Neuropsychologia 140 (2020) 107396

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: http://www.elsevier.com/locate/neuropsychologia

Review article

Audiovisual temporal integration: Cognitive processing, neural mechanisms, developmental trajectory and potential interventions

Han-yu Zhou [a,b], Eric F.C. Cheung [c], Raymond C.K. Chan [a,b,*]

[a] *Neuropsychology and Applied Cognitive Neuroscience Laboratory, CAS Key Laboratory of Mental Health, Institute of Psychology, Beijing, China*
[b] *Department of Psychology, University of Chinese Academy of Sciences, Beijing, China*
[c] *Castle Peak Hospital, Hong Kong, China*

ARTICLE INFO

Keywords:
Temporal binding window
Cognitive processing
Neural mechanisms
Developmental trajectories
Neurodevelopmental disorders
Perceptual and music training

ABSTRACT

To integrate auditory and visual signals into a unified percept, the paired stimuli must co-occur within a limited time window known as the Temporal Binding Window (TBW). The width of the TBW, a proxy of audiovisual temporal integration ability, has been found to be correlated with higher-order cognitive and social functions. A comprehensive review of studies investigating audiovisual TBW reveals several findings: (1) a wide range of top-down processes and bottom-up features can modulate the width of the TBW, facilitating adaptation to the changing and multisensory external environment; (2) a large-scale brain network works in coordination to ensure successful detection of audiovisual (a)synchrony; (3) developmentally, audiovisual TBW follows a U-shaped pattern across the lifespan, with a protracted developmental course into late adolescence and rebounding in size again in late life; (4) an enlarged TBW is characteristic of a number of neurodevelopmental disorders; and (5) the TBW is highly flexible via perceptual and musical training. Interventions targeting the TBW may be able to improve multisensory function and ameliorate social communicative symptoms in clinical populations.

We live in a world where a multitude of information is coded in different sensory modalities. One challenge of our brain is to appropriately integrate multisensory inputs from a noisy background to form coherent perceptual representations. Low-level physical stimulus characteristics including location, time and intensity are crucial determinants of the possibility and strength of multisensory integration (Meredith and Stein, 1986a, 1986b; Meredith et al., 1987). For example, a pair of audiovisual stimuli is likely to be integrated only if they are presented close in time and fall within a limited time window known as the Temporal Binding Window (TBW) (Dixon and Spitz, 1980). This audiovisual temporal integration is indexed by faster and more accurate detection at the behavioural level and non-linear and super-additive neural responses at the neural level (Calvert et al., 2004). Given the difference in the speed of sound and light and the time for neural transmission between the two senses, some tolerance of audiovisual asynchrony has great adaptive value for constructing a unified perceptual world. At the same time, to avoid confusing and improperly filtered representations, it is also important to appropriately segregate auditory and visual information from separate events/objects (Wallace and Stevenson, 2014).

Audiovisual temporal integration is a rudimentary process and can scaffold cognitive, language and social development. Infants utilize temporal relationships between sound and visual features to learn cross-modal object-sound associations (Bahrick, 1988), detect tempo or rhythm changes (Bahrick and Lickliter, 2000; Bahrick et al., 2002) and discriminate different prosody (Bahrick et al., 2019) and emotions (Flom and Bahrick, 2007). Integrating auditory and visual signals can also facilitate language acquisition (Gogate et al., 2001; Lewkowicz and Hansen-Tift, 2012), person recognition (Robertson and Schweinberger, 2010) and emotion processing (Walker-Andrews, 1997). Failure to correctly detect audiovisual (a)synchrony or identify temporal order is manifested as an abnormally enlarged TBW, which leads to inappropriate combination of unrelated audiovisual stimuli that should otherwise be separated (e.g., in autistic and dyslexic individuals and children with language impairment, Francisco et al., 2017; Kaganovich, 2017; Stevenson et al., 2016, 2014a). This inaccurate representation may result in less audiovisual fusion (i.e., the McGurk effect) (Stevenson et al., 2012b) and may undermine speech perception (Conrey and Pisoni, 2006), making it difficult for individuals to interact with the world around them (Baum et al., 2015). Since the integration of visual speech signals (e.g., mouth movement, facial expressions and gestures) and auditory speech sounds is important for effective social

* Corresponding author. Institute of Psychology, Chinese Academy of Sciences, 16 Lincui Road, Beijing 100101, China.
E-mail address: rckchan@psych.ac.cn (R.C.K. Chan).

https://doi.org/10.1016/j.neuropsychologia.2020.107396
Received 17 October 2019; Received in revised form 14 February 2020; Accepted 15 February 2020
Available online 20 February 2020
0028-3932/

communication, it is not surprising that a widened audiovisual TBW has been shown to be correlated with impaired verbal (Woynaroski et al., 2013) and non-verbal (Noel et al., 2018a) communication skills. Furthermore, a narrower TBW, reflecting higher accuracy in multisensory integration, has been found to be associated with enhanced problem solving abilities, suggesting that sharpened multisensory temporal sensitivity may mirror precision of abstract thinking and thus the ability of complex reasoning (Zmigrod and Zmigrod, 2016).

This paper aimed to comprehensively review the concept of audiovisual TBW. We first reviewed relevant experimental paradigms and theoretical models investigating the TBW (Section 1). Next, the factors affecting the width of the TBW (Section 2) and the neural mechanisms underlying successful audiovisual temporal integration (Section 3) were reviewed. We also conducted an activation likelihood estimation (ALE) meta-analysis to examine the neural activities engaged by synchronous versus asynchronous audiovisual processing. In Section 4, we reviewed how audiovisual temporal integration changes across the lifespan. Given its fundamental role in cognitive functions and social interactions, manifestations of multisensory temporal abnormalities in various neurodevelopmental and neurodegenerative disorders were then discussed in Section 5. Finally, studies examining the flexibility of TBW, which could have important implications in clinical populations with abnormal audiovisual TBW, were reviewed in Section 6.

1. Literature search

A comprehensive search was conducted using the following search terms in PubMed, Elsevier and PsycInfo databases: ("temporal binding" OR "temporal window" OR "time window" OR "temporal acuity" OR "temporal processing" OR "temporal integration" OR "temporal order" OR "simultaneity" OR "*synchrony") AND ("multisensory" OR "crossmodal" OR "audiovisual" OR "sensory integration"). We only included studies in English involving human subjects. The final search was completed in November 26, 2019 for studies published in peer reviewed journals or accessed online. In addition, the references cited by the review articles on this topic were scrutinized to include other relevant studies. A total of 15 review/meta-analysis papers were included and the details of these studies can be found in the supplementary text.

There were two inclusion criteria for eligible studies. First, studies had to focus on human ability of using temporal cues to integrate sensory information. That is, studies should use some relevant paradigms (which are summarized in section 1 and Table 1) to measure behavioural or neural underpinnings of multisensory temporal integration. Secondly, studies included must be concerned with both auditory and visual modalities. Studies measuring unisensory temporal acuity or investigating integration of other senses (e.g., auditory-tactile, Kitagawa et al., 2005; and visual-tactile, Chen et al., 2018) were not considered.

We identified 738 unique articles after removing duplicates. Titles and abstracts were first screened to examine whether these studies met inclusion criteria. Then, the remaining articles were further considered for full-text review. The detailed procedure of study selection and the reasons for study exclusion are illustrated in Fig. 1. A total of 302 papers were included, which were further divided into seven sub-groups according to their topics. The number of studies included in the "Relationship with other abilities" was very small (N = 4) and have already been mentioned in the opening section (Conrey and Pisoni, 2006; Robertson and Schweinberger, 2010; 2012b; Zmigrod and Zmigrod, 2016). Studies in other sub-topics of audiovisual temporal integration are discussed separately in the following sections.

2. Paradigms and models of audiovisual temporal integration

2.1. Paradigms to measure audiovisual temporal integration

To measure the width of the TBW, a variety of tasks could be used. Table 1 shows all the relevant paradigms, including some specially designed for infants. These tasks could also be classified into explicit (e.g., explicit judgement of simultaneity) and implicit (e.g., implicit fusion of incongruent audiovisual speech stimuli, McGurk fusion) ones. It is unclear whether these tasks capture the same process of multisensory temporal integration. For instance, both the Simultaneity Judgement (SJ) Task and Temporal Order Judgement (TOJ) Task are explicit paradigms, but different size of the TBWs may be obtained from these two tasks (Stevenson and Wallace, 2013). It has been reported that the SJ task should be preferred over the TOJ task when the primary goal is to measure audiovisual synchrony perception (van Eijk et al., 2008). In addition, they may measure slightly different cognitive processes (Vatakis et al., 2008b; Vroomen and Keetels, 2010), with the TOJ task requiring an extra stage of "order" processing apart from the (a)synchrony perception (Love et al., 2013) and thus eliciting stronger activity in several regions in the left hemisphere (Binder, 2015; Love et al., 2018). Another example is that explicit synchrony judgement is not a prerequisite for implicit fusion of audiovisual stimuli (Soto-Faraco and Alsius, 2009; Zmigrod and Hommel, 2011), and they may involve separate mechanisms that are not tightly correlated (Tsilionis and Vatakis, 2016) or even negatively correlated (Freeman et al., 2013; Ipser et al., 2018). However, all the tasks listed in Table 1 share a core concept and reflect the ability to appropriately bind or separate audiovisual stimuli based on temporal cues.

It should be noted that we did not include "target detection" paradigms in this review. By comparing participants' response times (RT) between unimodal (visual and auditory) and bimodal targets and using a "Race Model", researchers can find a temporal window of multisensory behavioural facilitation with faster RTs (Miller, 1986, 1982). This speeding by bimodal cues is above and beyond what would be expected by statistical summation of the unimodal RTs. In other words, the temporal window calculated by the Race Model is due to benefit from multisensory redundancy. Although both of them are called "Temporal binding window", the TBW defined by the Race Model, which is a time span with RT benefits, is an entirely different concept from audiovisual temporal precision/acuity. This review focuses on the latter.

2.2. Models of audiovisual temporal integration

Some researchers have put forward theoretical models to explain how audiovisual synchrony perception is achieved despite the arrival-time differences of two information streams, and tried to fit their models to data from relevant TBW paradigms. Here, we make a brief introduction to these models. Future studies may benefit from adopting these models to obtain interpretable parameters and tap into the cognitive and neural basis of audiovisual temporal processing.

The first class of models share a core framework in which signals from two stimuli arrive at a central mechanism with randomly distributed arrival times; and based on the arrival-time difference, a judgement of synchrony or temporal order is made by adopting some decision rules. For example, the Time-Window-of-Integration (TWIN) Model, which initially applies to RT tasks, has recently been extended to TOJ tasks (Diederich and Colonius, 2015). Here, participants' TOJ is determined based on the first-stage unisensory (auditory and visual) processing speed and the width of the TBW - a numerical parameter which could be estimated. Whenever the detection times for a stimulus pair fall within the TBW, participants bind them together and can only guess which stimulus occurs first with a fixed response bias. Similarly, Garcia-Perez and Alcala-Quintana (2012) also included parameters of sensory processing (e.g., arrival time difference and sensory processing variability) and response bias in their unified model for SJ and TOJ tasks. Moreover, their model adds decisional parameters such as response errors due to wrong motor response (Garcia-Perez and Alcala-Quintana, 2012). These estimated parameters all have empirical meanings and can be used to reflect certain underlying processes.

Such a framework has also been adopted to develop neural models. The main idea of the multisensory correlation detector (MCD) model, for

Table 1
Paradigms relating to audiovisual temporal binding window (AV TBW).

Paradigms	Brief introductions	Indexes
Paradigms for infants and young children		
Habituation and test Example study Lewkowicz (2010)	Habituate infants to a synchronous audiovisual event and then test for their detection of different levels of asynchrony. In the test phase, if infants look longer at the asynchronous stimuli compared to the initial synchronous pair, it indicates they successfully detect the asynchrony.	The looking time of the asynchronous ('new') stimuli Threshold for AV asynchrony detection measured at **group level**
Preferential looking Example study Grossman et al. (2015)	Participants are presented with two identical visual videos at the same time. Only one of them is in synchrony with the corresponding audio track, but the other has a certain degree of temporal offset. If infants show preference for the temporally-congruent video, it suggests they successfully discriminate the synchronous from the asynchronous one.	The percentage of looking time of the synchronous video Threshold for AV asynchrony detection measured at **group level**
Paradigms for older children, adolescents and adults		
● Explicit Judgement tasks		
Temporal Order Judgement (TOJ) Example study Hillock et al. (2011)	Participants are required to report the temporal order of auditory and visual stimuli across a range of Stimulus Onset Asynchronies (SOAs). ("Which came first?")	Just Noticeable Difference (JND) measured at **individual level**
Simultaneity Judgement (SJ) Example study (Hillock-Dunn and Wallace, 2012)	Participants are required to report the simultaneity of auditory and visual stimuli across a range of SOAs. ("Synchronous or not?") Alternatively, two audiovisual stimulus pairs (one synchronous and the other asynchronous) can be presented on each trial and participants are asked select the most synchronous (Yarrow et al., 2016).	(1) TBW, within which participants are highly likely to report simultaneity. (2) Point of subjective simultaneity (PSS), the point at which participants maximally fused the two stimuli. Both indexes measured at **individual level**
● Implicit fusion tasks		
McGurk task with different audio-visual SOAs Example study Woynaroski et al. (2013)	Participants are presented with auditory "ba" and visual "ga" stimuli across a range of SOAs, and are required to report the stimuli they perceive ("ba", "ga", or McGurk fusion "da"). For each individual, mean rates of fusion across SOAs were normalized to his or her maximum value of perceived fusion.	TBW, the width of the stimulus offset range at which participants are highly likely to fuse the visual "ga" and auditory "ba" stimulus pair (e.g., McGurk fusion on 75% or more of the trials) TBW measured at **individual level**
Sound Induced Flash Illusion (SiFi) Example study Foss-Feig et al. (2010)	In illusory conditions, two beeps are presented with a single flash (2B1F). One beep is always presented with an onset coincident with the flash, and the other beep has a certain SOA with the flash. In control conditions, one beep and one flash are presented at the same time (IBIF). In all trials, participants are asked to report the number of flashes perceived. Beeps presented in close temporal proximity to the flash are more likely to produce illusory flashes (i.e., report seeing two flashes).	TBW is defined as the temporal span within which participants have significantly higher possibility of reporting an illusory flash compared to the control condition (1B1F) TBW measured at **group level**
Multisensory TOJ (temporal ventriloquism: visual TOJ with irrelevant auditory stimuli) Example study De Boer-Schellekens et al. (2014b)	Two visual stimuli are presented at variable SOAs, and participants are required to make visual TOJ. The visual stimuli are flanked by two auditory stimuli. The presence of irrelevant auditory stimuli can improve participants' performance on visual TOJ task, as if they pull the flashes in temporally-opposite directions. The first-sound-and–first-light and second-light-and-second-sound SOAs are varied in order to measure the TBW.	TBW is defined as the audiovisual SOA span where participants' visual TOJ performances (i.e., JND) are significantly better than the performances without auditory stimuli (the visual-only control condition). TBW measured **at group level**
Stream/Bounce Illusion Example study Donohue et al. (2015)	Two visual objects (e.g., circle disks) move toward each other, overlap, and then move away from each other. This pattern of motion can be perceived either as two objects streaming through each other or as two objects bouncing off each other. Participants tend to report "streaming" without additional auditions. However, when a sound is presented near the time of overlap, participants are more likely to perceive the objects as bouncing off of each other. This bouncing illusion is regarded as a result of audio-visual integration. By manipulating the temporal offsets between the auditory stimulus and intersection of the visual stimulus pair, potential TBW can be calculated.	TBW is defined as the audiovisual SOA span where participants have significantly higher possibility of reporting "bouncing" illusion compared to the control condition (without additional sound). TBW measured at **group level**
Pip and pop effect **No example study yet**	Participants are asked to search for a horizontal or a vertical target bar among oblique distractor bars. Both the target and distractors changed their color randomly in a pre-determined cycle. In the tone present condition, a spatially uninformative sound which is synchronized with color changes of the visual target could substantially decrease the search time (i.e., in order of seconds). This facilitation effect due to automatic audiovisual integration is referred to as "pip-and-pop" effect (van der Burg et al., 2008). Although no study to date has used this paradigm to measure the width of audiovisual TBW, we could manipulate the tone-target intervals to find a temporal window within which an auditory cue could significantly enhance the search of a visual target compared with the tone-absent condition.	TBW is defined as the temporal span within which a nonspatial sound could significantly reduce the search time of a visual target compared with the sound-absent condition. TBW could be measured at **individual level**.

instance, is that multisensory neurons first receive inputs in separate sensory channels and then detect correlation and lag between the two stimuli to optimally integrate related inputs (Parise and Ernst, 2016). Luu et al. (2008) has proposed a similar cross-correlational model where a peripheral stage of parallel processing in separate sensory channels is followed by a secondary stage of cross-correlation calculation. The output of the cross-correlator is an array of values decoding activation patterns to yield the best estimate of the lag between auditory and visual

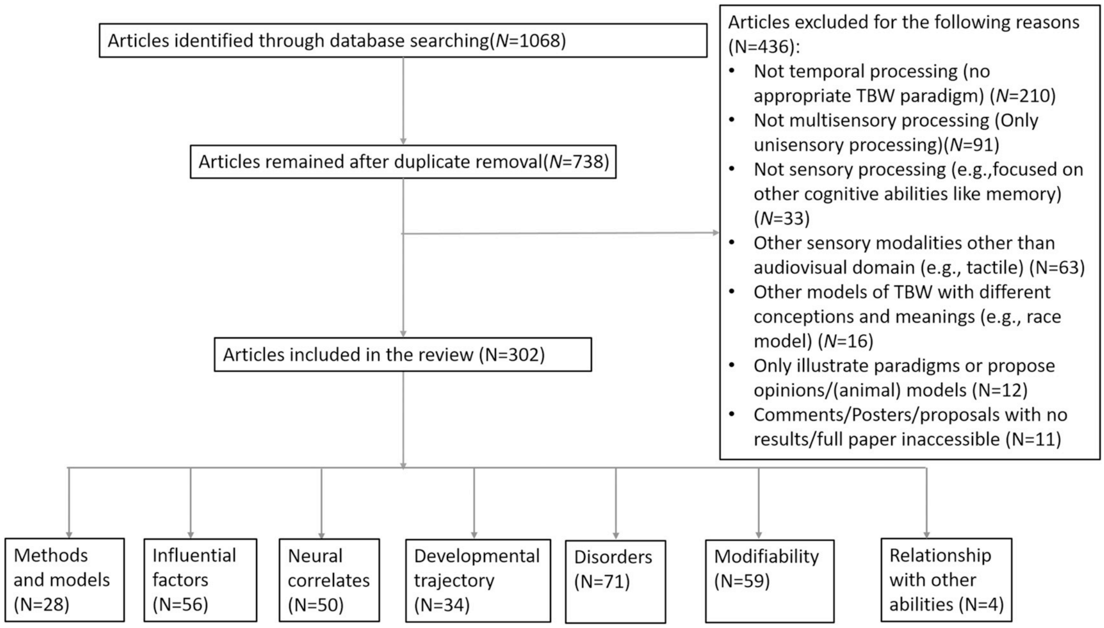

Fig. 1. The flow chart of literature search.

streams.

