a

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid recalibration to audiovisual asynchrony. *Journal of Neuroscience*, *33*(37), 14633–14637. https://doi.org/10.1523/JNEUROSCI.1182-13.2013

Van der Burg, E., Alais, D., & Cass, J. (2015). Audiovisual temporal recalibration occurs independently at two different time scales. *Scientific Reports*, *5*(1), 14526. https://doi.org/10.1038/srep14526

Van der Burg, E., Orchard-Mills, E., & Alais, D. (2015). Rapid temporal recalibration is unique to audiovisual stimuli. *Experimental Brain Research*, *233*(1), 53–59. https://doi.org/10.1007/s00221-014-4085-8

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*, *45*(3), 598–607. https://doi.org/10.1016/j.neuropsychologia.2006.01.001

Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the "unity assumption" using audiovisual speech stimuli. *Perception & Psychophysics*, *69*(5), 744–756. https://doi.org/10.3758/BF03193776

Vatakis, A., & Spence, C. (2008). Evaluating the influence of the 'unity assumption' on the temporal perception of realistic audiovisual stimuli. *Acta Psychologica*, *127*(1), 12–23. https://doi.org/10.1016/j.actpsy.2006.12.002

Vihman, M. M. (1993). *Phonological development: The origins of language in the child*. Blackwell Publishing.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, *72*(4), 871–884. https://doi.org/10.3758/APP.72.4.871

Weatherhead, D., & White, K. S. (2017). Read my lips: Visual speech influences word processing in infants. *Cognition*, *160*, 103–109. https://doi.org/10.1016/j.cognition.2017.01.002

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667. https://doi.org/10.1037/0033-2909.88.3.638

Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child Development*, *60*(3), 611–624. https://doi.org/10.1111/j.1467-8624.1989.tb02742.x

Chapter 3

Developmental Changes in Audiotactile Event Perception

Brendan M. Stanley[a], Yi-Chuan Chen[b],

Daphne Maurer[a], Terri L. Lewis[a], and David I. Shore[a,c]

Author Note

[a] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Canada

[b] Department of Medicine, Mackay Medical College, New Taipei City, Taiwan

[c]Multisensory Perception Laboratory, Division of Multisensory Mind Inc., Hamilton, Canada

This chapter has been published:

Stanley, B. M., Chen, Y.-C., Maurer, D., Lewis, T. L., & Shore, D. I. (2023). Developmental changes in audiotactile event perception. *Journal of Experimental Child Psychology*, *230*, 105629. https://doi.org/10.1016/j.jecp.2023.105629

Correspondence concerning this article to be addressed to David I. Shore, Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, L8S 4K1, Canada. Email: dshore@mcmaster.ca

Abstract

The fission and fusion illusions provide measures of multisensory integration. The sound-induced tap fission illusion occurs when a tap is paired with two distractor sounds, resulting in the perception of two taps; the sound-induced tap fusion illusion occurs when two taps are paired with a single sound, resulting in the perception of a single tap. Using these illusions, we measured integration in three groups of children (9-, 11-, and 13-year-olds) and compared them with a group of adults. Based on accuracy, we derived a measure of magnitude of illusion and used a signal detection analysis to estimate perceptual discriminability and decisional criterion. All age groups showed a significant fission illusion, whereas only the three groups of children showed a significant fusion illusion. When compared with adults, the 9-year-olds showed larger fission and fusion illusions (i.e., reduced discriminability and greater bias), whereas the 11-year-olds were adult-like for fission but showed some differences for fusion: significantly worse discriminability and marginally greater magnitude and criterion. The 13-year-olds were adult-like on all measures. Based on the pattern of data, we speculate that the developmental trajectories for fission and fusion differ. We discuss these developmental results in the context of three non-mutually exclusive theoretical frameworks: sensory dominance, maximum likelihood estimation, and causal inference.

Keywords

Multisensory; Development; Audition; Touch; Fission; Fusion

Acknowledgements

David I. Shore was funded by a Discovery grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Daphne Maurer was funded by the James S. McDonnell Foundation and NSERC. We thank Sally Stafford and Miroslav Cika for their administrative and technical contributions, respectively. We also thank Sehar Arfeen, Katherine Silang, Vicki Kam, Sheila Wang, Michelle Ho, and Abir Shamim for all their help with recruitment and data collection.

Introduction

The perceptual system tends to integrate sensory signals from different modalities when they are likely originating from the same event (Shams & Beierholm, 2010; Welch & Warren, 1980). Such multisensory integration allows us to perceive, interact with, and navigate through the world more precisely (Gori et al., 2008; Nardini et al., 2008; see Ernst & Bülthoff, 2004, for an early review). Within this context, the integration of audition and touch provides an important perspective. More specifically, touch may provide a scaffold for the development of the other spatial senses because of its direct and proximal perception. However, touch and especially its interactions with audition are underrepresented in developmental psychology (Bremner & Spence, 2017). Thus, to help fill that gap, we examined audiotactile integration across middle childhood.

The somatosensory sense develops during gestation (~1–2 months; see Bremner et al., 2012), receiving stimulation as the fetus bumps into itself, the uterus, or the umbilical cord (e.g., Hooker, 1958). This is long before the auditory and visual systems become functional and receive stimulation (~7 months gestational age and after birth, respectively). Near birth, the fetus shows sensitivity to cross-modal auditory–vibrotactile stimulation (Kisilvesky & Muir, 1991), whereas functional visuotactile connections appear to develop postnatally (Begum Ali et al., 2015; Held et al., 2011; Maurer et al., 1999). From birth onward, multisensory integration refines and follows unique trajectories depending on the type of task (e.g., Gori et al., 2021) and modality pairings (e.g., Stanley et al., 2019 or Chapter 2, Figure 3). Most often, adult-like performance is reached around middle to late childhood (see Burr & Gori, 2012), but development may be more protracted when one of the modalities is touch (e.g., (Gori et al., 2008; Petrini et al., 2014; Scheller et al., 2021; Stanley et al., 2019 or Chapter 2).

The few developmental cross-modal studies specifically examining the integration of hearing and touch find a late maturity. Although there are no studies yet examining the development of integration with passive tactile perception, two studies examined the integration of audition and active haptic size discrimination using the same task (Petrini et al., 2014; Scheller et al., 2021). Using a child-friendly design, participants actively patted a comparison ball either before or after patting a standard ball; the task was to indicate which was larger. On bimodal trials, a pre-recorded sound of a ball hitting a table was presented. The sound was either size congruent (the loudness of the sound corresponded to the size of the comparison ball) or size incongruent (the size of the comparison ball and the loudness of the sound averaged to match the size of the standard). Using maximum likelihood estimation, these authors determined the relative reliability of the auditory and haptic cues and determined if and when observers combine multisensory cues optimally. Scheller et al. (2021) demonstrated that adult-like optimal integration emerged between 13 and 15 years of age. This appears to be the latest age of maturation reported to date in the developmental literature for all possible pairings of cross-modal integration (see Scheller et al., 2021, Figure 9). This late development of audiohaptic integration may have to do with the active nature of this task; it may index processes beyond simple touch, including proprioception, efference copy, attention, and other higher-order cognitive processes such as the prior probability of cue combination (see Discussion).

Stanley et al. (2019; Chapter 2) charted the development of simultaneity perception for audition and passive touch. Although audiotactile integration was not measured in this study, simultaneity perception provides one of several cues that the perceptual system relies on to support multisensory integration (see Chen et al., 2016, for discussion). In a cross-modal simultaneity judgment task, participants are presented with two stimuli from different modalities either coincidently or separated by one of several stimulus onset asynchronies (SOAs); participants' task is to report whether the two stimuli were simultaneous or not (Chen et al., 2016, 2018; Machulla et al., 2016; Stone et al., 2001; Zampini et al., 2005). Based on the proportion of simultaneous responses at each SOA, the temporal window of simultaneity is estimated. With the pairing of audition and passive tactile stimulation, Stanley et al. (2019; Chapter 2) demonstrated an adult-like maturity by 11 years of age; the same age as visuotactile (Chen et al., 2018), and approximately two years later than for audiovisual (Chen et al., 2016) simultaneity perception. Although both Stanley et al. (2019; Chapter 2) and Scheller et al. (2021) agree that interactions between audition and touch mature late, clear questions emerge as to when integration between audition and passive tactile touch becomes adult-like and whether it is comparable to the age of maturation for audiohaptic integration.

The fission and fusion illusions provide strong measures of multisensory integration. These illusions occur when conflicting numbers of stimuli are presented passively to two separate modalities close in time. First described in the audiovisual domain, fission occurs when a single visual flash is perceived as two distinct flashes when paired with two auditory stimuli (Shams et al., 2000, 2002), whereas fusion occurs when two visual flashes are perceived as a single flash when paired with a single auditory stimulus (Andersen et al., 2004). Mounting evidence suggests that these two illusions are driven by different mechanisms (see Bolognini et al., 2011, 2016; Chen et al., 2017; Mishra et al., 2007, 2008; but see Hirst et al., 2020), including the finding of different developmental trajectories (Innes-Brown et al., 2011). It is suggested that fusion is influenced by decisional factors more than fission, and the latter appears to be more perceptual in nature (Chen et al., 2017). Regardless, these illusions have been demonstrated for audiotactile pairings in which audition influences the perception of touch (sound-induced tap illusion) (Bresciani et al., 2005; Bresciani & Ernst, 2007; Bresciani et al., 2008; Hötting & Röder, 2004; Wozny et al., 2008). The reciprocal tap-induced sound illusion, to our knowledge, has only been reported in one study (Bresciani & Ernst, 2007), which required careful balancing of the relative reliabilities of the auditory and tactile signals.

Predictions regarding when audiotactile integration becomes adult-like are difficult because of the paucity of relevant data; there are few studies of audiotactile integration and even fewer of its development (e.g., Bremner & Spence, 2017; Occelli et al., 2011). Given the early development of touch (Bremner & Spence, 2017) and hints of integration occurring with audition prenatally (Kisilvesky & Muir, 1991), one might expect audiotactile integration to mature earlier than other modality pairings. However, our own work on the perception of multisensory simultaneity found later development associated with touch; as mentioned previously, audiovisual perception was adult-like by 9 years of age, but visuotactile and audiotactile perception was adult-like by 11 years of age (Chen et al., 2016, 2018; Stanley et al., 2019 or Chapter 2). Hence, we predicted that audiotactile fission and fusion illusions would mature between 9 and 13 years of age. Note that we could not rule out an even later age of maturation based on the results from the audiohaptic size discrimination task (Petrini et al., 2014; Scheller et al., 2021).

In the current study, we measured developmental changes for the integration of audition and passive touch as indexed by the fission and fusion illusions. Given that fission and fusion plausibly occur at different levels of processing—fission appears to be more perceptual, whereas fusion appears to be more influenced by decisional factors (Chen et al., 2017)—we predicted an earlier maturation for fission than for fusion. Each participant completed both sound-induced tap and tap-induced sound tasks. On each trial, either one or two target taps (or sounds) were presented concurrently with zero, one, or two distractor sounds (or taps). Participants reported the number of taps (or sounds) perceived. We measured the magnitude of each illusion based on accuracy in the conflict conditions; the more errors made, the stronger the illusion. In addition, we applied a signal detection analysis to disentangle the influence of distractors on perceptual discriminability from response criterion when reporting the number of taps (or sounds). Our main question concerned when performance became adult-like. As such, we compared each group of children against the adult control group separately for each illusion. In addition to expecting all groups to demonstrate measurable fission and fusion illusions, we also expected children to demonstrate larger illusions than adults (e.g., Adams, 2016). All analyses reflect these a priori planned directional hypotheses (see "Analysis" section in Method for details)

Method

Participants

In total, 20 participants were tested from each of four age groups: 9-year-olds (M_{age} = 9.1 years, *SD* = 0.1), 11-year-olds (M_{age} = 11.1 years, *SD* = 0.1), 13-year-olds (M_{age} = 13.0 years, *SD* = 0.1), and adults (M_{age} = 20.0 years, *SD* = 1.9). Based on past research (Stanley et al., 2019 or Chapter 2), we set the recruitment criterion to ±3 months for the three groups of children. All groups had an equal split of male and female participants (defined by sex at birth). All participants were right-handed and reported normal auditory and tactile acuity. As is standard in our lab, visual ability was assessed using the Randot test of stereoacuity in which a minimum of 40 s of arc was required to participate.

An additional 18 participants were tested but excluded because they did not meet the minimum inclusion criteria in the practice blocks (2 nine-year-olds and 1 thirteen-year-old), they did not reach 80% correct on the catch trials (6 nine-year-olds and 4 eleven-year-olds), they failed the vision screening (1 thirteen-year-old), or the tap machine malfunctioned (2 nine-year-olds and 2 eleven-year-olds).

Children were recruited from a database of parents who, at the time of their children's birth, consented to be contacted about participation in developmental research. Parents provided written consent for their children to participate, and children provided verbal assent to participate after being read a child-friendly consent form. In exchange for participation, children were rewarded with a book or toy and a junior scientist certificate. Adult participants were recruited from the undergraduate study pool at McMaster

University and received a course credit in exchange for their participation. This study was approved by the McMaster Research Ethics Board and adhered to the Tri-Council Statement on Ethical Conduct of Research Involving Humans (TCPS2; Canada).

Apparatus and stimuli

The auditory stimulus was a 10-ms peak-to-peak white noise burst with a flat amplitude envelope and 2-ms onset and offset ramping (henceforth referred to as a beep). The beeps were presented via Sennheiser HDA-200 (closed-ear) headphones at 107 dB SPL (measurements obtained with Brüel & Kjær artificial ear Type 4152 and Brüel & Kjær sound level meter Type 2270). White noise was played in free field at 73 dB to reduce the possibility of perceiving extraneous noise emitted by the tactile stimulators. The tactile stimulus was generated from below the participant's hand comprising a 10ms tap to the right index finger (see Figure 1A). The taps were generated by one of two custom-built tactile stimulator machines that were mounted in a noise-attenuating box. These machines indented the skin by approximately 3 mm, which to an observer was perceived as a light touch. The original machine used a dull metal pin mounted on a mechanical solenoid. This machine was replaced approximately halfway through data collection when it began missing trials; this change did not affect the results (see supplementary material below for additional details and analyses). The replacement machine was an electromagnetic solenoid mounted in a wooden block. When activated, a dull plastic peg protruded from the base and tapped the finger. In total, 8 nine-year-olds, 12 eleven-year-olds, 15 thirteen-year-olds, and 16 adults were tested on the original machine.

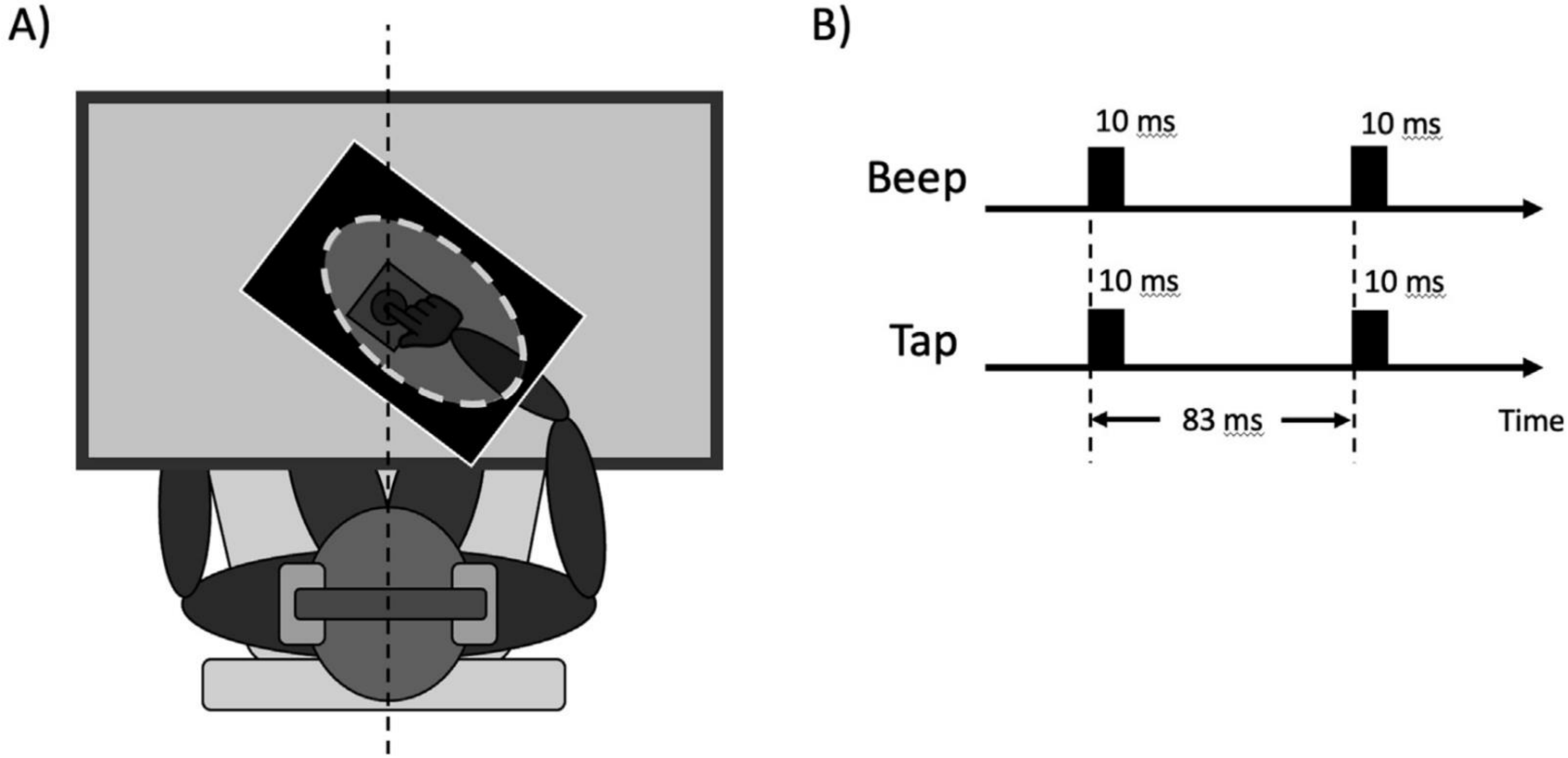

Figure 1. Experimental setup and a sample trial. (A) Overhead view of a participant seated facing the table. Headphones were used to present the auditory stimuli, and a tactor mounted in a sound-attenuating box delivered the tactile stimulus to the right index finger. The tactor was aligned with the participant's midline. (B) The temporal profile of a trial consisting of two beeps and two taps with their onsets separated by an 83-ms delay.

Both the auditory and tactile stimuli were generated and temporally controlled by MATLAB (MathWorks, Natick, MA, USA) with the Psychtoolbox-3 package (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) installed on an Apple Mac Mini. All timings were verified with an oscilloscope. The apparatus and stimuli were identical to those used in Stanley et al. (2019; Chapter 2) that measured audiotactile simultaneity perception.

Design

The experiment consisted of two tasks; one task measured the sound-induced tap illusion, and the other task measured the tap-induced sound illusion. Each participant completed both tasks in a counterbalanced order[2]. The tap-induced sound task revealed only a weak fission illusion and no fusion illusion, along with no developmental changes. These results did not provide any meaningful theoretical contributions, and as such were excluded from the main body of this article (see supplementary material for methods and results).

Sound-induced tap illusions

There were four types of trials: unimodal, congruent, incongruent, and catch. A unimodal trial consisted of either one tap (1T0B) or two taps (2T0B). A congruent trial was either a single tap presented simultaneously with a single beep (1T1B) or two taps

[2] Inclusion of the factor of task order did not change the pattern of significance reported below and, critically, did not interact with age (see online supplementary material).

presented simultaneously with two beeps (2T2B). The configurations of the two types of incongruent trials consisted of different numbers of taps and beeps: one tap with two beeps (1T2B) or two taps with one beep (2T1B). The fission illusion occurred in the 1T2B condition when two taps were reported. The fusion illusion occurred in the 2T1B condition when a single tap was reported. The first tap and beep were always presented simultaneously. The second stimulus, if presented, occurred 83 ms after the onset of the first stimulus (see Figure 1B) (delay based on previous work; see Chen et al., 2018).

Catch trials were included to determine whether the fission and fusion illusions were truly a product of auditory–tactile integration. These trials contained identical stimuli as incongruent trial types (i.e., 1T2B; 2T1B); however, the beep(s) preceded the tap(s) by 300 ms. This time value of 300 ms was chosen because it is beyond the simultaneity window in which audiotactile events are judged to have occurred together (Stanley et al., 2019; Chapter 2). Hence, if audiotactile fission and fusion are indeed the products of integration, then we expected no illusions in the catch trials. Inclusion of these catch trials also served two additional purposes: firstly, it ensured that participants were engaged with the task, and secondly, that participants were responding to the correct (i.e., target) modality. Participants were excluded if their accuracy on the catch trials fell below 80%.

There were 140 trials divided across five blocks of 28 trials. In each block, there were eight unimodal trials (4 1T0B and 4 2T0B), eight congruent trials (4 1T1B and 4 2T2B), and eight incongruent trials (4 1T2B and 4 2T1B), plus four catch trials (2 1T2B and 2 2T1B), presented in a completely random order.

Procedure

The participant sat at a desk in front of the tactile stimulator in a dimly lit room. The experimenter sat at the computer keyboard to the right side of the participant. The participant's right hand was inserted into the opening of the sound-attenuating box with their right index finger resting gently over a hole on the tactor device. When activated, a dull pin protruded through the hole to tap the finger. The tactile stimulator was positioned approximately 40 cm away from the participant centered at the midline of their body. Each trial was initiated by the experimenter by pressing the Enter key. The presentation of the stimulus/stimuli occurred 500 to 1500 ms later (six random foreperiods separated by 200-ms intervals). The participant's task was to respond verbally whether they perceived one or two targets, and the experimenter keyed in the response by pressing either the 1 or 2 key. If participants missed the trial, the experimenter keyed in a 0 (zero). There was no time limit to respond.

Prior to the main experiment, participants completed four practice sessions. The first two were unimodal practice sessions: one for taps and one for beeps. These sessions ensured that participants could distinguish between one and two targets accurately. Each of these sessions included eight unimodal trials: four with one target and four with two targets. Two additional practice sessions were completed, each prior to the corresponding main task. The purpose of these sessions was to ensure that participants were responding correctly to the target modality. Each of these sessions consisted of 16 trials: eight unimodal, four congruent, and four catch. These sessions did not include any incongruent trials in which the illusion may occur. An accuracy of 85% was required in all practice tasks to participate in the main experiment. A mandatory break intervened between the two tasks to reduce the possibility of lapses in attention (especially for children who tend to get fidgety) and, more importantly, to minimize the switch cost from one target modality in the first task to the other target modality in the second task. During the break, all participants completed the Randot stereoacuity task and the handedness task. Children were offered snacks and time to play for about 15 minutes. The average duration of the entire experimental protocol was approximately 50 minutes.

Analysis

The accuracy for each participant was first calculated as the proportion of correct responses in each condition. Missed trials (0.36% of all trials) were removed from these calculations. The mean accuracy (sound-induced tap presented in Figure 2; analyses for both sound-induced tap and tap-induced sound provided in supplementary material) was used to compute the magnitude of illusion and used in the signal detection analysis.

