Check for
updates

The Ventriloquist Illusion as a Tool to Study Multisensory Processing: An Update

Patrick Bruns*

Biological Psychology and Neuropsychology, University of Hamburg, Hamburg, Germany

Ventriloquism, the illusion that a voice appears to come from the moving mouth of a puppet rather than from the actual speaker, is one of the classic examples of multisensory processing. In the laboratory, this illusion can be reliably induced by presenting simple meaningless audiovisual stimuli with a spatial discrepancy between the auditory and visual components. Typically, the perceived location of the sound source is biased toward the location of the visual stimulus (the ventriloquism effect). The strength of the visual bias reflects the relative reliability of the visual and auditory inputs as well as prior expectations that the two stimuli originated from the same source. In addition to the ventriloquist illusion, exposure to spatially discrepant audiovisual stimuli results in a subsequent recalibration of unisensory auditory localization (the ventriloquism aftereffect). In the past years, the ventriloquism effect and aftereffect have seen a resurgence as an experimental tool to elucidate basic mechanisms of multisensory integration and learning. For example, recent studies have: (a) revealed top-down influences from the reward and motor systems on cross-modal binding; (b) dissociated recalibration processes operating at different time scales; and (c) identified brain networks involved in the neuronal computations underlying multisensory integration and learning. This mini review article provides a brief overview of established experimental paradigms to measure the ventriloquism effect and aftereffect before summarizing these pathbreaking new advancements. Finally, it is pointed out how the ventriloquism effect and aftereffect could be utilized to address some of the current open questions in the field of multisensory research.

Keywords: cross-modal, multisensory, recalibration, space, ventriloquism

INTRODUCTION

Ventriloquism, literally meaning to speak with the stomach, has a long cultural history that dates back to the ancient Greeks (Connor, 2000). Modern-day ventriloquists entertain their audiences by exploiting the illusion that their voice, produced without overt lip movements, is perceived to originate from the moving lips of a puppet. This visual capture

of the perceived auditory location has become one of the most frequently studied examples of multisensory processing in the scientific literature (Stratton, 1897; Klemm, 1909; Thomas, 1941; Jackson, 1953; Thurlow and Jack, 1973; Bertelson and Radeau, 1981; Bertelson and Aschersleben, 1998; Alais and Burr, 2004).

In a typical experimental procedure, participants are presented with a synchronous but spatially discrepant audiovisual stimulus. When asked to localize the sound source, participants usually perceive the auditory stimulus closer to the visual stimulus than it actually is (Bertelson and Radeau, 1981). Although this effect is often tested with simple meaningless stimuli such as tones and light flashes, it has become widely known as the *ventriloquism effect* (Howard and Templeton, 1966). The strength of the ventriloquism effect depends on the relative reliability of the auditory and visual stimuli (Alais and Burr, 2004) as well as on the prior (or expectation) that the two stimuli originated from the same event (Van Wanrooij et al., 2010). This flexible multisensory integration seen at the behavioral level is well-described by Bayesian causal inference models in which the spatial estimates obtained under the assumption of a common vs. separate causes are combined (Körding et al., 2007; Rohe and Noppeney, 2015b). Recent findings suggest that human observers tend to put overly high emphasis on the visual cue in this process (Arnold et al., 2019; Meijer et al., 2019). In addition to the immediate visual influence on auditory localization seen in the ventriloquism effect, exposure to audiovisual stimuli with a consistent audiovisual spatial disparity results in a subsequent recalibration of unisensory auditory

spatial perception known as the *ventriloquism aftereffect* (Canon, 1970; Radeau and Bertelson, 1974; Recanzone, 1998). The aftereffect represents an instance of cross-modal learning that can be dissociated from multisensory integration seen in the ventriloquism effect (Bruns et al., 2011a; Zaidel et al., 2011).

The ventriloquism effect and aftereffect are both highly reliable effects that have been replicated in dozens of studies (see **Table 1**). Both effects are not specific for audiovisual processing but have been demonstrated for audio-tactile and visuo-tactile stimulus pairings as well (Pick et al., 1969; Caclin et al., 2002; Bruns and Röder, 2010; Bruns et al., 2011b; Samad and Shams, 2016, 2018). This robustness and versatility make them ideal experimental paradigms to study basic mechanisms of multisensory integration and learning. The extensive literature on the ventriloquism effect and aftereffect has been summarized in several excellent reviews (Bertelson and de Gelder, 2004; Woods and Recanzone, 2004; Recanzone, 2009; Chen and Vroomen, 2013). However, since the last comprehensive review by Chen and Vroomen (2013), several new lines of research have emerged that have helped clarifying the role of the reward and motor systems in cross-modal binding, the time scales involved in recalibration, and the neural mechanisms underlying multisensory integration and learning. The aim of the present review article is to provide an update on these exciting recent developments which are summarized in **Table 1**. In addition, the following section describes some of the standard procedures to measure the ventriloquism effect and aftereffect to encourage more researchers to utilize these

TABLE 1 | Key studies on the ventriloquism effect and aftereffect published since 2013.