The second class of models falls into the framework of the Bayesian model. The Causal Inference model of audiovisual speech (Magnotti et al., 2013) explains synchrony perception as an inference about the causal relationship between auditory and visual speech signals. In other words, participants utilize the temporal relationship between cues to determine the likelihood that the speech arises from a single speaker or multiple speakers. A 'Bayes-optimal synchrony window' can be derived from this model and this approach allows for the examination of different reasons leading to enlarged TBWs (Noel et al., 2018d). Abnormally enlarged TBWs can be a result of either a-priori bias to ascribe common source or poor fidelity in sensory representation (i.e., high levels of sensory noise). Compared to the first class of independent-channel models, the Bayesian framework emphasizes more on top-down cognitive processes (e.g., prior assumptions) in addition to bottom-up stimulus characteristics.

3. Factors affecting audiovisual TBW

Both top-down processes, such as attention and expectations, and bottom-up features, like stimulus types and intensity, modulate the likelihood of whether unisensory stimuli would be combined. The following section reviews relevant studies investigating factors affecting the width of the TBW.

3.1. Top-down processes

A handful of studies have investigated the correlations between attention and audiovisual temporal integration. For instance, reduced attentional resources induced by distractors (Vatakis et al., 2007; Vatakis and Spence, 2006a) or dual tasks (Dean et al., 2017) can lead to worse performance in the audiovisual TOJ task. On the other hand, heightened attention capacity is associated with a narrower TBW and increased temporal acuity (Donohue et al., 2015; Li et al., 2018). Using warning cues to induce a transient alerting effect, the intensity of attention can be increased, which in turn contributes to enhanced multisensory temporal precision (Li et al., 2018). Finally, selective attention also has a prior-entry effect so that the attended stimulus comes into consciousness more quickly than unattended ones, with the Point of Subjective Simultaneity (i.e., PSS, the peak point of audiovisual Stimulus Onset Asynchrony (SOA) where participants have the highest probability of integrating stimuli) shifting towards the unattended sensory modality (Barrett and Krumbholz, 2012; Zampini et al., 2005).

Prior experiences and unity assumption is another important top-down factor affecting the width of the TBW. Unity assumption occurs in conditions where individuals believe unisensory stimuli originate from the same object/event due to their prior experiences (either short-term contextual experiences or long-term memory) (Welch and Warren, 1980). Unity assumption can facilitate audiovisual integration. That is to say, individuals allow for larger temporal offsets and a wider TBW for matched compared with mismatched audiovisual events (Chuen and Schutz, 2016; Habets et al., 2017; Margiotoudi et al., 2014; ten Oever et al., 2013; van Wassenhove et al., 2007; Vatakis and Spence, 2007). Stevenson et al. (2014b) has put forward a "parallel accumulator model" to explain this unity effect. In their model, both temporal relationship and cognitive factors (e.g., semantic congruency) can add information about whether or not to integrate two sensory inputs in parallel. Therefore, when audiovisual stimuli are semantically matching and induce a strong unity assumption, less additional evidence from temporal proximity is needed to reach a judgement that two inputs originate from the same source and should be integrated. This in turn leads to greater tolerance of temporal misalignment for congruent stimulus pairs. However, this unity effect is only consistently found in human speech stimuli and does not necessarily extend to other kinds of stimulus (Chen and Spence, 2017), such as monkey calls (Vatakis et al., 2008a), object actions or musical videos (Vatakis and Spence, 2008). One hypothesis is that unity assumption needs extensive prior experience to form. If there is a large amount of exposure to musical stimuli, it is possible to demonstrate a unity effect with stimuli in the area of expertise (Petrini et al., 2009a).

Similar to the influence of attention, prior experiences could also shift the PSS of audiovisual TBW (Navarra et al., 2010). Specifically, visual speech stream has to precede the auditory stream by a larger interval in the participants' native language than in their non-native language for synchrony perception. A possible reason is that the anticipatory effect of visual information on upcoming auditory speech is stronger in one's native language, which may speed up the processing of the auditory stimuli and decrease the apparent temporal separation between visual and auditory speech signals (Navarra et al., 2010).

Other top-down processes including predictability and emotion have also received some research attention. Higher predictability learned by action-outcome associations (Arikan et al., 2017; Desantis and Haggard, 2016a, 2016b) or obtained by special pitch/temporal patterns in stimuli sequence (Cook et al., 2011) can facilitate audiovisual binding, resulting in a wider TBW. As for emotions, induced positive emotions in individuals with low depressive traits promote audiovisual integration and thus broaden the TBW as indicated by stream/bounce illusion (Kitamura et al., 2016).

3.2. Bottom-up features

In addition to top-down processes, audiovisual temporal processing is also influenced by bottom-up stimulus characteristics.

Spatial correspondence and the intensity of stimuli, in addition to temporal proximity, are two other factors that determine the possibility and strength of multisensory integration (Calvert et al., 2004). However, these factors do not take effect independently, and they interact with each other to facilitate appropriate binding or separation of sensory inputs from different modalities. With regard to multisensory temporal integration, spatial information of audiovisual pairs can affect the width of TBW, with increased temporal acuity (i.e., smaller TBW) for stimuli from different spatial locations (Bertelson and Aschersleben, 2003; Zampini et al., 2003a, 2003b). Even for audiovisual stimuli presented at the same location, the exact position of the targets also affects individuals' tolerance of temporal offsets. For example, individuals are more likely to judge stimuli as synchronous when they occur in the periphery (Noesselt et al., 2005; Stevenson et al., 2012a). There is also a tendency to bind audiovisual stimuli of large temporal disparity within the peripersonal space (surrounding the body) compared with those appearing in the distal space (Noel et al, 2016b, 2018c). As proposed by Noel et al. (2016b), this effect is due to an automatic remapping of stimuli into a body-centred reference frame, which is necessary to further integrate external stimuli with tactile bodily stimulation. Due to the difference in the transmission speed of light and sound, individuals can shift their PSS adaptively when the distance of audiovial targets changes (Lewald and Guski, 2004; Silva et al., 2013). In terms of the effect of stimulus intensity/effectiveness, no consistent conclusion has been reached with regard to whether salient (Shahin et al., 2017) or low-intensity (Krueger Fister et al., 2016) audiovisual stimuli are associated with a larger TBW. More research is needed in this area.

Besides space and intensity, the intrinsic characteristics of the stimulus also plays a part in shaping the audiovisual TBW. In general, research findings indicate that the TBW becomes larger with increasing complexity of the stimulus (e.g., non-speech object actions < syllable < speech < musical events) (Schwartz and Savariaux, 2014; Stevenson and Wallace, 2013; Vatakis and Spence, 2006b, 2006c), with some exceptions (Maier et al., 2011; Vroomen and Stekelenburg, 2011). For simple stimuli like flash and beeps, the stimulus duration has little effect as the width of the TBW is mainly determined by stimulus onset (Tiippana and Salmela, 2018). However, for more complex speech stimuli, longer durations can offer more useful timing cues, which in turn could lead to better performance in detecting audiovisual asynchrony (Eg and Behne, 2015).

More details about studies focusing on other bottom-up features (e.g., synaesthetic congruency of audiovisual stimuli, reverberation levels, frame rate of video clips, and invertion of human face) can be found in Supplementary Table 1. In short, a wide range of low-level stimulus characteristics could influence perceptual synchrony.

4. Neural mechanisms of audiovisual TBW

4.1. Findings from fMRI/PET studies

Detecting audiovisual temporal correlation requires coordinated work of a large-scale brain network, including the auditory (van Atteveldt et al., 2007a, 2007b) and visual (Macaluso et al., 2004) cortices, the fronto-parietal dorsal attention network (Binder, 2015) and some regions crucial for multisensory integration, such as the superior colliculi (Calvert et al., 2001), the insula (Bushara et al., 2001; Lamichhane et al., 2016), the inferior parietal cortex (Adhikari et al., 2013; Dhamala et al., 2007) and the superior temporal sulcus (Marchant et al., 2012; Noesselt et al., 2012; Stevenson et al., 2011, 2010). Moreover, if the stimuli are speech-related, some prefrontal areas (e.g., the inferior frontal gyrus), which are related to speech production, are also involved (Baumann et al., 2018; Biau et al., 2016; Romanski and Hwang, 2012). Among these brain regions, the superior temporal sulcus (STS) is regarded as the hub for multisensory integration, as well as language comprehension and social cognition. Previous research has revealed two anatomically distinct sub-regions within the STS, one sensitive to low-level temporal synchrony, and the other responsible for processing audiovisual perceptual fusion at higher cognitive level (Stevenson et al., 2011).

Different brain networks are engaged when individuals are exposed to synchronous versus asynchronous audiovisual events (see meta-analysis, Erickson et al., 2014). Our ALE meta-analysis of synchronous audiovisual integration included 114 foci from 10 experiments, and the ALE meta-analysis of asynchronous condition included 163 foci from six experiments (Table 2). Details on paper selection and statitistical procedures can be found in the supplementary text. We found that the STS plays a pivotal role in audiovisual integration regardless of temporal (a) synchrony, with an anterior-posterior topographic pattern corresponding to synchrony and asynchrony perception (Fig. 2). Processing synchronous audiovisual pairs, compared with asynchronous conditions, is more likely to activate the bilateral primary auditory cortex (Fig. 2; Table 3), suggesting that redundant visual input may boost multisensory validation and increase activities in neurons receiving consistent auditory information. Such reinforcement in the auditory cortex can in turn facilitate fusion and correct perception. On the other hand, perception of asynchrony needs to harness a different functional network, including the prefrontal, the posterior superior temporal and cerebellar areas (Fig. 2; Table 3).

Based on findings about audiovisual asynchrony processing, we depicted a brain network model responsible for conflicting audiovisual temporal integration (Fig. 3). First, at the primary sensory level to encode sensory information, asynchrony may elicit prediction errors in the visual and auditory cortices (Lee and Noppeney, 2014; Wiggins and Hartley, 2015). These prediction error signals are further transmitted to higher associatve regions, leading to a stronger functional connection between the prefrontal cortex and the STS (Noesselt et al., 2012) and a significant coactivation between the anterior insula (AI) and the dorsal anterior cingulate cortex (dACC) within the Salience Network (SN), which in turn may facilitate the detection of temporal misalignment (Lamichhane and Dhamala, 2015; Lamichhane et al., 2016). These higher-order prefrontal regions may be involved in error monitoring and conflict resolution on encountering inconsistent multisensory inputs (Erickson et al., 2014). The AI is responsible for integrating sensory information, and the functional connection from the AI to the dACC guides the selection of appropriate choices (Lamichhane and Dhamala, 2015). More importantly, activation of the AI and the dACC has been shown to be stronger when audiovisual stimuli are asynchronous, highlighting that regions in the SN are part of a system supporting difficult sensory decision-making tasks (Lamichhane et al., 2016). In addition to the prefrontal cortex, audiovisual asynchrony perception recruits brain regions related to motor planning and action production (Lee and Noppeney, 2011). Specifically, asynchronous stimuli significantly enhance the connectivity from the premotor cortex and the cerebellum to the STS. To interpret this process of sensorimotor integration, Noppeney and Lee (2018) have proposed an internal forward model that "maps from the motor plan of intended action (e.g., speaking) onto its sensory consequences (e.g., concurrent auditory sounds)", which is thought to be a supplementary mechanism enabling more precise temporal prediction of audiovisual events. Finally, apart from its function of motor planning, the cerebellum also contributes to time perception, especially the detection of stimulus onset and consequent error correction (Petter et al., 2016). Thus, the robust activation of the cerebellum in TBW-related tasks (Lee and Noppeney, 2014, 2011; Petrini et al., 2011) suggests that precise timing may be a prerequisite for intact multisensory temopal integration.

To summarize, the brain model here (Fig. 3) provides potential neurocognitive mechanisms underlying audiovisual temporal processing, including low-level perceptual encoding, high-level cognitive control, supplementary motor prediction and time perception. Deficits in any phase could result in reduced asynchrony sensitivity and an enlarged TBW. It is important for future research to separate these components and investigate their roles in TBW-related tasks.

4.2. Evidence from EEG/MEG studies

Apart from fMRI/PET findings, electroencephalographic (EEG) and magnetoencephalographic (MEG) studies provide converging evidence of the neural mechanisms underlying audiovisual temporal integration. Multisensory integration as indexed by the difference between event-related potentials (ERPs) evoked by audiovisual stimuli and the summation of ERPs evoked by unisensory stimuli is strongest when auditory and visual stimuli are synchronous, with longer SOAs weakening the integration effect (Liu et al., 2011). Source analysis of EEG data has revealed that enhanced frontal activity and decreased activity in the auditory and visual cortices, indexing superior top-down regulation, are correlated with fewer sound-induced visual illusion and narrower TBWs (Bidelman and Heath, 2019b). Additionally, given the methodological advantage of EEG and MEG (i.e., higher temporal resolution compared with fMRI), these studies also provide information on the different processing stages of multisensory integration. The following section reviews the main EEG/MEG findings regarding brain oscillations and ERPs of audiovisual temporal processing, covering the whole sensory and cognitive process, starting from pre-stimulus preparation and ending with final sensory decision-making.

First, pre-stimulus local and network activity can influence upcoming multisensory perception (Ikumi et al., 2019; Keil et al., 2014).

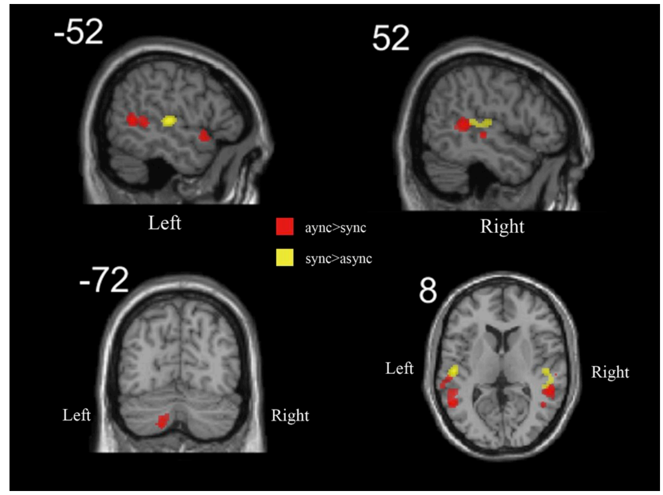

Fig. 2. Significant clusters for asynchronous vs. synchronous audiovisual integration. Notes: Synchronous audiovisual stimuli activate bilateral primary auditory cortex more strongly than asynchronous conditions (shown in yellow). In contrast, asynchronous audiovisual processing recruits posterior superior temporal cortex, left cerebellum and the inferior frontal cortex (shown in red). Threshold: $p < 0.001$ for cluster-forming method and $p < 0.05$ for cluster-level inference. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 2
Studies contrasting synchronous vs. asynchronous audiovisual integration included in the ALE meta-analysis.

Study	Sample size	audiovisual stimuli	Task	N of foci	Summary of regions
Synchronous > Asynchronous (10 studies, 114 foci)					
Bishop and Miller (2009)	25	vowel-consonant-vowel tokens	Speech identification	22	Bilateral temporal-occipital boundary; intraparietal sulcus; precentral and postcentral gyrus; caudate and putamen
Calvert et al. (2001)	10	Flash and beeps	Passive viewing and listening	15	superior colliculus; insula; frontal regions
Lewis and Noppeney (2010)	16	rotated visual patterns with auditory clicks	Shape/motion discrimination	14	Auditory and visual areas
Macaluso et al. (2004)	8	Words/nouns	Target detection	9	ventral and dorsolateral occipital cortex; left superior temporal sulcus (STS)
Marchant et al. (2012)	16	flash(chessboard) and beeps	Target detection	9	STS; supplementary motor area; thalamus and basal gangalia
Noesselt et al. (2007)	24	flash and beep streams	Target detection to ensure attention	19	STS; visual and auditory areas
Stevenson et al. (2010)	8	monosyllabic words/nouns	Semantic categorization	8	superior colliculus; visual and auditory areas
van Atteveldt et al. (2007a)	8	single letters and speech sounds	Passive viewing and listening	4	Posterior and anterior auditory association cortex
Miller and D`Esposito (2005)	11	vowel-consonant-vowel tokens	Simultaneity Judgement	2	STS
Noesselt et al. (2012)	11	sentences	Simultaneity Judgement	12	STS; prefrontal cortex
Asynchronous > Synchronous (6 studies, 163 foci)					
Baumann et al. (2018)	24	Non-speech visual and auditory streams	Target detection to ensure attention	3	Motor areas; dorsal lateral prefrontal cortex
Bushara et al. (2001)	12	Flash and beeps	Simultaneity Judgement	9	right insula; inferior frontal gyrus; inferior parietal lobe; left cerebellum
Lee and Noppeney (2011)	37	Short sentences and music	Passive viewing and listening	33	bilateral STS; cerebellum
Lee and Noppeney (2014)	37	speech, sinewave speech and music	Passive viewing and listening	61	STS; visual and auditory areas; cerebellum
Miller and D`Esposito (2005)	11	vowel-consonant-vowel tokens	Simultaneity Judgement	15	STS; supplementary motor areas; intraparietal sulcus; prefrontal cortex
Noesselt et al. (2012)	11	sentences	Simultaneity Judgement	42	STS; prefrontal cortex; insula

Table 3
ALE analysis results of asynchronous vs. synchronous audiovisual integration.