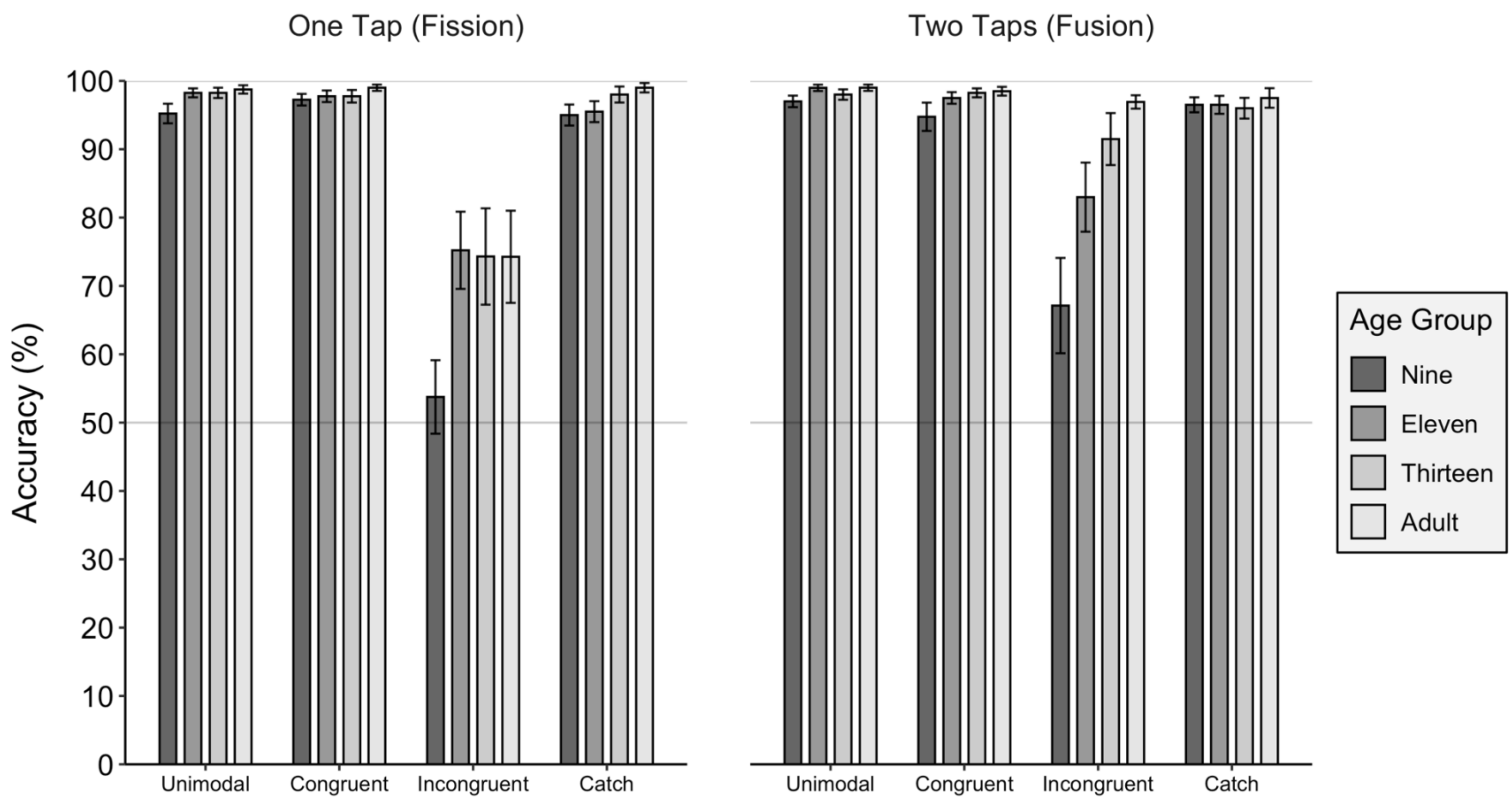

Figure 2. The mean accuracy of reporting the number of taps when paired with beeps in the four age groups tested. The horizontal line at 50% correct represents chance performance. Error bars are +/- one standard error of the mean.

The magnitude of the fission and fusion illusions was calculated for each participant by subtracting the accuracy in the incongruent condition from the accuracy in the corresponding congruent condition. Specifically, for sound-induced tap fission, the 1T2B accuracy was subtracted from the 1T1B accuracy. For sound-induced tap fusion, the 2T1B accuracy was subtracted from the 2T2B accuracy. We evaluated outliers (>3 standard deviations beyond the mean within each age group) based on the magnitude of the fission and fusion illusions, and one adult and one 13-year-old were identified. To keep equal group sizes, we replaced the adult with another participant who already completed the task, but because of the COVID-19 pandemic, we could not recruit another 13-year-old[3].

We applied an analysis based on signal detection theory to separate perceptual discriminability (*d'*) and decisional criterion © (see Chen et al., 2017; McCormick & Mamassian, 2008). Responses based on multisensory processing are affected by stimulus parameters (typically held constant in an experiment), the integration mechanism itself, and decisional processes (Ernst, 2008). The advantage of using a signal detection analysis is that it separates estimates of the influence of perceptual and decisional components on participants' performance as indexed by *d'* and *c*, respectively.

To calculate *d'* and *c*, the accuracy data were transformed into hits and false alarms by defining the two-tap conditions as targets and the one-tap conditions as noise;

---

[3] All analyses presented include the 13-year-old outlier; however, the pattern of results does not change when this outlier is removed.

thus, the proportion of hits was taken as the accuracy of performance in the two-tap conditions, and the proportion of false alarms was taken as one minus the accuracy (i.e., the error rate) of performance in the one-tap condition. This was done separately for the zero-beep (unimodal trials), one-beep, and two-beep conditions. The *d’* was computed as the *z*-score of hits minus the *z*-score of false alarms. The *c* was computed as -0.5 multiplied by the *z*-score of hits plus the *z*-score of false alarms. Given that the *z*-score for 0 or 1 is infinity, we replaced these values with half the smallest unit of measure (i.e., 1 trial of 20 in each condition). So, a value of 0 was replaced with 0.5/20 = 0.025, and a value of 1 was replaced with–1 - (0.5/20) = 0.975. For the zero-beep condition, the values index unimodal *d’* and *c* of taps; for the one-beep condition, the change from the zero-beep condition in the *d’* and *c* index the influence of the presentation of one beep, thereby associated with the fusion illusion; and for the two-beep condition, the change from the zero-beep condition in the *d’* and *c* index the influence of the presentation of two beeps, thereby associated with the fission illusion.

The fission and fusion conditions were analyzed separately; th©cluded the measures of magnitudes of fission and fusion illusions and *d’* and *c* in the zero-, one-, and two-beep conditions. All eight dependent measures were submitted separately to a one-way analysis of variance (ANOVA) with the single between-participant factor of age (9-year-olds, 11-year-olds, 13-year-olds, or adults). Comparisons associated with an effect of age were conducted using a Dunnett’s test in which each child group was compared independently to the adult (i.e., control) group; the comparisons were one-tailed because previous studies consistently demonstrate that children show larger fission and fusion illusions than adults (Adams, 2016; Innes-Brown et al., 2011; Nava & Pavani, 2013). Finally, all dependent measures for each age group were compared against zero using one-sample t tests. Statistically marginal effects were reported only when below $p < .10$ and in the predicted direction. Any violations of sphericity were corrected with the Greenhouse–Geisser correction to the degrees of freedom. Effect sizes are reported as partial eta-squared (η_p^2 ANOVA) and Cohen's *d* (*t*-test).

Results: Sound-induced tap illusion

Magnitude of illusion

The fission illusion was significantly different from zero for all age groups, all *t*s(19) > 3.34, *p*s < .002, *d*s > 0.74 (see Figure 3). The fusion illusion was significantly different from zero for children, all *t*s(19) > 1.36, *p*s < .04, *d*s > 0.45, but not for adults, *t*(19) = 1.36, *p* = .09, *d* = 0.30.

The one-way ANOVA for the magnitude of fission revealed only a marginal main effect of age, $F(3, 76) = 2.57$, $p = .06$, $\eta_p^2 = .09$. The Dunnett's test showed that 9-year-olds had a significantly greater fission illusion than adults, *t*(19) = 2.12, *p* = .047, *d* = 0.48[4], but 11- and 13-year-olds did not, both *t*s (19) < 0.25, *p*s > .80, *d*s < 0.06. The one-way ANOVA for the magnitude of fusion revealed a significant main effect of age, $F(3, 76) = 6.41$, $p < .001$, $\eta_p^2 = .20$. The Dunnett's test revealed that 9-year-olds showed a significantly greater fusion illusion than adults, *t*(19) = 4.12, *p* < .001, *d* = 0.88, whereas 11-year-olds showed a marginally greater fusion illusion than adults, *t*(19) = 2.05, *p* = .06, *d* = 0.57. The 13-yearolds did not differ from adults, *t*(19) = 0.82, *p* = .40, *d* = 0.04. In summary, the magnitude of the fission illusion was statistically adult-like by 11 years of age, whereas the magnitude of the fusion illusion may still not be fully adult-like until beyond age 11.

[4] Significance confirmed by independent bootstrap analysis (*p* = .03).

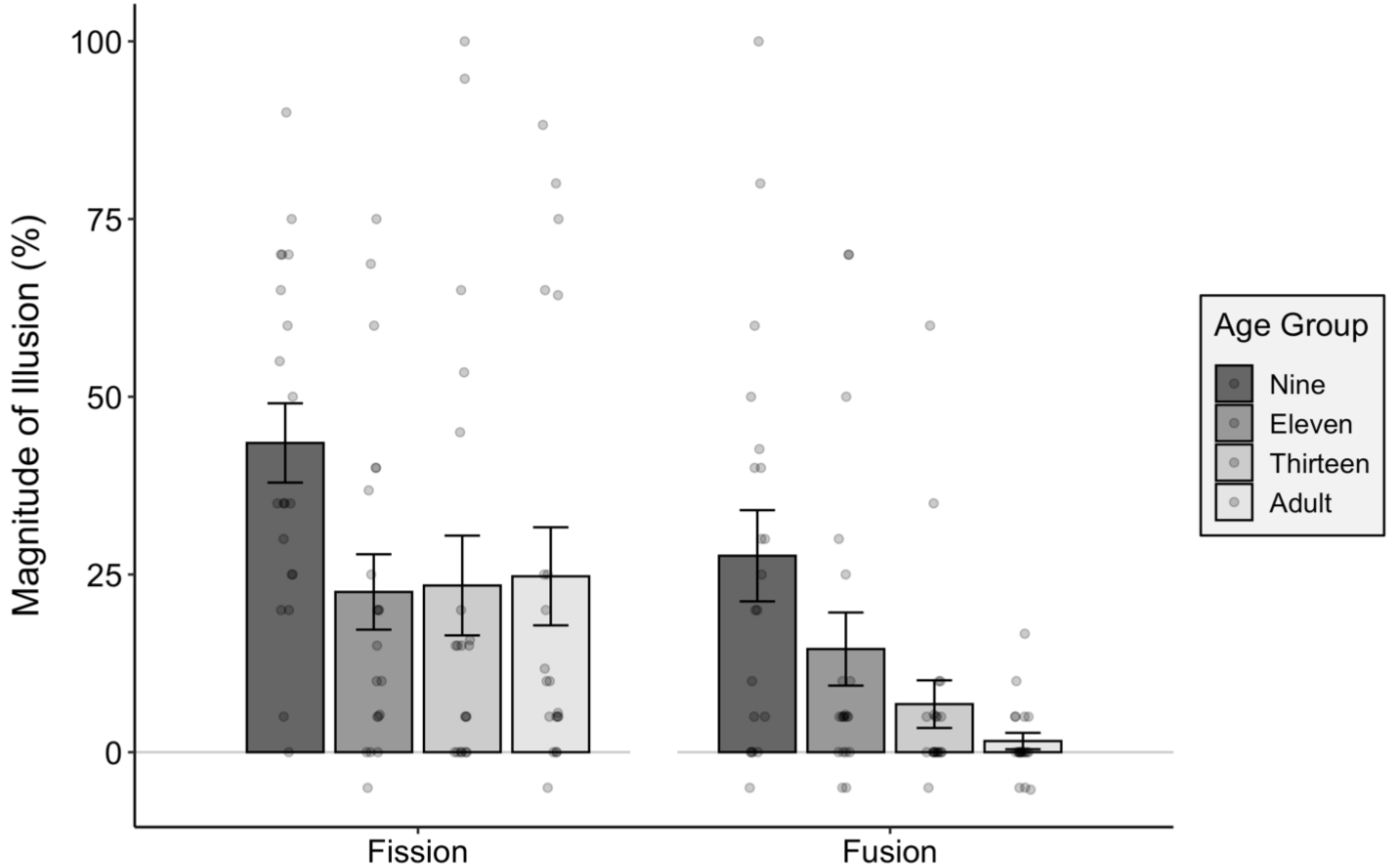

Figure 3. Mean magnitudes of the fission (left) and fusion (right) illusions for the four age groups tested. The magnitude of both illusions was calculated by subtracting the accuracy of the incongruent condition from the accuracy of the congruent condition: fission (1T1B – 1T2B) and fusion (2T2B – 2T1B). Individual data points represent the performance of each participant. Error bars are +/- one standard error of the mean.

Signal detection analysis

Discriminability (d'). The *d'* values were greater than zero for all beep and age conditions, all $ts(19) > 37.18$, $ps < .001$, $ds > 2.18$ (see Figure 4A), confirming that all age groups were able to discriminate between one and two taps.

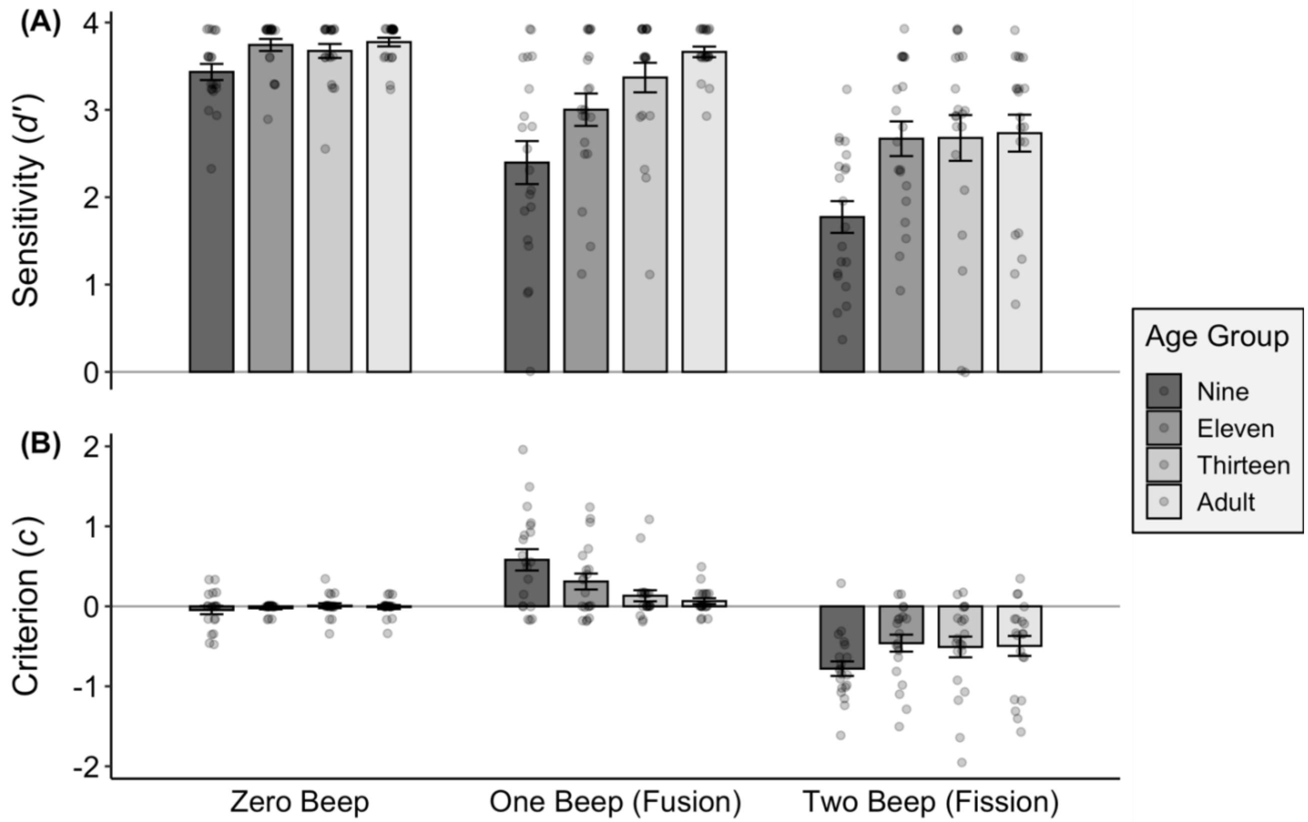

Figure 4. Signal De–ection - Discriminability (d') and Cr©rion (c) Associated with Sound-induced Tap Illusion. **(A)** Mean *d'* scores for the four age groups tested. The *d'* represents the ability to distinguish two taps (target) from one tap (noise) when paired with either zero (unimodal), one, or two beeps. **(B)** Mean *c* scores for the four age groups tested. A shift toward a positive criterion indicates a bias to report "one tap", whereas a shift toward a negative criterion indicates a bias to report "two taps". Individual data points represent the mean *d'* and *c* of each participant in **(A)** and **(B)**, respectively. Fission is represented in the two-beep condition (right), whereas fusion is represented in the one-beep condition (middle). Error bars are +/- one standard error of the mean.

The effect of age was significant for all three one-way ANOVAs [zero beeps: $F(3, 76) = 4.31$, $p = .007$, $\eta_p^2 = .15$; one beep: $F(3, 76) = 9.37$, $p < .001$, $\eta_p^2 = .27$; two beeps: $F(3, 76) = 4.58$, $p = .005$, $\eta_p^2 = .15$][5]. The Dunnett's tests showed that 9-year-olds had a lower *d'* than adults in all beep conditions [zero beeps: $t(19) = -3.25$, $p = .002$, $d = 0.71$; one beep: $t(19) = -5.02$, $p < .001$, $d = 1.17$; two beeps: $t(19) = -3.15$, $p = .003$, $d = 0.74$). The 11-year-olds had a lower *d'* than the adults in the one-beep condition, $t(19) = -5.02$, $p = .01$, $d = 0.85$; however, they did not differ significantly from adults in the zero-beep and two-beep conditions, both ts(19) < 0.31, ps > .63, ds < 0.09. The 13-year-olds were adult-like for all three beep conditions, all ts(19) < 1.16, ps > .26, ds < 0.35).©***iterion (c).*** For the zero-beep condition, none of the age groups differed from zero, all ts(19) < 0.41, ps > .38, ds < 0.20 (two-tailed), suggesting that they had no bias to report one or two taps in the absence of sound. In the one-beep condition, all age groups had criterion scores greater than zero, all ts (19) > 1.80, ps < .04, ds > 0.40, indicating a bias to report one tap, consistent with the characteristic of the fusion illusion.

In the two-beep condition, all age groups demonstrated criteria below zero, all ts(19) > -3.92, ps < .001, ds > 0.88, representing an overall bias to report two taps, consistent with the characteristic of the fission illusion. For each level of beep, the effect of age was significant in the one-beep condition, $F(3, 76) = 6.27$, $p < .001$, $\eta_p^2 = .20$, but

---

[5] Although these one-way ANOVAs for each level of beep were planned a priori (see ''Analysis" section for rationale), we conducted a two-way ANOVA with the factor of beep (zero, one, or two beeps) and age (9-year-olds, 11-year-olds, 13-year-olds, or adults), which confirmed a significant two-way interaction ($p = .04$).

not in the zero-beep condition, $F(3, 76) = 0.49$, $p = .69$, $\eta_p^2 = .02$, or the two-beep condition, $F(3, 76) = 1.67$, $p = .18$, $\eta_p^2 = .06$.5 The Dunnett's test revealed that in the one-beep condition, 9-year-olds had a significantly larger bias to report one tap compared with adults, $t(19) = 3.96$, $p < .001$, $d = 0.78$, whereas 11-year-olds had a bias that was only marginally larger than that of adults, $t(19) = 1.88$, $p = .08$, $d = 0.58$. The 13-year-olds did not differ from the adults, $t(19) = 0.51$, $p = .54$, $d = 0.18$. In summary, the 9-year-olds and some 11-year-olds were more biased than the adults in the one-beep condition representative of the fusion illusion.

Discussion

We tested three groups of children (9-, 11-, and 13-year-olds) and one group of adults on the sound-induced tap illusion and tap-induced sound illusion. For the sound-induced tap illusion, we observed fission in all four groups and fusion in only the three groups of children. The magnitude of both illusions was greatest in 9-year-olds when compared with adults. The fission illusion is reduced to an adult-like magnitude by 11 years of age, whereas the fusion illusion might not be fully adult-like until 13 years of age. The signal detection analysis revealed that 9-year-olds had poorer discriminability than adults for both the fission and fusion illusions and had a significant bias to report one tap in the fusion illusion. The 11-year-olds also had poorer discriminability than adults and a marginal bias to report one tap in the fusion illusion. Thus, the fission and fusion illusions show different developmental trajectories; fission is adult-like by 11 years of age, which is primarily explained by children's improved discriminability, whereas fusion requires an additional two years of development, which can be attributed to both improved discriminability and a less biased criterion.

A transition from immature during late childhood (<10 years of age) to mature during early adolescence (~11–13 years of age) for the sound-induced tap illusions mirrors the maturation of audiotactile simultaneity perception (Stanley et al., 2019 or Chapter 2). Both studies used identical stimuli (with minor differences in associated parameters, such as temporal intervals between stimuli) and similar experimental context but differed in the type of temporal task (event counting and simultaneity perception), thereby providing convergent evidence that 9 to 13 years of age appears to be a critical transition period of maturation in audiotactile perception. This age range is also similar to other tasks involving visual flash illusions in which children begin to show adult-like multisensory integration: approximately 8 to 10 years for audiovisual (Adams, 2016; Nava & Pavani, 2013) and approximately 10 to 12 years for visuotactile (O'Dowd et al., 2021). For audiohaptic integration, Scheller et al. (2021) reported a later maturation age (adult-like by ~15 years of age). Whereas both the current study and their study supported a late maturation, the later age observed In their study could be attributed to different tasks and associated information processing; Scheller et al.'s task required estimation of object size using active touch (i.e., haptic), whereas the task in the current study required numerosity judgments using passive touch (i.e., tactile) (see Heller, 1984 and Simões-Franklin et al., 2011). In addition, it has been proposed that spatial abilities mature after temporal abilities (see Scheller et al., 2021). Finally, mature integration in Scheller et al.'s (2021) study was determined by optimal cue integration; although the fission and fusion illusions can be explained by optimal integration (Andersen et al., 2004; Bresciani & Ernst, 2007; Shams et al., 2005; Wozny et al., 2008), we did not measure optimality, which may mature later than the benchmark for adult-like perception used in the current study. Regardless, the development of cross-modal abilities involving active touch (e.g., size estimation) appears to stretch into late adolescence, whereas we see adult-like audiotactile abilities in the temporal domain slightly earlier by 11 to 13 years of age.

To account for the developmental trends observed, we can consider several theories within multisensory perception. During early development, children demonstrate sensory dominance (e.g., Gori et al., 2008, 2012; Nava & Pavani, 2013); sensory signals are not integrated, and the modality with greater precision drives perception. In the current study, audition appears to be the dominant sense because we observed only a weak tap-induced sound fission illusion and no tap-induced sound fusion illusion (see supplementary material for discussion), implying that the tactile signal was less salient (i.e., less reliable) than the auditory signal. Sensory dominance gives way to multisensory integration as children age; the transition, based on prior research with audiovisual and visuotactile tasks, appears to reach completion by around 10 years of age (Adams, 2016; Burr & Gori, 2012; Gori et al., 2008, 2021; Nava & Pavani, 2013; see Burr & Gori, 2012, for a review). This could account for the larger illusions observed in the 9- and 11-year-olds in the current study if we assume that this transition occurs later for audiotactile pairings (plausibly after age 11), which is consistent with protracted development when touch is involved (see Introduction).