Contrast Type	Brain Region	Volume (mm^3)	ALE value	MNI x	MNI y	MNI z	Brodmann areas	Contributing experiments
Sync > Async	Right STG	1496	0.0148	50	−20	10	Brodmann area 41	Lewis and Noppeney (2010)
			0.0112	56	−18	2	Brodmann area 22	Noesselt et al. (2007, 2012)
			0.0108	54	−30	8	Brodmann area 41	Stevenson et al. (2010)
			0.0102	48	−36	10	Brodmann area 41	
			0.0102	40	−36	16	Brodmann area 41	
	Left STG	896	0.0183	−52	−20	6	Brodmann area 22	Lewis and Noppeney (2010); Marchant et al. (2012); Noesselt et al. (2007)
Async > Sync	Right MTG	2120	0.0202	54	−42	8	Brodmann area 21	Lee and Noppeney (2011, 2014);
			0.0189	46	−58	4	Brodmann area 37	Miller and D'Esposito (2005)
	Left STG/STS/MTG	1936	0.0214	−52	−54	8	Brodmann area 39	Lee and Noppeney (2011, 2014);
			0.0209	−54	−42	4	Brodmann area 22	Noesselt et al. (2012)
	Left cerebellum	1536	0.0189	−16	−72	−44	Inferior Semi-Lunar Lobule	Lee and Noppeney (2011, 2014); Bushara et al. (2001)
			0.0166	−22	−66	−46	Cerebellar Tonsil	
			0.0166	−16	−72	−40	Inferior Semi-Lunar Lobule	
			0.0161	−26	−66	−50	Cerebellar Tonsil	
			0.0102	−12	−80	−32	Uvula	
	Left STG/STS/MTG	1224	0.0139	−66	−38	16	Brodmann area 22	Lee and Noppeney (2011, 2014);
			0.0131	−58	−28	8	Brodmann area 41	Noesselt et al. (2012)
			0.0128	−64	−36	8	Brodmann area 22	
			0.0128	−62	−32	6	Brodmann area 22	
	Right STG/STS	752	0.0189	54	−24	−2	Brodmann area 21	Lee and Noppeney (2011, 2014)
			0.0166	58	−26	2	Brodmann area 22	
	Left STG/IFG	576	0.0177	−52	12	−9	Brodmann area 22	Lee and Noppeney (2011, 2014)

Notes: Sync = synchronous; Async = asynchronous; STG = superior temporal gyrus; STS = superior temporal sulcus; MTG = middle temporal gyrus; IFG = inferior frontal gyrus.

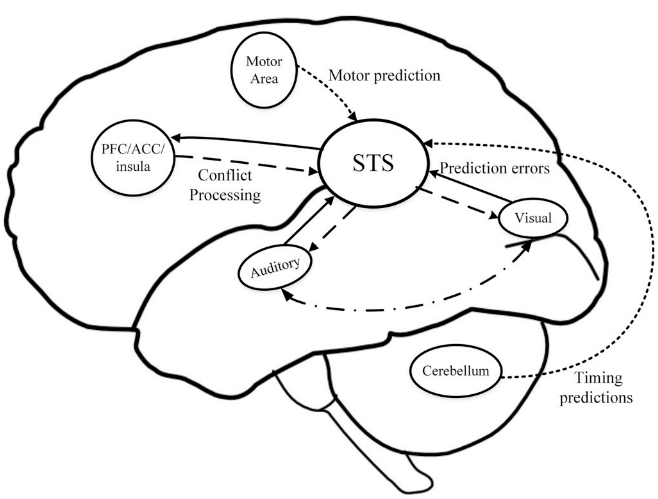

Fig. 3. A hypothetical model of the brain network responsible for audiovisual asynchrony detection. **Notes:** PFC = prefrontal cortex; ACC = Anterior Cingulate Cortex; STS = superior temporal sulcus (1) ·-· communications between unisensory areas (e.g., phase resetting) (2) — bottom-up processes: Sensory inputs and prediction errors are transmitted from the primary sensory cortex to the multisensory STS, and further to the prefrontal cortex. (3) — top-down modulations: The prefrontal cortex send conflict-solving signals to the STS, which further reduce prediction errors in the auditory and visual cortex. (4) ···· supplementary mechanisms for more accurate audiovisual temporal integration.

For example, a down-modulation of pre-stimulus alpha power in the auditory cortex (an index of decreased inhibition of ongoing information) is linked to better performance in temporal judgement (Grabot et al., 2017). Baseline fluctuation in the beta and gamma band before stimulus onset could predict perceived simultaneity (Yuan et al., 2016). However, it should be noted that synchrony perception for auditory-leading and visual-leading sequences is correlated with opposite changes in pre-stimulus oscillatory power (enhanced vs. decreased prestimulus beta and gamma power respectively for AV and VA synchrony perception) (Yuan et al., 2016) and may recruit separate neural pathways (Cecere et al., 2017).

Consistent with fMRI findings, the primary sensory cortices have already shown sensitivity to the temporal offset of multisensory information, as MEG source analysis has shown that temporal asynchrony could affect the latency of peak activation in the primary visual and auditory cortices (Franciotti et al., 2011). Multisensory interactions occur in the early stages of cortical processing (i.e., auditory gammar-band responses (GBRs) in a post-stimulus time window of 30–80 ms and visual GBRs after 60–120 ms), but this early multisensory effect can only be observed when auditory and visual stimuli are presented with the highest temporal proximity (SOA < 25 ms) (Senkowski et al., 2007). Other evidence supporting temporal synchrony that could influence early-stage sensory processing includes the reduction of auditory evoked potentials by congruent visual stimuli (Pilling, 2009). This attenuation effect is observed as early as six months of age (Kopp, 2014; Kopp and Dietrich, 2013), and is maximal when the visual signal precedes the auditory one by a small SOA or the audiovisual pairings occur simultaneously, both of which correspond to the greatest possibility of synchrony perception (Simon and Wallace, 2018). In contrast, out-of-synch perception (i.e., the segragation of irrelevant audiovisual pairs) is associated with larger amplitude auditory ERPs (e.g., more positive voltage between 210 and 270 ms following sound onset (Kaganovich and Schumaker, 2016) and larger N1–P2–N2 auditory sequences (Bhat et al., 2015)) and also more persistent low theta-band oscillations (3.5–5 Hz) well after the auditory stimulus (Simon and Wallace, 2018). Enhanced auditory potentials may reflect a more detailed neural encoding for acoustic onsets, which in turn conribute to higher sensitivity for audiovisual temporal asynchrony (Kaganovich and Schumaker, 2016).

Apart from the modulations of auditory ERPs by multisensory stimuli (Musacchia and Schroeder, 2009), phase resetting is another important mechanism that could explain audiovisual temporal integration in the primary sensory cortices. This is a phenomenon where a preceding sensory stimulus may reset the phase of ongoing oscillations in the cortex that receives the second sensory input, leading to stronger phase coherence between two separate sensory areas (Kambe et al., 2015). Such processes, coupled with earlier pre-stimulus oscillatory modulation, can prepare the reset brain region for subsequent processing of the second stimulus, integrating it with the first sensory input, thus contributing to temporal binding and synchrony perception (Kambe et al., 2015).

To facilitate audiovisual temporal integration, enhancement of phase coherence not only occurs between the primary auditory and visual cortices (Kambe et al., 2015; Nozaradan et al., 2012), but also involves higher-order and multisensory brain regions, such as the middle temporal gyrus (Keil et al., 2014) and the prefrontal-parietal network (Doesburg et al., 2008). In fact, a large scale functional network, indicated by heightened global coherence (especially in the gamma band linked to feature integration and associative learning), has been found for both synchronous (Kumar et al., 2016) and asynchronous audiovisual perception (Doesburg et al., 2008). According to the communication-through-coherence hypothesis, neural populations communicate more efficiently through long-range phase synchronization (see review, Senkowski et al., 2008). These coherently oscillating neurons may form a task-relevant network responsible for explicit simultaneity judgement or implicit multisensory fusion, enabling either a unified percept for temporally-aligned audiovisual pairs or a segragation of temporally-distant audiovisual stimuli. Using fixed audiovisual presentations in which participants perceived simultaneity/fusion in approximately 50% of trials, it would be informative to assess how and to what extent oscillatory activities could predict different responses (i.e., synchronous vs. unsynchronous; illusion vs. non-illusion) for identical stimuli. Previous studies using this method have demonstrated that phase-resetting only ocurrs when participants make simultaneity responses (Kambe et al., 2015) and pre-stimulus phase coupling between multimodal temporal area and the auditory cortex only precedes multisensory illusion (Keil et al., 2014). More studies are needed to disentangle the effect of ongoing (both pre-stimulus and post-stimulus) oscillations on varying perception for invariant stimuli.

4.3. Neuroregulatory studies of TBW to explore causal links

All the above findings are based on correlational studies. To determine the causal links between brain activities and audiovisual TBW, non-invasive brain stimulation methods such as transcranial direct current stimulation (tDCS) and transcranial alternating current stimulation (tACS) can be used to modulate neural excitability or brain oscillations in specified cortical regions. In this review, we found two relevant neuro-modulatory studies. Zmigrod and Zmigrod (2015) found that audiovisual TBW substantially narrowed when anodal tDCS was applied to the posterior parietal cortex (PPC) (enhancing its excitibility), indicating a possible causal link between the PPC and the flexibility of the TBW. However, since the PCC is responsible for a wide range of other cognitive abilities (e.g., attention, memory retrieval) (Ciaramelli et al., 2008), it is difficult to pinpoint which stage of cognitive process supported by the PCC is facilitated to cause TBW narrowing. Cecere et al. (2015) modulated occipital oscillatory activitives by tACS while participants were performing the Sound Induced Flash Illusion (SiFi) Task. Driving oscillations in the alpha band towards slower versus faster frequencies leads to wider versus shorter TBWs. Therefore, the duration of the alpha cycle may be a neural temporal unit to bind auditory and visual inputs (Cecere et al., 2015). More importantly, the relationship between occipital oscillations and the width of the TBW shows a frequency-specific effect for different sensory modalities, with auditory–visual interaction in the alpha band and somatosensory–visual connectivity in the beta band (Cooke et al., 2019). This finding indicates modality-specifc neural mechanisms when integrating sensory inputs from different kinds of crossmodal stimulus pairs.

5. Audiovisual TBW across the lifespan

The final product of multisensory integration is shaped by the relative reliance on low-level stimulus characteristics including temporal relationships and experienced-related learned associations (see review, Murray et al., 2016). Developmentally, there is a shift from heavy dependence on stimulus characteristics in early life to increasing emphasis on learned associations as one progresses into adulthood (Murray et al., 2016). Specifically for the ability of utilizing temporal cues to integrate or segregate audiovisual information, the size of the audiovisual TBW follows a U-shaped pattern across the lifespan and it is the narrowest and most mature in intermediate ages (Noel et al., 2016a; Stevenson et al., 2018a).

Starting at birth, young infants show preference for synchronous audiovisual videos (Curtindale et al., 2019; Morrongiello et al., 1998) and are able to detect audiovisual asynchrony for both non-speech (Hannon et al., 2017; Lewkowicz, 1996) and speech (Lewkowicz, 2010, 2003; 2000; Pons and Lewkowicz, 2014) stimuli after habituating to audio-visually synchronous pairings, but their sensitivity to multisensory temporal relationships is considerably poorer than adults (i.e., requiring large temporal intervals to detect audiovisual asynchrony) (Lewkowicz, 2010, 1996). Temporal synchrony is thus considered a basic and simple cue for multisensory integration, helping infants broadly bind corresponding audiovisual information, including non-human primate faces and vocalizations (Lewkowicz et al., 2010), point lights of biological motion (Falck-Ytter et al., 2011), native audible and visible speech stimuli and non-native multisensory inputs (Hillairet de Boisferon et al., 2017; Lewkowicz et al., 2015). In addition, the mechanism underlying asynchrony perception in infants may be domain-general and based on the sensitivity to stimulus-energy onsets and offsets rather than speech-specific characteristics such as lip movement (Hollich et al., 2005) or acoustic speech signals (Lewkowicz, 2010).

With accumulating perceptual experience and accelerating neural conduction during childhood, the audiovisual TBW becomes more fine-tuned but still remains larger than adults (Kaganovich, 2016; Lewkowicz and Flom, 2014). In fact, for simple and non-speech stimuli (e.g., flashbeeps), this developmental course is rather protracted, extending beyond the first decade (Hillock et al., 2011) and throughout the teenage years (Hillock-Dunn and Wallace, 2012). However, , 2016b) found that 9-year-olds could perceive flashbeep simultaneity as efficiently as adults. For more complex audiovisual speech stimuli, the TBW seems to mature much earlier (Hillock-Dunn et al., 2016). Such findings indicate that multisensory integration is an experience-dependent process, with temporal associations of more ecologically relevant communicative signals being learnt first.

It is unclear whether Auditory-Visual (A-V) and Visual-Auditory (V-A) asynchrony sensitivities have the same developmental trajectory (Hillock-Dunn and Wallace, 2012) or mature at different ages (Hillock et al., 2011; Kaganovich, 2016). As visual-leading circumstances are predominantly represented in the real world, a typical TBW of a mature adult is often asymmetrical, with a flat slope on the visual-leading side (i.e., more tolerance of visual-leading conditions) (Love et al., 2013; Stevenson et al., 2012b). Future studies are required to determine when and how the different detection sensitivity of A-V and V-A asynchrony are established.

In normal ageing, the TBW tends to increase in size again where the elderly typically requires longer temporal offsets to detect asynchrony (Chan et al., 2014), have difficulty in discriminating temporal order (Bedard and Barnett-Cowan, 2016; De Boer-Schellekens and Vroomen, 2014; Setti et al., 2011b) and become more vulnerable to flash fission illusions induced by well-separated sounds (Hernandez et al., 2019; McGovern et al., 2014; Setti et al., 2011a). However, Fiacconi et al. (2013) found no age effect on performance in the TOJ task. Older adults also show altered auditory N1 and visual P1 amplitude when inputs are separated by long temporal intervals, suggesting impaired attention switching between different sensory modalities and an extended TBW among the elderly (Basharat et al., 2018; Setti et al., 2011b). Even if some older adults perform no worse than their younger counterparts in audiovisual synchrony tasks, they need to recruit more widespread frontal and parietal brain regions to maintain a similar level of performance, as indicated by the differential spatial-temporal EEG patterns between participants in two different age groups (Chan et al., 2017). More importantly, such impaired audiovisual asynchrony perception cannot be explained by a decline in unisensory detectability alone (Chan et al., 2014). In other words, even if visual and auditory stimuli are scaled according to individual detection thresholds, a broader audiovisual TBW is still found in older adults (Chan et al., 2014).

The U-shaped pattern of audiovisual TBW across the lifespan is thought to have its corresponding neural underpinnings. Animal studies with invasive recordings suggest that the number of multisensory neurons (i.e., neurons activated by stimuli from multiple sensory modalities) increases gradually after birth and they mature over an extended developmental period (Wallace and Stein, 1997; Wallace et al., 2006). The late maturation of the TBW may thus coincide with protracted maturation of multisensory brain networks. In contrast to multisensory development, much less is known about how multisensory neurons and their function change with age. However, evidence from human behavioural measures suggests stronger multisensory facilitation in older adults, which may serve as a compensatory mechanism for general perceptual slowing (Laurienti et al., 2006; Peiffer et al., 2007). Consistent with this viewpoint, larger TBWs in older adults are likely to reflect the need for greater accumulation of evidence to counteract deteriorating peripheral sensory processing efficiency and degraded sensory representations (Diederich et al., 2008). Therefore, larger TBWs during development and ageing may be caused by different mechanisms which warrant future investigations.

6. Atypical audiovisual TBW in various mental and physical disorders

6.1. Neurodevelopmental disorders

A newly published review article summarized literature on a broader scope of multisensory processing not restricted to audiovisual domain and temporal factors (Wallace et al., 2020). They have found altered multisensory function in several neurodevelopmental and neuropsychiatric conditions and such atypical multisensory integration has been shown to correlate with higher-order cognitive and social abilities (Wallace et al., 2020).

Specifically for the temporal aspect of multisensory processing, an 'appropriate' size of TBW is critical for forming a coherent perception of the external world. On the one hand, It should be narrow enough to separate temporally disparate events, On the other hand, it should also be wide enough to tolerate small offsets and allow multisensory information from the same source to be bound taking into consideration the different transmission speed in auditory and visual system. In other words, abnormally widened or shortened TBW can be equally maladaptive and result in decreased integration. In fact, an enlarged TBW has been found to be characteristic of a wide range of psychiatric and neurological disorders (see Fig. 4). Among all the disorders investigated, neurodevelopmental disorders such as autism, dyslexia and schizophrenia have received most research attention (Wallace and Stevenson, 2014; Zhou et al., 2018). We found a total of 22 studies investigating audiovisual TBW in autistic individuals. More details about the study characteristics and main findings can be found in Table 4. To summarize, the overwhelming majority of evidence from children and adolescents support the presence of impaired audiovisual temporal integration in individuals with autism spectrum disorders (ASD), regardless of the tasks used to measure the size of the TBW. Several theoretical models including predictive coding and weak coherence theories have been proposed to explain impaired temporal integration of audiovisual stimuli in ASD (see reviews, Chan et al., 2016; Chan and Naumer, 2014). In contrast, autistic adults seem to have intact temporal acuity for audiovisual stimuli, at least for non-speech and arbitrary flash and beeps (Poole et al., 2017; Turi et al., 2016). These findings indicate delayed development of multisensory processing in ASD (see a recent review, Beker et al., 2018), and children with ASD may have the potential to catch up to their typically developing peers during adulthood. Our results are consistent with a recent systematic review and meta-analysis suggesting that ASD individuals show general impairment in audiovisual integration and the abnormality is more prominent in younger populations (Feldman et al., 2018). Further longitudinal studies are needed to track the developmental trajectories of the TBW in ASD individuals.

As individuals with ASD demonstrate severe social and communication dysfunctions, one controversial question is whether the extended audiovisual TBW is restricted to speech-related events (Bebko et al., 2006; Noel et al., 2017b; Stevenson et al., 2016, 2014a) or generalized to non-speech stimuli as well (de Boer-Schellekens et al., 2013a; Foss-Feig et al., 2010; Kwakye et al., 2011; Noel et al., 2018b). However, the TBW for non-speech stimuli can predict the strength of audiovisual speech integration (indicated by McGurk fusion), which is in turn correlated with the capacity of speech comprehension in noise (Stevenson et al., 2018b). Sensitivity to audiovisual synchrony has also been found to be correlated with language and communicative skills in ASD children (Noel et al., 2018a; Patten et al., 2016; Righi et al., 2018). Extending to the subclinical population, adults with elevated autistic traits show a strong bias to perceive auditory-leading pairs as simultaneous (Donohue et al., 2012) and recalibrate less after exposure to the visual-leading adaptation phase (Stevenson et al., 2017b). In addition, adults showing higher tendency to focus on sensory details may have a narrower TBW, while those reporting more restricted and rigid behaviours are more likely to integrate audiovisual speech stimuli over wider TBWs (van Laarhoven et al., 2019). These results suggest that altered sensory perception may contribute to autistic symptomatology in the general population.

Not unique to ASD, children with other developmental disorders such as specific language impairment similarly tend to bind auditory and visual information over abnormally long temporal intervals (Kaganovich et al., 2014; Pons et al., 2013), which is a predictor of their language skills (Kaganovich, 2017). In addition, decreased temporal acuity seems to be a general feature of dyslexia in children and adults, ranging from unisensory (Cacace et al., 2000; Chen et al., 2016a) to multisensory processing (Hairston et al., 2005; Virsu et al., 2003), and also from non-speech to speech stimuli (Francisco et al., 2017). As reading depends on rapid and accurate integration of multisensory cues, such an enlarged TBW over which even irrelevant auditory and visual elements of words would be integrated may cause inappropriate graphene-phoneme associations (Hairston et al., 2005; Blomert and Froyen, 2010; Hahn et al., 2014), consequently impairing phonological awareness (Laasonen et al., 2002), and the speed and accuracy of reading (Laasonen et al., 2012) in dyslexic individuals. However, it is uncertain whether graphene-phoneme correspondences are temporally dependent. The hypothesis that reading deficits may be a result of abnormally widened TBWs warrants further investigation. Despite the impairment in temporal processing, dyslexic children could utilize synaesthetic-congruent auditory cues (i.e., auditory pitch-visual size matching) to separate visual stimuli (Chen et al., 2016a). Future interventions could target at improving dyslexic children's performance on synaesthetic associations (reducing reaction times and enhancing audiovisual congruent mapping), and utilize this improved synaesthetic integration to further enhance their performance in visual TOJ tasks in which visual stimuli are flanked by congruent beeps. Finally, the sharpened temporal acuity may help to improve reading abilities (Chen et al., 2016a).

On the other hand, patients with schizophrenia are more likely to integrate temporally separate multisensory stimuli, indicated by the SiFi

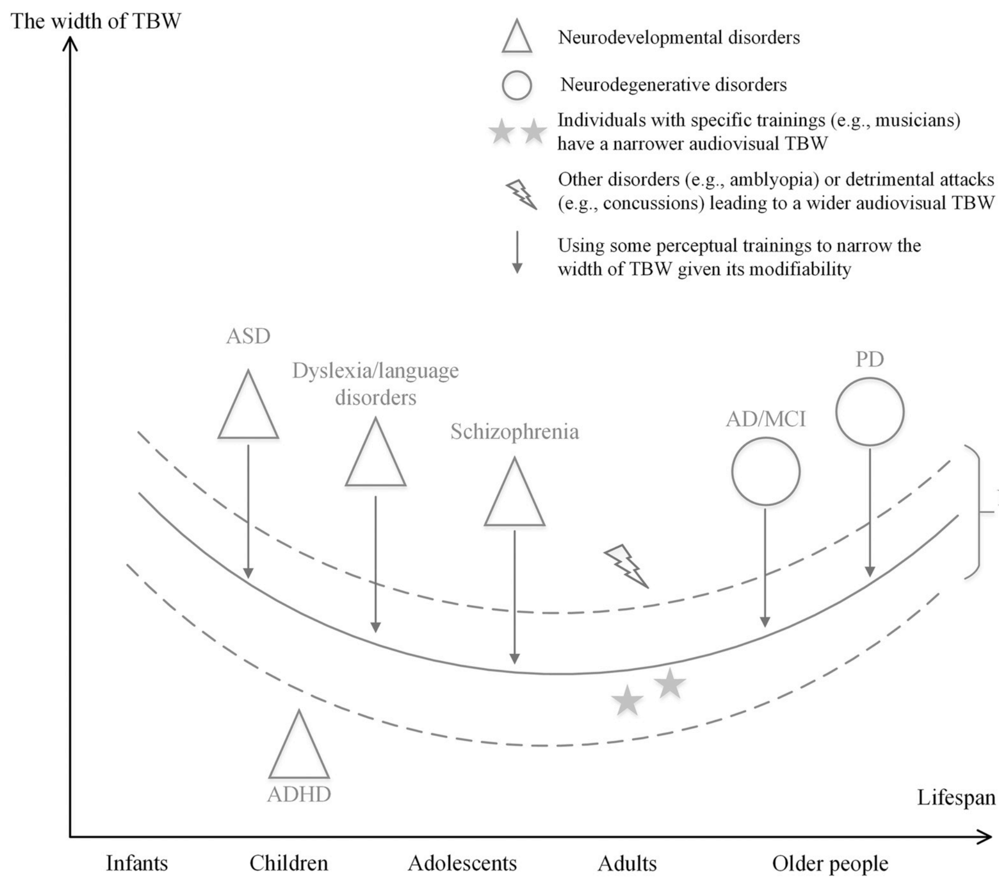

Fig. 4. The developmental trajectory of audiovisual TBW across the lifespan and its abnormality in different kinds of disorders. **Notes:** ASD = Autism Spectrum Disorders; ADHD = Attention Deficit/Hyperactive Disorders; AD = Alzheimer Disorders; MCI = Mild Cognitive Impairment; PD = Parkinson Disorders.

task (Haβ et al., 2017) and stream/bounce illusion (Zvyagintsev et al., 2017). They find it hard to explicitly detect audiovisual speech asynchrony (Martin et al., 2013) or judge visual temporal orders (De Boer-Schellekens et al., 2014a). Such reduced temporal acuity has also been found in individuals with high levels of schizotypal traits (Ferri et al., 2018, 2017). Indeed, similar to individuals with dyslexia, impaired temporal integration in schizophrenia patients generalizes from visual, auditory to bimodal domains, which may cascade into clinical symptoms such as disorganization and auditory hallucinations (Foucher et al., 2007; Stevenson et al., 2017a). Some may argue that widening of the audiovisual TBW is only a reflection of slower and impaired within-modality processing. However, previous research has suggested that multisensory TBW anomaly in schizophrenia cannot be fully explained by unisensory processing, and contributes uniquely to aberrant perception in schizophrenia (Stevenson et al., 2017a).

Evidence from autism, reading/language disorders and schizophrenia all suggests an enlarged TBW in neurodevelopmental disorders. However, Attention Deficit Hyperactivity Disorder (ADHD), another common neurodevelopmental disorder, may be an exception (Panagiotidi et al., 2017). Adults with high levels of ADHD-like traits show a much smaller TBW than those with low levels of ADHD-like traits. This abnormally shortened TBW may reflect a failure to integrate multimodal stimuli in ADHD. In other words, subclinical ADHD adults have a bias to separate audiovisual pairings even if they are only misaligned for a small and tolerable temporal span, which could lead to distractibility, a core symptom of ADHD (Panagiotidi et al., 2017). The aforementioned findings indicate the potential of atypical TBW as a gateway to investigate the similarities and differences between different neurodevelopmental disorders.

Few studies have tried to disentangle the neural substrates of less efficient multisensory synchrony perception in neurodevelopmental disorders. Only one study using task-based fMRI (Sanfratello et al., 2018) found decreased activity at the intraparietal sulcus (dorsal visual stream) under multisensory conditions in patients with schizophrenia, which is correlated with the severity of positive symptoms. However, there is accumulating evidence suggesting altered brain structure and function underlying multisensory integration and temporal processing in these clinical populations. For example, the posterior STS, which in our ALE meta-analysis (Section 3) is considered a hub region for audiovisual temporal integration, has been found to have reduced grey matter volume, atypical activation pattern and altered structural and functional connectivity in individuals with autism, dyslexia and schizophrenia (see review, Wallace and Stevenson, 2014). Moreover, the prefrontal cortex and the Salience Network (SN), involved in conflict detection for asynchronous audiovisual stimuli, have consistently been reported to show abnormality in individuals with ASD (Vissers et al., 2012), schizophrenia (Glahn et al., 2008) and ADHD (De La Fuente et al., 2013). Finally, distortion in time perception is a common feature of these neurodevelopmental disorders (e.g., ASD, schizophrenia and ADHD, Allman and Meck, 2012; dyslexia, Farmer and Klein, 1995). The cerebellum identified in our ALE analysis plays a key role in temporal processing (Petter et al., 2016). Thus, impaired function of the cerebellum may be one of the neural substrates for altered TBW in these disorders.

Although an atypical TBW has been found in several neurodevelopmental disorders and has been linked to perceptual and communicative impairments, little is known about how altered audiovisual temporal integration finally cascades into clinical symptoms. One hypothesis is that enlarged TBWs may be a reflection of problems in combining irrelevant sensory information, which results in an unpredictable and confusing environment. This may explain the link between atypical TBWs and hallucinations in schizophrenia (Stevenson et al.,

Table 4
Summary of studies about audiovisual temporal integration in ASD.

Authors(year)	Paradigm	stimuli type	main findings
Infants and young children (<6 years)			
Bebko et al. (2006)	preferential looking	non-linguistic, simple and complex linguistic videos(3s AV temporal offset)	ASD failed to show preference for congruent linguistic videos
Falck-Ytter et al. (2018)	preferential looking	point light displays of biological motion	Ten-month-old infants who later received an autism diagnosis did not orient to audiovisual synchrony expressed within biological motion.
Falck-Ytter et al. (2013)	preferential looking	point light displays of biological motion	Three-year-old children with autism orient to neither audiovisual synchronous video nor to biological motion
Righi et al. (2018)	preferential looking/eye-tracking	speech videos	1 ASD failed to detect V-A 1s asynchrony while TD can detect 0.6s 2 ASD looked less at eyes and mouth of the dynamic face. 3 Individual differences in sensitivity to audiovisual asynchronies and individual differences in orientation to relevant facial features were both correlated with scores on a standardized measure of language abilities
Patten et al. (2016)	preferential looking; language-based paradigm	speech with face obscured or not(audiovisual videos)	Children with low language abilities were significantly worse at detecting synchrony when the stimuli include an unobscured face than when the face was obscured.
◆Patten et al. (2014)◆	preferential looking	speech **without face** (linguistic stimuli paired with movement of related toys in the absence of faces(700 ms out of syn))	1 Children with autism demonstrated the ability to detect audio-visual synchrony when faces were obscured. 2 The amount of time they attended to the synchronous condition was positively correlated with receptive language.
Children and adolescents (High functioning ASD)			
Kwakye et al. (2011)	Multisensory TOJ (temporal ventriloquism; auditory effects on visual TOJ)	flashbeep	1 Enlarged TBW in ASD 2 Impaired auditory TOJ in ASD
Foss-Feig et al. (2010)	SiFi	flashbeep	Enlarged TBW in ASD
Stevenson et al. (2018b)	TOJ, McGurk, speech in noise	flashbeep for TOJ/TBW	1 No difference of flashbeep TBW. 2 Larger audiovisual TBW contributed to impaired speech perception in noise, and mediated by McGurk fusions.
Stevenson et al. (2014a)	SJ and TOJ	flashbeep, tool, syllable	1 Speech-specific enlargement of TBW in ASD 2 Enlarged TBW was associated with less McGurk effects.
Noel et al., 2017b	SJ	flashbeep, tool, syllable	1 Speech-specific enlargement of TBW in ASD 2 Less rapid temporal recalibration for non-speech audiovisual stimuli in ASD
◆Smith et al. (2017)◆	SJ	speech (i.e., consonant-vowel utterances) and object stimuli (i.e., a bouncing ball)	1 While controls showed similar tolerance of asynchrony for the simple speech and object stimuli, individuals with ASD showed less tolerance of asynchrony for speech stimuli compared to object stimuli. 2 Decreased tolerance for asynchrony in speech stimuli was associated with higher ratings of autism symptom severity.
Grossman et al. (2015)	preferential looking	speech(audiovisual videos)	1 ASD looked at the in-synch video less than TD peers and did not increase their gaze time as much as TD participants in the explicit task. 2 ASD looked significantly less at the mouth than their TD peers, and significantly more at non-face regions of the image.
Noel et al. (2018a)	SJ	speech (syllable)	1 Enlarged TBW in the ASD group. 2 ASD individuals performed less complex head and hand movement and showed less nonverbal synchrony with an interacting partner. 3 Audiovisual temporal acuity significantly predicted synchrony in hand and head movements between TD participants and the experimenter (i.e., narrower the TD participants' TBW, the greater their interpersonal synchrony), but not between the ASD individuals and the experimenter.
Noel et al. (2018d)	SJ	speech(syllable)	1 Enlarged TBW in ASD and schizophrenia. 2 While the wider TBWs in ASD most prominently results from atypical priors, the wider TBWs in SZ results from a trend toward changes in prior and weaknesses in the sensory representations.
Feldman et al., 2019	McGurk with varying audiovisual SOAs	speech(syllable)	1 Enlarged TBW in ASD 2 Enlarged TBW and impaired multisensory speech perception were correlated with atypical sensory responsiveness (i.e., hyporesponsiveness and sensory seeking) across the ASD and TD group.
Woynaroski et al. (2013)	McGurk with varying audiovisual SOAs	speech(syllable)	1 Enlarged TBW in ASD 2 ASD displayed deficits in visual only and matched audiovisual speech perception
Adolescents and young adults (High functioning ASD)			
de Boer-Schellekens et al. (2013a)	TOJ	flashbeep/handclap/speaking face	general impairment in audiovisual TOJ in ASD
Noel et al. (2018b)	SJ	flashbeep; heartbeep	1 Enlarged audiovisual TBW in ASD 2 Four-folds larger cardio-visual TBW in ASD
de Boer-Schellekens et al. (2013b)	Multisensory TOJ (temporal ventriloquism; auditory effects on visual TOJ)	flashbeep	1 impaired visual TOJ in ASD 2 intact low-level audiovisual integration or the phasic alerting by abrupt sounds
Adults (High functioning ASD)			
◆Turi et al. (2016)◆	SJ	flashbeep	1 No difference in TBW

(continued on next page)

Table 4 (*continued*)

Authors(year)	Paradigm	stimuli type	main findings
			2 No rapid audiovisual recalibration in ASD
◆Poole et al. (2017)◆	TOJ	flashbeep; tactile	1 No group differences in temporal acuity for crossmodal stimuli. 2 Visual-tactile temporal acuity and bias towards vision when presented with visual-auditory information were both predictors of self-reported sensory reactivity.
Trait adults			
Stevenson et al. (2017b)	SJ and adaptation	flashbeep	1 Following exposure to the visual-leading adaptation phase, participants' perception of synchrony was biased towards visual-leading presentations. 2 The strength of adaptation was significantly related to the level of autistic traits, especially the Attention to detail
Donohue et al. (2012)	SJ	flashbeep	Auditory-leading PSS is related to high autistic traits
van Laarhoven et al. (2019)	SJ	Speech (pseudoword)	1 Increased difficulties in attentional switching (which reflected more rigid and restricted behaviour) was correlated with a wider TBW. 2 An increased tendency to focus on local aspects of sensory inputs was correlated to a narrower TBW.

Notes. TBW = temporal binding window; SJ = simultaneity judgement; TOJ = temporal order judgement; SiFi = sound induced flash illusion.
Studies with "◆" did not find enlarged audiovisual TBW in ASD.

2017a). In addition, an atypical TBW may also reflect impaired ability to correctly integrate relevant multimodal stimuli (Stevenson et al., 2018b, 2012b), which could undermine a vast array of abilities relying on multisensory integration. For example, the failure to utilize temporal cues to match facial vocal information with auditory speech signals may cause ASD individuals to experience difficulty in communicating effectively with others (Stevenson et al., 2018b). Similarly, weakened graphene-phoneme integration is a potential mechanism underlying deficits in reading and language comprehension (Hairston et al., 2005). Given the scarcity of studies investigating the pathway linking TBW with higher-order abilities, more research is needed to fill this gap.

6.2. Neurodegenerative disorders

Compared with developmental disorders, relatively few studies have investigated audiovisual temporal acuity in neurodegenerative disorders. Individuals with Parkinson's disease show abnormal audiovisual temporal order judgement where auditory processing is considerably delayed and thus the PSS shifts towards the auditory-leading side (Lewald et al., 2006). People with mild cognitive impairment (MCI) tend to integrate audiovisual information across a wider TBW compared with healthy elderly people (Chan et al., 2015). Although no study to date has specifically examined the temporal factor of multisensory integration in Alzheimer's disease (AD), patients with AD do show reduced audiovisual speech integration and report less McGurk perception (Delbeuck et al., 2007; Festa et al., 2013). Moreover, extended TBWs are linked to increased susceptibility to falls in the elderly, indicating the potential effect of multisensory integration on balance maintenance (Setti et al., 2011a). Compromised audiovisual temporal acuity may also undermine general cognitive abilities (Hernandez et al., 2019) and language and reading skills in older adults (Virsu et al., 2003).

6.3. Other disorders

A number of studies have also examined audiovisual temporal integration in different neurological and physical disorders. For example, individuals with visual impairment (e.g., amblyopia) have been reported to have an enlarged TBW, which may serve as a compensatory mechanism of enhanced multisensory integration (Narinesingh et al., 2017; Richards et al., 2018, 2017). Adults with one eye surgically removed early in life, however, have normal TBW, suggesting adaptive changes for optimal audiovisual integration due to the loss of visual input during early development (Moro and Steeves, 2018). Other conditions associated with altered TBW include obesity (Scarpina et al., 2016), concussion (Wise and Barnett-Cowan, 2018) and psychogenic non-epileptic events (Noel et al., 2017a). More details about studies of conditions besides neurodevelopmental and neurodegenerative disorders can be found in Supplementary Table 2.

These findings suggest that altered audiovisual temporal integration is not unique to one particular mental disorder but shared by various disorders. Therefore, it is impractical to rely on altered TBW alone to differentiate one disorder from another. However, the width of the TBW combined with age may provide unique information for detection of different clinical groups. For example, Falck-Ytter et al. (2018) have found that a failure to orientate to audiovisual synchrony during infancy can predict future autism. More importantly, future research could target the potentially modifiable TBW to alleviate symptoms such as language and social communicative impairment in autism, dyslexia and other relevant disorders.

7. Modifiability of the audiovisual TBW

Multisensory temporal integration has a high degree of plasticity even in adult brains. On the one hand, the width of the TBW can be significantly narrowed with perceptual training (Powers et al., 2009). On the other hand, the PSS can be shifted towards the auditory or visual-leading side after transient or intensive exposure to temporally-asynchronous pairs (Fujisaki et al., 2004; Van der Burg and Goodbourn, 2015), which is a process called temporal recalibration. As this review has an emphasis on the width of the TBW and temporal recalibration is more of an adaptation of the PSS, this section will only cover studies exploring the modifiability of the width of the TBW. More details on studies concerning temporal recalibration could be found in the supplementary text.

7.1. Perceptual training

Two studies have investigated the effect of unisensory training (Alais and Cass, 2010; Stevenson et al., 2013). Stevenson et al. (2013) have found that visual temporal order training can improve temporal acuity across modalities, manifesting as narrowing of the audiovisual TBW. However, Alais and Cass (2010)'s findings suggest that unisensory (visual or auditory) learning does not transfer to other sensory modalities. Only cross-modal TOJ learning could be transferred to visual (but not auditory) temporal tasks. Taken together, it is unclear whether there is any transfer effect of unisensory training, but audiovisual-and-visual transfer, which is found in both studies, does indicate some degree of sensory interactions and central supramodal timing process (i.e., timing mechanisms independent of sensory modalities).