Alternatively, 9- and 11-year-olds may have already transitioned to multisensory integration, in which case perception would be driven by remaining immaturities in a multisensory integration process such as optimal cue combination (see Alais & Burr, 2004; Ernst & Banks, 2002; Ernst & Bülthoff, 2004). To date, several developmental studies have charted the development of optimal integration, which appears to emerge between late childhood and late adolescence depending on the task (Adams, 2016; Burr & Gori, 2012; Gori et al., 2008, 2021; Nardini et al., 2008; Petrini et al., 2014; Scheller et al., 2021). Optimal integration is often modeled using maximum likelihood estimation (MLE) in which the final integrative percept is determined by the relative weights of the unimodal sensory signals in terms of their reliabilities. Recall that the 9-year-olds had significantly worse unimodal tactile perception than adults (see Figure 4A); when combined with the finding of 9-year-olds' adult-like unimodal auditory accuracy (see Figure S3 in supplementary material), we can assume that audition had a greater relative reliability within this age group given the stimulus parameters used. This greater reliability of audition would cause an MLE-based integration mechanism to produce the larger magnitudes of fission and fusion observed in the 9-year-olds. However, the 11-year-olds had an immature fusion illusion even though they demonstrated adult-like unimodal tactile accuracy, which thus requires additional theoretical considerations. In addition, the standard MLE model assumes weighted integration occurs on every trial and fails to consider circumstances in which signals should be segregated (Shams et al., 2005; Wozny et al., 2008). Specifically, sometimes the illusions were not perceived (i.e., accuracy in the incongruent trials was greater than 0), which means that the auditory and tactile signals were likely processed independently.

The Bayesian causal inference framework incorporates a weighted MLE process and adds a weighted prior for one versus two sources based on experience, expectancy, task context, and stimulus parameters (Körding et al., 2007; Wozny et al., 2010). The addition of the weighted prior allows for sensory segregation (see Shams & Beierholm, 2010, Figure 2), which is important to consider here because the magnitudes of fission and fusion are derived from the proportion of trials in which integration did or did not occur. This prior, at least for adults, is assumed to weight the one- and two-source hypotheses appropriately given the context. The greater illusion in 9-year-olds therefore could result from an immature prior that favors a single source (i.e., integration). At this age, they are just beginning to experience the perceptual advantage produced by an MLE-based integration process (i.e., multisensory gains: superior precision for integrated percepts than for either of the unisensory percepts alone; see Ernst & Bülthoff, 2004). However, children at this age and younger may lack the experience with perceptual errors that result from spurious integrations which ultimately balance the prior between one and two sources. Gaining this experience may also contribute to the development of an adult-like ability to rapidly recalibrate (e.g., Han et al., 2022; Rohlf et al., 2020). Nevertheless, given that 11-year-olds show only immature fusion but adult-like fission and show ceiling unimodal accuracy, this account alone—a biased prior for one source—cannot account for all these data. Clearly, these ideas are speculative because studies addressing the nature of priors remain in their infancy (Shams & Beierholm, 2022), and we are not aware of any studies exploring changes in priors across development. To disentangle these multiple accounts for our data will require future research specifically designed to independently estimate the priors and the MLE function, and the possibility of sensory dominance.

Finally, we must also entertain the possibility that the different developmental trends for fission and fusion require independent explanations. Although the fission and fusion illusions are generally perceptual in nature, there are suggestions in the literature that different mechanisms may underlie the two illusions (e.g., Bolognini et al., 2011, 2016; Chen et al., 2017; Mishra et al., 2007, 2008; see Hirst et al., 2020, for a review of the audiovisual data). Specifically, there are hints that the fusion illusion may be more affected by decisional factors than the fission illusion (see Chen et al., 2017, for a discussion). Qualitative examination of the magnitude of illusions (Figure 3) suggests that fission develops in a step-like function after 9 years of age, whereas fusion has a more gradual refinement across development. Decisional processes, especially those involved with inhibition of irrelevant signals, appear to have a protracted development (Hooper et al., 2004), and could result in the later maturation of the fusion illusion.

Conclusion

Audiotactile integration measured with the current task was completely adult-like by 13 years of age. The fission and fusion illusions, used to measure integration, appear to follow different developmental trends. Based on the magnitude of illusions and the signal detection analysis, fission reaches maturity by 11 years of age, whereas fusion is not fully mature until 13 years of age. Theoretical accounts of these data considered sensory dominance, optimal integration mechanisms (i.e., MLE), and the prior for one or two sources associated with causal inference. Further research is needed to specify the relative contributions of these accounts.

References

Adams, W. J. (2016). The development of audio-visual integration for temporal judgements. *PLoS Computational Biology*, *12*(4), 1–17. https://doi.org/10.1371/journal.pcbi.1004865

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*(3), 301–308. https://doi.org/10.1016/j.cogbrainres.2004.06.004

Begum Ali, J., Spence, C., & Bremner, A. J. (2015). Human infants' ability to perceive touch in external space develops postnatally. *Current Biology*, *25*(20), R978–R979. https://doi.org/10.1016/j.cub.2015.08.055

Bolognini, N., Convento, S., Casati, C., Mancini, F., Brighina, F., & Vallar, G. (2016). Multisensory integration in hemianopia and unilateral spatial neglect: Evidence from the sound induced flash illusion. *Neuropsychologia*, *87*, 134–143. https://doi.org/10.1016/j.neuropsychologia.2016.05.015

Bolognini, N., Rossetti, A., Casati, C., Mancini, F., & Vallar, G. (2011). Neuromodulation of multisensory perception: A tDCS study of the sound-induced flash illusion. *Neuropsychologia*, *49*(2), 231–237. https://doi.org/10.1016/j.neuropsychologia.2010.11.015

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Bremner, A. J., Holmes, N. P., & Spence, C. (2012). The development of multisensory representations of the body and of the space around the body. In *Multisensory Development* (113–136). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199586059.003.0005

Bremner, A. J., & Spence, C. (2017). The development of tactile perception. In *Advances in Child Development and Behavior* (Vol. 52). Elsevier Inc. https://doi.org/10.1016/bs.acdb.2016.12.002

Bresciani, J. P., & Ernst, M. O. (2007). Signal reliability modulates auditory-tactile integration for event counting. *NeuroReport*, *18*(11), 1157–1161. https://doi.org/10.1097/WNR.0b013e3281ace0ca

Bresciani, J. P., Ernst, M. O., Drewing, K., Bouyer, G., Maury, V., & Kheddar, A. (2005). Feeling what you hear: Auditory signals can modulate tactile tap perception. *Experimental Brain Research*, *162*(2), 172–180. https://doi.org/10.1007/s00221-004-2128-2

Bresciani, J.-P., Dammeier, F., & Ernst, M. O. (2008). Tri-modal integration of visual, tactile and auditory signals for the perception of sequences of events. *Brain Research Bulletin*, *75*, 753–760. https://doi.org/10.1016/j.brainresbull.2008.01.009

Burr, D., & Gori, M. (2012). Multisensory integration develops late in humans. In *The Neural Bases of Multisensory Processes* (pp. 345–362). CRC Press/Taylor & Francis. https://doi.org/10.1201/b11092-23

Chen, Y. C., Lewis, T. L., Shore, D. I., Spence, C., & Maurer, D. (2018). Developmental changes in the perception of visuotactile simultaneity. *Journal of Experimental Child Psychology*, *173*, 304–317. https://doi.org/10.1016/j.jecp.2018.04.014

Chen, Y. C., Maurer, D., Lewis, T. L., Spence, C., & Shore, D. I. (2017). Central–peripheral differences in audiovisual and visuotactile event perception. *Attention, Perception, and Psychophysics*, *79*(8), 2552–2563. https://doi.org/10.3758/s13414-017-1396-4

Chen, Y. C., Shore, D. I., Lewis, T. L., & Maurer, D. (2016). The development of the perception of audiovisual simultaneity. *Journal of Experimental Child Psychology*, *146*, 17–33. https://doi.org/10.1016/j.jecp.2016.01.010

Ernst, M. O. (2008). Multisensory integration: A late bloomer. *Current Biology*, *18*(12), 519–521. https://doi.org/10.1016/j.cub.2008.05.003

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

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

Gori, M., Campus, C., & Cappagli, G. (2021). Late development of audio-visual integration in the vertical plane. *Current Research in Behavioral Sciences*, *2*(April), 100043. https://doi.org/10.1016/j.crbeha.2021.100043

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

Gori, M., Giuliana, L., Sandini, G., & Burr, D. (2012). Visual size perception and haptic calibration during development. *Developmental Science*, *15*(6), 854–862. https://doi.org/10.1111/j.1467-7687.2012.01183.x

Han, S., Chen, Y.-C., Maurer, D., Shore, D. I., Lewis, T. L., Stanley, B. M., & Alais, D. (2022). The development of audio–visual temporal precision precedes its rapid recalibration. *Scientific Reports*, *12*(1), 21591. https://doi.org/10.1038/s41598-022-25392-y

Held, R., Ostrovsky, Y., Degelder, B., Gandhi, T., Ganesh, S., Mathur, U., & Sinha, P. (2011). The newly sighted fail to match seen with felt. *Nature Neuroscience*, *14*(5), 551–553. https://doi.org/10.1038/nn.2795

Heller, M. A. (1984). Active and passive touch: The influence of exploration time on form recognition. *The Journal of General Psychology*, *110*, 243–249.

Hirst, R. J., McGovern, D. P., Setti, A., Shams, L., & Newell, F. N. (2020). What you see is what you hear: Twenty years of research using the Sound-Induced Flash Illusion. *Neuroscience and Biobehavioral Reviews*, *118*, 759–774. https://doi.org/10.1016/j.neubiorev.2020.09.006

Hooker, D. (1958). *Evidence of prenatal function of the central nervous system in man (James Arthur lecture on the evolution of the human brain, no. 26, 1957).* American Museum of Natural History.

Hooper, C. J., Luciana, M., Conklin, H. M., & Yarger, R. S. (2004). Adolescents' performance on the iowa gambling task: Implications for the development of decision making and ventromedial prefrontal cortex. *Developmental Psychology*, *40*(6), 1148–1158. https://doi.org/10.1037/0012-1649.40.6.1148

Hötting, K., & Röder, B. (2004). Hearing cheats touch, but less in congenitally blind than in sighted individuals. In *Psychological science* (Vol. 15, Issue 1, pp. 60–64). https://doi.org/10.1111/j.0963-7214.2004.01501010.x

Innes-Brown, H., Barutchu, A., Shivdasani, M. N., Crewther, D. P., Grayden, D. B., & Paolini, A. (2011). Susceptibility to the flash-beep illusion is increased in children compared to adults. *Developmental Science*, *14*(5), 1089–1099. https://doi.org/10.1111/j.1467-7687.2011.01059.x

Kisilvesky, B. S., & Muir, D. W. (1991). Human fetal and subsequent newborn responses to sound and vibration. *Infant Behavior and Development*, *14*(1), 1–26. https://doi.org/10.1016/0163-6383(91)90051-S

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & Cornelissen, F. (2007). *What's new in Psychtoolbox-3? A free cross-platform toolkit for Psychophysics with Matlab & GNU/Octave*. http://www.psychtoolbox.org

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9). https://doi.org/10.1371/journal.pone.0000943

Machulla, T.-K., Di Luca, M., & Ernst, M. O. (2016). The consistency of crossmodal synchrony perception across the visual, auditory, and tactile Senses. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 1026–1038. https://doi.org/10.1037/xhp0000191

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

McCormick, D., & Mamassian, P. (2008). What does the illusory-flash look like? *Vision Research*, *48*(1), 63–69. https://doi.org/10.1016/j.visres.2007.10.010

Mishra, J., Martinez, A., & Hillyard, S. A. (2008). Cortical processes underlying sound-induced flash fusion. *Brain Research*, *1242*, 102–115. https://doi.org/10.1016/j.brainres.2008.05.023

Mishra, J., Martinez, A., Sejnowski, T. J., & Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *Journal of Neuroscience*, *27*(15), 4120–4131. https://doi.org/10.1523/JNEUROSCI.4912-06.2007

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

Nava, E., & Pavani, F. (2013). Changes in sensory dominance during childhood: Converging evidence from the Colavita Effect and the Sound-Induced Flash Illusion. *Child Development*, *84*(2), 604–616. https://doi.org/10.1111/j.1467-8624.2012.01856.x

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in temporal perception. *Psychonomic Bulletin & Review*, *18*(3), 429–454. https://doi.org/10.3758/s13423-011-0070-4

O'Dowd, A., Cooney, S. M., Sorgini, F., O' Rourke, E., Reilly, R. B., Newell, F. N., & Hirst, R. J. (2021). The development of visuotactile congruency effects for sequences of events. *Journal of Experimental Child Psychology*, *207*, 105094. https://doi.org/10.1016/j.jecp.2021.105094

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Petrini, K., Remark, A., Smith, L., & Nardini, M. (2014). When vision is not an option: Children's integration of auditory and haptic information is suboptimal. *Developmental Science*, *17*(3), 376–387. https://doi.org/10.1111/desc.12127

Rohlf, S., Li, L., Bruns, P., & Röder, B. (2020). Multisensory integration develops prior to crossmodal recalibration. *Current Biology*, *30*(9), 1726-1732.e7. https://doi.org/10.1016/j.cub.2020.02.048

Scheller, M., Proulx, M. J., de Haan, M., Dahlmann-Noor, A., & Petrini, K. (2021). Late- but not early-onset blindness impairs the development of audio-haptic multisensory integration. *Developmental Science*, *24*(1), 1–17. https://doi.org/10.1111/desc.13001

Shams, L., & Beierholm, U. (2022). Bayesian causal inference: A unifying neuroscience theory. *Neuroscience & Biobehavioral Reviews*, *104619*.

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., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152. https://doi.org/10.1016/S0926-6410(02)00069-1

Shams, L., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport*, *16*(17), 1923–1927. https://doi.org/10.1097/01.wnr.0000187634.68504.bb

Shams, L., Yukiyasu Kamitani, & Shinsuke Shimojo. (2000). What you see is what you hear. *Nature*, *408*(6814), 788–788.

Simões-Franklin, C., Whitaker, T. A., & Newell, F. N. (2011). Active and passive touch differentially activate somatosensory cortex in texture perception. *Human Brain Mapping*, *32*(7), 1067–1080. https://doi.org/10.1002/hbm.21091

Stanley, B. M., Chen, Y. C., Lewis, T. L., Maurer, D., & Shore, D. I. (2019). Developmental changes in the perception of audiotactile simultaneity. *Journal of Experimental Child Psychology*, *183*, 208–221. https://doi.org/10.1016/j.jecp.2019.02.006

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., Port, M., & Porter, N. R. (2001). When is now? Perception of simultaneity. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1462), 31–38. https://doi.org/10.1098/rspb.2000.1326

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667. https://doi.org/10.1037/0033-2909.88.3.638

Wozny, D. R., Beierholm, U. R., & Shams, L. (2008). Human trimodal perception follows optimal statistical inference. *Journal of Vision*, *8*(3), 1–11. https://doi.org/10.1167/8.3.24

Wozny, D. R., Beierholm, U. R., & Shams, L. (2010). Probability matching as a computational strategy used in perception. *PLoS Computational Biology*, *6*(8). https://doi.org/10.1371/journal.pcbi.1000871

Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio–visual simultaneity judgments. *Perception & Psychophysics*, *67*(3), 531–544.

Supplementary Materials

Sound-induced Tap Illusions

Accuracy. Accuracy for each participant was calculated as the percent of correct responses in each condition. Missed trials were removed from analyses. A two-way Analysis of Variance (ANOVA) was conducted separately for the one-tap and two-tap conditions. Both included the factors of Age (9-, 11-, and 13-year-olds, and adults) and Congruency (unimodal, congruent, incongruent, and catch). For the one-tap condition, both the main effects of Age ($F(3,76) = 1.88$, $p = .01$, $\eta_p^2 = .13$) and Congruency ($F(1.12,85.36) = 79.73$, $p < .001$, $\eta_p^2 = .51$) were significant; however the interaction between Age and Congruency was only marginally significant ($F(3.37, 85.36) = 2.28$, $p = .08$, $\eta_p^2 = .08$) and therefore was not explored further.

For the two-tap condition, the main effects of Age ($F(3,76) = 7.37$, $p < .001$, $\eta_p^2 = .23$) and Congruency ($F(1.30,99.07) = 27.70$, $p < .001$, $\eta_p^2 = .27$), and their interaction ($F(3.91, 99.08) = 6.30$, $p < .001$, $\eta_p^2 = .20$) were all significant. Nine-year-olds performed worse than adults in the unimodal ($t(19) = -2.16$, $p = .04$, $d = .43$), the congruent ($t(19) = -2.19$, $p = .04$, $d = .44$), and the incongruent ($t(19) = -4.45$, $p < .001$, $d = .93$) conditions, but not in the catch condition ($t(19) = -0.52$, $p = .53$, $d = .14$). Eleven-year-olds performed marginally worse only in the incongruent condition ($t(19) = -2.08$, $p = .052$, $d = .65$), but they were adult like in the unimodal, congruent, and catch conditions (all $t(19) < 0.57$, ps $> .51$, ds $< .21$). Thirteen-year-olds did not differ from adults in any condition (all $t(19) < 1.08$, ps $< .29$, ds $< .30$).

Tap-induced Sound Illusions

Accuracy. The tap-induced sound illusion phase had the same design as the sound-induced tap illusion described in the Methods section but now the target modality changed—instead of reporting the number of taps, participants reported the number of beeps. Hence, the 2T1B trials constitute the condition where the fission of tap-induced sound illusion may occur, whereas the 1T2B trials constitute the condition where the fusion of tap-induced sound illusion may occur. The participant now responded to the number of beeps (1 or 2) in the unimodal condition. The catch trials were identical to the sound-induced tap illusion except that the taps preceded the beeps by 300 ms.

The analysis of the115ccuracycy data for the tap-induced sound task was identical to the analysis of the sound-induced tap task described previously. Two 2-way ANOVAs were conducted separately for the one-beep and two-beep conditions. Both included the factors of Age (9-, 11-, and 13-year-olds, and adults) and Congruency (unimodal, congruent, incongruent, and catch).There were no effects of Age (1-beep: $F(3,76) = 1.41$, $p = .25$, $\eta_p^2 = .05$; 2-beep: $F(3,76) = 1.00$, $p = .40$, $\eta_p^2 = .04$; see Figure S1), and no Age by Congruency interactions (1-beep: $F(5.06, 128.59) = 1.49$, $p = .20$, $\eta_p^2 = .06$; 2-beep: $F(7.76,196.68) = 0.54$, $p = .83$, $\eta_p^2 = .02$). The effect of Congruency was only significant for the 1-beep condition ($F(1.69, 128.59) = 25.52$, $p < .001$, $\eta_p^2 = .25$), with the accuracy in the incongruent condition being lower than the unimodal ($t(79) = 5.57$, $p < .001$, $d = .89$), congruent ($t(79) = 5.26$, $p < .001$, $d = .76$), and catch ($t(79) = 5.96$, $p < .001$, $d = .79$) conditions.

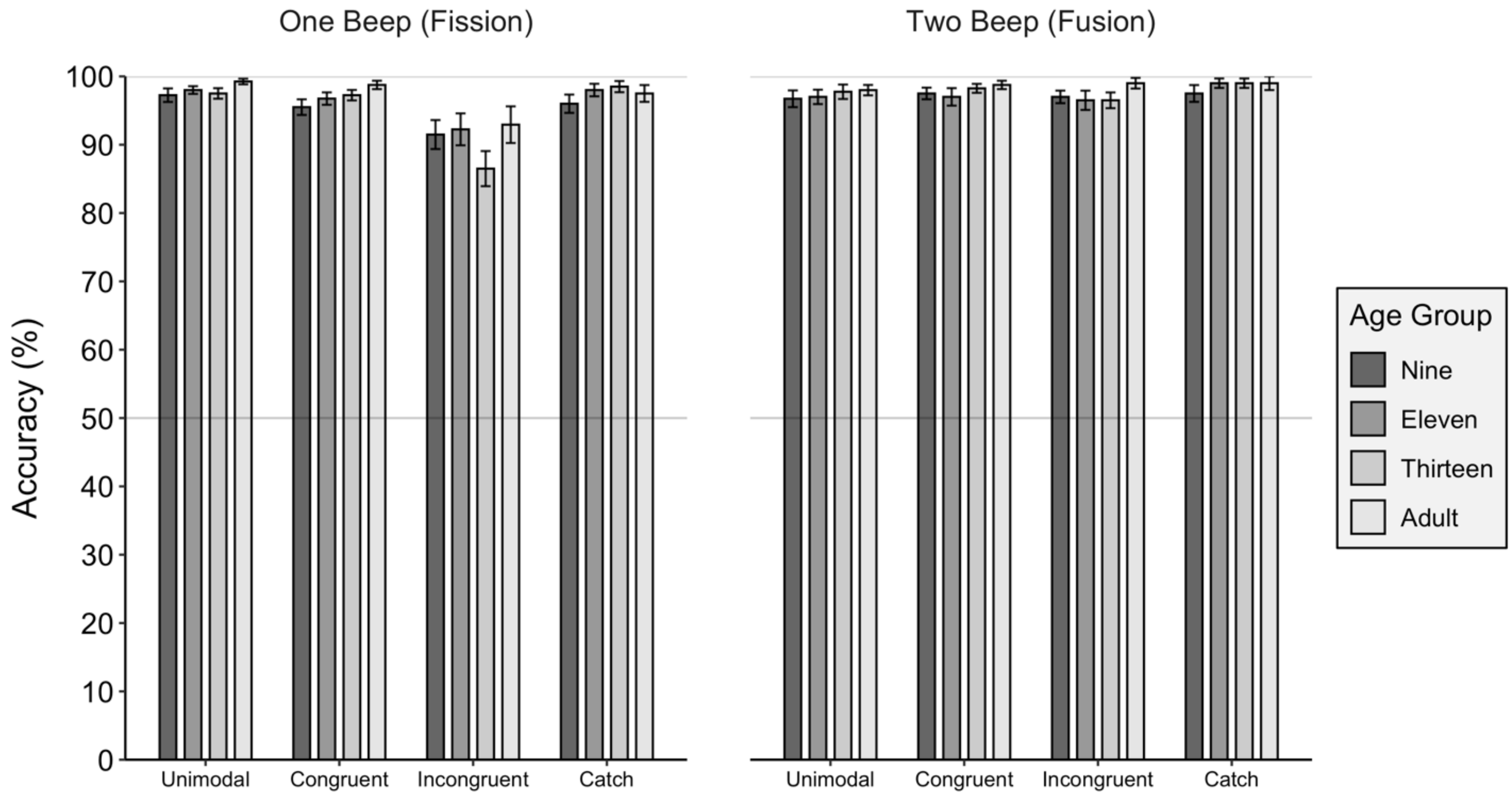

Figure S1. Accuracy of Tap-Induced Sound Illusion. The mean accuracy of reporting the number of beeps when paired with taps in the four age groups tested. The horizontal line at 50% correct represents chance performance for the unimodal, congruent, and catch conditions. The incongruent condition indexes an illusion and thus performance at or below 50% is meaningful. Error bars are +/- one standard error of the mean.

Magnitude of Illusions. Two 1-way ANOVAs with the single factor of Age (9-, 11-, and 13-year-olds, and adults) were conducted separately for fission and fusion (see Figure S2). Neither of the one-way ANOVAs for fission and fusion produced a main effect of Age (fission: $F(3,76) = 1.72$, $p = .17$, $\eta_p^2 = .06$; fusion: $F(3,76) = 0.89$, $p = .45$, $\eta_p^2 = .03$). When compared against zero, all age groups showed a small but significant fission illusion (all $t(19) > 4.51$, *p*s < .05, *d*s > .39). None of the groups demonstrated a significant fusion illusion (all $t(19) < 1.45$, *p*s > .08, *d*s < 1.01).