Most other studies employ multisensory simultaneity or temporal order training. For example, using staircase procedures to maximize the chance of effective training, both audiovisual SJ (McGovern et al., 2016) and TOJ (Setti et al., 2014) training with feedback can significantly narrow the TBW. Similarly, using the method of constant stimulus, the audiovisual SJ task with feedback successfully improves multisensory temporal acuity (Cecere et al., 2016; De Niear et al., 2018; Powers et al., 2016, 2009), the effect of which could last for at least one week (Powers et al., 2009). However, difficult stimuli with small SOAs (i.e., audiovisual pairs for which it is hard to discriminate between synchrony and asynchrony) should be employed in the training protocol to optimize the training effect (Niear et al., 2016). As for the neural correlates of multisensory perceptual training, the functional connectivity between the posterior STS (a hub for multisensory integration) and areas of the auditory and visual cortex has been found to be significantly increased after training (Powers et al., 2012). Narrowing of the TBW is also accompanied by enhanced beta-band activity under asynchronous conditions (Theves et al., 2019). This training effect on beta oscillations has been shown to exist across distributed cortical areas and persist well after the onset of auditory stimuli (>400 ms), which suggests that multisensory learning takes place in a large network involved in asynchrony detection (Theves et al., 2019).

These initial findings suggest the potential of using perceptual training to improve multisensory temporal precision, but it should be noted that such effect seems to be modality-specific (Virsu et al., 2008), and could not be generalized across different levels of stimulus complexity (De Niear et al., 2018). It also remains inconclusive whether narrowing of the TBW in a specific task could benefit other sensory tasks (Schlesinger et al., 2014; Zerr et al., 2019; Powers et al., 2016). Moreover, the modifiability of the TBW shows an asymmetrical pattern, with the visual-leading side appearing more trainable than the auditory-leading side (Cecere et al., 2016; Powers et al., 2009).

No study to date has utilized unisensory or multisensory temporal training in clinical groups with abnormally enlarged TBWs. Given the feasibility of such perceptual training and its effectiveness in narrowing the TBW, it may be possible to develop interventions to improve audiovisual integration ability in clinical populations such as autism and schizophrenia. However, some crucial questions concerning the transferability of such perceptual training effect to other tasks and higher-order abilities (e.g., social communicative functions) should be answered first before applying this intervention in clinical settings.

7.2. Musical training

In addition to laboratory-based and short-term perceptual training, long-term musical training may provide an alternative way to enhance audiovisual temporal precision. Experience-dependent plasticity in audiovisual temporal processing may also apply to bilingual language experiences, with bilinguals demonstrating a shorter TBW (Bidelman and Heath, 2019a, 2019b). However, given the scarcity of research in this field, we only focus on musical training in this review. A series of studies have converged to show that musicians are more sensitive to audiovisual asynchrony and have a much smaller TBW compared with non-musicians (see review, Noppeney and Lee, 2018). Specifically, Petrini and colleagues were among the first to show a narrowing of the TBW due to expertise with drums (Petrini et al., 2010, 2009b; 2009a). Musical practice could help drummers more efficiently detect asynchrony for drumming point-light displays even when the point-lights were rotated to a less recognizable orientation (Petrini et al., 2010) or when important information were missing (Petrini et al., 2009b). These findings indicate the potential of musical training to eliminate the influence of low-level stimulus properties when integrating multisensory information. Furthermore, the beneficial effect of musical training has been found to emerge at a very early age and drumming experience could effectively heighten young infants' sensitivity to multimodal asynchrony (Gerson et al., 2015). Stronger activation in the temporal-parietal-occipital area during TOJ tasks (Hodges et al., 2005), reduced brain activity in the bilateral cerebellum when viewing synchronous displays (Petrini et al., 2011), and more effective connectivity within the sensorimotor integrative network (i.e., the superior temporal sulcus-premotor-cerebellar circuitry) to detect asynchronous conditions (Lee and Noppeney, 2011), may serve as the neural substrates underlying musicians' superior multisensory temporal functions.

As for the generalization effect of musical practice, it remains inconclusive whether enhanced temporal sensitivity among musicians is restricted to musical events (Lee and Noppeney, 2011) or could be generalized to simple flashbeeps (Bidelman, 2016; Jicol et al., 2018) and naturalistic speech stimuli (Lu et al., 2014). In addition, some evidence suggests that enhancement of multisensory temporal integration may be instrument-specific (Bishop and Goebl, 2014). For example, pianists show significantly better performance for piano stimuli compared with other musical instruments. Although different types of musical training seem to have a similar narrowing effect on the TBW, drumming experiences, which emphasize rhythm maintenance and coordination, have a unique advantage over other melodic training (Jicol et al., 2018). In fact, synchronizing movement to auditory sequences could substantially improve timing prediction (Manning and Schutz, 2013). As the training of percussion instruments involves a lot of auditory-motor integration, it is not surprising that percussionists show superior movement timing abilities (perfect synchronization at tapping tasks) and higher sensitivity in perceptual tasks to detect the timing of rhythmic stimuli (Manning and Schutz, 2015). This may have important implications for future interventions where priority should be given to rhythmic trainings as they may be more relevant to multisensory integration at least in the temporal domain.

Given its positive effect on social interactions and cognitive development in both clinical (e.g., ASD, Kim et al., 2009; LaGasse, 2017; Sharda et al., 2018) and non-clinical children populations (Miendlarzewska and Trost, 2014), music may be a potential intervention to enhance multisensory integration and alleviate social-communicative symptoms in ASD.

8. Conclusions

This paper systematically reviews studies investigating audiovisual TBW. We have summarized the neural mechanisms of different cognitive stages underlying audiovisual temporal integration, indicating a widely distributed brain network responsible for such a basic multisensory function. The width of the TBW follows a U-shaped pattern across the lifespan, which does not mature until late adolescence and rebound in size again in old age. An enlarged TBW, reflecting impaired multisensory integration, is found in several neurodevelopmental disorders, and may also play an important role in neurodegenerative disorders. Such abnormal audiovisual temporal integration is associated with aberrant perceptual experiences and compromises social communicative skills. Given the plasticity of the TBW via modulations by both top-down and bottom-up factors, and also the potential of perceptual and musical training to narrow the TBW, future interventional research may have important implications in clinical populations.

Declaration of competing interest

No conflict of interest to be declared.

Acknowledgments

This study was supported by a grant from the National Science Foundation China (31970997), Beijing Training Project for Leading Talents in S&T (Z151100000315020), and the CAS Key Laboratory of Mental Health, Institute of Psychology.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2020.107396.

References

Adhikari, B.M., Goshorn, E.S., Lamichhane, B., Dhamala, M., 2013. Temporal-order judgment of audiovisual events involves network activity between parietal and prefrontal cortices. Brain Connect. 3, 536–545. https://doi.org/10.1089/brain.2013.0163.

Alais, D., Cass, J., 2010. Multisensory perceptual learning of temporal order: audiovisual learning transfers to vision but not audition. PloS One 5, e11283. https://doi.org/10.1371/journal.pone.0011283.

Allman, M.J., Meck, W.H., 2012. Pathophysiological distortions in time perception and timed performance. Brain 135, 656–677. https://doi.org/10.1093/brain/awr210.

Arikan, B.E., van Kemenade, B.M., Straube, B., Harris, L.R., Kircher, T., 2017. Voluntary and involuntary movements widen the window of subjective simultaneity. Iperception 8. https://doi.org/10.1177/2041669517719297.

Bahrick, L.E., 1988. Intermodal learning in infancy: learning on the basis of two kinds of invariant relations in audible and visible events. Child Dev. 59, 197–209.

Bahrick, L.E., Flom, R., Lickliter, R., 2002. Intersensory redundancy facilitates discrimination of tempo in 3-month-old infants. Dev. Psychobiol. 41, 352–363. https://doi.org/10.1002/dev.10049.

Bahrick, L.E., Lickliter, R., 2000. Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. Dev. Psychol. 36, 190–201. https://doi.org/10.1037//0012-1649.36.2.190.

Bahrick, L.E., McNew, M.E., Pruden, S.M., Castellanos, I., 2019. Intersensory redundancy promotes infant detection of prosody in infant-directed speech. J. Exp. Child Psychol. 183, 295–309. https://doi.org/10.1016/j.jecp.2019.02.008.

Barrett, D.J.K., Krumbholz, K., 2012. Evidence for multisensory integration in the elicitation of prior entry by bimodal cues. Exp. Brain Res. 222, 11–20. https://doi.org/10.1007/s00221-012-3191-8.

Basharat, A., Adams, M.S., Staines, W.R., Barnett-Cowan, M., 2018. Simultaneity and temporal order judgments are coded differently and change with age: an event-related potential study. Front. Integr. Neurosci. 12 https://doi.org/10.3389/fnint.2018.00015.

Baum, S.H., Stevenson, R.A., Wallace, M.T., 2015. Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. Prog. Neurobiol. 134, 140–160. https://doi.org/10.1016/j.pneurobio.2015.09.007.

Baumann, O., Vromen, J.M.G., Cheung, A., McFadyen, J., Ren, Y., Guo, C.C., 2018. Neural correlates of temporal complexity and synchrony during audiovisual correspondence detection. eNeuro 5. https://doi.org/10.1523/ENEURO.0294-17.2018.

Bebko, J.M., Weiss, J.A., Demark, J.L., Gomez, P., 2006. Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. J. Child Psychol. Psychiatry Allied Discip. 47, 88–98. https://doi.org/10.1111/j.1469-7610.2005.01443.x.

Bedard, G., Barnett-Cowan, M., 2016. Impaired timing of audiovisual events in the elderly. Exp. Brain Res. 234, 331–340. https://doi.org/10.1007/s00221-015-4466-7.

Beker, S., Foxe, J.J., Molholm, S., 2018. Ripe for solution: delayed development of multisensory processing in autism and its remediation. Neurosci. Biobehav. Rev. 84, 182–192. https://doi.org/10.1016/j.neubiorev.2017.11.008.

Bertelson, P., Aschersleben, G., 2003. Temporal ventriloquism: crossmodal interaction on the time dimension. 1. Evidence from auditory-visual temporal order judgment. Int. J. Psychophysiol. 50, 147–155.

Bhat, J., Miller, L.M., Pitt, M.A., Shahin, A.J., 2015. Putative mechanisms mediating tolerance for audiovisual stimulus onset asynchrony. J. Neurophysiol. 113, 1437–1450. https://doi.org/10.1152/jn.00200.2014.

Biau, E., Moris Fernandez, L., Holle, H., Avila, C., Soto-Faraco, S., 2016. Hand gestures as visual prosody: BOLD responses to audio-visual alignment are modulated by the communicative nature of the stimuli. Neuroimage 132, 129–137. https://doi.org/10.1016/j.neuroimage.2016.02.018.

Bidelman, G.M., 2016. Musicians have enhanced audiovisual multisensory binding: experience-dependent effects in the double-flash illusion. Exp. Brain Res. 234, 3037–3047. https://doi.org/10.1007/s00221-016-4705-6.

Bidelman, G.M., Heath, S.T., 2019a. Enhanced temporal binding of audiovisual information in the bilingual brain. Biling. Lang. Cognit. 22, 752–762. https://doi.org/10.1017/S1366728918000408.

Bidelman, G.M., Heath, S.T., 2019b. Neural correlates of enhanced audiovisual processing in the bilingual brain. Neuroscience 401, 11–20. https://doi.org/10.1016/j.neuroscience.2019.01.003.

Binder, M., 2015. Neural correlates of audiovisual temporal processing-comparison of temporal order and simultaneity judgments. Neuroscience 300, 432–447. https://doi.org/10.1016/j.neuroscience.2015.05.011.

Bishop, C.W., Miller, L.M., 2009. A multisensory cortical network for understanding speech in noise. J. Cognit. Neurosci. 21, 1790–1805. https://doi.org/10.1162/jocn.2009.21118.

Bishop, L., Goebl, W., 2014. Context-specific effects of musical expertise on audiovisual integration. Front. Psychol. 5 https://doi.org/10.3389/fpsyg.2014.01123.

Blomert, L., Froyen, D., 2010. Multi-sensory learning and learning to read. Int. J. Psychophysiol. 77, 195–204. https://doi.org/10.1016/j.ijpsycho.2010.06.025.

Bushara, K.O., Grafman, J., Hallett, M., 2001. Neural correlates of auditory-visual stimulus onset asynchrony detection. J. Neurosci. 21, 300–304.

Cacace, A.T., McFarland, D.J., Ouimet, J.R., Schrieber, E.J., Marro, P., 2000. Temporal processing deficits in remediation-resistant reading-impaired children. Audiol. Neuro. Otol. 5, 83–97. https://doi.org/10.1159/000013871.

Calvert, G.A., Hansen, P.C., Iversen, S.D., Brammer, M.J., 2001. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. Neuroimage 14, 427–438. https://doi.org/10.1006/nimg.2001.0812.

Calvert, G.A., Spence, C., Stein, B.E., 2004. The Handbook of Multisensory Processes. MIT Press.

Cecere, R., Gross, J., Thut, G., 2016. Behavioural evidence for separate mechanisms of audiovisual temporal binding as a function of leading sensory modality. Eur. J. Neurosci. 43, 1561–1568. https://doi.org/10.1111/ejn.13242.

Cecere, R., Gross, J., Willis, A., Thut, G., 2017. Being first matters: topographical representational similarity analysis of ERP signals reveals separate networks for audiovisual temporal binding depending on the leading sense. J. Neurosci. 37, 5274–5287. https://doi.org/10.1523/JNEUROSCI.2926-16.2017.

Cecere, R., Rees, G., Romei, V., 2015. Individual differences in alpha frequency drive crossmodal illusory perception. Curr. Biol. 25, 231–235. https://doi.org/10.1016/j.cub.2014.11.034.

Chan, J.S., Kaiser, J., Brandl, M., Matura, S., Prvulovic, D., Hogan, M.J., Naumer, M.J., 2015. Expanded temporal binding windows in people with mild cognitive impairment. Curr. Alzheimer Res. 12, 61–68.

Chan, J.S., Langer, A., Kaiser, J., 2016. Temporal integration of multisensory stimuli in autism spectrum disorder: a predictive coding perspective. J. Neural. Transm. 123, 917–923. https://doi.org/10.1007/s00702-016-1587-5.

Chan, J.S., Naumer, M.J., 2014. Explaining autism spectrum disorders: central coherence vs. predictive coding theories. J. Neurophysiol. 112, 2669–2671. https://doi.org/10.1152/jn.00242.2014.

Chan, Y.M., Pianta, M.J., Bode, S., McKendrick, A.M., 2017. Spatio-temporal patterns of event-related potentials related to audiovisual synchrony judgments in older adults. Neurobiol. Aging 55, 38–48. https://doi.org/10.1016/j.neurobiolaging.2017.03.011.

Chan, Y.M., Pianta, M.J., McKendrick, A.M., 2014. Older age results in difficulties separating auditory and visual signals in time. J. Vis. 14, 13. https://doi.org/10.1167/14.11.13.

Chen, L., Zhang, M., Ai, F., Xie, W., Meng, X., 2016a. Crossmodal synesthetic congruency improves visual timing in dyslexic children. Res. Dev. Disabil. 55, 14–26. https://doi.org/10.1016/j.ridd.2016.03.010.

Chen, Y.C., Shore, D.I., Lewis, T.L., Maurer, D., 2016b. The development of the perception of audiovisual simultaneity. J. Exp. Child Psychol. 146, 17–33. https://doi.org/10.1016/j.jecp.2016.01.010.

Chen, Y.C., Lewis, T.L., Shore, D.I., 2018. Developmental changes in the perception of visuotactile simultaneity. J. Exp. Chid Psychol. 173, 304–317. https://doi.org/10.1016/j.jecp.2018.04.014.

Chen, Y.C., Spence, C., 2017. Assessing the role of the "unity assumption" on multisensory integration: a review. Front. Psychol. 8, 1–22. 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. Atten. Percept. Psychophys. 78, 1512–1528. https://doi.org/10.3758/s13414-016-1088-5.

Ciaramelli, E., Grady, C.L., Moscovitch, M., 2008. Top-down and bottom-up attention to memory: a hypothesis (AtoM) on the role of the posterior parietal cortex in memory retrieval. Neuropsychologia 46, 1828–1851. https://doi.org/10.1016/j.neuropsychologia.2008.03.022.

Conrey, B., Pisoni, D.B., 2006. Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. J. Acoust. Soc. Am. 119, 4065–4073.

Cook, L.A., Van Valkenburg, D.L., Badcock, D.R., 2011. Predictability affects the perception of audiovisual synchrony in complex sequences. Atten. Percept. Psychophys. 73, 2286–2297. https://doi.org/10.3758/s13414-011-0185-8.

Cooke, J., Poch, C., Gillmeister, H., Costantini, M., Romei, V., 2019. Oscillatory properties of functional connections between sensory areas mediate cross-modal illusory perception. J. Neurosci. 39, 5711–5718. https://doi.org/10.1523/JNEUROSCI.3184-18.2019.

Curtindale, L.M., Bahrick, L.E., Lickliter, R., Colombo, J., 2019. Effects of multimodal synchrony on infant attention and heart rate during events with social and nonsocial stimuli. J. Exp. Child Psychol. 178, 283–294. https://doi.org/10.1016/j.jecp.2018.10.006.

de Boer-Schellekens, L., Eussen, M., Vroomen, J., 2013a. Diminished sensitivity of audiovisual temporal order in autism spectrum disorder. Front. Integr. Neurosci. 7, 8. https://doi.org/10.3389/fnint.2013.00008.

de Boer-Schellekens, L., Keetels, M., Eussen, M., Vroomen, J., 2013b. No evidence for impaired multisensory integration of low-level audiovisual stimuli in adolescents and young adults with autism spectrum disorders. Neuropsychologia 51, 3004–3013. https://doi.org/10.1016/j.neuropsychologia.2013.10.005.

De Boer-Schellekens, L., Stekelenburg, J.J., Maes, J.P., Van Gool, A.R., Vroomen, J., 2014a. Sound improves diminished visual temporal sensitivity in schizophrenia. Acta Psychol. (Amst). 147, 136–142. https://doi.org/10.1016/j.actpsy.2013.06.013.

De Boer-Schellekens, L., Stekelenburg, J.J., Maes, J.P., Van Gool, A.R., Vroomen, J., 2014b. Sound improves diminished visual temporal sensitivity in schizophrenia. Acta Psychol. (Amst). 147, 136–142. https://doi.org/10.1016/j.actpsy.2013.06.013.

De Boer-Schellekens, L., Vroomen, J., 2014. Multisensory integration compensates loss of sensitivity of visual temporal order in the elderly. Exp. Brain Res. 232, 253–262. https://doi.org/10.1007/s00221-013-3736-5.

De Niear, M.A., Gupta, P.B., Baum, S.H., Wallace, M.T., 2018. Perceptual training enhances temporal acuity for multisensory speech. Neurobiol. Learn. Mem. 147, 9–17. https://doi.org/10.1016/j.nlm.2017.10.016.

Dean, C.L., Eggleston, B.A., Gibney, K.D., Aligbe, E., Blackwell, M., Kwakye, L.D., 2017. Auditory and visual distractors disrupt multisensory temporal acuity in the crossmodal temporal order judgment task. PloS One 12, e0179564. https://doi.org/10.1371/journal.pone.0179564.

De La Fuente, A., Xia, S., Branch, C., Li, X., 2013. A review of attention-deficit/hyperactivity disorder from the perspective of brain networks. Front. Hum. Neurosci. 7, 192. https://doi.org/10.3389/fnhum.2013.00192.

Delbeuck, X., Collette, F., Van der Linden, M., 2007. Is Alzheimer's disease a disconnection syndrome?. Evidence from a crossmodal audio-visual illusory experiment. Neuropsychologia 45, 3315–3323. https://doi.org/10.1016/j.neuropsychologia.2007.05.001.

Desantis, A., Haggard, P., 2016a. Action-outcome learning and prediction shape the window of simultaneity of audiovisual outcomes. Cognition 153, 33–42. https://doi.org/10.1016/j.cognition.2016.03.009.

Desantis, A., Haggard, P., 2016b. How actions shape perception: learning action-outcome relations and predicting sensory outcomes promote audio-visual temporal binding. Sci. Rep. 6, 39086. https://doi.org/10.1038/srep39086.

Dhamala, M., Assisi, C.G., Jirsa, V.K., Steinberg, F.L., Kelso, J.A.S., 2007. Multisensory integration for timing engages different brain networks. Neuroimage 34, 764–773. https://doi.org/10.1016/j.neuroimage.2006.07.044.

Diederich, A., Colonius, H., 2015. The time window of multisensory integration: relating reaction times and judgments of temporal order. Psychol. Rev. 122, 232–241. https://doi.org/10.1037/a0038696.

Diederich, A., Colonius, H., Schomburg, A., 2008. Assessing age-related multisensory enhancement with the time-window-of-integration model. Neuropsychologia 46, 2556–2562.

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

Doesburg, S.M., Emberson, L.L., Rahi, A., Cameron, D., Ward, L.M., 2008. Asynchrony from synchrony: long-range gamma-band neural synchrony accompanies perception of audiovisual speech asynchrony. Exp. Brain Res. 185, 11–20. https://doi.org/10.1007/s00221-007-1127-5.

Donohue, S.E., Darling, E.F., Mitroff, S.R., 2012. Links between multisensory processing and autism. Exp. Brain Res. 222, 377–387. https://doi.org/10.1007/s00221-012-3223-4.

Donohue, S.E., Green, J.J., Woldorff, M.G., 2015. The effects of attention on the temporal integration of multisensory stimuli. Front. Integr. Neurosci. 9 https://doi.org/10.3389/fnint.2015.00032.

Eg, R., Behne, D.M., 2015. Perceived synchrony for realistic and dynamic audiovisual events. Front. Psychol. 6.

Erickson, L.C., Heeg, E., Rauschecker, J.P., Turkeltaub, P.E., 2014. An ALE meta-analysis on the audiovisual integration of speech signals. Hum. Brain Mapp. 35, 5587–5605. https://doi.org/10.1002/hbm.22572.

Falck-Ytter, T., Bakker, M., von Hofsten, C., 2011. Human infants orient to biological motion rather than audiovisual synchrony. Neuropsychologia 49, 2131–2135. https://doi.org/10.1016/j.neuropsychologia.2011.03.040.

Falck-Ytter, T., Rehnberg, E., Bolte, S., 2013. Lack of visual orienting to biological motion and audiovisual synchrony in 3-year-olds with autism. PloS One 8, e68816. https://doi.org/10.1371/journal.pone.0068816.

Falck-Ytter, T., Nyström, P., Gredebäck, G., Gliga, T., Bölte, S., 2018. Reduced orienting to audiovisual synchrony in infancy predicts autism diagnosis at 3 years of age. JCPP (J. Child Psychol. Psychiatry) 59, 872–880. https://doi.org/10.1111/jcpp.12863.

Farmer, M.E., Klein, R.M., 1995. The evidence for a temporal processing deficit linked to dyslexia: a review. Psychon. Bull. Rev. 2, 460–493. https://doi.org/10.3758/BF03210983.

Feldman, J.I., Dunham, K., Cassidy, M., Wallace, M.T., Liu, Y., Woynaroski, T.G., 2018. Audiovisual multisensory integration in individuals with autism spectrum disorder: a systematic review and meta-analysis. Neurosci. Biobehav. Rev. 95, 220–234. https://doi.org/10.1016/j.neubiorev.2018.09.020.

Feldman, J.I., Kuang, W., Conrad, J.G., Tu, A., Santapuram, P., Simon, D.M., Foss-Feig, J. H., Kwakye, L.D., Stevenson, R.A., Wallace, M.T., Woynaroski, T.G., 2019. Brief Report: Differences in Multisensory Integration Covary with Sensory Responsiveness in Children with and without Autism Spectrum Disorder. Journal of Autism and Developmental Disorders 49 (1), 397–403. https://doi.org/10.1007/s10803-018-3667-x.

Ferri, F., Nikolova, Y.S., Perrucci, M.G., Costantini, M., Ferretti, A., Gatta, V., Huang, Z., Edden, R.A.E., Yue, Q., D'Aurora, M., Sibille, E., Stuppia, L., Romani, G.L., Northoff, G., 2017. A neural "tuning curve" for multisensory experience and cognitive-perceptual schizotypy. Schizophr. Bull. 43, 801–813. https://doi.org/10.1093/schbul/sbw174.

Ferri, F., Venskus, A., Fotia, F., Cooke, J., Romei, V., 2018. Higher proneness to multisensory illusions is driven by reduced temporal sensitivity in people with high schizotypal traits. Conscious. Cognition 65, 263–270. https://doi.org/10.1016/j.concog.2018.09.006.

Festa, E., Katz, A., Ott, B., Heindel, W., 2013. Perceptual binding of audiovisual speech information is disrupted at the preclinical stage of Alzheimer's disease. Alzheimer's Dementia 9, 452. https://doi.org/10.1016/j.jalz.2013.05.903.

Fiacconi, C.M., Harvey, E.C., Sekuler, A.B., Bennett, P.J., 2013. The influence of aging on audiovisual temporal order judgments. Exp. Aging Res. 39, 179–193. https://doi.org/10.1080/0361073X.2013.761896.

Flom, R., Bahrick, L.E., 2007. The development of infant discrimination of affect in multimodal and unimodal stimulation: the role of intersensory redundancy. Dev. Psychol. 43, 238–252. https://doi.org/10.1037/0012-1649.43.1.238.

Foss-Feig, J.H., Kwakye, L.D., Cascio, C.J., Burnette, C.P., Kadivar, H., Stone, W.L., Wallace, M.T., 2010. An extended multisensory temporal binding window in autism spectrum disorders. Exp. Brain Res. 203, 381–389. https://doi.org/10.1007/s00221-010-2240-4.

Foucher, J.R., Lacambre, M., Pham, B.T., Giersch, A., Elliott, M.A., 2007. Low time resolution in schizophrenia. Lengthened windows of simultaneity for visual, auditory and bimodal stimuli. Schizophr. Res. 97, 118–127. https://doi.org/10.1016/j.schres.2007.08.013.

Franciotti, R., Brancucci, A., Della Penna, S., Onofrj, M., Tommasi, L., 2011. Neuromagnetic responses reveal the cortical timing of audiovisual synchrony. Neuroscience 193, 182–192. https://doi.org/10.1016/j.neuroscience.2011.07.018.

Francisco, A.A., Jesse, A., Groen, M.A., McQueen, J.M., 2017. A general audiovisual temporal processing deficit in adult readers with dyslexia. J. Speech Lang. Hear. Res. 60, 144–158. https://doi.org/10.1044/2016_JSLHR-H-15-0375.

Freeman, E.D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert, C., Leff, A., Driver, J., 2013. Sight and sound out of synch: fragmentation and renormalisation of audiovisual integration and subjective timing. Cortex 49, 2875–2887. https://doi.org/10.1016/j.cortex.2013.03.006.

Fujisaki, W., Shimojo, S., Kashino, M., Nishida, S., 2004. Recalibration of audiovisual simultaneity. Nat. Neurosci. 7, 773–778. https://doi.org/10.1038/nn1268.

Garcia-Perez, M.A., Alcala-Quintana, R., 2012. On the discrepant results in synchrony judgment and temporal-order judgment tasks: a quantitative model. Psychon. Bull. Rev. 19, 820–846. https://doi.org/10.3758/s13423-012-0278-y.

Gerson, S.A., Schiavio, A., Timmers, R., Hunnius, S., 2015. Active drumming experience increases infants' sensitivity to audiovisual synchrony during observed drumming actions. PloS One 10, e0130960. https://doi.org/10.1371/journal.pone.0130960.

Glahn, D.C., Laird, A.R., Ellison-Wright, I., Thelen, S.M., Robinson, J.L., Lancaster, J.L., Bullmore, E., Fox, P.T., 2008. Meta-analysis of gray matter anomalies in schizophrenia: application of anatomic likelihood estimation and network analysis. Biol. Psychiatr. 64, 774–781. https://doi.org/10.1016/j.biopsych.2008.03.031.

Gogate, L.J., Walker-Andrews, A.S., Bahrick, L.E., 2001. The intersensory origins of word comprehension: an ecological-dynamic systems view. Dev. Sci. 4, 1–18. https://doi.org/10.1111/1467-7687.00143.

Grabot, L., Kosem, A., Azizi, L., van Wassenhove, V., 2017. Prestimulus alpha oscillations and the temporal sequencing of audiovisual events. J. Cognit. Neurosci. 29, 1566–1582. https://doi.org/10.1162/jocn_a_01145.

Grossman, R.B., Steinhart, E., Mitchell, T., Mcilvane, W., 2015. "Look who's talking!" gaze patterns for implicit and explicit audio-visual speech synchrony detection in children with high-functioning autism. Autism Res. 8, 307–316. https://doi.org/10.1002/aur.1447.

Habets, B., Bruns, P., Roder, B., 2017. Experience with crossmodal statistics reduces the sensitivity for audio-visual temporal asynchrony. Sci. Rep. 7, 1486. https://doi.org/10.1038/s41598-017-01252-y.

Hahn, N., Foxe, J.J., Molholm, S., 2014. Impairments of multisensory integration and cross-sensory learning as pathways to dyslexia. Neurosci. Biobehav. Rev. 47, 384–392. https://doi.org/10.1016/j.neubiorev.2014.09.007.

Hairston, W.D., Burdette, J.H., Flowers, D.L., Wood, F.B., Wallace, M.T., 2005. Altered temporal profile of visual-auditory multisensory interactions in dyslexia. Exp. Brain Res. 166, 474–480. https://doi.org/10.1007/s00221-005-2387-6.

Hannon, E.E., Schachner, A., Nave-Blodgett, J.E., 2017. Babies know bad dancing when they see it: older but not younger infants discriminate between synchronous and asynchronous audiovisual musical displays. J. Exp. Child Psychol. 159, 159–174. https://doi.org/10.1016/j.jecp.2017.01.006.

Haß, K., Sinke, C., Reese, T., Roy, M., Wiswede, D., Dillo, W., Oranje, B., Szycik, G.R., Hass, K., Sinke, C., Reese, T., Roy, M., Wiswede, D., Dillo, W., Oranje, B., Szycik, G. R., Haß, K., Sinke, C., Reese, T., Roy, M., Wiswede, D., Dillo, W., Oranje, B., Szycik, G.R., 2017. Enlarged temporal integration window in schizophrenia indicated by the double-flash illusion. Cognit. Neuropsychiatry 22, 145–158. https://doi.org/10.1080/13546805.2017.1287693.

Hernandez, B., Setti, A., Kenny, R.A., Newell, F.N., 2019. Individual differences in ageing, cognitive status, and sex on susceptibility to the sound-induced flash illusion: a large-scale study. Psychol. Aging 34, 978–990. https://doi.org/10.1037/pag0000396.

Hillairet de Boisferon, A., Tift, A.H., Minar, N.J., Lewkowicz, D.J., 2017. Selective attention to a talker's mouth in infancy: role of audiovisual temporal synchrony and linguistic experience. Dev. Sci. 20, e12381 https://doi.org/10.1111/desc.12381.

Hillock-Dunn, A., Grantham, D.W., Wallace, M.T., 2016. The temporal binding window for audiovisual speech: children are like little adults. Neuropsychologia 88, 74–82. https://doi.org/10.1016/j.neuropsychologia.2016.02.017.

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

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

Hodges, D.A., Hairston, W.D., Burdette, J.H., 2005. Aspects of multisensory perception: the integration of visual and auditory information in musical experiences. Ann. N. Y. Acad. Sci. 1060, 175–185. https://doi.org/10.1196/annals.1360.012.

Hollich, G., Newman, R.S., Jusczyk, P.W., 2005. Infants' use of synchronized visual information to separate streams of speech. Child Dev. 76, 598–613. https://doi.org/10.1111/j.1467-8624.2005.00866.x.

Ikumi, N., Torralba, M., Ruzzoli, M., Soto-Faraco, S., 2019. The phase of pre-stimulus brain oscillations correlates with cross-modal synchrony perception. Eur. J. Neurosci. 49, 150–164. https://doi.org/10.1111/ejn.14186.

Ipser, A., Karlinski, M., Freeman, E., 2018. Correlation of individual differences in audiovisual asynchrony across stimuli and tasks: new constraints on temporal