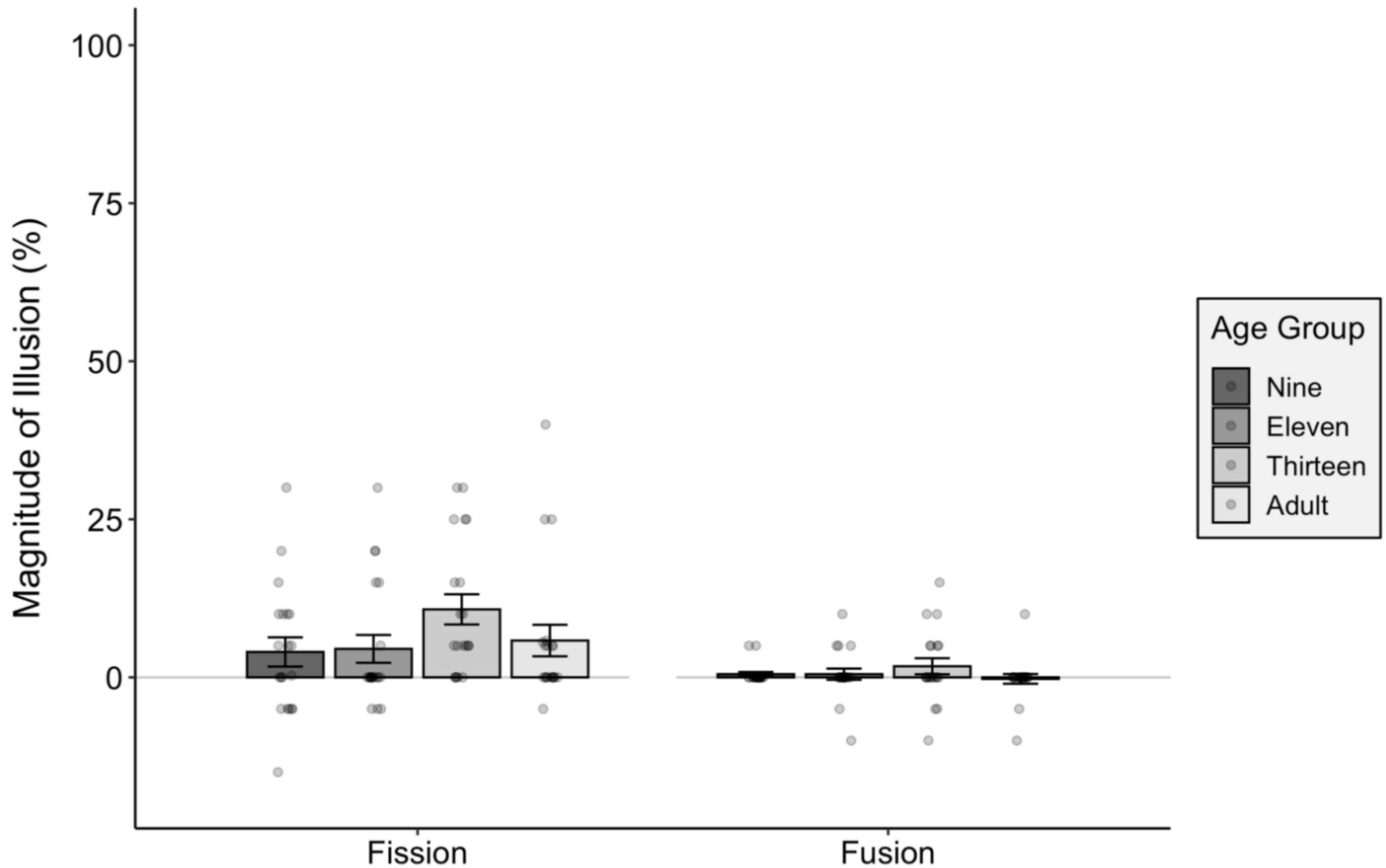

Figure S2. Mean magnitudes of the tap-induced sound fission (left) and fusion (right) illusions for the four age groups tested. The magnitude of both illusions was calculated by subtracting the accuracy of the incongruent condition from the accuracy of the congruent condition: fission (1B1T – 1B2T) and fusion (2B2T – 2B1T). Individual data points represent the performance of each participant. Error bars are +/- one standard error of the mean.

Signal Detection Analysis. The analysis of the *d'* data revealed that all age groups at every level of taps (0, 1, and 2) were significantly different from zero (all $t(19) > 20.82$, *p*s < .001, *d*s > 4.67; see Figure S3). None of any level of tap produced a significant effect of Age (all $F(3,76) < 2.02$, *p*s > .05, η_p^2's < .09).

For *c,* none of the age groups differed from zero in the 0-tap condition (all $t(19) < 1.46$, *p*s > .16, *d*s < .33). In the 1-tap condition, in which a significant positive bias is associated with the fusion illusion, none of the age groups differed from zero (all $t(19) < 0.85$, *p*s > .27, *d*s < .19). In the 2-tap condition, in which a significant negative bias is associated with the fission illusion, all age groups were significantly lower than zero (all $t(19) > -1.98$, *p*s < .04, *d*s > .44). None of the three levels of tap produced a significant effect of Age (all $F(3,76) < 2.02$, *p*s > .11, η_p^2's < .16).

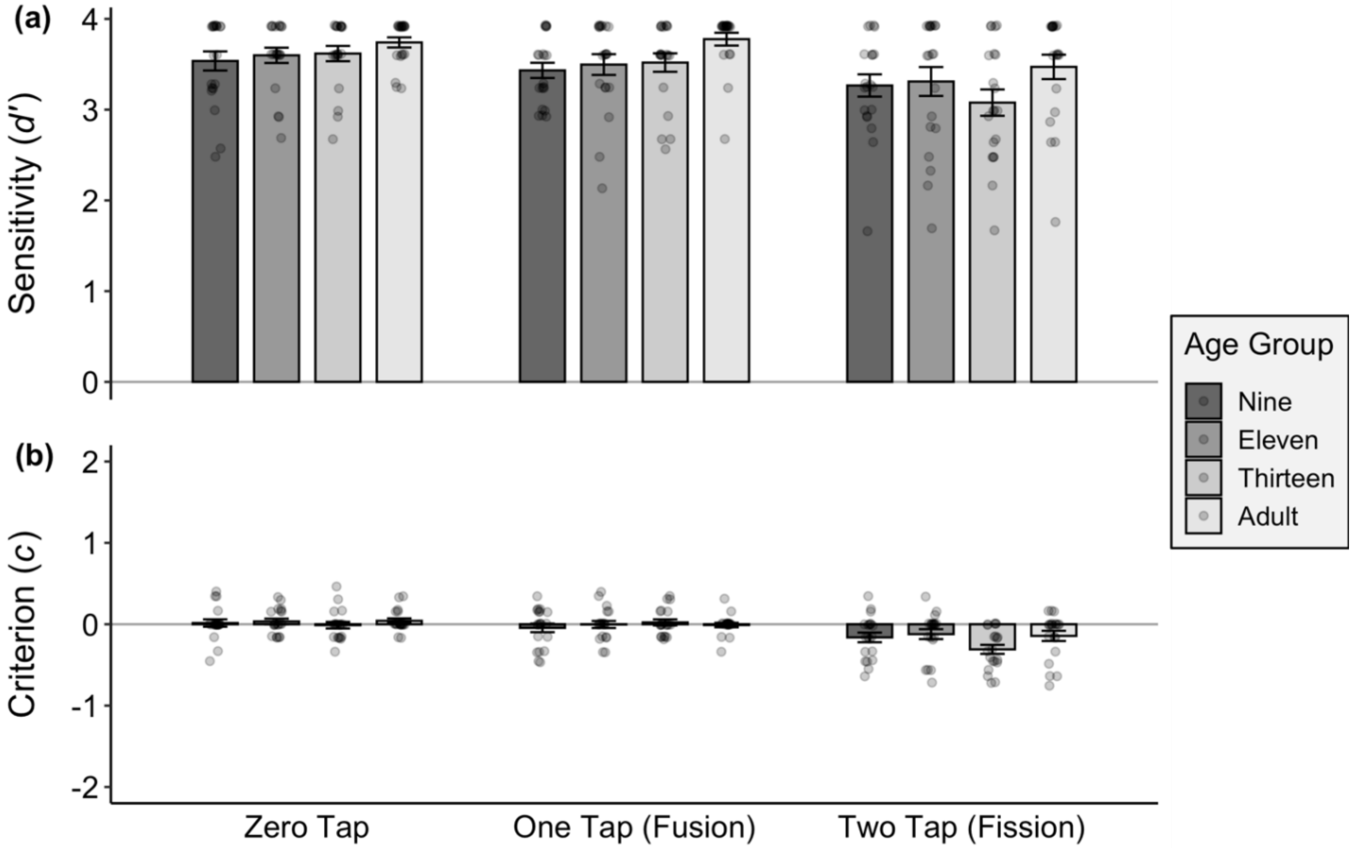

Figure S3. Signal detection – discriminability (d') and criterionI). **(a)** Mean *d'* scores for the four age groups tested. The *d'* represents the ability to distinguish two beeps (target) from one beep (noise) when paired with either zero (unimodal), one, or two taps. **(b)** Mean *c* scores for the four age groups tested. A shift toward a positive criterion indicates a bias to report "one beep", whereas a shift toward a negative criterion indicates a bias to report "two beeps". Individual data points represent the mean *d'* and *c* of each participant in **(a)** and **(b)**, respectively. Fission is represented in the two-tap condition (right), whereas fusion is represented in the one-tap condition (middle). Error bars are +/- one standard error of the mean.

Tap-induced Sound Task Conclusions

Given the absence of the fusion illusion, the weak fission illusion, and its lack of change across age groups, we chose not to include the tap-induced sound illusion results in the main body of this study. In comparing the magnitudes of fission and fusion observed for the sound-induced tap experiment, the very small magnitude of the fission illusion and absence of a fusion illusion were unexpected given that audition and touch typically demonstrate a more balanced bidirectional influence compared to other modality pairings (Hecht & Reiner, 2009; Occelli et al., 2011). This can likely be attributed to the relative reliabilities of the auditory and tactile signals. For the fission and fusion illusions to occur, the more reliable sense influences the perception of the less reliable sense, resulting in the illusory percept. In the current design, it was likely that the auditory signal was substantially more reliable than the tactile signal. This explanation is supported by a study that was able to produce a tap-induced sound fission illusion by reducing the amplitude of the auditory signal—and thus its reliability—relative to the tactile signal (Bresciani & Ernst, 2007; see Andersen et al., 2004 for an audiovisual study).

For future work, it would be interesting to equate the relative reliabilities of the auditory and tactile signals in adults before using those parameters to test groups of children as that could reveal differential developmental patterns between the sound-induced tap and tap-induced sound illusions. Given that the temporal resolution of unimodal audition matures earlier (~ age 4; Wightman et al., 1989) than that of unimodal touch (~ age 6; Röder et al., 2013), the earlier stabilization of audition may have greater reliability, and thus influence touch rather than the reverse, at the critical age of maturation.

Effect of Task Order

All participants completed both the sound-induced tap illusion task and the tap-induced sound illusion task. To control for any potential influences of completing one task prior to another, the order of performing the tasks was counterbalanced between participants. Although there were no a priori expectations that task order would matter, when included as a factor in the analyses of magnitude of illusions, *d'*, and *c*, some interactions emerged.

Results when Task Order is included. Two ANOVAs, each with the factors of Age (9-, 11-, 13-year-olds, and adults) and Task Order, were conducted separately on the magnitude of fission and fusion data from the sound-induced tap illusion task. For fission, the main effect of Task Order was not significant ($F(1,72) = 0.26$, $p = .61$, $\eta_p^2 < .01$) nor did it interact with Age ($F(3,72) = 0.44$, $p = .73$, $\eta_p^2 = .02$). For fusion, the main effect of Task Order was significant ($F(1,72) = 11.28$, $p = .001$, $\eta_p^2 = .14$) but not the Age by Task Order interaction ($F(3,72) = 1.06$, $p = .37$, $\eta_p^2 = .04$). Those who completed the tap-induced sound illusion task prior to the sound-induced tap illusion task showed a significantly larger fusion illusion ($t(55.86) = 3.01$, $p = .004$, $d = .67$) than those who completed the two tasks in the reverse order. This was not the case for the fission illusion ($t(77.56) = 0.50$, $p = .62$, $d = .11$).

Three two-way ANOVAs were conducted for each level of Beep on the *d'* scores with the factors of Age and Task Order. For the 0-beep and 2-beep conditions, the effect of Task Order was not significant (0-beep: $F(1,72) = 1.06$, $p = .31$, $\eta_p^2 = .01$; 2-beep: $F(1,72) = 0.11$, $p = .74$, $\eta_p^2 < .01$) nor was the interaction with Age (0-beep: $F(3,72) = 0.92$, $p = .44$, $\eta_p^2 = .04$; 2-beep: $F(3,72) = 0.43$, $p = .73$, $\eta_p^2 = .02$). For the 1-beep condition, however, the main effect of Task Order was significant ($F(1,72) = 13.50$, $p < .001$, $\eta_p^2 = .16$), but not the Age by Task Order interaction ($F(3,72) = 1.32$, $p = .27$, $\eta_p^2 = .05$). Those who completed the tap-induced sound illusion task prior to the sound-induced tap illusion task had poorer discriminability only in the 1-beep condition (representative of the fusion illusion, $t(64.6) = 3.11$, $p = .003$, $d = .69$).

Finally, three two-way ANOVAs were conducted for each level of Beep on the *c* scores with the factors of Age and Task Order. In the 0-beep condition, both the main effect of Task Order ($F(1,72) = 4.03$, $p = .049$, $\eta_p^2 = .05$) and its interaction with Age ($F(1,72) = 3.40$, $p = .02$, $\eta_p^2 = .12$) were significant. However, further decomposition of Task Order did not reveal a significant effect of Age either when completing the sound-induced tap illusion task first ($F(3,36) = 2.72$, $p = .06$, $d = .18$) or completing the tap-induced sound illusion task first ($F(3,36) = 1.07$, $p = .38$, $d = .08$). In the 1-beep condition, the effect of Task Order was significant ($F(1,72) = 7.48$, $p = .008$, $\eta_p^2 = .09$), but not the interaction with Age ($F(3,72) = 0.77$, $p = .51$, $\eta_p^2 = .03$). Those who completed the tap-induced sound illusion task first were significantly more biased in the 1-beep condition than those who completed tap-induced sound illusion task second ($t(61.88) = -2.48$, $p = .02$, $d = .56$). For the 2-beep condition, significance was not reached for the main effect of Task Order, Age, or the interaction between Task Order and Age (all *p*s > .19).

In summary, those who completed the tap-induced sound illusion task prior to the sound-induced tap illusion task showed a greater magnitude of the sound-induced tap *fusion* illusion, overall poorer discriminability (i.e., *d'*) in the *1-beep* condition, and a stronger bias (i.e., *c*) also in the *1-beep* condition. Based on the results of the magnitude of fusion illusion, and t'e *d'* and *c* in the 1-beep condition, they suggest that fusion was more susceptible to previous context than was fission. We excluded the factor of Task Order in the results section presented in the manuscript for two reasons: 1) the order in which the tasks were conducted did not interact with Age for any of the three measures (magnitude of illusions, *d'*, and *c*), and 2) inclusion of this factor did not change the overall pattern of significance across all the statistics, and thus did not change the main conclusions drawn in the study.

Tactor Machine Switch

As noted in the Method, the original tactor device began to fail midway through data collection and a replacement device was used to complete data collection. To determine if switching to a new machine midway through data collection influenced performance, the accuracy of two groups of adults were compared. In total, there were 30 adults tested (16 on the original machine and 14 on the replacement machine). To balance this comparison, we selected the first 14 tested on the original machine and all the 14 tested on the replacement machine. The mixed ANOVA included a between-subject factor of Machine (old vs. new), and two within-subject factors of Condition (unimodal, congruent, incongruent, and catch) and Target (1 tap vs. 2 taps). The main effect of Machine was not significant ($F(1,26) = 1.03$, $p = .32$, $\eta_p^2 = .04$), and it did not interact with either within-subject factors (Target: $F(1,26) = 0.91$, $p = .35$, $\eta_p^2 = .03$; Condition: $F(3,78) = 0.43$, $p = .60$, $\eta_p^2 = .02$). As such, the factor of Machine was not considered in any further analyses. This result is consistent with the null effect of machine on adult performance from Stanley et al. (2019; Chapter 2) in which the same two machines were used to measure judgments of simultaneity. The data included in the result section consist of all 16 participants using the original machine and 4 participants using the replacement machine.

Supplementary Materials References

Andersen, T. S., Tiippana, K., & Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cognitive Brain Research*, *21*(3), 301–308. https://doi.org/10.1016/j.cogbrainres.2004.06.004

Bresciani, J. P., & Ernst, M. O. (2007). Signal reliability modulates auditory-tactile integration for event counting. *NeuroReport, 18*(11), 1157–1161. https://doi.org/10.1097/WNR.0b013e3281ace0ca

Hecht, D., & Reiner, M. (2009). Sensory dominance in combinations of audio, visual and haptic stimuli. *Experimental Brain Research, 193*(2), 307–314. https://doi.org/10.1007/s00221-008-1626-z

Occelli, V., Spence, C., & Zampini, M. (2011). Audiotactile interactions in temporal perception. *Psychonomic Bulletin & Review, 18*(3), 429–454. https://doi.org/10.3758/s13423-011-0070-4

Röder, B., Pagel, B., & Heed, T. (2013). The implicit use of spatial information develops later for crossmodal than for intramodal temporal processing. *Cognition, 126*(2), 301–306. https://doi.org/10.1016/j.cognition.2012.09.009

Stanley, B. M., Chen, Y. C., Lewis, T. L., Maurer, D., & Shore, D. I. (2019). Developmental changes in the perception of audiotactile simultaneity. *Journal of Experimental Child Psychology, 183*, 208–221. https://doi.org/10.1016/j.jecp.2019.02.006

Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child Development, 60*(3), 611–624. https://doi.org/10.1111/j.1467-8624.1989.tb02742.x

Chapter 4

Auditory–tactile simultaneity perception in adults treated for bilateral congenital cataracts

Brendan M. Stanley[a], Yi-Chuan Chen[b],

Terri L. Lewis[a,c], Daphne Maurer[a,c], and David I. Shore[a,d]

Author Note

[a] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Canada

[b] Department of Medicine, Mackay Medical College, New Taipei City, Taiwan

[c]Department of Opthamology and Vision Sciences, The Hospital for Sick Children, Toronto, Canada

[d]Multisensory Perception Laboratory, Division of Multisensory Mind Inc., Hamilton, Canada

This chapter has not been published in an academic journal.

Abstract

Patients treated for bilateral cataracts demonstrate lifelong deficits for both unimodal vision and cross-modal interactions involving vision. In fact, such patients show abnormal audiovisual, but not visuotactile, simultaneity judgment (Chen et al., 2017). Two hypotheses were proposed: first, according to the cross-modal calibration hypothesis, during normal development vision calibrates audition for audiovisual pairings, and touch calibrates vision for visuotactile pairings. However, the abnormal visual development in the bilateral cataract patients prevented the proper calibration of audition by vision. Alternatively, it may have been the development of superior auditory processing responsible for the abnormal audiovisual but not visuotactile simultaneity perception. The present study attempted to discern which of these two hypotheses are more probable by measuring audiotactile simultaneity perception in these patients. Cross-modal calibration would predict normal audiotactile simultaneity perception (given vision is not involved), whereas superior auditory processing would predict abnormal audiotactile perception. Preliminary results show that patients have normal audiotactile simultaneity perception, thus supporting the cross-modal calibration hypotheses. However, because of significant and unplanned limitations, further testing is required.

Acknowledgements

We thank Sally Stafford and Miroslav Cika for their administrative and technical contributions, respectively.

Keywords

Visual deprivation; Cataract; Temporal perception; Audition; Touch

Introduction

Congenital cataracts deprive the infant of patterned visual input until the cataract(s) are removed surgically shortly after birth and vision is restored with the use of contact lenses. Both the visual system and its interactions with other modalities experience life-long permanent effects from the absence of appropriate visual stimulation immediately after birth (see Maurer, 2017 and Röder et al., 2021, for reviews). Recent work by Chen et al. (2017) revealed that adult patients treated for bilateral congenital cataracts as infants demonstrated abnormal audiovisual but normal visuotactile simultaneity perception, despite both modality pairings involving the formerly deprived visual modality. The curve representing the patient's audiovisual simultaneity judgments was wider on the vision-leading side compared to typically developed controls indicating poorer sensitivity (i.e., wider temporal simultaneity window), and a shift in the point of subjective simultaneity (PSS) toward the vision-leading side. Two hypotheses were posited to explain this finding; the transient period of visual deprivation may have interfered with the normal calibration of the auditory modality by vision, or the absence of vision allowed for the auditory modality to develop superior ability. However, it remains unclear which hypothesis accounts for the abnormal audiovisual temporal perception observed in patients treated for bilateral congenital cataracts. The present study aims to decipher between these two hypotheses by testing a third modality pairing, specifically audiotactile simultaneity perception.

Bilateral Congenital Cataract Patients

Early visual deprivation profoundly affects life-long vision. Despite cataract reversal surgery within the first year of life and a lifetime of corrected vision, adults born with bilateral cataracts show deficits in visual acuity (Maurer & Lewis, 2001), contrast sensitivity (Ellemberg et al., 1999, 2000), global motion (Ellemberg et al., 2002; Hadad et al., 2012), and face perception (De Heering & Maurer, 2014; Robbins et al., 2010). Most of these deficits are experienced as sleeper effects (Maurer et al., 2007) in which the deficit does not emerge until much later in development. Both the immediate and the late onset visual deficits support the hypothesis for a *sensitive period* during which proper visual stimulation is necessary for normal visual development. It is likely that the visual deprivation experienced by the infant prevents the typical development of the neural architecture to support certain low- and high-level visual functions, leading to these lifelong deficits. This holds true not only for unimodal visual functions, but also for cross-modal interactions involving the visual modality.

Bilateral congenital cataract patients also exhibit deficits in cross-modal perception. Normal multisensory perception relies on accurate sensory signals and perceptual experience with the environment (Gori et al., 2021). When a sensory modality is deprived or the signal is degraded during early development, multisensory perception can develop abnormally. This suggests that cross-modal perception involving vision also has a sensitive period after birth like that of normal unimodal visual development. For example, adult patients treated for bilateral congenital cataracts demonstrate behavioural (Chen et al., 2017, 2014; Putzar et al., 2007, 2010), electrophysiological (Röder et al., 2013; Segalowitz et al., 2017), and functional/structural (Collignon et al., 2015; Feng et al., 2021) differences/deficits when compared to typically developed controls. Some of these are likely a product of cross-modal recruitment (Bavelier & Neville, 2002; Pascual-Leone et al., 2005) of the brain areas typically dedicated to visual processing by competing modalities (such as audition or touch) during the period of deprivation. This is supported by electrophysiological and functional/structural imaging studies both in the blind (e.g., Büchel et al., 1998; Dietrich et al., 2013; Sadato et al., 1996) and in patients treated for bilateral congenital cataracts (Collignon et al., 2015; Saenz et al., 2008). In other cases, it is possible that multisensory perception may be altered because of suppressed visual processing in the primary visual cortex (Guerreiro et al., 2015), possibly because of reorganization of the visual cortex in response to the noisy visual signal during the period of deprivation. Although there are several multisensory studies in cataract patients where vision is one of the modalities tested, there are virtually no studies examining multisensory performance for audiotactile interactions. The present work addresses this gap in the literature.

Cross-modal Temporal Simultaneity Judgments

Temporal coincidence provides a fundamental cue for the perceptual system to decide whether cross-modal signals should be integrated or segregated. When stimuli arrive to the observer with large temporal asynchronies (e.g., a beep presented one second before a tactile tap), the perceptual system considers the two signals as originating from independent events and thus does not integrate. On the other hand, when stimuli arrive at the same time or within short temporal asynchronies, these are likely to be integrated into a single percept and be attributed to a singular event.

A simultaneity judgment task (e.g., Chen et al., 2016, 2017, 2018; Stanley et al., 2019 or Chapter 2) can be used to estimate the temporal limits in which the observer perceives temporal coincidence—which we consider as a proxy for the temporal window of integration. This task involves presenting two stimuli—each from different modalities—either together (veridically synchronous) or separated by a temporal delay. We compute the percent of trials at each of a range of stimulus onset asynchronies (SOAs) in which the observer reports "simultaneous". The percent of simultaneous responses as a function of SOA typically resembles a normal (i.e., Gaussian) distribution with the peak comprised of the highest percentage of simultaneous responses. As the SOA increases. The distribution tapers off with the lowest percentage of simultaneous responses being reflected in the tails of the distribution. From this distribution, two parameters of interest can be extracted: the point of subjective simultaneity (PSS) which denotes the SOA corresponding to the highest percentage of simultaneous responses (typically the peak of the response curve), and the temporal simultaneity window which reflects the width of the curve (one standard deviation of the mean was used in the current study)[6]. These two measures provide a proxy for cross-modal temporal perception that can be used as a comparator across populations.

---

[6] The criteria for measuring the temporal simultaneity window can differ depending on preferences of the research group. Most other studies define the temporal simultaneity window as the width of the distribution at 50% perceived simultaneous which is when the observer is most uncertain (e.g., Chen et al., 2016, 2018;

Purpose of Present Study

A previous study by our research group examined cross-modal temporal perception in patients treated for unilateral or bilateral congenital cataracts for both audiovisual and visuotactile modality parings using the simultaneity judgment task (Chen et al., 2017). For audiovisual stimuli, the simultaneity judgment curve of the patients treated for bilateral cataract was abnormal when compared to the typically developed controls. Specifically, patients made more simultaneous responses when the visual stimulus led the auditory stimulus, resulting in a greater shift of the PSS toward the visual leading side. As for visuotactile simultaneity perception, no deficits were observed in that the patient's simultaneity curve was identical to that of the typically developed controls. Considering that the deficit experienced by the patients affected their visual modality, and both the audiovisual and visuotactile simultaneity judgment tasks used the same visual stimulus, this suggests that the transient period of visual deprivation immediately after birth differentially affects modality pairings in cross-modal temporal perception (Chen et al., 2017).

The abnormal audiovisual but not visuotactile temporal perception observed in patients treated for bilateral cataract is best explained by the cross-modal calibration hypothesis (see Burr & Gori, 2012) in that the more accurate modality calibrates the less

Machulla et al., 2016; Stanley et al., 2019) whereas others use the 75% perceived simultaneous (e.g., Hillock-Dunn & Wallace, 2012; Stevenson et al., 2014, 2018) or the standard deviation of the simultaneity curve (e.g., Zampini et al., 2005).

accurate modality during development. For audiovisual perception, vision typically provides the more accurate signal and thus typically calibrates the auditory modality (e.g., Gori et al., 2012), whereas for visuotactile perception, touch is considered the more accurate modality and thus calibrates vision (e.g., Gori et al., 2008). So, for the patients treated for bilateral cataracts in which the visual modality is presumed to be adversely affected, it is possible that vision was not able to calibrate the auditory modality accurately, producing abnormal development of audiovisual temporal perception (Chen et al., 2017). However, for visuotactile interactions, vision would have been calibrated appropriately by touch, thus sparing typical development of visuotactile temporal perception (Chen et al., 2017).

An alternative hypothesis also needs to be considered. Perhaps the auditory modality itself is responsible for the abnormal audiovisual simultaneity judgment curve observed in patients treated for bilateral cataracts. This explanation would support the normal visuotactile simultaneity perception given the absence of audition in this condition. As discussed in Chen et al. (2017), it is possible that short-term visual deprivation following birth allowed for the development of faster-than-normal auditory processing (e.g., functional compensation, see Kupers & Ptito, 2014). This could occur from either an attentional mechanism (i.e., prior entry; Shore & Spence, 2005; Spence et al., 2001), or an unrefined ability to adapt to distance-dependent arrival time differences for audiovisual signals because of the absence of early-life visual input (see Chen et al., 2017 for details).

The present study aimed to determine if the abnormal audiovisual, but not visuotactile, temporal perception observed previously in patients treated for bilateral congenital cataract (Chen et al., 2017) was a product of cross-modal calibration or the abnormal development of faster-than-normal auditory temporal processing. To do so, we tested adult patients treated for bilateral congenital cataract using an audiotactile simultaneity judgment task. If the audiotactile simultaneity judgement curve in patients is found to be normal, then the results of Chen et al., (2017) are more likely attributed to the cross-modal calibration hypothesis given that deficits in the visual modality had no influence in calibrating the interactions between audition and touch. On the other hand, if the audiotactile simultaneity judgment curve is abnormal in the patients, especially on the tactile leading side, then the account that short-term visual deprivation at birth results in the development of faster-than-normal auditory processing is supported.

Method

Participants

Fifteen patients (mean age = 27.4 years; age range = 18–38 years) treated for dense bilateral congenital cataracts took part in the study (see Table 1 for a history of the patients' visual deprivation and treatments). Twenty-seven control participants (mean age = 22.0 years; age range = 18–37) were taken from a previous study (Stanley et al., 2019 or Chapter 2) because the COVID-19 pandemic prevented the planned data collection of two age- and gender-matched control participants per patient. These control participants were tested under nearly identical conditions (see Design below for minor differences). All patients and controls had normal self-reported auditory and tactile acuity. We required a minimum of 40 s of arc on the Randot test of stereoacuity for control participants to participate. This task was chosen as a benchmark of normal visual function because binocular vision is most susceptible to abnormal visual experience during early life. Both patients and control participants provided written consent to participate. The study was cleared by the Research Ethics Boards of McMaster University and The Hospital for Sick Children (SickKids Hospital) and conformed to the Tri-Council statement on Ethical Conduct of Research Involving Humans (TCPS2; Canada).

Table 1
Clinical details of the bilateral congenital cataract patients.

Patient	Age (yrs)	Diag. (days)	Contact Lens[a] (days)	Snellen acuity[b]	Stereo acuity[c] (arc sec)	Binocular fusion[d]	Nystagmus[e]	Additional details
MD	38	0	129	OD: 20/40 OS: 20/100	>400	Alternator	Latent OU	Capsular membrane needling OD at 3 months old; Secondary membranes & Elschnig's pearls removed OU at 9.7 months old
JS1	27	124	152	OD: 20/40 OS: 20/63	400	Diplopia	Manifest	Strabismus surgery OU at 1 year old; Pupil enlarged and pupillary membrane removed OD at 2 years of age; Glaucoma OU
JB	22	82	98	OD: 20/125 OS: 20/50	>400	Diplopia	Latent OD	Strabismus surgery OU at 10 months old
WS	23	0	9	OD: Finger counting at 153 cm OS: 20/25	>400	Suppresses OD	Manifest	Ahmed Tube OD; Glaucoma OU
ZC	32	69	142	OD: 20/100 OS: 20/125	400	Diplopia	Latent OU	Glaucoma OU; Glaucoma procedure at 4 years of age; Hallermann Streiff Syndrome
JS2	25	61	92	OD: 20/100 OS: 20/25	>400	Diplopia	Latent OD	Microcornea OU; left esotropia surgery at 1.5 years old; Strabismus

								surgery OU at 3 years old; Glaucoma OU
AB	26	0	106	OD: 20/80 OS: 20/100 at 50 cm	>400	Suppresses OS	Latent OS	Microcornea OU; Left membranectomy at 9 months old; Right pupilectomy & membranectomy at 2.5 years old; Glaucoma OU
CB	34	38	91	OD: 20/32 OS: 20/40	>400	Diplopia	Latent OU	Left esotropia surgery at 1.5 years of age
CP	37	142	187	OD: 20/40 OS: 20/80	>400	Diplopia	Latent OU	Left esotropia surgery at 1.7 years old; lens implant OU at 26 years old
NA	27	104	134	OD: 20/100 OS: 20/80	>400	Diplopia	Intermittent nystagmus OU with an esotropia	Microcornea OU; Strabismus surgery OU at 3.9 years old; Glaucoma OU
JO	18	0	17	OD: 20/40 OS: 20/32	140	Fused		Glaucoma OU
MM	26	12	48	OD: 20/32 OS: 20/40	>400	Fused	Manifest	Glaucoma OU
SA	27	113	147	OD: 20/32 OS: 20/40	>400	Diplopia		Strabismus surgery OS at 1 year of age
VO	23	1	34	OD: 20/63 OS: 20/40	>400	Diplopia	Latent OU	Removed scar tissue at 2 months old; Strabismus surgery OU at 1 year of age; Strabismus surgery for left esotropia at 3.8 years old; Glaucoma OS

CR	26	188	238	OD: 20/32 OS: 20/200	>400	Fused	Latent OS	Strabismus surgery OU at 9 months old; excision of left orbital dermoid at 1.8 years old; Glaucoma OU

Note. Table format replicated from Chen et al. (2017)
OD: Right eye; OS: Left eye; OU: Both eyes
[a]Age at time of first optical correction after cataract surgery (defined as duration of deprivation)
[b]Measured at time of test
[c]Measured using Randot at time of test
[d]Measured using Worth 4 Dot at time of test
[e]History of nystagmus based on records since first optical correction

Apparatus and Stimuli

Patient data were collected at SickKids Hospital and control data were collected at McMaster University. At both locations, the experiment was conducted in a dimly lit room. The auditory stimulus consisted of a 10 ms peak-to-peak white noise burst (with 2 ms onset and offset ramping) delivered via closed-ear headphones (Sennheiser HDA-200) at 107 dB SPL. The tactile stimulus consisted of a 3 mm indentation to the participants right index finger for 10 ms delivered by one of two custom-built tap-devices[7]. The stimuli were generated and controlled using MATLAB (Mathworks) and the PsychToolBox-3 package (Brainard, 1997; Kleiner et al., 2007).

To reduce the possibility of extraneous noises interfering with the task, we used closed-ear headphones, played free-field white noise (at 73 dB SPL measured at the ear while wearing the headphones), and enclosed the tap device in a custom-built sound-attenuating box constructed of sound-dampening ceiling tile and lined with shag carpet.

Design and Procedure

Each group (patients and controls) were tested using a different stimulus onset asynchronies (SOAs). For the patient group, 15 SOAs were tested (+/- 500 ms, +/- 400 ms, +/- 300 ms, +/- 200 ms, +/- 150 ms, +/- 100 ms, +/- 50 ms, and 0 ms), whereas for the adult control group, 13 SOAs were tested (+/- 1200 ms, +/- 800 ms, +/- 400 ms, +/- 300 ms, +/- 200 ms, +/- 100 ms, and 0 ms). The negative values indicate trials in which the

[7] Switching devices did not interact with performance in Stanley et al. (2019) and Stanley et al. (2023)

auditory stimulus preceded the tactile stimulus. For both experiments, each SOA was tested twice per block across 10 blocks resulting in 300 trials for patients and 260 trials for control participants. All stimulus durations and SOAs were verified with an oscilloscope.

Participants sat at a table facing forward with their right arm inside the sound attenuating box and their index finger resting gently over the tactile device (centered on the midline approximately 40 cm distal to the body). Control participants were instructed to keep their eyes closed during the experiment, whereas cataract patients wore a blindfold. On each trial, the task was to say "yes" verbally if the two stimuli were perceived as simultaneous, or "no" if the two stimuli were perceived separately. All responses were keyed in by the experimenter.

Both the control participants and cataract participants were required to complete two practice sessions, each with slightly different parameters. The first practice session consisted of 8 trials: 4 at large SOAs (controls: -1200, -800, 800, and 1200 ms; cataract patients: -500, -300, 300, and 500 ms) and 4 simultaneous trials (0 ms). Participants in both groups were required to achieve at least seven correct of eight trials (up to 3 three attempts) to participate in the study. The second practice session familiarized the participants with the SOAs included in the main experiment; one trial at each SOA was presented, and there were no accuracy requirements. Accuracy feedback was not provided for either the practice sessions or the main experiment. The duration of the experiment was less than 60 minutes for both groups, and breaks were encouraged as needed.

Analysis

All analyses were conducted in R Studio running R version 4.1.2 (R Core Team., 2021). Given that each group was tested using different SOAs, only the SOAs that overlapped were included in the present analysis (+/- 400 ms, +/- 300 ms, +/- 200 ms, +/- 100 ms, and 0 ms). For each participant, a mean "percentage simultaneous" score was calculated for each SOA. An outlier check was performed to determine if any participant's percentage simultaneous scores were more than three standard deviations beyond their respective group mean for each SOA.

Curve Fitting. Each participant's data were fitted using a non-linear least squares regression procedure. This iterative procedure obtained parameters of a best-fit Gaussian distribution by minimizing the residuals of the sum of squares (i.e., by minimizing the differences between the observed means at each SOA and the values determined by the model). Two parameters of interest are obtained from the fitted Gaussian distribution: the mean of the curve (or the SOA corresponding to the peak of the distribution), and of particular interest, the standard deviation of the curve (or, *sensitivity*, which is half of the width of the temporal simultaneity window). The Gaussian formula used was:

$$ae^{-0.5\left(\frac{x-m}{s}\right)^2} + b$$

where *m* represents the mean, *s* represents the standard deviation, *b* represents the minimum value, and *a* represents the maximum value. These four parameters were unconstrained, although the underlying iterative fitting algorithm does require an approximate estimate of starting values to be able to fit the data (used in the current analysis: m = 0, s = 100, b = 0, a = 1).

The sum of squares analysis was conducted by using the minpack.lm package (Elzhov et al., 2016) in R that solves non-linear least-squares analyses using an iterative algorithm. This package uses a modified Levenberg-Marquardt adaptive algorithm which combines two commonly used algorithms to optimize the fitting procedure (see Gavin, 2019, for a review). The gradient-descent algorithm was used when the estimated parameters are far from the optimal values (e.g., during the initial iterations of the fitting procedure); by updating parameters on each iteration in a "downhill" direction, the fitting can converge quickly on the optimal parameters. Once the model converged broadly on the optimal parameters, the Gauss-Newton algorithm was used to fine tune the fit by minimizing the sum-of-squares function.

Statistical Analysis. The percentage simultaneous data were submitted to a two-way mixed analysis of variance (ANOVA) with a between-subject factor of Group (cataract patients vs. controls) and a within-subject factor of SOA. The effect sizes were reported as partial eta-squared, and the degrees of freedom were corrected using the Greenhouse-Geisser correction when sphericity was violated.

From the least-squares curve fitting routine, the means and standard deviations were submitted to the Shapiro-Wilk test to determine if the data were normally distributed. For normally distributed data, an independent samples t-test was used, and for non-homogenous data, a non-parametric Wilcoxon test was used to compare the cataract patients to controls. In addition, because of the unequal sample sizes between the two groups, bootstrapped independent samples t-tests were conducted on both the mean and standard deviation scores. This allows for better estimates of the sampling distributions for each group using an iterative random sampling technique. The bootstraps were conducted using the wBoot package (Weiss, 2016) in R. This bootstrap function from the wBoot package also incorporates a bias-correction and acceleration (BCa) interval that corrects for bias and skewness in the bootstrap estimates, providing better estimates of the confidence intervals. Each bootstrap had 9999 iterations and the confidence interval level was 95%. Participants whose data could not be fit were excluded from analysis. All statistical tests were two-tailed, and the effect sizes are reported Cohen's d (for t-tests).

Results

Outliers and Fit Failures

Of the initial 15 cataract patients and 27 control participants, 5 patients and 6 controls were excluded. These exclusions resulted from either an outlier analysis on the percentage simultaneous scores at each SOA (1 cataract patient and 2 controls), or because of a failure of the least-squares analysis to fit the data (4 patients and 4 controls).

Percentage Simultaneous

The cataract patients and control participants produced percentage simultaneous scores that did not differ at any SOA tested (see Figure 1; see Appendix Figure A1 and Figure A3 for individual plots of patients and controls, respectively). A mixed ANOVA revealed a main effect of SOA ($F(3.56,106.80) = 178.14$, $p < .001$, $\eta_p^2 = .86$). There was no main effect of Group (cataract patients vs. controls; $F(1,30) = 0.13$, $p = .72$, $\eta_p^2 < .01$), nor an SOA by Group interaction ($F(3.56,106.80) = 0.51$, $p = .74$, $\eta_p^2 = .02$).

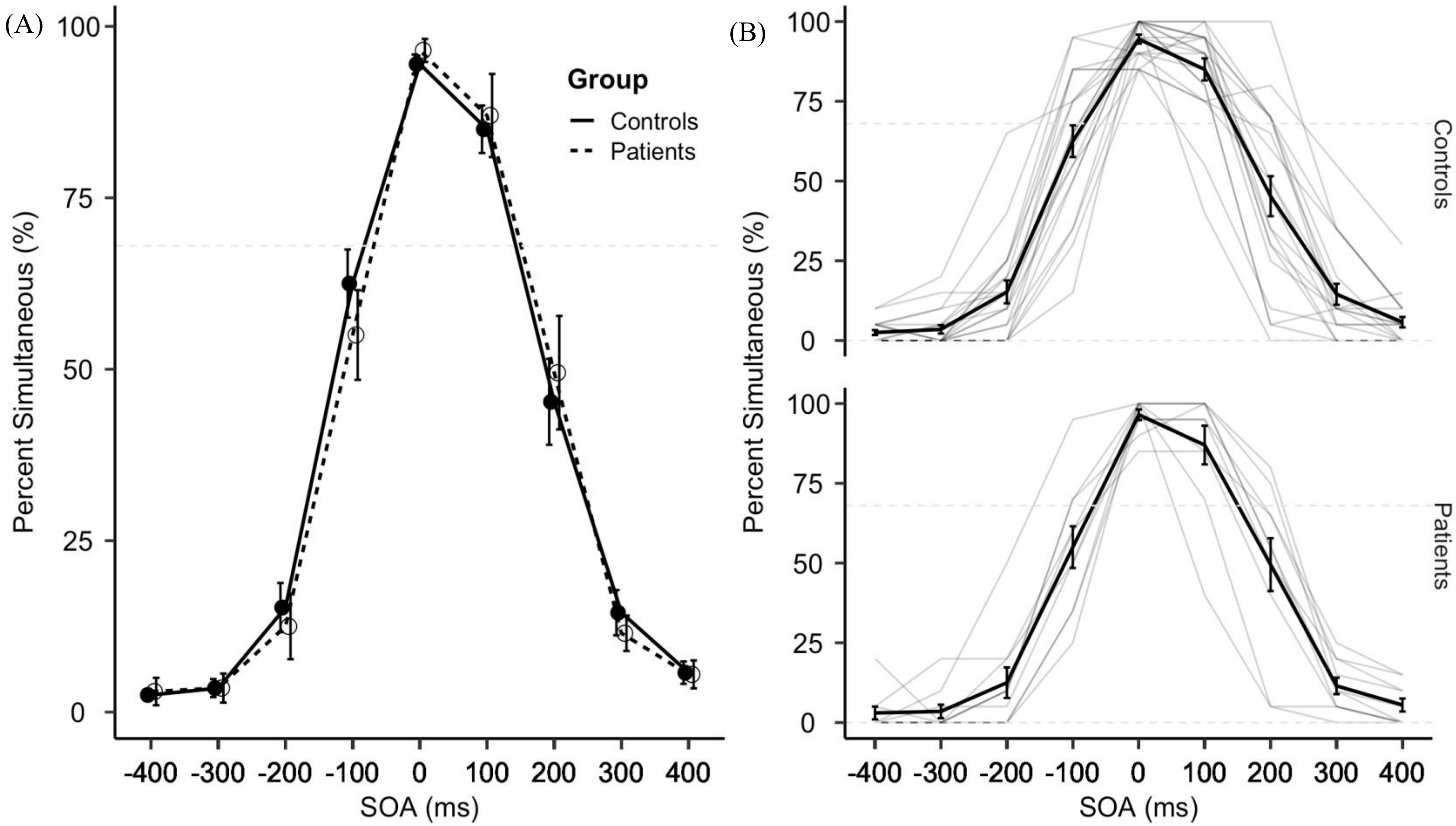

Figure 1. (A) The mean percentage of simultaneous responses for bilateral cataract patients (black dotted line) and control participants (solid line) across SOA. Error bars are +/- one standard error of the mean. (B) Group (black) and individual (grey) curves from the simultaneity judgment task.

Least-Squares Regression

PSS (i.e., estimated mean/peak of the fitted curve). The means of the estimated PSS values obtained from the least-squares fitting procedure did not differ between cataract patients and controls ($t(20.25) = 0.58$, $p = .57$, $d = .21$; see Table 2 and Figure 2; see Appendix Figure A2 and Figure A4 for individual plots of patients and controls, respectively). However, the Shapiro-Wilk normality test revealed that the cataract patient's data were not normally distributed ($p = .03$). As such, the non-parametric Wilcoxon Rank Sum test was also performed in which the two groups also did not differ ($p = .60$). Finally, given the unequal group sizes, and the violation of the assumption of normality for the data from cataract patients, the PSS scores were submitted to a bias-corrected bootstrap routine. Again, the patient group did not differ from the control group ($p = .58$; see Table 3).

Sensitivity (i.e., standard deviation of the curve or half the width of the temporal window). The means of the estimated sensitivities of the curves did not differ between cataract patients and controls ($t(25.7) = 0.60$, $p = .55$, $d = .20$; see Table 2 and Figure 2). Again, given the unequal group sizes, these scores were also submitted to the bias-corrected bootstrap routine. The results of the bootstrap also did not reach significance ($p = .52$; see Table 3).

Table 2.
Results of t-tests for least-squared estimated PSS and sensitivity

	Patients (n = 10)		**Controls** (n = 21)						
	M	*SD*	*M*	*SD*	CI	*t*	df	*p*	Cohen's *d*
PSS	41.3	25.2	35.3	28.9	[-15.2, 27.1]	0.58	20.25	.57	.21
Sensitivity (i.e., standard deviation)	123.2	30.1	131.6	45.8	[-36.7, 20.0]	0.60	25.70	.55	.20

Table 3.
Results of bootsrapped t-tests for least-squared estimated PSS and sensitivity

Bootstrapped *t* (9999 replications)	**Patients** (n = 10)	**Controls** (n = 21)					
	M (observed)	*M* (observed)	*M* Difference (observed)	*M* Difference (bootstrapped)	Bias (%)	CI	*p*
PSS	41.3	35.3	-5.94	-6.10	2.75	[-23.1, 15.2]	.58
Sensitivity (i.e., standard deviation)	123.2	131.6	8.34	8.57	2.79	[-16.4, 35.7]	.52

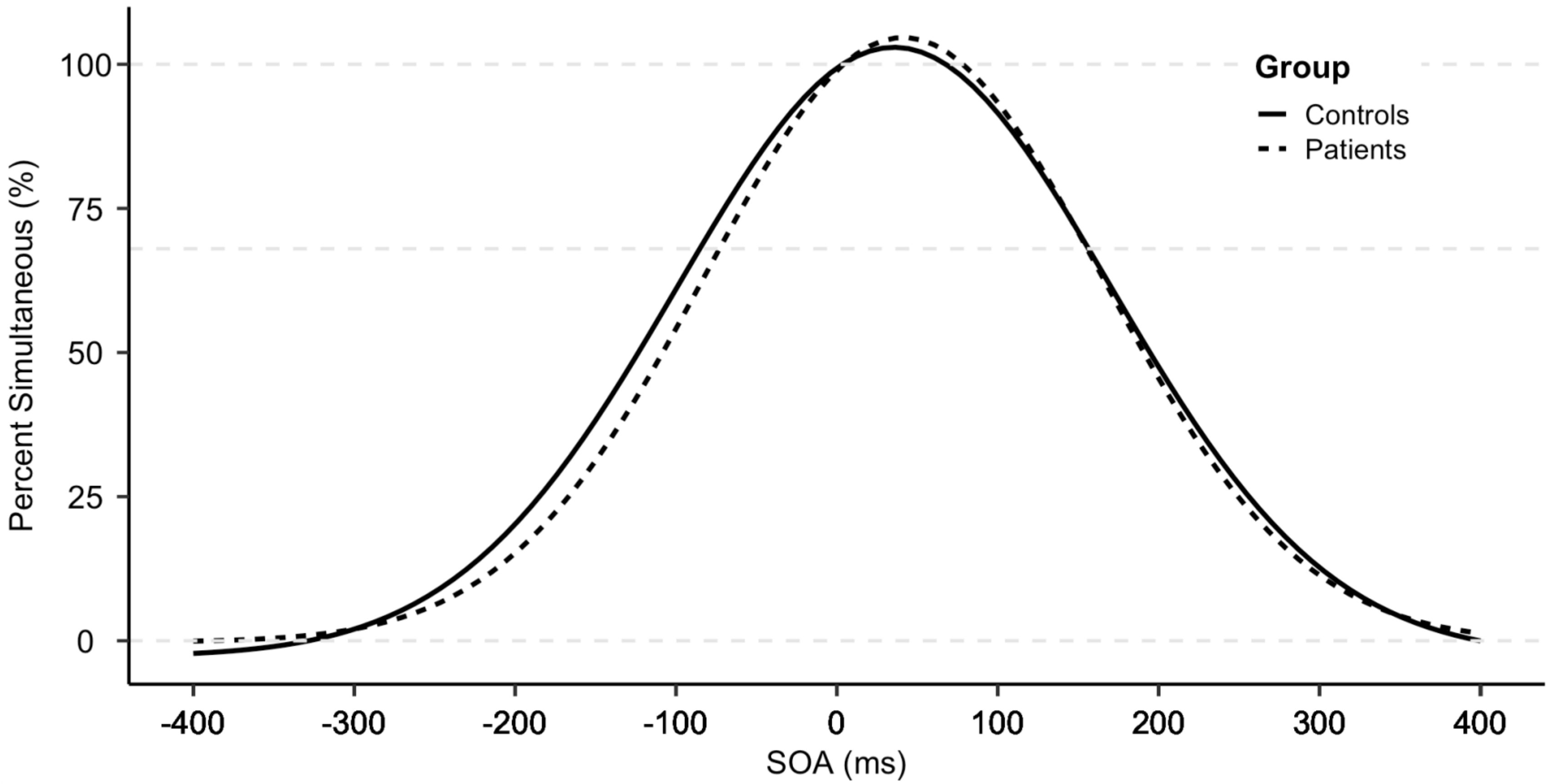

Figure 2. Results of the least squares curve fitting for bilateral cataract patients (black dotted line) and control participants (solid line).