renormalization theory. J. Exp. Psychol. Hum. Percept. Perform. 44, 1283–1293. https://doi.org/10.1037/xhp0000535.
Jicol, C., Proulx, M.J., Pollick, F.E., Petrini, K., 2018. Long-term music training modulates the recalibration of audiovisual simultaneity. Exp. Brain Res. 236, 1869–1880. https://doi.org/10.1007/s00221-018-5269-4.
Kaganovich, N., 2017. Sensitivity to audiovisual temporal asynchrony in children with a history of specific language impairment and their peers with typical development: a replication and follow-up study. J. Speech Lang. Hear. Res. 60, 2259–2270. https://doi.org/10.1044/2017_JSLHR-L-16-0327.
Kaganovich, N., 2016. Development of sensitivity to audiovisual temporal asynchrony during midchildhood. Dev. Psychol. 52, 232–241. https://doi.org/10.1037/dev0000073.
Kaganovich, N., Schumaker, J., 2016. Electrophysiological correlates of individual differences in perception of audiovisual temporal asynchrony. Neuropsychologia 86, 119–130. https://doi.org/10.1016/j.neuropsychologia.2016.04.015.
Kaganovich, N., Schumaker, J., Leonard, L.B., Gustafson, D., Macias, D., 2014. Children with a history of SLI show reduced sensitivity to audiovisual temporal asynchrony: an ERP study. J. Speech Lang. Hear. Res. 57, 1480–1502. https://doi.org/10.1044/2014_JSLHR-L-13-0192.
Kambe, J., Kakimoto, Y., Araki, O., 2015. Phase reset affects auditory-visual simultaneity judgment. Cogn. Neurodyn. 9, 487–493. https://doi.org/10.1007/s11571-015-9342-4.
Keil, J., Muller, N., Hartmann, T., Weisz, N., 2014. Prestimulus beta power and phase synchrony influence the sound-induced flash illusion. Cerebr. Cortex 24, 1278–1288. https://doi.org/10.1093/cercor/bhs409.
Kim, J., Wigram, T., Gold, C., 2009. Emotional, motivational and interpersonal responsiveness of children with autism in improvisational music therapy. Autism 13, 389–409. https://doi.org/10.1177/1362361309105660.
Kitagawa, N., Zampini, M., Spence, C., 2005. Audiotactile interactions in near and far space. Exp. Brain Res. 166, 528–537. https://doi.org/10.1007/s00221-005-2393-8.
Kitamura, M.S., Watanabe, K., Kitagawa, N., 2016. Positive emotion facilitates audiovisual binding. Front. Integr. Neurosci. 9 https://doi.org/10.3389/fnint.2015.00066.
Kopp, F., 2014. Audiovisual temporal fusion in 6-month-old infants. Dev. Cogn. Neurosci. 9, 56–67. https://doi.org/10.1016/j.dcn.2014.01.001.
Kopp, F., Dietrich, C., 2013. Neural dynamics of audiovisual synchrony and asynchrony perception in 6-month-old infants. Front. Psychol. 4 https://doi.org/10.3389/fpsyg.2013.00002.
Krueger Fister, J., Stevenson, R.A., Nidiffer, A.R., Barnett, Z.P., Wallace, M.T., 2016. Stimulus intensity modulates multisensory temporal processing. Neuropsychologia 88, 92–100. https://doi.org/10.1016/j.neuropsychologia.2016.02.016.
Kumar, G.V., Halder, T., Jaiswal, A.K., Mukherjee, A., Roy, D., Banerjee, A., 2016. Large scale functional brain networks underlying temporal integration of audio-visual speech perception: an EEG study. Front. Psychol. 7.
Kwakye, L.D., Foss-Feig, J.H., Cascio, C.J., Stone, W.L., Wallace, M.T., 2011. Altered auditory and multisensory temporal processing in autism spectrum disorders. Front. Integr. Neurosci. 4, 129. https://doi.org/10.3389/fnint.2010.00129.
Laasonen, M., Service, E., Virsu, V., 2002. Crossmodal temporal order and processing acuity in developmentally dyslexic young adults. Brain Lang. 80, 340–354. https://doi.org/10.1006/brln.2001.2593.
Laasonen, M., Virsu, V., Oinonen, S., Sandbacka, M., Salakari, A., Service, E., 2012. Phonological and sensory short-term memory are correlates and both affected in developmental dyslexia. Read. Writ. An Interdiscip. J. 25, 2247–2273. https://doi.org/10.1007/s11145-011-9356-1.
LaGasse, B., 2017. Social outcomes in children with autism spectrum disorder: a review of music therapy outcomes. Patient Relat. Outcome Meas. 8, 23–32. https://doi.org/10.2147/prom.s106267.
Lamichhane, B., Adhikari, B.M., Dhamala, M., 2016. Salience network activity in perceptual decisions. Brain Connect. 6, 558–571. https://doi.org/10.1089/brain.2015.0392.
Lamichhane, B., Dhamala, M., 2015. The salience network and its functional architecture in a perceptual decision: an effective connectivity study. Brain Connect. 5, 362–370. https://doi.org/10.1089/brain.2014.0282.
Laurienti, P.J., Burdette, J.H., Maldjian, J.A., Wallace, M.T., 2006. Enhanced multisensory integration in older adults. Neurobiol. Aging 27, 1155–1163. https://doi.org/10.1016/j.neurobiolaging.2005.05.024.
Lee, H., Noppeney, U., 2014. Temporal prediction errors in visual and auditory cortices. Curr. Biol. 24, R309–R310. https://doi.org/10.1016/j.cub.2014.02.007.
Lee, H., Noppeney, U., 2011. Long-term music training tunes how the brain temporally binds signals from multiple senses. Proc. Natl. Acad. Sci. Unit. States Am. 108, E1441–E1450. https://doi.org/10.1073/pnas.1115267108.
Lewald, J., Falkenstein, M., Schwarz, M., 2006. Abnormal auditory-visual crossmodal temporal-order judgments in Parkinson's disease. Cognit. Process. 7 https://doi.org/10.1007/s10339-006-0103-5, 134–134.
Lewald, J., Guski, R., 2004. Auditory-visual temporal integration as a function of distance: no compensation for sound-transmission time in human perception. Neurosci. Lett. 357, 119–122. https://doi.org/10.1016/j.neulet.2003.12.045.
Lewis, R., Noppeney, U., 2010. Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. J. Neurosci. 30, 12329–12339. https://doi.org/10.1523/jneurosci.5745-09.2010.
Lewkowicz, D.J., 2010. Infant perception of audio-visual speech synchrony. Dev. Psychol. 46, 66–77. https://doi.org/10.1037/a0015579.
Lewkowicz, D.J., 2003. Learning and discrimination of audiovisual events in human infants: the hierarchical relation between intersensory temporal synchrony and rhythmic pattern cues. Dev. Psychol. 39, 795–804.
Lewkowicz, D.J., 2000. Infants' perception of the audible, visible, and bimodal attributes of multimodal syllables. Child Dev. 71, 1241–1257.
Lewkowicz, D.J., 1996. Perception of auditory-visual temporal synchrony in human infants. J. Exp. Psychol. Hum. Percept. Perform. 22, 1094–1106.
Lewkowicz, D.J., Flom, R., 2014. The audiovisual temporal binding window narrows in early childhood. Child Dev. 85, 685–694. https://doi.org/10.1111/cdev.12142.
Lewkowicz, D.J., Hansen-Tift, A.M., 2012. Infants deploy selective attention to the mouth of a talking face when learning speech. Proc. Natl. Acad. Sci. Unit. States Am. 109, 1431–1436. https://doi.org/10.1073/pnas.1114783109.
Lewkowicz, D.J., Leo, I., Simion, F., 2010. Intersensory perception at birth: newborns match nonhuman primate faces and voices. Infancy 15, 46–60. https://doi.org/10.1111/j.1532-7078.2009.00005.x.
Lewkowicz, D.J., Minar, N.J., Tift, A.H., Brandon, M., 2015. Perception of the multisensory coherence of fluent audiovisual speech in infancy: its emergence and the role of experience. J. Exp. Child Psychol. 130, 147–162. https://doi.org/10.1016/j.jecp.2014.10.006.
Li, Q., Liu, P., Huang, S., Huang, X., 2018. The influence of phasic alerting on multisensory temporal precision. Exp. Brain Res. 236, 3279–3296. https://doi.org/10.1007/s00221-018-5372-6.
Liu, B., Jin, Z., Wang, Z., Gong, C., 2011. The influence of temporal asynchrony on multisensory integration in the processing of asynchronous audio-visual stimuli of real-world events: an event-related potential study. Neuroscience 176, 254–264. https://doi.org/10.1016/j.neuroscience.2010.12.028.
Love, S.A., Petrini, K., Cheng, A., Pollick, F.E., 2013. A psychophysical investigation of differences between synchrony and temporal order judgments. PloS One 8, e54798. https://doi.org/10.1371/journal.pone.0054798.
Love, S.A., Petrini, K., Pernet, C.R., Latinus, M., Pollick, F.E., 2018. Overlapping but divergent neural correlates underpinning audiovisual synchrony and temporal order judgments. Front. Hum. Neurosci. 12 https://doi.org/10.3389/fnhum.2018.00274.
Lu, Y., Paraskevopoulos, E., Herholz, S.C., Kuchenbuch, A., Pantev, C., 2014. Temporal processing of audiovisual stimuli is enhanced in musicians: evidence from magnetoencephalography (MEG). PloS One 9, e90686. https://doi.org/10.1371/journal.pone.0090686.
Luu, S., Wong, W., Noordin, H., 2008. Low-level audiovisual synchrony: experiments and model. Jpn. Psychol. Res. 50, 214–222. https://doi.org/10.1111/j.1468-5884.2008.00377.x.
Macaluso, E., George, N., Dolan, R., Spence, C., Driver, J., 2004. Spatial and temporal factors during processing of audiovisual speech: a PET study. Neuroimage 21, 725–732. https://doi.org/10.1016/j.neuroimage.2003.09.049.
Magnotti, J.F., Ma, W.J., Beauchamp, M.S., 2013. Causal inference of asynchronous audiovisual speech. Front. Psychol. 4, 798. https://doi.org/10.3389/fpsyg.2013.00798.
Maier, J.X., Di Luca, M., Noppeney, U., 2011. Audiovisual asynchrony detection in human speech. J. Exp. Psychol. Hum. Percept. Perform. 37, 245–256. https://doi.org/10.1037/a0019952.
Manning, F.C., Schutz, M., 2013. "Moving to the beat" improves timing perception. Psychon. Bull. Rev. 20, 1133–1139. https://doi.org/10.3758/s13423-013-0439-7.
Manning, F.C., Schutz, M., 2015. Trained to keep a beat: movement-related enhancements to timing perception in percussionists and non-percussionists. Psychol. Res. 80, 532–542. https://doi.org/10.1007/s00426-015-0678-5.
Marchant, J.L., Ruff, C.C., Driver, J., 2012. Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. Hum. Brain Mapp. 33, 1212–1224. https://doi.org/10.1002/hbm.21278.
Margiotoudi, K., Kelly, S., Vatakis, A., 2014. Audiovisual temporal integration of speech and gesture. Procedia - Soc. Behav. Sci. 126, 154–155. https://doi.org/10.1016/j.sbspro.2014.02.351.
Martin, B., Giersch, A., Huron, C., van Wassenhove, V., 2013. Temporal event structure and timing in schizophrenia: preserved binding in a longer "now". Neuropsychologia 51, 358–371. https://doi.org/10.1016/j.neuropsychologia.2012.07.002.
McGovern, D.P., Roudaia, E., Newell, F.N., Roach, N.W., 2016. Perceptual learning shapes multisensory causal inference via two distinct mechanisms. Sci. Rep. 6, 24673. https://doi.org/10.1038/srep24673.
McGovern, D.P., Roudaia, E., Stapleton, J., McGinnity, T.M., Newell, F.N., 2014. The sound-induced flash illusion reveals dissociable age-related effects in multisensory integration. Front. Aging Neurosci. 6 https://doi.org/10.3389/fnagi.2014.00250.
Meredith, M.A., Nemitz, J.W., Stein, B.E., 1987. Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. J. Neurosci. 7, 3215–3229. https://doi.org/10.1097/00005072-199808000-00008.
Meredith, M.A., Stein, B.E., 1986a. Spatial factors determine the activity of multisensory neurons in cat superior colliculus. Brain Res. 365, 350–354. https://doi.org/10.1016/0006-8993(86)91648-3.
Meredith, M.A., Stein, B.E., 1986b. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. J. Neurophysiol. 56, 640–662. https://doi.org/10.1152/jn.1986.56.3.640.
Miendlarzewska, E.A., Trost, W.J., 2014. How musical training affects cognitive development: rhythm, reward and other modulating variables. Front. Neurosci. 7, 279. https://doi.org/10.3389/fnins.2013.00279.
Miller, J., 1986. Timecourse of coactivation in bimodal divided attention. Percept. Psychophys 40, 331–343. https://doi.org/10.3758/BF03203025.
Miller, J., 1982. Divided attention: evidence for coactivation with redundant signals. Cognit. Psychol. 14, 247–279. https://doi.org/10.1016/0010-0285(82)90010-X.
Miller, L.M., D'Esposito, M., 2005. Perceptual fusion and stimulus coincidence in the cross-modal integration of speech. J. Neurosci. 25, 5884–5893. https://doi.org/10.1523/JNEUROSCI.0896-05.2005.

Moro, S.S., Steeves, J.K.E., 2018. Normal temporal binding window but no sound-induced flash illusion in people with one eye. Exp. Brain Res. 236, 1825–1834. https://doi.org/10.1007/s00221-018-5263-x.

Morrongiello, B.A., Fenwick, K.D., Nutley, T., 1998. Developmental changes in associations between auditory-visual events. Infant Behav. Dev. 21, 613–626. https://doi.org/10.1016/S0163-6383(98)90033-9.

Murray, M.M., Lewkowicz, D.J., Amedi, A., Wallace, M.T., 2016. Multisensory processes: A balancing act across the lifespan. Trend. Neurosci. 39 (8), 567–579. https://doi.org/10.1016/j.tins.2016.05.003.

Musacchia, G., Schroeder, C.E., 2009. Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. Hear. Res. 258, 72–79. https://doi.org/10.1016/j.heares.2009.06.018.

Narinesingh, C., Goltz, H.C., Wong, A.M.F., 2017. Temporal binding window of the sound-induced flash illusion in amblyopia. Invest. Ophthalmol. Vis. Sci. 58, 1442–1448. https://doi.org/10.1167/iovs.16-21258.

Navarra, J., Alsius, A., Velasco, I., Soto-Faraco, S., Spence, C., 2010. Perception of audiovisual speech synchrony for native and non-native language. Brain Res. 1323, 84–93. https://doi.org/10.1016/j.brainres.2010.01.059.

Niear, M.A., Koo, B., Wallace, M.T., De Niear, M.A., Koo, B., Wallace, M.T., 2016. Multisensory perceptual learning is dependent upon task difficulty. Exp. Brain Res. 234, 3269–3277. https://doi.org/10.1007/s00221-016-4724-3.

Noel, J.-P., De Niear, M., Lazzzara, N.S., Wallace, M.T., 2018a. Uncoupling between multisensory temporal function and nonverbal turn-taking in autism spectrum disorder. IEEE T. Cogn. Dev. Syst. 10, 973–982. https://doi.org/10.1109/TCDS.2017.2778141.

Noel, J.-P., De Niear, M., Van der Burg, E., Wallace, M.T., 2016a. Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan. PloS One 11, e0161698. https://doi.org/10.1371/journal.pone.0161698.

Noel, J.-P., Kurela, L., Baum, S.H., Yu, H., Neimat, J.S., Gallagher, M.J., Wallace, M., 2017a. Multisensory temporal function and EEG complexity in patients with epilepsy and psychogenic nonepileptic events. Epilepsy Behav. 70, 166–172. https://doi.org/10.1016/j.yebeh.2017.02.018.

Noel, J.-P., Łukowska, M., Wallace, M., Serino, A., 2016b. Multisensory simultaneity judgment and proximity to the body. J. Vis. 16.

Noel, J., Lytle, M., Cascio, C., Wallace, M.T., 2018b. Disrupted integration of exteroceptive and interoceptive signaling in autism spectrum disorder. Autism Res. 11, 194–205. https://doi.org/10.1002/aur.1880.

Noel, J.P., De Niear, M.A., Stevenson, R., Alais, D., Wallace, M.T., 2017b. Atypical rapid audio-visual temporal recalibration in autism spectrum disorders. Autism Res. 10, 121–129. https://doi.org/10.1002/aur.1633.

Noel, J.P., Modi, K., Wallace, M.T., Van der Stoep, N., 2018c. Audiovisual integration in depth: multisensory binding and gain as a function of distance. Exp. Brain Res. 236, 1939–1951. https://doi.org/10.1007/s00221-018-5274-7.

Noel, J.P., Stevenson, R.A., Wallace, M.T., 2018d. Atypical audiovisual temporal function in autism and schizophrenia: similar phenotype, different cause. Eur. J. Neurosci. 47, 1230–1241. https://doi.org/10.1111/ejn.13911.

Noesselt, T., Bergmann, D., Heinze, H.J., Münte, T., Spence, C., 2012. Coding of multisensory temporal patterns in human superior temporal sulcus. Front. Integr. Neurosci. 6 https://doi.org/10.3389/fnint.2012.00064.

Noesselt, T., Fendrich, R., Bonath, B., Tyll, S., Heinze, H.-J., 2005. Closer in time when farther in space-spatial factors in audiovisual temporal integration. Brain Res. Cogn. Brain Res. 25, 443–458. https://doi.org/10.1016/j.cogbrainres.2005.07.005.

Noesselt, T., Rieger, J.W., Schoenfeld, M.A., Kanowski, M., Hinrichs, H., Heinze, H.-J., Driver, J., 2007. Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. J. Neurosci. 27, 11431–11441. https://doi.org/10.1523/JNEUROSCI.2252-07.2007.

Noppeney, U., Lee, H.L., 2018. Causal inference and temporal predictions in audiovisual perception of speech and music. Ann. N. Y. Acad. Sci. 1423, 102–116. https://doi.org/10.1111/nyas.13615.

Nozaradan, S., Peretz, I., Mouraux, A., 2012. Steady-state evoked potentials as an index of multisensory temporal binding. Neuroimage 60, 21–28. https://doi.org/10.1016/j.neuroimage.2011.11.065.

Panagiotidi, M., Overton, P.G., Stafford, T., 2017. Multisensory integration and ADHD-like traits: evidence for an abnormal temporal integration window in ADHD. Acta Psychol. (Amst) 181, 10–17. https://doi.org/10.1016/j.actpsy.2017.10.001.

Parise, C., Ernst, M., 2016. Correlation detection as a general mechanism for multisensory integration. Nat. Commun. 7, 11543. https://doi.org/10.1038/ncomms11543.

Patten, E., Labban, J.D., Casenhiser, D.M., Cotton, C.L., 2016. Synchrony detection of linguistic stimuli in the presence of faces: neuropsychological implications for language development in ASD. Dev. Neuropsychol. 41, 362–374. https://doi.org/10.1080/87565641.2016.1243113.

Patten, E., Watson, L.R., Baranek, G.T., 2014. Temporal synchrony detection and associations with language in young children with ASD. Autism Res. Treat. 2014 1–8. https://doi.org/10.1155/2014/678346.

Peiffer, A.M., Mozolic, J.L., Hugenschmidt, C.E., Laurienti, P.J., 2007. Age-related multisensory enhancement in a simple audiovisual detection task. Neuroreport 18, 1077–1081. https://doi.org/10.1097/WNR.0b013e3281e72ae7.

Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C.H., Avanzini, F., Puce, A., Pollick, F.E., 2009a. Multisensory integration of drumming actions: musical expertise affects perceived audiovisual asynchrony. Exp. Brain Res. 198, 339–352. https://doi.org/10.1007/s00221-009-1817-2.

Petrini, K., Holt, S.P., Pollick, F., 2010. Expertise with multisensory events eliminates the effect of biological motion rotation on audiovisual synchrony perception. J. Vis. 10, 2. https://doi.org/10.1167/10.5.2.

Petrini, K., Pollick, F.E., Dahl, S., McAleer, P., McKay, L., Rocchesso, D., Waadeland, C. H., Love, S., Avanzini, F., Puce, A., 2011. Action expertise reduces brain activity for audiovisual matching actions: an fMRI study with expert drummers. Neuroimage 56, 1480–1492. https://doi.org/10.1016/j.neuroimage.2011.03.009.

Petrini, K., Russell, M., Pollick, F., 2009b. When knowing can replace seeing in audiovisual integration of actions. Cognition 110, 432–439. https://doi.org/10.1016/j.cognition.2008.11.015.

Petter, E.A., Lusk, N.A., Hesslow, G., Meck, W.H., 2016. Interactive roles of the cerebellum and striatum in sub-second and supra-second timing: support for an initiation, continuation, adjustment, and termination (ICAT) model of temporal processing. Neurosci. Biobehav. Rev. 71, 739–755. https://doi.org/10.1016/j.neubiorev.2016.10.015.

Pilling, M., 2009. Auditory event-related potentials (ERPs) in audiovisual speech perception. J. Speech Lang. Hear. Res. 52, 1073–1081. https://doi.org/10.1044/1092-4388(2009/07-0276.

Pons, F., Andreu, L., Sanz-Torrent, M., Buil-Legaz, L., Lewkowicz, D.J., 2013. Perception of audio-visual speech synchrony in Spanish-speaking children with and without specific language impairment. J. Child Lang. 40, 687–700. https://doi.org/10.1017/S0305000912000189.

Pons, F., Lewkowicz, D.J., 2014. Infant perception of audio-visual speech synchrony in familiar and unfamiliar fluent speech. Acta Psychol. (Amst) 149, 142–147. https://doi.org/10.1016/j.actpsy.2013.12.013.

Poole, D., Gowen, E., Warren, P.A., Poliakoff, E., 2017. Brief report: which came first? Exploring crossmodal temporal order judgements and their relationship with sensory reactivity in autism and neurotypicals. J. Autism Dev. Disord. 47, 215–223. https://doi.org/10.1007/s10803-016-2925-z.

Powers, A.R., Hevey, M.A., Wallace, M.T., 2012. Neural correlates of multisensory perceptual learning. J. Neurosci. 32, 6263–6274. https://doi.org/10.1523/JNEUROSCI.6138-11.2012.

Powers, A.R., Hillock, A.R., Wallace, M.T., 2009. Perceptual training narrows the temporal window of multisensory binding. J. Neurosci. 29, 12265–12274. https://doi.org/10.1523/JNEUROSCI.3501-09.2009.

Powers, A.R., Hillock-Dunn, A., Wallace, M.T., 2016. Generalization of multisensory perceptual learning. Sci. Rep. 6, 23374. https://doi.org/10.1038/srep23374.

Richards, M.D., Goltz, H.C., Wong, A.M.F., 2018. Temporal ventriloquism reveals intact audiovisual temporal integration in amblyopia. Invest. Ophthalmol. Vis. Sci. 59, 637–643. https://doi.org/10.1167/iovs.17-22613.

Richards, M.D., Goltz, H.C., Wong, A.M.F., 2017. Alterations in audiovisual simultaneity perception in amblyopia. PloS One 12, e0179516. https://doi.org/10.1371/journal.pone.0179516.

Righi, G., Tenenbaum, E.J., McCormick, C., Blossom, M., Amso, D., Sheinkopf, S.J., 2018. Sensitivity to audio-visual synchrony and its relation to language abilities in children with and without ASD. Autism Res. 11, 645–653. https://doi.org/10.1002/aur.1918.

Robertson, D.M.C., Schweinberger, S.R., 2010. The role of audiovisual asynchrony in person recognition. Q. J. Exp. Psychol. 63, 23–30. https://doi.org/10.1080/17470210903144376.

Romanski, L.M., Hwang, J., 2012. Timing of audiovisual inputs to the prefrontal cortex and multisensory integration. Neuroscience 214, 36–48. https://doi.org/10.1016/j.neuroscience.2012.03.025.

Sanfratello, L., Aine, C., Stephen, J., 2018. Neuroimaging investigations of dorsal stream processing and effects of stimulus synchrony in schizophrenia. Psychiatry Res. Neuroimaging. 278, 56–64. https://doi.org/10.1016/j.pscychresns.2018.05.005.

Scarpina, F., Migliorati, D., Marzullo, P., Mauro, A., Scacchi, M., Costantini, M., 2016. Altered multisensory temporal integration in obesity. Sci. Rep. 6, 28382. https://doi.org/10.1038/srep28382.

Schlesinger, J.J., Stevenson, R.A., Shotwell, M.S., Wallace, M.T., 2014. Improving pulse oximetry pitch perception with multisensory perceptual training. Anesth. Analg. 118, 1249–1253. https://doi.org/10.1213/ANE.0000000000000222.

Schwartz, J.-L., Savariaux, C., 2014. No, there is no 150 ms lead of visual speech on auditory speech, but a range of audiovisual asynchronies varying from small audio lead to large audio lag. PLoS Comput. Biol. 10, e1003743 https://doi.org/10.1371/journal.pcbi.1003743.

Senkowski, D., Schneider, T.R., Foxe, J.J., Engel, A.K., 2008. Crossmodal binding through neural coherence: implications for multisensory processing. Trends Neurosci. 31, 401–409. https://doi.org/10.1016/j.tins.2008.05.002.

Senkowski, D., Talsma, D., Grigutsch, M., Herrmann, C.S., Woldorff, M.G., 2007. Good times for multisensory integration: effects of the precision of temporal synchrony as revealed by gamma-band oscillations. Neuropsychologia 45, 561–571. https://doi.org/10.1016/j.neuropsychologia.2006.01.013.

Setti, A., Burke, K.E., Kenny, R.A., Newell, F.N., 2011a. Is inefficient multisensory processing associated with falls in older people? Exp. Brain Res. 209, 375–384. https://doi.org/10.1007/s00221-011-2560-z.

Setti, A., Finnigan, S., Sobolewski, R., McLaren, L., Robertson, I.H., Reilly, R.B., Anne Kenny, R., Newell, F.N., 2011b. Audiovisual temporal discrimination is less efficient with aging: an event-related potential study. Neuroreport 22, 554–558. https://doi.org/10.1097/WNR.0b013e328348c731.

Setti, A., Stapleton, J., Leahy, D., Walsh, C., Kenny, R.A., Newell, F.N., 2014. Improving the efficiency of multisensory integration in older adults: audio-visual temporal discrimination training reduces susceptibility to the sound-induced flash illusion. Neuropsychologia 61, 259–268. https://doi.org/10.1016/j.neuropsychologia.2014.06.027.

Shahin, A.J., Shen, S., Kerlin, J.R., 2017. Tolerance for audiovisual asynchrony is enhanced by the spectrotemporal fidelity of the speaker's mouth movements and speech. Lang. Cogn. Neurosci. 32, 1102–1118. https://doi.org/10.1080/23273798.2017.1283428.

Sharda, M., Tuerk, C., Chowdhury, R., Jamey, K., Foster, N., Custo-Blanch, M., Tan, M., Nadig, A., Hyde, K., 2018. Music improves social communication and auditory-motor connectivity in children with autism. Transl. Psychiatry 8, 231. https://doi.org/10.1038/s41398-018-0287-3.

Silva, C.C., Mendonça, C., Mouta, S., Silva, R., Campos, J.C., Santos, J., 2013. Depth cues and perceived audiovisual synchrony of biological motion. PloS One 8, e80096. https://doi.org/10.1371/journal.pone.0080096.

Simon, D.M., Wallace, M.T., 2018. Integration and temporal processing of asynchronous audiovisual speech. J. Cognit. Neurosci. 30, 319–337. https://doi.org/10.1162/jocn_a_01205.

Soto-Faraco, S., Alsius, A., 2009. Deconstructing the McGurk–MacDonald illusion. J. Exp. Psychol. Hum. Percept. Perform. 35, 580–587. https://doi.org/10.1037/a0013483.

Smith, E., Zhang, S., Bennetto, L., 2017. Temporal synchrony and audiovisual integration of speech and object stimuli in autism. Res. Autism Spectr. Disord. 39, 11–19. https://doi.org/10.1016/j.rasd.2017.04.001.

Stevenson, R.A., Altieri, N.A., Kim, S., Pisoni, D.B., James, T.W., 2010. Neural processing of asynchronous audiovisual speech perception. Neuroimage 49, 3308–3318. https://doi.org/10.1016/j.neuroimage.2009.12.001.

Stevenson, R.A., Baum, S.H., Krueger, J., Newhouse, P.A., Wallace, M.T., 2018a. Links between temporal acuity and multisensory integration across life span. J. Exp. Psychol. Hum. Percept. Perform. 44, 106–116. https://doi.org/10.1037/xhp0000424.

Stevenson, R.A., Fister, J.K., Barnett, Z.P., Nidiffer, A.R., Wallace, M.T., 2012a. Interactions between the spatial and temporal stimulus factors that influence multisensory integration in human performance. Exp. Brain Res. 219, 121–137. https://doi.org/10.1007/s00221-012-3072-1.

Stevenson, R.A., Park, S., Cochran, C., McIntosh, L.G., Noel, J.P., Barense, M.D., Ferber, S., Wallace, M.T., 2017a. The associations between multisensory temporal processing and symptoms of schizophrenia. Schizophr. Res. 179, 97–103. https://doi.org/10.1016/j.schres.2016.09.035.

Stevenson, R.A., Segers, M., Ferber, S., Barense, M.D., Camarata, S., Wallace, M.T., 2016. Keeping time in the brain: autism spectrum disorder and audiovisual temporal processing. Autism Res. 9, 720–738. https://doi.org/10.1002/aur.1566.

Stevenson, R.A., Segers, M., Ncube, B.L., Black, K.R., Bebko, J.M., Ferber, S., Barense, M. D., 2018b. The cascading influence of multisensory processing on speech perception in autism. Autism 22, 609–624. https://doi.org/10.1177/1362361317704413.

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

Stevenson, R.A., Toulmin, J.K., Youm, A., Besney, R.M.A., Schulz, S.E., Barense, M.D., Ferber, S., 2017b. Increases in the autistic trait of attention to detail are associated with decreased multisensory temporal adaptation. Sci. Rep. 7, 1–10. https://doi.org/10.1038/s41598-017-14632-1.

Stevenson, R.A., VanDerKlok, R.M., Pisoni, D.B., James, T.W., 2011. Discrete neural substrates underlie complementary audiovisual speech integration processes. Neuroimage 55, 1339–1345. https://doi.org/10.1016/j.neuroimage.2010.12.063.

Stevenson, R.A., Wallace, M.T., 2013. Multisensory temporal integration: task and stimulus dependencies. Exp. Brain Res. 227, 249–261. https://doi.org/10.1007/s00221-013-3507-3.

Stevenson, R.A., Wallace, M.T., Altieri, N., 2014b. The interaction between stimulus factors and cognitive factors during multisensory integration of audiovisual speech. Front. Psychol. 5, 352. https://doi.org/10.3389/fpsyg.2014.00352.

Stevenson, R.A., Wilson, M.M., Powers, A.R., Wallace, M.T., 2013. The effects of visual training on multisensory temporal processing. Exp. Brain Res. 225, 479–489. https://doi.org/10.1007/s00221-012-3387-y.

Stevenson, R.A., Zemtsov, R.K., Wallace, M.T., 2012b. Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. J. Exp. Psychol. Hum. Percept. Perform. 38, 1517–1529. https://doi.org/10.1037/a0027339.

ten Oever, S., Sack, A.T., Wheat, K.L., Bien, N., van Atteveldt, N., 2013. Audio-visual onset differences are used to determine syllable identity for ambiguous audio-visual stimulus pairs. Front. Psychol. 4, 331. https://doi.org/10.3389/fpsyg.2013.00331.

Theves, S., Chan, J.S., Naumer, M.J., Kaiser, J., 2019. Improving audio-visual temporal perception through training enhances beta-band activity. NeuroImage. https://doi.org/10.1016/j.neuroimage.2019.116312.

Tiippana, K., Salmela, V.R., 2018. Stimulus duration has little effect on auditory, visual and audiovisual temporal order judgement. Exp. Brain Res. 236, 1273–1282. https://doi.org/10.1007/s00221-018-5218-2.

Tsilionis, E., Vatakis, A., 2016. Multisensory binding: is the contribution of synchrony and semantic congruency obligatory? Curr. Opin. Behav. Sci. 8, 7–13. https://doi.org/10.1016/j.cobeha.2016.01.002.

Turi, M., Karaminis, T., Pellicano, E., Burr, D., 2016. No rapid audiovisual recalibration in adults on the autism spectrum. Sci. Rep. 6, 21756. https://doi.org/10.1038/srep21756.

van Atteveldt, N.M., Formisano, E., Blomert, L., Goebel, R., 2007a. The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. Cerebr. Cortex 17, 962–974. https://doi.org/10.1093/cercor/bhl007.

van Atteveldt, N.M., Formisano, E., Goebel, R., Blomert, L., 2007b. Top-down task effects overrule automatic multisensory responses to letter-sound pairs in auditory association cortex. Neuroimage 36, 1345–1360. https://doi.org/10.1016/j.neuroimage.2007.03.065.

Van der Burg, E., Goodbourn, P.T., 2015. Rapid, generalized adaptation to asynchronous audiovisual speech. Proc. Biol. Sci. 282, 20143083 https://doi.org/10.1098/rspb.2014.3083.

van der Burg, E., Olivers, C.N., Bronkhorst, A.W., Theeuwes, J., 2008. Pip and Pop: nonspatial auditory signals improve spatial visual search. J. Exp. Psychol. Hum. Percept. Perform. 34, 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053.

van Eijk, R.L., Kohlrausch, A., Juola, J.F., van de Par, S., 2008. Audiovisual synchrony and temporal order judgments: effects of experimental method and stimulus type. Percept. Psychophys 70, 955–968. https://doi.org/10.3758/pp.70.6.955.

van Laarhoven, T., Stekelenburg, J.J., Vroomen, J., 2019. Increased sub-clinical levels of autistic traits are associated with reduced multisensory integration of audiovisual speech. Sci. Rep. 9 (1), 9535 https://doi.org/10.1038/s41598-019-46084-0.

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

Vatakis, A., Bayliss, L., Zampini, M., Spence, C., 2007. The influence of synchronous audiovisual distractors on audiovisual temporal order judgments. Percept. Psychophys 69, 298–309. https://doi.org/10.3758/BF03193751.

Vatakis, A., Ghazanfar, A.A., Spence, C., 2008a. Facilitation of multisensory integration by the "unity effect" reveals that speech is special. J. Vis. 8, 14.1–14.11. https://doi.org/10.1167/8.9.14.

Vatakis, A., Navarra, J., Soto-Faraco, S., Spence, C., 2008b. Audiovisual temporal adaptation of speech: temporal order versus simultaneity judgments. Exp. Brain Res. 185, 521–529. https://doi.org/10.1007/s00221-007-1168-9.

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

Vatakis, A., Spence, C., 2007. Crossmodal binding: evaluating the "unity assumption" using audiovisual speech stimuli. Percept. Psychophys. 69, 744–756. https://doi.org/10.3758/BF03193776.

Vatakis, A., Spence, C., 2006a. Temporal order judgments for audiovisual targets embedded in unimodal and bimodal distractor streams. Neurosci. Lett. 408, 5–9. https://doi.org/10.1016/j.neulet.2006.06.017.

Vatakis, A., Spence, C., 2006b. Audiovisual synchrony perception for speech and music assessed using a temporal order judgment task. Neurosci. Lett. 393, 40–44. https://doi.org/10.1016/j.neulet.2005.09.032.

Vatakis, A., Spence, C., 2006c. Audiovisual synchrony perception for music, speech, and object actions. Brain Res. 1111, 134–142. https://doi.org/10.1016/j.brainres.2006.05.078.

Vissers, M.E., Cohen, M.X., Geurts, H.M., 2012. Brain connectivity and high functioning autism: a promising path of research that needs refined models, methodological convergence, and stronger behavioral links. Neurosci. Biobehav. Rev. 36, 604–625. https://doi.org/10.1016/j.neubiorev.2011.09.003.

Virsu, V., Lahti-Nuuttila, P., Laasonen, M., 2003. Crossmodal temporal processing acuity impairment aggravates with age in developmental dyslexia. Neurosci. Lett. 336, 151–154. https://doi.org/10.1016/S0304-3940(02)01253-3.

Virsu, V., Oksanen-Hennah, H., Vedenpaa, A., Jaatinen, P., Lahti-Nuuttila, P., 2008. Simultaneity learning in vision, audition, tactile sense and their cross-modal combinations. Exp. Brain Res. 186, 525–537. https://doi.org/10.1007/s00221-007-1254-z.

Vroomen, J., Keetels, M., 2010. Perception of intersensory synchrony: a tutorial review. Atten. Percept. Psychophys. 72, 871–884. https://doi.org/10.3758/APP.72.4.871.

Vroomen, J., Stekelenburg, J.J., 2011. Perception of intersensory synchrony in audiovisual speech: not that special. Cognition 118, 75–83. https://doi.org/10.1016/j.cognition.2010.10.002.

Walker-Andrews, A.S., 1997. Infants' perception of expressive behaviors: differentiation of multimodal information. Psychol. Bull. 121, 437–456. https://doi.org/10.1037/0033-2909.121.3.437.

Wallace, M.T., Carriere, B.N., Perrault, T.J.J., Vaughan, J.W., Stein, B.E., 2006. The development of cortical multisensory integration. J. Neurosci. 26, 11844–11849. https://doi.org/10.1523/JNEUROSCI.3295-06.2006.

Wallace, M.T., Stein, B.E., 1997. Development of multisensory neurons and multisensory integration in cat superior colliculus. J. Neurosci. 17, 2429–2444.

Wallace, M.T., Stevenson, R.A., 2014. The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. Neuropsychologia 64, 105–123. https://doi.org/10.1016/j.neuropsychologia.2014.08.005.

Wallace, M.T., Woynaroski, T.G., Stevenson, R.A., 2020. Multisensory integration as a window into orderly and disrupted cognition and communication. Annu. Rev. Psychol. 71, 15.1–15.27. https://doi.org/10.1146/annurev-psych-010419-051112.

Welch, R.B., Warren, D.H., 1980. Immediate perceptual response to intersensory discrepancy. Psychol. Bull. 88, 638–667.

Wiggins, I.M., Hartley, D.E.H., 2015. A synchrony-dependent influence of sounds on activity in visual cortex measured using functional near-infrared spectroscopy (fNIRS). PloS One 10, e0122862. https://doi.org/10.1371/journal.pone.0122862.

Wise, A., Barnett-Cowan, M., 2018. Perceived simultaneity and temporal order of audiovisual events following concussion. Front. Hum. Neurosci. 12 https://doi.org/10.3389/fnhum.2018.00139.

Woynaroski, T.G., Kwakye, L.D., Foss-Feig, J.H., Stevenson, R.A., Stone, W.L., Wallace, M.T., 2013. Multisensory speech perception in children with autism spectrum disorders. J. Autism Dev. Disord. 43, 2891–2902. https://doi.org/10.1007/s10803-013-1836-5.

Yarrow, K., Martin, S.E., Di Costa, S., Solomon, J.A., Arnold, D.H., 2016. A roving dual-presentation simultaneity-judgment task to estimate the point of subjective simultaneity. Front. Psychol. 7.

Yuan, X., Li, H., Liu, P., Yuan, H., Huang, X., 2016. Pre-stimulus beta and gamma oscillatory power predicts perceived audiovisual simultaneity. Int. J. Psychophysiol. 107, 29–36. https://doi.org/10.1016/j.ijpsycho.2016.06.017.

Zampini, M., Shore, D.I., Spence, C., 2005. Audiovisual prior entry. Neurosci. Lett. 381, 217–222. https://doi.org/10.1016/j.neulet.2005.01.085.

Zampini, M., Shore, D.I., Spence, C., 2003a. Audiovisual temporal order judgments. Exp. Brain Res. 152, 198–210. https://doi.org/10.1007/s00221-003-1536-z.

Zampini, M., Shore, D.I., Spence, C., 2003b. Multisensory temporal order judgments: the role of hemispheric redundancy. Int. J. Psychophysiol. 50, 165–180.

Zerr, M., Freihorst, C., Schütz, H., Sinke, C., Müller, A., Bleich, S., Münte, T.F., Szycik, G. R., 2019. Brief sensory training narrows the temporal binding window and enhances long-term multimodal speech perception. Front. Psychol. 10, 2489. https://doi.org/10.3389/fpsyg.2019.02489.

Zhou, H., Cai, X., Weigl, M., Bang, P., Cheung, E.F.C., Chan, R.C.K., 2018. Multisensory temporal binding window in autism spectrum disorders and schizophrenia spectrum disorders: a systematic review and meta-analysis. Neurosci. Biobehav. Rev. 86, 66–76. https://doi.org/10.1016/j.neubiorev.2017.12.013.

Zmigrod, L., Zmigrod, S., 2016. On the temporal precision of thought: individual differences in the multisensory temporal binding window predict performance on verbal and nonverbal problem solving tasks. Multisensory Res. 29, 679–701. https://doi.org/10.1163/22134808-00002532.

Zmigrod, S., Hommel, B., 2011. The relationship between feature binding and consciousness: evidence from asynchronous multi-modal stimuli. Conscious. Cogn. An Int. J. 20, 586–593. https://doi.org/10.1016/j.concog.2011.01.011.

Zmigrod, S., Zmigrod, L., 2015. Zapping the gap: reducing the multisensory temporal binding window by means of transcranial direct current stimulation (tDCS). Conscious. Cognit. 35, 143–149. https://doi.org/10.1016/j.concog.2015.05.012.

Zvyagintsev, M., Parisi, C., Mathiak, K., 2017. Temporal processing deficit leads to impaired multisensory binding in schizophrenia. Cognit. Neuropsychiatry 22, 361–372. https://doi.org/10.1080/13546805.2017.1331160.