Discussion

The current study compared audiotactile temporal perception in patients treated for bilateral cataracts to typically developed controls using a simultaneity judgment task. This was to determine if the abnormal audiovisual but not visuotactile simultaneity perception observed in patients treated for bilateral cataracts from Chen et al. (2017) could be best explained by the cross-modal calibration hypothesis, or faster-than-normal auditory processing. Results from the current study provided no evidence that the audiotactile simultaneity judgement curve for patients differed from typically developed controls. In fact, the omnibus ANOVA did not reveal a group difference in percent simultaneous responses across SOAs; nor did the bootstrapped t-tests reveal any differences for the PSS and sensitivity of the temporal simultaneity window estimated by the least-squares analysis. These results favour the cross-modal calibration hypothesis over the speeded auditory processing hypothesis in explaining abnormal audiovisual and normal visuotactile perception observed in the patients treated for bilateral cataracts in Chen et al. (2017). However, given several significant limitations introduced by the COVID-19 pandemic, these results should be interpreted with utmost caution, and be considered only a preview to future more rigorous testing.

Implications

Had significant limitations not been encountered (see Limitations below), this observation of normal audiotactile temporal perception in patients treated for bilateral cataracts supports a cross-modal calibration hypothesis. Chen et al. (2017) proposed two hypotheses to explain the observation that patients treated for bilateral cataracts showed abnormal audiovisual but normal visuotactile temporal perception. One hypothesis was based on the cross-modal calibration hypothesis in that the more accurate sense calibrates the less accurate sense with vision typically calibrating audition, and touch typically calibrating vision. Given that the patients treated for bilateral cataracts experienced abnormal visual development, vision could not calibrate audition optimally, whereas the highly precise sense of touch was still able to calibrate vision. The other hypothesis was that vision was not directly responsible for abnormal audiovisual perception, however, it was a faster-than-normal auditory processing that had developed in response to the absence of early visual experience (i.e., functional compensation). Atypical auditory processing and normal visual processing can explain why audiovisual but not visuotactile was abnormal.

To disentangle these two hypotheses, we examined audiotactile temporal perception in bilateral cataract patients. We hypothesized that if audiotactile simultaneity judgments did not differ from typically developed visually normal controls, then the likely hypothesis to explain the findings of Chen et al. (2017) is that of cross-modal calibration. Like visuotactile interactions, touch provided the more accurate signal compared to audition, and as such, touch would calibrate audition during development. Because the auditory-leading side of the simultaneity judgment curve for patients treated for bilateral cataracts in the current study does not differ from the controls, it suggests that the speed of auditory processing likely developed normally. However, this cannot be determined quantitatively because we were unable obtain estimates of processing times of each signal because of use of a different modelling routine than that of Chen et al. (2017) (see Limitations below).

Limitations - Data Collection

A full dataset of audiotactile simultaneity judgements was collected from patients treated for bilateral cataracts prior to the onset of the COVID-19 pandemic, however, restrictions on in-person data collection prevented the testing of age- and gender-matched visually normal controls. As such, the control data used in the present study was the group of adults previously tested in a developmental study by our group (Stanley et al., 2019 or Chapter 2) that used nearly identical methodology, stimuli, and testing parameters. However, one critical difference that prevented confidence in our conclusions concerns the different range of SOAs tested.

Limitations - Data Analysis and Modelling

In Chen et al. (2017), the parameters of interest (PSS and the sensitivity), were estimated using a bootstrapped curve fitting routine designed for simultaneity judgment tasks (Alcalá-Quintana & García-Pérez, 2013). This procedure, in addition to estimating the parameters of interest, also estimates the probability of participant error (e.g., attentional lapses or key-press mistakes) and estimates of processing times for each signal. For between-study consistency, this routine would have been the ideal choice to estimate the parameters of interest. However, the ad hoc control group used presented a significant challenge to this model. This routine estimates the probability of error at 0 ms

SOA, and at the two most extreme SOAs, which are then factored into the estimates PSS and the sensitivity. This is done to increase the precision of these estimates by minimizing the contribution of undesired error such as attentional lapses or key-press mistakes. Because the control group in the current study was obtained from a developmental study (Stanley et al., 2019 or Chapter 2), only the overlapping SOAs from that study and current study were used, which ranged from +/- 500 ms. This range turned out to be narrower than ideal for this model because some participants' performance (mostly in the control group) did not reach the floor at either of the extreme SOAs. Because the model considers above floor performance at the tails of the distribution as participant error, the estimates of the PSS and sensitivity of the simultaneity window may be contaminated.

An alternative fitting procedure that was considered was Maximum Likelihood Estimation (MLE; Ernst & Bülthoff, 2004). This inference-based procedure estimates parameters by maximizing a likelihood distribution based on the observed data. This is done by comparing a distribution defined by sets of parameter values (i.e., PSS and width of the simultaneity window) against the distribution of the observed data to determine which estimated parameters best describe the observed distribution. This fitting procedure has been used previously to estimate the PSS and temporal simultaneity window from a simultaneity judgment curve (e.g., Stone et al., 2001; Vroomen et al., 2004; Zampini et al., 2005). This model was deemed unsuitable for the current study again because of the overlapped SOAs we were forced to use. The MLE procedure for simultaneity judgment tasks assumes a Gaussian distribution within a probability range from 0 to 1 when estimating the parameters of interest. When enough SOAs and a sufficiently wide range of SOAs are tested, a simultaneity judgement curve strongly resembles a Gaussian distribution, and the data are typically fit well using MLE. However, in cases when these criteria are not met, which was the case for several patients and control participants, the MLE model fails to fit the curve entirely or yields parameters that may not reflect the true shape of the curve. For the current study, one way to circumvent having a narrow SOA range and a less-than-ideal number of SOAs tested is to not restrict the upper boundary of the probability distribution. By not restricting the curve, the fit can be elongated to provide better fits, and thus better estimates of the parameters that describe the observed data. However, due to the nature of the underlying mathematical computations of MLE (which is beyond the scope of this paper), this was not feasible.

A Least Squares Analysis was chosen to estimate the PSS and temporal simultaneity window to overcome the limitations presented by the two routines discussed previously. This regression-based routine estimates the parameters of a Gaussian distribution that best fits the observed data by minimizing the squared error between the predicted and observed curves. The estimated parameters consist of the mean (i.e., PSS) and the standard distribution (used as a proxy of the temporal simultaneity window) of the best fit distribution. Unlike MLE, Least Squares Analysis can fit Gaussian distributions that are not restricted to a probability range between 0 and 1, thus providing overall better fitted Gaussian curves, and thus more accurate estimates of PSS and the temporal simultaneity window. Although this method was found to be superior of the three candidates, and thus used in the current study, significant limitations were still incurred. First, this routine does not account for participant error, unlike the simultaneity judgment routine created by Alcala-Quintana & Garcia-Perez (2013), which could contaminate the best estimates of the parameters of interest. Second, this routine was still unable to fit eight participants data (four patients and four controls). Given the difficulties faced in collecting data from special populations, excluding 27% of the patient data is not ideal. Third, the use of the standard deviation of the best fit Gaussian distribution is not ideal because it does not reflect the true width of the window when the participant is most uncertain (typically the 50% probability of simultaneous response mark).

Conclusion

The findings from the current study suggest that patients treated for bilateral cataracts do not differ from typically sighted controls on audiotactile simultaneity perception, but this finding should be taken lightly because of significant and unplanned limitations because of the COVID-19 pandemic. Two of the primary limitations were the inability to collect a proper control group because of restrictions on in-person data collection and having to use a less-than-ideal routine to estimate the PSS and width of the temporal simultaneity window because of the restricted SOA range we were forced to use. The findings of this study warrant the future collection of proper control data so that stronger conclusions can be drawn, and these conclusions can help elucidate the probable mechanism underlying the abnormal audiovisual but normal visuotactile simultaneity judgment data observed in patients treated for bilateral cataracts in Chen et al. (2017).

References

Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods*, *45*(4), 972–998. https://doi.org/10.3758/s13428-013-0325-2

Bavelier, D., & Neville, H. J. (2002). Cross-modal plasticity: Where and how? *Nature Reviews Neuroscience*, *3*(6), 443–452. https://doi.org/10.1038/nrn848

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Büchel, C., Price, C., Frackowiak, R. S. J., & Friston, K. (1998). Different activation patterns in the visual cortex of late and congenitally blind subjects. *Brain*, *121*(3), 409–419. https://doi.org/10.1093/brain/121.3.409

Burr, D., & Gori, M. (2012). Multisensory integration develops late in humans. In *The Neural Bases of Multisensory Processes* (pp. 345–362). CRC Press/Taylor & Francis. https://doi.org/10.1201/b11092-23

Chen, Y. C., Lewis, T. L., Shore, D. I., Spence, C., & Maurer, D. (2018). Developmental changes in the perception of visuotactile simultaneity. *Journal of Experimental Child Psychology*, *173*, 304–317. https://doi.org/10.1016/j.jecp.2018.04.014

Chen, Y. C., Shore, D. I., Lewis, T. L., & Maurer, D. (2016). The development of the perception of audiovisual simultaneity. *Journal of Experimental Child Psychology*, *146*, 17–33. https://doi.org/10.1016/j.jecp.2016.01.010

Chen, Y. C., Lewis, T. L., Shore, D. I., & Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology*, *27*(4), 583–589. https://doi.org/10.1016/j.cub.2017.01.009

Chen, Y. C., Lewis, T. L., Shore, D. I., & Maurer, D. (2014). Role of early visual experience in the development of audiovisual integration: Evidence from cataract-reversal patients. *XIX Biennial International Conference on Infant Studies*.

Collignon, O., Dormal, G., De Heering, A., Lepore, F., Lewis, T. L., & Maurer, D. (2015). Long-lasting crossmodal cortical reorganization triggered by brief postnatal visual deprivation. *Current Biology*, *25*(18), 2379–2383. https://doi.org/10.1016/j.cub.2015.07.036

De Heering, A., & Maurer, D. (2014). Face memory deficits in patients deprived of early visual input by bilateral congenital cataracts. *Developmental Psychobiology*, *56*(1), 96–108. https://doi.org/10.1002/dev.21094

Dietrich, S., Hertrich, I., & Ackermann, H. (2013). Ultra-fast speech comprehension in blind subjects engages primary visual cortex, fusiform gyrus, and pulvinar—A functional magnetic resonance imaging (fMRI) study. *BMC Neuroscience*, *14*(1), 1. https://doi.org/10.1186/1471-2202-14-74

Ellemberg, D., Lewis, T. L., Hong Liu, C., & Maurer, D. (1999). Development of spatial and temporal vision during childhood. *Vision Research*, *39*(14), 2325–2333. https://doi.org/10.1016/S0042-6989(98)00280-6

Ellemberg, D., Lewis, T. L., Maurer, D., Brar, S., & Brent, H. P. (2002). Better perception of global motion after monocular than after binocular deprivation. *Vision Research*, *42*(2), 169–179. https://doi.org/10.1016/S0042-6989(01)00278-4

Ellemberg, D., Lewis, T. L., Maurer, D., & Brent, H. P. (2000). Influence of monocular deprivation during infancy on the later development of spatial and temporal vision. *Vision Research*, *40*(23), 3283–3295. https://doi.org/10.1016/S0042-6989(00)00165-6

Elzhov, T. V., Mullen, K. M., Spiess, A.-N., & Bolker, B. (2016). *minpack.lm: R Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK, Plus Support for Bounds*.

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

Feng, Y., Collignon, O., Maurer, D., Yao, K., & Gao, X. (2021). Brief postnatal visual deprivation triggers long-lasting interactive structural and functional reorganization of the human cortex. *Frontiers in Medicine*, *8*(November), 1–9. https://doi.org/10.3389/fmed.2021.752021

Gavin, H. P. (2019). The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems. *Duke University*, 1–19.

Gori, M., Campus, C., & Cappagli, G. (2021). Late development of audio-visual integration in the vertical plane. *Current Research in Behavioral Sciences*, *2*(April), 100043. https://doi.org/10.1016/j.crbeha.2021.100043

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

Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in Integrative Neuroscience*, *6*(SEPTEMBER), 1–8. https://doi.org/10.3389/fnint.2012.00077

Guerreiro, M. J. S., Erfort, M. V., Henssler, J., Putzar, L., & Röder, B. (2015). Increased visual cortical thickness in sight-recovery individuals. *Human Brain Mapping*, *36*(12), 5265–5274. https://doi.org/10.1002/hbm.23009

Hadad, B. S., Maurer, D., & Lewis, T. L. (2012). Sparing of sensitivity to biological motion but not of global motion after early visual deprivation. *Developmental Science*, *15*(4), 474–481. https://doi.org/10.1111/j.1467-7687.2012.01145.x

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science*, *15*(5), 688–696. https://doi.org/10.1111/j.1467-7687.2012.01171.x

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & Cornelissen, F. (2007). *What's new in Psychtoolbox-3? A free cross-platform toolkit for Psychophysics with Matlab & GNU/Octave*. http://www.psychtoolbox.org

Kupers, R., & Ptito, M. (2014). Compensatory plasticity and cross-modal reorganization following early visual deprivation. *Neuroscience and Biobehavioral Reviews*, *41*, 36–52. https://doi.org/10.1016/j.neubiorev.2013.08.001

Machulla, T.-K., Di Luca, M., & Ernst, M. O. (2016). The consistency of crossmodal synchrony perception across the visual, auditory, and tactile Senses. *Journal of*

Experimental Psychology: Human Perception and Performance, *42*(7), 1026–1038. https://doi.org/10.1037/xhp0000191

Maurer, D. (2017). Critical periods re-examined: Evidence from children treated for dense cataracts. *Cognitive Development*, *42*, 27–36. https://doi.org/10.1016/j.cogdev.2017.02.006

Maurer, D., & Lewis, T. L. (2001). Visual acuity: The role of visual input in inducing postnatal change. *Clinical Neuroscience Research*, *1*(4), 239–247. https://doi.org/10.1016/S1566-2772(01)00010-X

Maurer, D., Mondloch, C. J., & Lewis, T. L. (2007). Effects of early visual deprivation on perceptual and cognitive development. *Progress in Brain Research*, *164*, 87–104.

Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annual Review of Neuroscience*, *28*(1), 377–401. https://doi.org/10.1146/annurev.neuro.27.070203.144216

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

Putzar, L., Hötting, K., & Röder, B. (2010). Early visual deprivation affects the development of face recognition and of audio-visual speech perception. *Restorative Neurology and Neuroscience*, *28*(2), 251–257. https://doi.org/10.3233/RNN-2010-0526

R Core Team. (2021). *R: a language and environment for statistical computing*.

Robbins, R. A., Nishimura, M., Mondloch, C. J., Lewis, T. L., & Maurer, D. (2010). Deficits in sensitivity to spacing after early visual deprivation in humans: A comparison of human faces, monkey faces, and houses. *Developmental Psychobiology*, *52*(8), 775–781. https://doi.org/10.1002/dev.20473

Röder, B., Kekunnaya, R., & Guerreiro, M. J. S. (2021). Neural mechanisms of visual sensitive periods in humans. *Neuroscience and Biobehavioral Reviews*, *120*(October 2020), 86–99. https://doi.org/10.1016/j.neubiorev.2020.10.030

Röder, B., Ley, P., Shenoy, B. H., Kekunnaya, R., & Bottari, D. (2013). Sensitive periods for the functional specialization of the neural system for human face processing. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 16760–16765. https://doi.org/10.1073/pnas.1309963110

Sadato, N., Pascual-Leone, A., Grafman, J., & Ibanez, V. (1996). Activation of the primary visual cortex by Braille reading in blind subjects. *Nature*, *380*(11), 526–528.

Saenz, M., Lewis, L. B., Huth, A. G., Fine, I., & Koch, C. (2008). Visual motion area MT + /V5 responds to auditory motion in human sight-recovery subjects. *Journal of Neuroscience*, *28*(20), 5141–5148. https://doi.org/10.1523/JNEUROSCI.0803-08.2008

Segalowitz, S. J., Sternin, A., Lewis, T. L., Dywan, J., & Maurer, D. (2017). Electrophysiological evidence of altered visual processing in adults who experienced visual deprivation during infancy. *Developmental Psychobiology*, *59*(3), 375–389. https://doi.org/10.1002/dev.21502

Shore, D. I., & Spence, C. (2005). Prior entry. In *Neurobiology of attention* (pp. 89–95). Elsevier.

Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, *130*(4), 799–832. https://doi.org/10.1037/0096-3445.130.4.799

Stanley, B. M., Chen, Y. C., Lewis, T. L., Maurer, D., & Shore, D. I. (2019). Developmental changes in the perception of audiotactile simultaneity. *Journal of Experimental Child Psychology*, *183*, 208–221. https://doi.org/10.1016/j.jecp.2019.02.006

Stanley, B. M., Chen, Y.-C., Maurer, D., Lewis, T. L., & Shore, D. I. (2023). Developmental changes in audiotactile event perception. *Journal of Experimental Child Psychology*, *230*, 105629. https://doi.org/10.1016/j.jecp.2023.105629

Stevenson, R. A., Baum, S. H., Krueger, J., Newhouse, P. A., & Wallace, M. T. (2018). Links between temporal acuity and multisensory integration across life span. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 106–116. https://doi.org/10.1037/xhp0000424

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M., & Wallace, M. T. (2014). Multisensory temporal integration in autism spectrum disorders. *Journal of Neuroscience*, *34*(3), 691–697. https://doi.org/10.1523/JNEUROSCI.3615-13.2014

Stone, J. V., Hunkin, N. M., Porrill, J., Wood, R., Keeler, V., Beanland, M., Port, M., & Porter, N. R. (2001). When is now? Perception of simultaneity. *Proceedings of the*

Royal Society B: Biological Sciences, *268*(1462), 31–38. https://doi.org/10.1098/rspb.2000.1326

Vroomen, J., Keetels, M., De Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research*, *22*(1), 32–35. https://doi.org/10.1016/j.cogbrainres.2004.07.003

Weiss, N. A. (2016). *wBoot: Bootstrap methods*.

Zampini, M., Guest, S., Shore, D. I., & Spence, C. (2005). Audio–visual simultaneity judgments. *Perception & Psychophysics*, *67*(3), 531–544.

Appendix

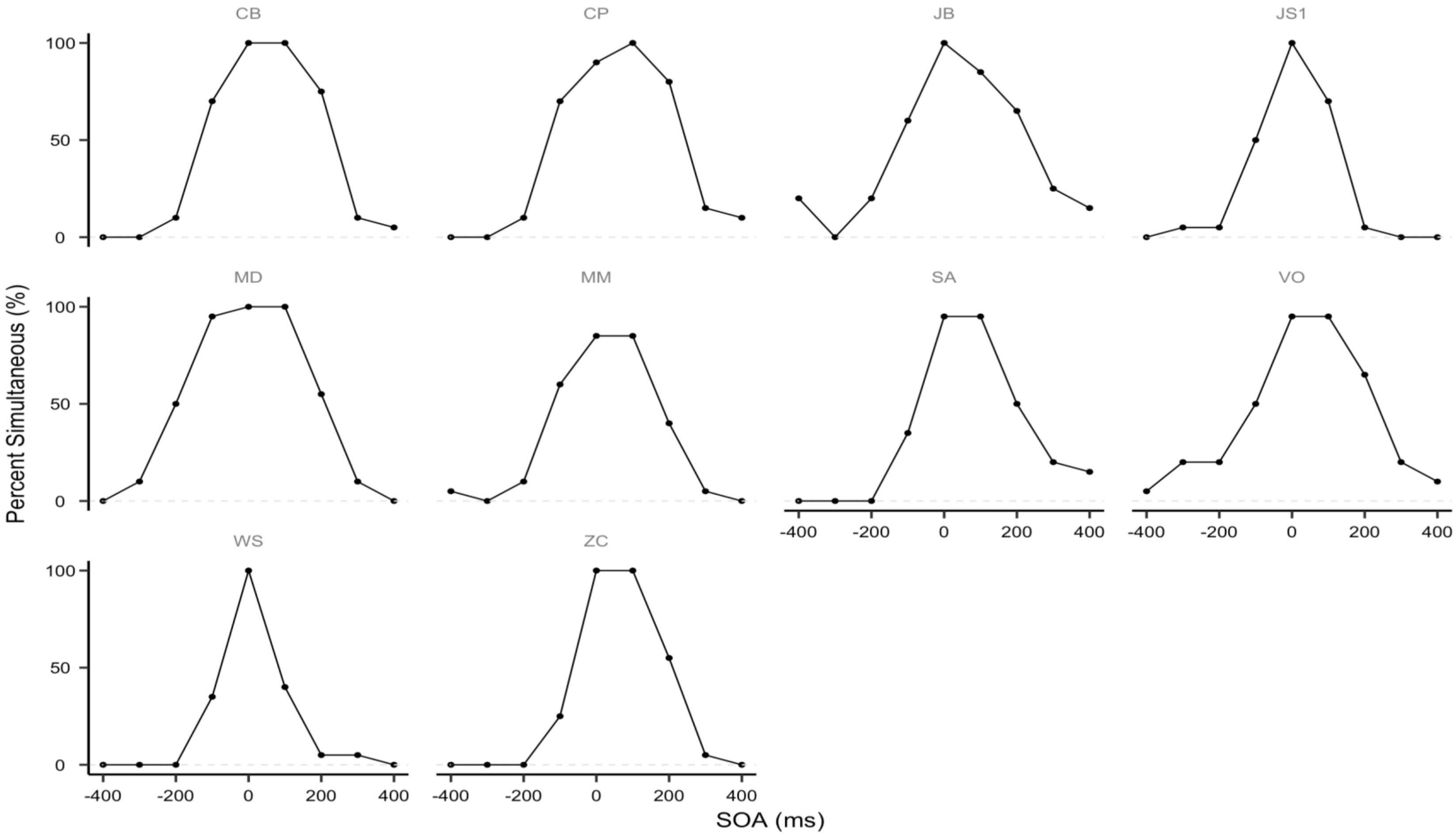

Figure A1 The mean percentage of simultaneous responses for individual bilateral cataract patients across SOA. Error bars are +/- one standard deviation of the mean.

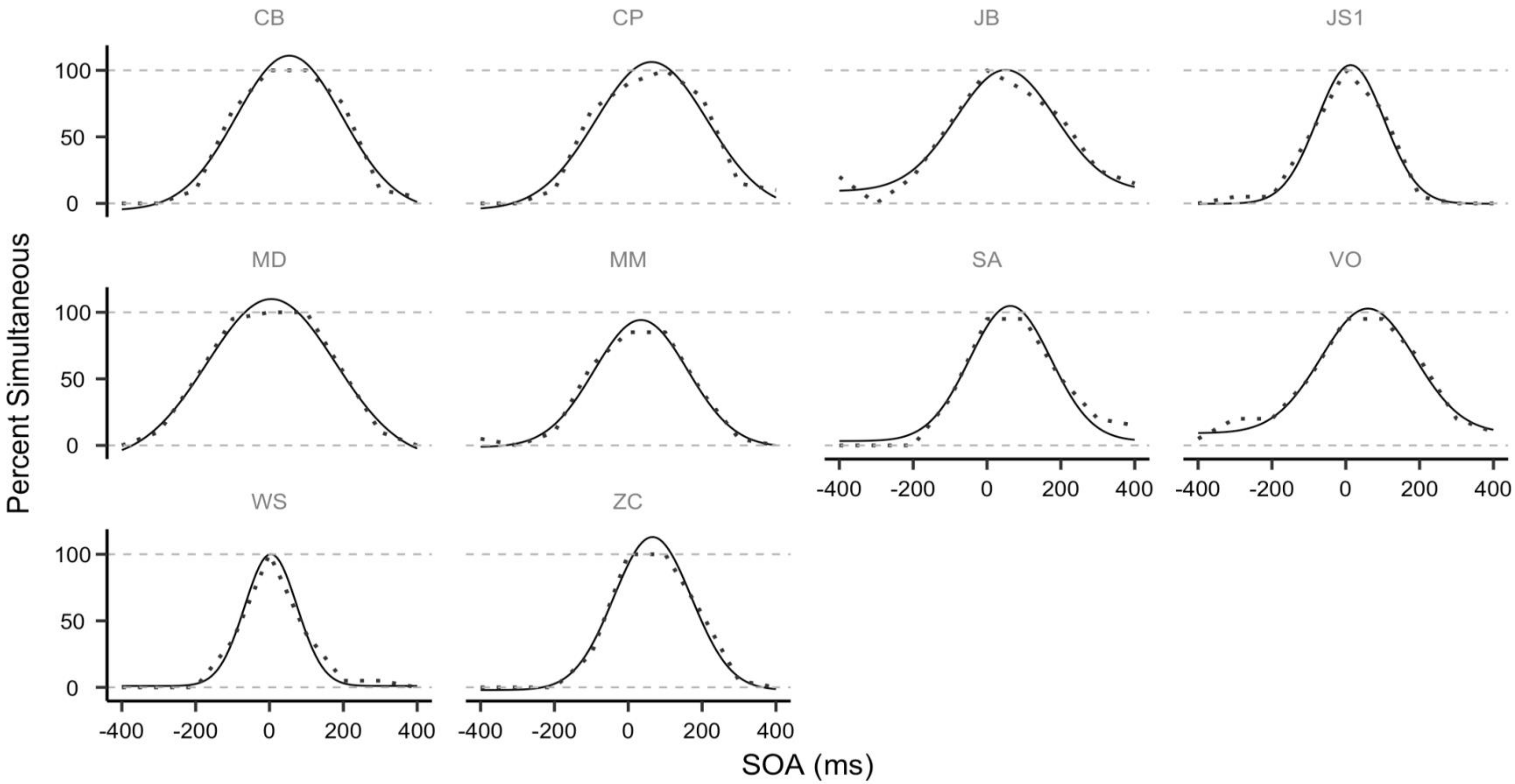

Figure A2 Curve fitting from the least squares estimation procedure (solid line) and percentage of simultaneous responses across SOA (dotted line) for individual bilateral cataract patients.

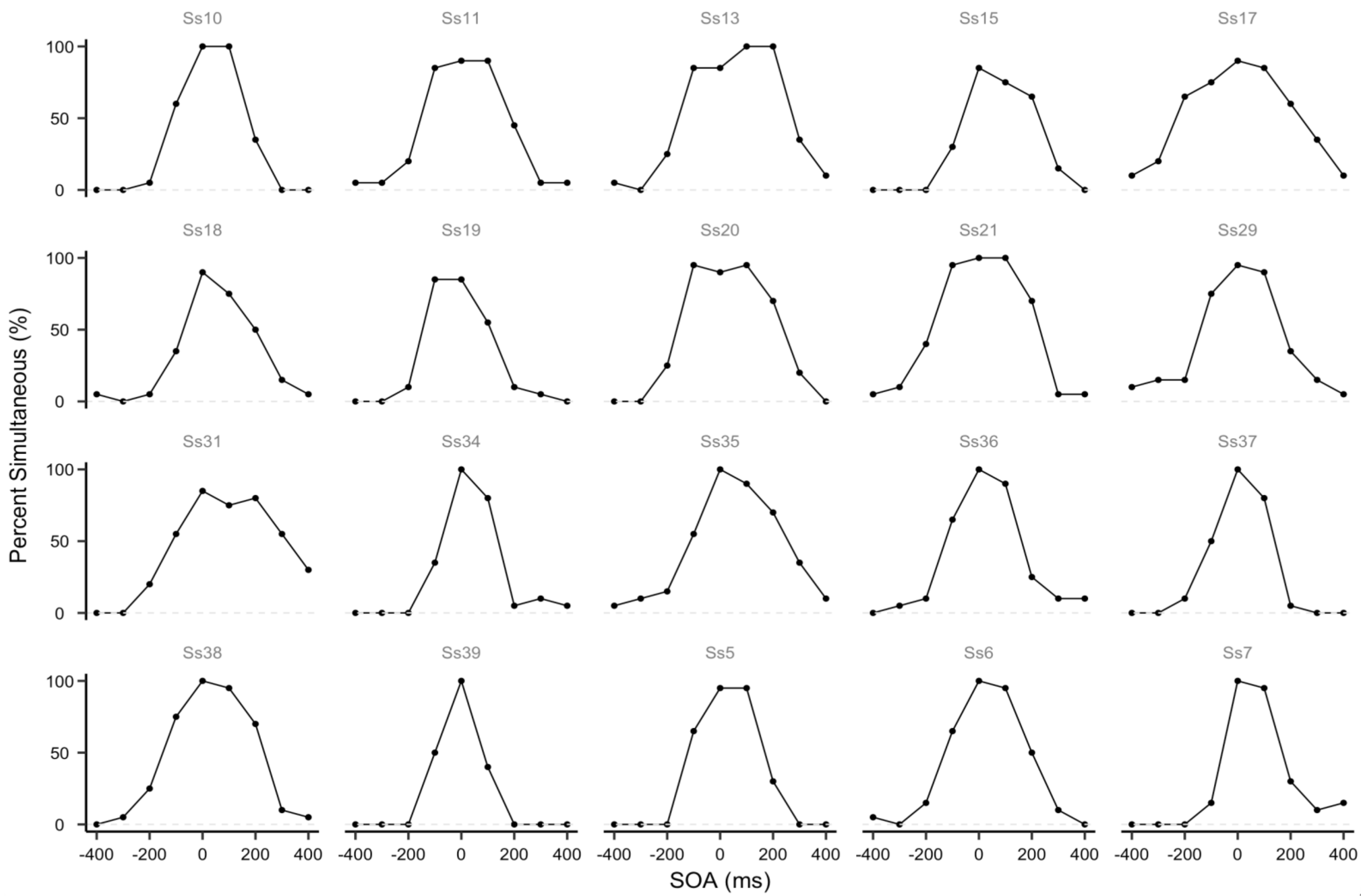

Figure A3 The mean percentage of simultaneous responses for individual control participants across SOA. Error bars are +/- one standard deviation of the mean.

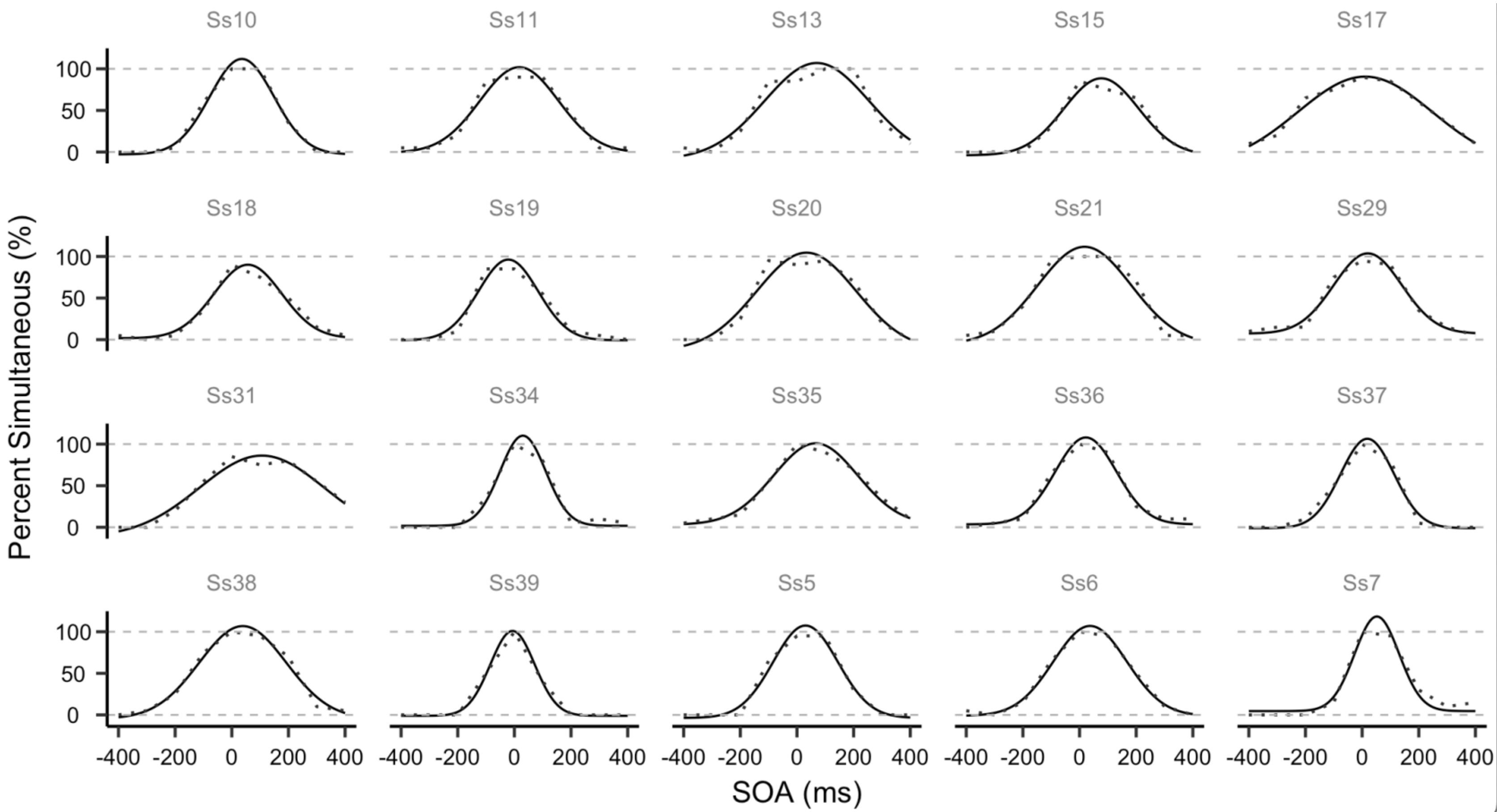

Figure A4 Curve fitting from the least squares estimation procedure (solid line) and percentage of simultaneous responses across SOA (dotted line) for individual control participants.

Chapter 5

General Discussion

This thesis examined the development of audiotactile temporal perception. The first data chapter (Chapter 2; Stanley et al ., 2019) charted the development of simultaneity perception in late-childhood to early-adolescence and compared children of those ages to a group of adults. Of the three groups of children tested (7-, 9-, and 11-year-olds), only the 11-year-olds were completely adult-like; 7-year-olds demonstrated a wider temporal simultaneity window, and both 7- and 9-year-olds made more response errors. The second data chapter (Chapter 3; Stanley et al., 2023) examined when children (9-, 11-, and 13-year-olds) became adult-like on measures of audiotactile integration by using the fission and fusion illusions. Here, only the 13-year-olds were completely adult-like; when contrasted against adults, 9-year-olds demonstrated larger magnitudes of illusion for both fission and fusion. Eleven-year-olds, on the other hand, appeared to be in a transitional stage between immature and adult-like performance; while they were adult-like for all measures indexing the fission illusion, they demonstrated a marginally larger fusion illusion, more biased criterion for fusion, and a significantly worse discriminability when compared to adults. The different pattern of results for fission and fusion in the 11-year-olds also suggests that the underlying mechanisms for fission and fusion may differ. Finally, the third data chapter (Chapter 4) examined audiotactile simultaneity perception in adults who had been treated for congenital bilateral cataracts (see Maurer, 2017 for review). Preliminary results suggest that transient early-life visual deprivation does not impact development of audiotactile simultaneity perception measured in adulthood. These results are preliminary because the COVID-19 pandemic produced significant limitations on data collection: the group of typically developed adults used for data comparison were not tested on the same SOAs and were not age- and gender-matched. Both separately and together, these three data chapters have implications for our understanding of multisensory development within the temporal domain.

The thesis focused on the development of audiotactile temporal perception. These data allow us to make inferences about the perceptual system as a whole, not only about this audiotactile perception specifically. For instance, Chapter 2, which charted the development of audiotactile simultaneity perception, completed a triad of studies measuring simultaneity perception across all the physical modality pairings (audiovisual: Chen et al., 2016; visuotactile: Chen et al., 2018; audiotactile: Chapter 2 or Stanley et al., 2019). The study design, stimuli, and participants were similar across the three studies (for greater detail, see Chapter 2). The similarities among these studies offers an opportunity to compare qualitatively the ages at which children reach adult-like maturity for each modality pairing. The comparison indicated that audiovisual simultaneity perception matures two years earlier than either visuotactile or audiotactile temporal perception, which suggests that audiovisual temporal perception may develop independently of the other two modality pairings (see specifically Figure 3 in Chapter 2). In this way, collecting data about audiotactile pairings allows us to make inferences about other modality pairings.

Audiovisual is Unique

The conclusion about distinctiveness of audiovisual temporal perception is further strengthened by considering the data from Chapter 4, which again completed the triad of sensory pairings, this time with patients treated for bilateral congenital cataracts. Recall, simultaneity perception for audiovisual and visuotactile pairings was previously tested in this patient population: patients demonstrated abnormal audiovisual simultaneity perception but completely normal visuotactile simultaneity perception (Chen et al., 2017). The new data presented here provide preliminary results (see limitations in Chapter 4), suggesting that the patients' perception of audiotactile simultaneity perception is also completely normal, like that of the visuotactile pairing. Comparisons across these six experiments (Chen et al., 2016; Chen, et al., 2017; Stanley et al. 2019 (or chapter 2); Chapter 4) highlight a unique development for the perception of audiovisual stimulus pairings. Here, we see earlier development that is more vulnerable to early visual deprivation. When considering the reasons for this unique developmental trajectory, many hypotheses have been developed and explored (see discussions in Chapter 2 and Chapter 4, as well as Chen et al., 2017). The most compelling hypothesis considers the premise that audiovisual signals are typically distal to the observer— they occur in the external environment.

Perceiving sound sources from different distances challenges the observer: sound travels significantly slower than light, with the arrival time between the two signals depending significantly on the distance of the source (Alais & Carlile, 2005; Han et al., 2022; Lewald & Guski, 2004). Because the distances of audiovisual events are continuously changing, the perceptual system maintains a wide and flexible audiovisual temporal window (Van der Burg, Alais, et al., 2015). With touch, in contrast, the stimulation always occurs on the surface of the body. Thus the width of the temporal window does not require the same degree of flexibility (Van der Burg, Orchard-Mills, et al., 2015), and appears narrower (see Chapter 2, Figure 3). In other words, touch stimulation is not distance dependent, and the only dynamic factor to consider is the neural transmission time between the location of the stimulation on the body and the brain (Alais et al., 2017). At a neural transmission speed of 60–100 metres per second (Bear et al., 2016), the maximum timing difference between being stimulated at opposite ends of the body (e.g., on the forehead versus the toe assuming a body height of 6ft) is a mere 30 ms. For comparison, if an audiovisual event occurs 10 metres from the observer, there will be an arrival time delay greater than 30 ms (sound travels at ~340 metres per second) between the auditory and visual signal; this delay continues to increase as the distance of the source of the event increases. Thus, the distal nature of audiovisual signals, which produces highly variable timing differences for different sources shapes the perceptual system to maintain a wide and flexible temporal window. When touch is involved, the tactile aspect of visuotactile and audiotactile interactions grounds perception onto the body, producing reliable and predictable relative timings, thus resulting in narrower temporal windows that are less flexible.

The distinctive flexibility of audiovisual, compared to both visuotactile and audiotactile temporal perception, further supports the literature on temporal recalibration. Multisensory temporal recalibration is the ability of the brain to restore and maintain perceptual coherence by minimizing temporal discrepancies between cross-modal sensory signals (Fujisaki et al., 2004; Hanson et al., 2008; Harrar & Harris, 2008; Keetels & Vroomen, 2012; Van der Burg et al., 2013). In the temporal domain, this is observed as a shift in the point of subjective simultaneity toward the leading stimulus after exposure to an asynchronous cross-modal stimulus pairing. Cross-modal temporal recalibration has been measured on two different timescales (e.g., Van der Burg, Alais, et al., 2015): slow recalibration is observed as a shift in the PSS after a lengthy adaptation phase consisting of repeated exposures to an asynchronous cross-modal event (Fujisaki et al., 2004; Vroomen et al., 2004), whereas rapid recalibration is observed as a shift in the PSS in as little as a single exposure to an asynchronous cross-model event (Van der Burg et al., 2013; Van der Burg, Orchard-Mills, et al., 2015). Interestingly, although slow recalibration has been measured reliably in all three physical modality pairings (audiovisual: Fujisaki et al., 2004; Vroomen et al., 2004; visuotactile: Hanson et al., 2008; Keetels & Vroomen, 2008; audiotactile: Hanson et al., 2008; Navarra et al., 2007), rapid recalibration is observed strongly only for audiovisual stimuli (Alais et al., 2017; Van der Burg, Orchard-Mills, et al., 2015). One mechanism proposed to drive rapid recalibration for audiovisual stimulation considers the temporal variability inherent in audiovisual events, relative to events that include touch (Van der Burg, Orchard-Mills, et al., 2015). To judge synchronicity of audiovisual events accurately, the perceptual system must remain flexible and estimate changes in distance of the sources. Stimulations involving touch always occur on the body; the source of variability is internal and thus minimal and predictable. As such the perceptual system does not require the ability to recalibrate on a rapid timescale. In collaboration with another research team, the developmental data for simultaneity perception for all three modality pairings (audiovisual: Chen et al., 2016; visuotactile: Chen et al., 2018; audiotactile: Stanley et al., 2019 or Chapter 2) were reanalyzed to determine if and when rapid recalibration would be observed during development (Han et al., 2022). In line with previous findings in the literature, rapid recalibration was observed only for the audiovisual pairing. Across development, audiovisual rapid recalibration emerged after audiovisual simultaneity perception reached adult-like maturity. Likely, a stabilized simultaneity window is a prerequisite for effective rapid recalibration.

The uniqueness of audiovisual integration holds true when considering the broader perspective of the other two modality pairings. While audiovisual simultaneity perception reaches maturity by age nine (Chen et al., 2016; Han et al., 2022), visuotactile and audiotactile perception take longer to mature, not reaching their full abilities until at least age 11 (Chen et al., 2018; Stanley et al., 2019 or Chapter 2; respectively). Temporal recalibration can provide a potential explanation for the earlier maturation of audiovisual simultaneity perception relative to the visuotactile and audiotactile pairings. Given that rapid recalibration is unique to the audiovisual modality pairing, one could hypothesize that this process allows for the earlier maturation of audiovisual simultaneity perception relative to other modality pairings. In other words, audiovisual rapid recalibration provides a degree of flexibility that can account for the distance-dependent nature of audiovisual stimuli, thus the temporal window can afford to crystalize at an earlier age. Since rapid adjustments are less ecologically relevant when touch is involved, the perceptual system can afford lengthier developmental trajectories for both visuotactile and audiotactile pairings, until changes in body growth stabilize in late childhood/early adolescence (e.g., Bremner et al., 2012). Irrespective of this speculative hypothesis, audiovisual rapid recalibration underscores the unique properties of audiovisual temporal perception compared to cross-modal temporal perception involving touch.

Audiotactile and Visuotactile Stimulation

The body-based influence of touch may be responsible for the similarities observed between audiotactile and visuotactile temporal perception. As discussed previously, visuotactile and audiotactile simultaneity perception both appear to reach maturity by age 11 (Chen et al., 2018; Stanley et al., 2019 or Chapter 2) and show comparable developmental patterns in refining of the temporal simultaneity window (Figure 3 in Stanley et al., 2019, or Chapter 2). Furthermore, recall that while bilateral congenital cataract patients showed abnormal audiovisual simultaneity perception (Chen et al., 2017), they were completely normal for both visuotactile (Chen et al., 2017) and audiotactile[8] (Chapter 4) pairings. As proposed in both Chapters 2 and 4, the similarities between these two pairings are likely attributable to the involvement of touch, which is inherently precise given its body-based nature. Although there are similarities that point toward a possible shared mechanism governed by touch for these two modality pairings, the studies discussed were not designed to tease these apart, and thus likely do not have

[8] This is assuming that the preliminary data presented in Chapter 4 is representative of the findings had the proper control group been collected and used as a comparison.

the necessary sensitivity to confirm or reject this hypothesis. In fact, there is a hint in the rapid recalibration literature that these two modality pairings are driven by independent mechanisms; while no evidence of rapid recalibration was found for the visuotactile pairing, a small-but-significant rapid recalibration effect was observed for the audiotactile pairing (Alais et al., 2017). This small effect appears to occur only when a causal relation exists between the auditory and tactile signals; the perceptual system allows for slight temporal variability only when the signals are co-located, likely because of the inherent variability in auditory timings relative to the highly predictable timings of touch. To reinforce this point, rapid recalibration has never been demonstrated for visuotactile stimuli because the only variability in timing is associated with where on the body was touched (the visual signal is constant because of the nearly instantaneous travel time of light). It would be interesting to examine differences in temporal based integration in the future; perhaps testing and comparing children on audiovisual and visuotactile fission and fusion illusions to the results of Chapter 3 (the development of audiotactile fission and fusion illusions) may add another piece to this puzzle. Despite the similarities between audiotactile and visuotactile temporal perception, it appears they may operate independently, although definitive conclusions cannot be drawn.

Audiotactile Development

Now setting visuotactile perception aside, this thesis presents two significant contributions to the understanding of the development of audiotactile perception. First, the results from the fission and fusion illusions (Chapter 3) showed that adult-like performance is achieved between 11 and 13 years of age. Previous work charting the development of integration between hearing and touch (Petrini et al., 2014; Scheller et al., 2021) demonstrated a considerably later maturation than found in Chapter 3: adult-like optimal integration was not demonstrated until 13 to 15 years of age (Scheller et al., 2021). Indeed, it is considerably later than all other studies charting the development and maturation of multisensory integration (see Figure 9 in Scheller et al., 2021). To reconcile this discrepancy, task differences and the nature of the tactile stimulation should be considered. In contrast to the simple task and stimuli used in Chapter 3, Scheller et al. used a complex task involving relative size estimations and the tactile stimulation required haptic exploration (Petrini et al., 2014; Scheller et al., 2021). As discussed in Chapter 3, active haptic tasks involve more complex processes such as attention, proprioception, motor control and planning, and other higher-order cognitive processes (Chapman, 1994; Heller, 1984; Simões-Franklin et al., 2011), which could extend the developmental trajectory. Furthermore, their proxy of adult-like performance required optimal integration, which involves appropriately weighting the reliability of each sensory signal (see Ernst & Banks, 2002), whereas we did not use optimality as our benchmark of adult-like performance. In other words, it is possible that children can match the performance of adults in terms of magnitude of illusions without necessarily integrating optimally (e.g., perhaps using sensory dominance). It is possible that optimal integration requires a longer developmental trajectory to reach maturity. These reasons may account for the discrepancy in the age of maturation. Critically, the data presented here (i.e., in Chapter 3) are consistent with the findings from other developmental research using simple stimuli and simple tasks.

The second significant contribution to the audiotactile literature lies in the comparison between the development of simultaneity perception (Chapter 2) and multisensory integration (Chapter 3). Recall that audiotactile simultaneity perception reached maturity by age nine, two years earlier than when maturity was reached for audiotactile integration. This suggests that integration maturation may require already developed temporal perception, at least for audiotactile stimuli. This appears to be a logical progression given the importance of temporal coincidence (i.e., the temporal rule) for multisensory integration (Welch & Warren, 1980). Other research, at least within the audiovisual domain, supports a link between the width of the temporal simultaneity window and temporal-based integration (Stevenson et al., 2018): temporal acuity predicts the degree to which complex speech information is integrated. As children develop and the temporal window narrows with experience (e.g., Chen et al., 2016, 2018; Stanley et al., 2019), there is less variability in the weighting of the sensory signals, resulting in refined temporal-based integrative abilities. These refinement processes—narrower temporal window, initial integration, and then optimal integration—mature roughly around the age when children transition into adolescence.

General Development

The protracted development of temporal perception is supported by the data in this thesis and the broader literature. For example, it was not until early adolescence that audiotactile temporal perception (Chapter 2) and integration (Chapter 3) reach adult-like levels. This is consistent with many other multisensory studies in the literature on both temporal perception and temporal-based integration in which maturation is not reached until late childhood into mid adolescence (Audiovisual: Adams, 2016; Chen et al., 2016; Gori et al., 2012; Visuotactile/haptic: Chen et al., 2018; Gori et al., 2008; O'Dowd et al., 2021; Audiohaptic: Petrini et al., 2014; Scheller et al., 2021). This late maturation can be attributed generally to the ongoing physical and neural changes experienced during development, and the necessary accrual of sensory experiences to shape and refine perception. A key factor that likely extends development into early adolescence for pairings including touch is the continuous growth of the body. Changes in body size, especially during disproportional periods of growth among the head, torso, and limbs (i.e., before adolescence; Bremner & Spence, 2017), will affect the relative timings between sensory signals involving touch as they reach the perceptual system. This may be why the temporal window remains wide during childhood, and only decreases and stabilizes once growth becomes proportional (i.e., relative changes amongst the head, torso, and limbs becomes linear) usually in early adolescence. This is what was observed for the audiotactile pairing (Chapter 2) and the visuotactile pairing (Chen et al., 2018). The relation between the temporal simultaneity window and integration abilities could explain why children appear to over-integrate relative to adults (as observed in Chapter 4 with the fission and fusion illusions); this strategy may be ideal because sensory events that are perceived close together in space and time are more likely to originate from the same event than from separate events (e.g., temporal and spatial rules, Welch & Warren, 1980). The cost of over-integrating signals that should have been segregated is likely less than the cost of segregating signals that should have been integrated.

These lengthy developmental trajectories in which cross-modal perception is calibrated by experience may be one of the contributing factors that spared audiotactile and visuotactile simultaneity perception in the congenital bilateral patients (Chapter 4). Recall that of the three physical modality pairings, only audiovisual appeared to be abnormal when these patients were tested as adults and compared to appropriate controls (Chen et al., 2017). Because maturation for temporal perception is not reached in typically developed observers until late childhood/early adolescence, this provides plenty of opportunity for touch to calibrate the less reliable modalities for visuotactile (vision is less reliable because of cataracts) and audition (audition is *slightly* less reliable because of distance dependency). Audiovisual stimulation, however, almost always occurs in the external environment which results in variable relative arrival times of the two signals (again, because of the distance dependency of audition), and as such, the unreliable visual signal cannot calibrate the normal development of audiovisual simultaneity (Chen et al., 2017). If this hypothesis is correct, the prolonged developmental trajectories observed for cross-modal temporal perception and integration serve to provide the most optimal perception possible to those who develop with normal visual input.

Multisensory Causal Inference

Considering these developmental results within the causal inference theory (see Shams & Beierholm, 2022) provides a potential comprehensive framework that accounts for the prolonged developmental trajectories observed for cross-modal perception. Recall from the general introduction that multisensory causal inference adds a weighted prior of a common cause (coined "p-common") to the standard maximum likelihood estimation (MLE) model (e.g., Ernst & Banks, 2002). This prior is a probability that the signals originated from a common cause (p-common = 1, or c = 1) or separate causes (p-common = 2, or c = 2). The final estimate is based on the MLE process but is also weighted by the prior for common cause; a strong bias for c = 1 increases the probability of integration whereas a strong bias for c = 2 decreases the probability of integration (i.e., segregation). The prior of a common cause is based on one's beliefs about the world at the time of stimulus presentation, which means that it is experience dependent. Given the strong emphasis on the importance of experience for normal development and for establishing an appropriate p-common, it is surprising that no one (as far as we are aware) has incorporated development into a causal inference framework.

Multisensory developmental researchers often ask when integration becomes optimal (i.e., when does the MLE process reach maturity; Adams, 2016; Gori et al., 2008, 2012; Nardini et al., 2008; Nava & Pavani, 2013; Petrini et al., 2014; Scheller et al., 2021); however, the possibility of a changing prior for common cause has yet to be explored. In terms of sensation and perception, developmental capabilities improve monotonically across development (i.e., regression is not typically observed); however, between infancy and adulthood, cross-modal integration appears to oscillate between tendencies to integrate and tendencies to segregate. Newborns and older infants appear to integrate indiscriminately (e.g., Maurer & Mondloch, 2005), whereas young children appear to not integrate at all and instead demonstrate sensory dominance (Adams, 2016; Gori et al., 2008, 2012; Nava & Pavani, 2013). In late childhood and early adolescence, children again begin to integrate, albeit to a greater extent than adults, and eventually refine to optimal integration. The results of the audiotactile integration task presented in Chapter 3 suggest that there may be a period of over-integration between sensory dominance and optimal integration (to the extent to which adult-like performance can be equated to integrating optimally). Given that the MLE process is unlikely to oscillate (unisensory noise should be relatively stable and thus the reliability of each signal should improve only across development), perhaps this suggests that the prior for common cause follows its own developmental trajectory and this can best explain these oscillations in integration. For example, the over-integration observed in newborns could be attributed to the absence of the p-common prior (given that the infant has no prior experience about the causal structure of the world), which means that their tendency to integrate is based solely on the likelihood function (all signals are integrated). This is in line with the infant synesthesia framework in which cross-modal influences are greater in newborns before experience-dependent pruning of excess connections occurs (e.g., Spector & Maurer, 2009). As the newborn develops into later infancy, p-common shifts toward $c = 2$, or sensory dominance, as now the emphasis is refinement of unimodal perception and not on integration (e.g., Adams, 2016; Nava & Pavani, 2013). Once the unimodal senses are refined enough to integrate meaningfully, p-common rebounds back to $c = 1$. This occurs around mid-to-late childhood in which a tendency to over-integrate is observed (e.g., O'Dowd et al., 2021, Chapter 4). A possible explanation for this rebound is that when learning the probabilities of casual structures, it would be less costly for the child to over-integrate than to over-segregate (see above). With continued perceptual experience, integration abilities are titrated until becoming optimal (or adult-like) in late-childhood or early-adolescence. By this time, p-common balances between c = 1 and c = 2 based on the accrual of perceptual experiences. While this hypothesis provides a comprehensive account of the fluctuations in integration observed across development, it has yet to be tested or supported. Given that this hypothesis emerged as a result of the work in the thesis, the data here cannot be used to test the hypothesis.

To test this hypothesis, experiments need to be designed to determine the relative contributions of *both* the likelihood and the prior (see Quintero et al., 2022). Quantifying and manipulating the likelihood function has already been done extensively; almost all studies measuring optimal integration use MLE to measure the outcome of systematic manipulations to the likelihood function by changing the relative weightings (i.e., reliabilities) of the individual signals (e.g., Alais & Burr, 2004; Bresciani & Ernst, 2007; Ernst & Banks, 2002; Scheller et al., 2021). In the literature, manipulating the prior (or "p-common") systematically is still in its infancy, and as mentioned previously, has yet to be done across development. Remember, p-common is essentially a bias based on the beliefs (or probabilities) about the causal relation between sensory signals, and this bias is shaped by experience and/or explicit knowledge about the context of the situation (Shams & Beierholm, 2022). As such, in an experimental setting, the ideal method to influence the prior is to provide explicit instruction, or change the context, regarding whether or not the stimuli should be integrated (e.g., Helbig & Ernst, 2007). This is achieved by using a

Bayesian Causal Inference (BCI) model that estimates both the unisensory estimates and the prior for common cause for each individual (Körding et al., 2007; Shams & Beierholm, 2010; Wozny et al., 2010; see Quintero et al., 2022 for review).

The challenge for future studies will be to develop experiments that are compatible with not only adults, but also newborns, older infants, and children. Paradigms that are ideal to use the BCI model such as numerosity judgments (e.g., Adams, 2016; Shams et al., 2005) or spatial tasks (e.g., Alais & Burr, 2004; Ernst & Banks, 2002) require a behavioural decision component that is beyond the capabilities of newborns and older infants. Multisensory perception, although, has been measured in newborns (Lewkowicz et al., 2010) and older infants (Lewkowicz, 2010) using face and voice matching, or audiovisual speech synchrony, respectively. However, these tasks are simplistic in their design, and modifying these to use BCI to determine the relative contributions of the prior and likelihood function would require 1) measuring the reliabilities of the unimodal signals, 2) systematically manipulating the weights of the signals for the likelihood function, and 3) manipulating the context of the situation to determine the contribution of the prior (explicit instructions are not an option for newborns and older infants). This would require newborns and older infants to respond reliably and appropriately at all levels of this task, which is unlikely given the inherent noise (e.g., inattention, fussiness, boredom, etc.) that plagues early developmental research. If these challenges can be overcome, demonstrating the potential role that a prior for causal structure plays during the development of multisensory perception will provide a novel and compelling explanation for the fluctuations in multisensory capabilities observed across development.

Limitations

Although each chapter addresses limitations in each study, there are three important limitations worth highlighting. The first has to do with the cross-sectional design used in Chapters 2 and 3. Developmental research seeks to understand the changes that occur as an individual grows from an infant into an adult. Given potentially large individual differences, the best designs are longitudinal, tracking participants as they change. However, these designs are costly and difficult to conduct. In comparison, a cross-sectional design is relatively easy to conduct. By taking a “snapshot” from different children at different ages, we can infer developmental milestones when groups differ on the measure of interest. By doing so, two potential issues arise; first, the precision regarding conclusions drawn about the age at which changes are observed is restricted by the age brackets used, and second, the rate of developmental changes *between* the age groups cannot be determined. For example, with our design, age groups were defined by two-year intervals (with +/- 3 months inclusion criteria) resulting in a 1.5-year gap existing between the age groups of children tested. Consequently, we must infer those changes between any two age groups occurred linearly and uniformly for the entire group. If there was a non-linear change between two age groups, we could not detect it. Future studies could avoid age brackets and instead populate the sample with children spanning the entire desired age range. Then, using regression or trend analyses, a more precise estimate of both the mean age, and rate in which the changes occurred, could be achieved. Yet, there are still two difficulties with this design: the number of participants needed across the age span of interest and the occurrence of individual differences.

Individual differences appear to be the rule more than the exception for most perceptual abilities—not all children behave the same. By comparing age groups, as we do in a cross-sectional design, one could easily neglect to consider that individual children develop at different rates. When a milestone is defined by a specific age, one could easily misclassify any deviations from these milestones as an indicator of abnormal development. Even though the variance within each group is represented by standard error-bars, the true extent of individual variability is not as obvious until individual datapoints are presented. For this very reason, individual differences were emphasized in this thesis by overlaying individual data points onto the group means on two of the key figures (Chapter 2, Figure 2, and Chapter 3, Figure 3). Designing research that considers the individual variation in development longitudinally requires a drastic change in the research paradigm embedded in higher education. Consider the average PhD takes four years to complete, and so students could not conduct a complete longitudinal experiment that spans a longer duration.

One final limitation of this thesis worth reiterating (discussed in Chapter 3) concerns the inability to collect the appropriate control population for the bilateral cataract patients. The initial program of research involved testing for both bilateral and unilateral congenital cataract reversal patients and their appropriate controls (two or three age- and sex-matched controls per patient) on the audiotactile simultaneity judgment task.

However, the onset of the COVID-19 pandemic halted data collection. By the onset of lockdowns, the only full dataset collected was from the bilateral patient group; only a handful of unilateral patients, and no control participants, had been collected. Given the uncertainty surrounding the duration of the shutdowns, and the need to complete my PhD program, decisions had to be made that traded off design for convenience. The first decision was to remove unilateral patients from this thesis. The second decision was to use the adult population who completed the same task in a previous developmental study (Stanley et al., 2019 or Chapter 2). Although most parameters were similar between the patients and the ad hoc control group, one critical difference that posed an issue were the SOA's tested. The bilateral patients were tested on 15 SOAs that ranged between -500 ms (audition-leading) and 500 ms (touch leading), whereas the adult controls from the developmental study were tested on 13 SOAs ranging from -1200 ms and 1200 ms. To compensate for this discrepancy, only data points from the nine overlapping SOAs between the two groups were used. This prevented the use of the proper modeling procedure (Alcalá-Quintana & García-Pérez, 2013) that estimates key parameters such as the PSS and the width of the temporal window because this model requires at least 13 SOAs to generate reliable estimates. Maximum likelihood estimation was considered as a candidate substitute to derive these parameters; however, although similar in appearance, simultaneity judgment curves often deviate from a Gaussian distribution shape, and inherent restrictions within the MLE model produced poor fits (or failed to fit altogether). As a last resort, least-squares regression was used to fit the data because it was not susceptible to the same restrictions as the MLE model (see Chapter 3 for more details).

However, this did not fare significantly better than MLE, as nearly a third of both patient and control participants could not be fitted. Because of the significant limitations incurred from the COVID-19 pandemic, only a preliminary foreshadow was provided to the possible outcome of the study described in Chapter 3. The obvious resolution to this limitation is to collect the appropriate control population so that the proper fitting-model can be used to extract the parameters of interest.

Conclusion

In summary, this thesis explored the development of audiotactile perception in the temporal domain. Chapters 2 and 3 showed that adult-like abilities for simultaneity perception and integration (respectively) were reached by late-childhood or early-adolescence in typically developing children. This was consistent with other findings in the literature that examined the development of multisensory abilities for both audiovisual and visuotactile pairings. Chapter 4 explored the potential for atypical cross-modal development by examining audiotactile temporal perception in those born without vision because of bilateral congenital cataracts. Although significant limitations were faced, preliminary data indicated that audiotactile, like visuotactile temporal perception, may be spared from lifelong deficits attributed to a transient period of early-life visual deprivation. Collectively, these findings contribute to the understanding of audiotactile temporal development, and more broadly, the development of multisensory perception, under both typical and atypical conditions.

References

Adams, W. J. (2016). The development of audio-visual integration for temporal judgements. *PLoS Computational Biology*, *12*(4), 1–17. https://doi.org/10.1371/journal.pcbi.1004865

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Alais, D., & Carlile, S. (2005). Synchronizing to real events: Subjective audiovisual alignment scales with perceived auditory depth and speed of sound. *Proceedings of the National Academy of Sciences*, *102*(6), 2244–2247. https://doi.org/10.1073/pnas.0407034102

Alais, D., Ho, T., Han, S., & Van Der Burg, E. (2017). A matched comparison across three different sensory pairs of cross-modal temporal recalibration from sustained and transient adaptation. *I-Perception*, *8*(4). https://doi.org/10.1177/2041669517718697

Alcalá-Quintana, R., & García-Pérez, M. A. (2013). Fitting model-based psychometric functions to simultaneity and temporal-order judgment data: MATLAB and R routines. *Behavior Research Methods*, *45*(4), 972–998. https://doi.org/10.3758/s13428-013-0325-2

Bremner, A. J., Holmes, N. P., & Spence, C. (2012). *The development of multisensory representations of the body and of the space around the body*.

Bremner, A. J., & Spence, C. (2017). The Development of tactile perception. In *Advances in Child Development and Behavior* (Vol. 52). Elsevier Inc. https://doi.org/10.1016/bs.acdb.2016.12.002

Bresciani, J. P., & Ernst, M. O. (2007). Signal reliability modulates auditory-tactile integration for event counting. *NeuroReport*, *18*(11), 1157–1161. https://doi.org/10.1097/WNR.0b013e3281ace0ca

Chapman, C. E. (1994). Active versus passive touch: Factors influencing the transmission of somatosensory signals to primary somatosensory cortex. *Canadian Journal of Physiology and Pharmacology*, *72*(5), 558–570. https://doi.org/10.1139/y94-080

Chen, Y. C., Lewis, T. L., Shore, D. I., Spence, C., & Maurer, D. (2018). Developmental changes in the perception of visuotactile simultaneity. *Journal of Experimental Child Psychology*, *173*, 304–317. https://doi.org/10.1016/j.jecp.2018.04.014

Chen, Y. C., Shore, D. I., Lewis, T. L., & Maurer, D. (2016). The development of the perception of audiovisual simultaneity. *Journal of Experimental Child Psychology*, *146*, 17–33. https://doi.org/10.1016/j.jecp.2016.01.010

Chen, Y.-C., Lewis, T. L., Shore, D. I., & Maurer, D. (2017). Early binocular input is critical for development of audiovisual but not visuotactile simultaneity perception. *Current Biology*, *27*(4), 583–589. https://doi.org/10.1016/j.cub.2017.01.009

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, *7*(7), 773–778. https://doi.org/10.1038/nn1268

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

Gori, M., Sandini, G., & Burr, D. (2012). Development of visuo-auditory integration in space and time. *Frontiers in Integrative Neuroscience*, *6*(SEPTEMBER), 1–8. https://doi.org/10.3389/fnint.2012.00077

Han, S., Chen, Y.-C., Maurer, D., Shore, D. I., Lewis, T. L., Stanley, B. M., & Alais, D. (2022). The development of audio–visual temporal precision precedes its rapid recalibration. *Scientific Reports*, *12*(1), 21591. https://doi.org/10.1038/s41598-022-25392-y

Hanson, J. V. M., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Experimental Brain Research*, *185*(2), 347–352. https://doi.org/10.1007/s00221-008-1282-3

Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research*, *186*(4), 517–524. https://doi.org/10.1007/s00221-007-1253-0

Helbig, H. B., & Ernst, M. O. (2007). Knowledge about a common source can promote visual-haptic integration. *Perception*, *36*(10), 1523–1533. https://doi.org/10.1068/p5851

Heller, M. A. (1984). Active and passive touch: The influence of exploration time on form recognition. *The Journal of General Psychology*, *110*, 243–249.

Keetels, M., & Vroomen, J. (2012). Perception of synchrony between the senses. In *The neural bases of multisensory processes*.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9). https://doi.org/10.1371/journal.pone.0000943

Lewald, J., & Guski, R. (2004). Auditory-visual temporal integration as a function of distance: No compensation for sound-transmission time in human perception. *Neuroscience Letters*, *357*(2), 119–122. https://doi.org/10.1016/j.neulet.2003.12.045

Lewkowicz, D. J. (2010). Infant perception of audio-visual speech synchrony. *Developmental Psychology*, *46*(1), 66–77. https://doi.org/10.1037/a0015579

Lewkowicz, D. J., Leo, I., & Simion, F. (2010). Intersensory perception at birth: Newborns match nonhuman primate faces and voices. *Infancy*, *15*(1), 46–60. https://doi.org/10.1111/j.1532-7078.2009.00005.x

Maurer, D. (2017). Critical periods re-examined: Evidence from children treated for dense cataracts. *Cognitive Development*, *42*, 27–36. https://doi.org/10.1016/j.cogdev.2017.02.006

Maurer, D., & Mondloch, C. (2005). Neonatal synesthesia: A reevaluation. *Synesthesia: Perspectives from Cognitive Neuroscience*, 193–213.

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

Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Adaptation to audiotactile asynchrony. *Neuroscience Letters*, *413*(1), 72–76. https://doi.org/10.1016/j.neulet.2006.11.027

O'Dowd, A., Cooney, S. M., Sorgini, F., O' Rourke, E., Reilly, R. B., Newell, F. N., & Hirst, R. J. (2021). The development of visuotactile congruency effects for sequences of events. *Journal of Experimental Child Psychology*, *207*, 105094. https://doi.org/10.1016/j.jecp.2021.105094

Petrini, K., Remark, A., Smith, L., & Nardini, M. (2014). When vision is not an option: Children's integration of auditory and haptic information is suboptimal. *Developmental Science*, *17*(3), 376–387. https://doi.org/10.1111/desc.12127

Quintero, S. I., Shams, L., & Kamal, K. (2022). Changing the tendency to integrate the senses. *Brain Sciences*, *12*(10), 1384. https://doi.org/10.3390/brainsci12101384

Scheller, M., Proulx, M. J., de Haan, M., Dahlmann-Noor, A., & Petrini, K. (2021). Late- but not early-onset blindness impairs the development of audio-haptic multisensory integration. *Developmental Science*, *24*(1), 1–17. https://doi.org/10.1111/desc.13001

Shams, L., & Beierholm, U. (2022). Bayesian causal inference: A unifying neuroscience theory. *Neuroscience & Biobehavioral Reviews*, *104619*.

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., Ma, W. J., & Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *NeuroReport*, *16*(17), 1923–1927. https://doi.org/10.1097/01.wnr.0000187634.68504.bb

Simões-Franklin, C., Whitaker, T. A., & Newell, F. N. (2011). Active and passive touch differentially activate somatosensory cortex in texture perception. *Human Brain Mapping*, *32*(7), 1067–1080. https://doi.org/10.1002/hbm.21091

Spector, F., & Maurer, D. (2009). Synesthesia: A new approach to understanding the development of perception. *Developmental Psychology*, *45*(1), 175–189. https://doi.org/10.1037/a0014171

Stanley, B. M., Chen, Y. C., Lewis, T. L., Maurer, D., & Shore, D. I. (2019). Developmental changes in the perception of audiotactile simultaneity. *Journal of Experimental Child Psychology*, *183*, 208–221. https://doi.org/10.1016/j.jecp.2019.02.006

Stanley, B. M., Chen, Y.-C., Maurer, D., Lewis, T. L., & Shore, D. I. (2023). Developmental changes in audiotactile event perception. *Journal of Experimental Child Psychology*, *230*, 105629. https://doi.org/10.1016/j.jecp.2023.105629

Stevenson, R. A., Baum, S. H., Krueger, J., Newhouse, P. A., & Wallace, M. T. (2018). Links between temporal acuity and multisensory integration across life span. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(1), 106–116. https://doi.org/10.1037/xhp0000424

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid recalibration to audiovisual asynchrony. *Journal of Neuroscience*, *33*(37), 14633–14637. https://doi.org/10.1523/JNEUROSCI.1182-13.2013

Van der Burg, E., Alais, D., & Cass, J. (2015). Audiovisual temporal recalibration occurs independently at two different time scales. *Scientific Reports*, *5*(1), 14526. https://doi.org/10.1038/srep14526

Van der Burg, E., Orchard-Mills, E., & Alais, D. (2015). Rapid temporal recalibration is unique to audiovisual stimuli. *Experimental Brain Research*, *233*(1), 53–59. https://doi.org/10.1007/s00221-014-4085-8

Vroomen, J., Keetels, M., De Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research*, *22*(1), 32–35. https://doi.org/10.1016/j.cogbrainres.2004.07.003

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667. https://doi.org/10.1037/0033-2909.88.3.638

Wozny, D. R., Beierholm, U. R., & Shams, L. (2010). Probability matching as a computational strategy used in perception. *PLoS Computational Biology*, *6*(8). https://doi.org/10.1371/journal.pcbi.1000871