Study	Main finding
Arnold et al. (2019) and Meijer et al. (2019)	Visual bias in VE is stronger than predicted by maximum likelihood integration
Bruns et al. (2014)	Monetary reward for accurate sound localization reduces the VE
Zierul et al. (2019)	Reduced VE for self-initiated audiovisual stimuli
Zaidel et al. (2013)	Feedback results in yoked recalibration of both cues in the same direction
Pages and Groh (2013)	VAE depends on visual feedback rather than on audiovisual synchrony
Berger and Ehrsson (2013) and Berger and Ehrsson (2018)	Imagined visual stimuli induce a VE and VAE
Delong et al. (2018)	Subliminal visual stimuli induce a (reduced) VE
Bruns and Röder (2015)	Immediate and cumulative VAE are dissociable processes
Bosen et al. (2017)	VAE accumulates with repetitions and decays over time
Bosen et al. (2018)	VAE consists of both a large and transient initial localization shift, as well as a smaller and more enduring shift
Mendonça et al. (2015)	Last audiovisual trial affects subsequent VAE the most
Watson et al. (2019)	VAE involves distinct recalibration mechanisms operating at different time scales
Bruns and Röder (2019)	Repeated training sessions enhance the VAE over days
Callan et al. (2015)	VE is associated with modulation of activity in space-sensitive auditory cortex
Bonath et al. (2014)	Separate but adjacent auditory regions code VE to synchronous and asynchronous stimuli
Rohe and Noppeney (2015a) and Rohe and Noppeney (2016)	Multisensory integration and causal inference are performed in parietal regions
Aller and Noppeney (2019)	Causal inference in the brain is accomplished by a dynamic encoding of multiple spatial estimates
Park and Kayser (2019)	VE and immediate VAE have a common neural substrate in parietal cortex
Cuppini et al. (2017)	Biologically inspired neural network model explains behavioral VE
Zierul et al. (2017)	VAE results in persistent adjustments of spatial representations in auditory cortex
Bruns and Röder (2017)	VAE depends on the sensory context
Odegaard et al. (2017)	Cross-modal binding (i.e., VE) increases after exposure to synchronous but spatially unrelated stimuli
Odegaard and Shams (2016)	Cross-modal binding (i.e., VE) is stable over time in adulthood

VE, ventriloquism effect; VAE, ventriloquism aftereffect.

effects in their quest to tackle the remaining open questions in multisensory research.

MEASURING THE VENTRILOQUISM EFFECT AND AFTEREFFECT

The ventriloquism effect and aftereffect have been reliably obtained with a large variety of different localization tasks. These tasks can be categorized into absolute (or continuous) localization measures and relative (or dichotomous) localization measures. In absolute localization tasks, participants directly localize the stimuli with a hand pointer (Lewald, 2002; Bruns and Röder, 2015, 2017, 2019) or by performing a finger (Frissen et al., 2003, 2005, 2012), head (Recanzone, 1998; Van Wanrooij et al., 2010), or eye movement (Kopco et al., 2009; Pages and Groh, 2013) toward the perceived stimulus location. Some studies have used categorical responses (e.g., left, center, or right) instead (Bonath et al., 2007, 2014; Bruns and Röder, 2010; Bruns et al., 2011a; Rohe and Noppeney, 2015a, 2016; Zierul et al., 2017). While categorical responses are less sensitive than continuous measures, they are preferable in studies involving electrophysiological or neuroimaging recordings to reduce motor noise. An alternative are relative localization tasks, in which stimulus location is judged relative to central fixation (i.e., left vs. right) or relative to a reference stimulus in a two-alternative forced choice (2AFC) manner (Bertelson and Aschersleben, 1998; Recanzone, 1998; Bruns et al., 2011b; Berger and Ehrsson, 2018). Some authors have also advocated two-interval forced choice (2IFC) procedures because they are less susceptible to response strategies (Alais and Burr, 2004; Vroomen and Stekelenburg, 2014).

The study design differs slightly depending on whether the ventriloquism effect or the ventriloquism aftereffect (or both) are to be measured (see **Figure 1**). To measure the ventriloquism effect, it is critical that different degrees and directions of cross-modal spatial disparity are presented in a random order to avoid cumulative recalibration effects during the test block (Bertelson and Radeau, 1981; Bertelson and de Gelder, 2004). In addition, baseline localization can be assessed in unimodal trials, either intermixed with the bimodal trials or in a separate pretest block. Aside from the size of the localization bias in the bimodal trials, the ventriloquism effect has been conceptualized as the percentage of trials in which participants perceive the (spatially disparate) cross-modal stimuli as originating from a common cause or the same location (Chen and Spence, 2017). Localization bias and perception of unity are usually correlated (Hairston et al., 2003; Wallace et al., 2004) but measure different aspects of cross-modal integration (Bertelson and Radeau, 1981; Bosen et al., 2016; Chen and Spence, 2017).

When assessing the ventriloquism aftereffect, a distinction needs to be made between immediate and cumulative recalibration effects (Bruns and Röder, 2015). In a study design in which unimodal trials are intermixed with bimodal trials (see **Figure 1B**), Wozny and Shams (2011) showed that localization responses in unimodal trials are systematically influenced by the cross-modal spatial disparity in the directly preceding bimodal trial, indicating an immediate or trial-by-trial recalibration effect. By contrast, the cumulative ventriloquism aftereffect requires exposure to a consistent cross-modal disparity (e.g., visual stimuli always 10° to the right of auditory stimuli). Typically, unisensory sound localization is measured before and after the exposure block (see **Figure 1C**), and the cumulative aftereffect is revealed by a shift in unisensory localization from pre- to post-test (Recanzone, 1998; Lewald, 2002; Frissen et al., 2003; Bruns and Röder, 2017).
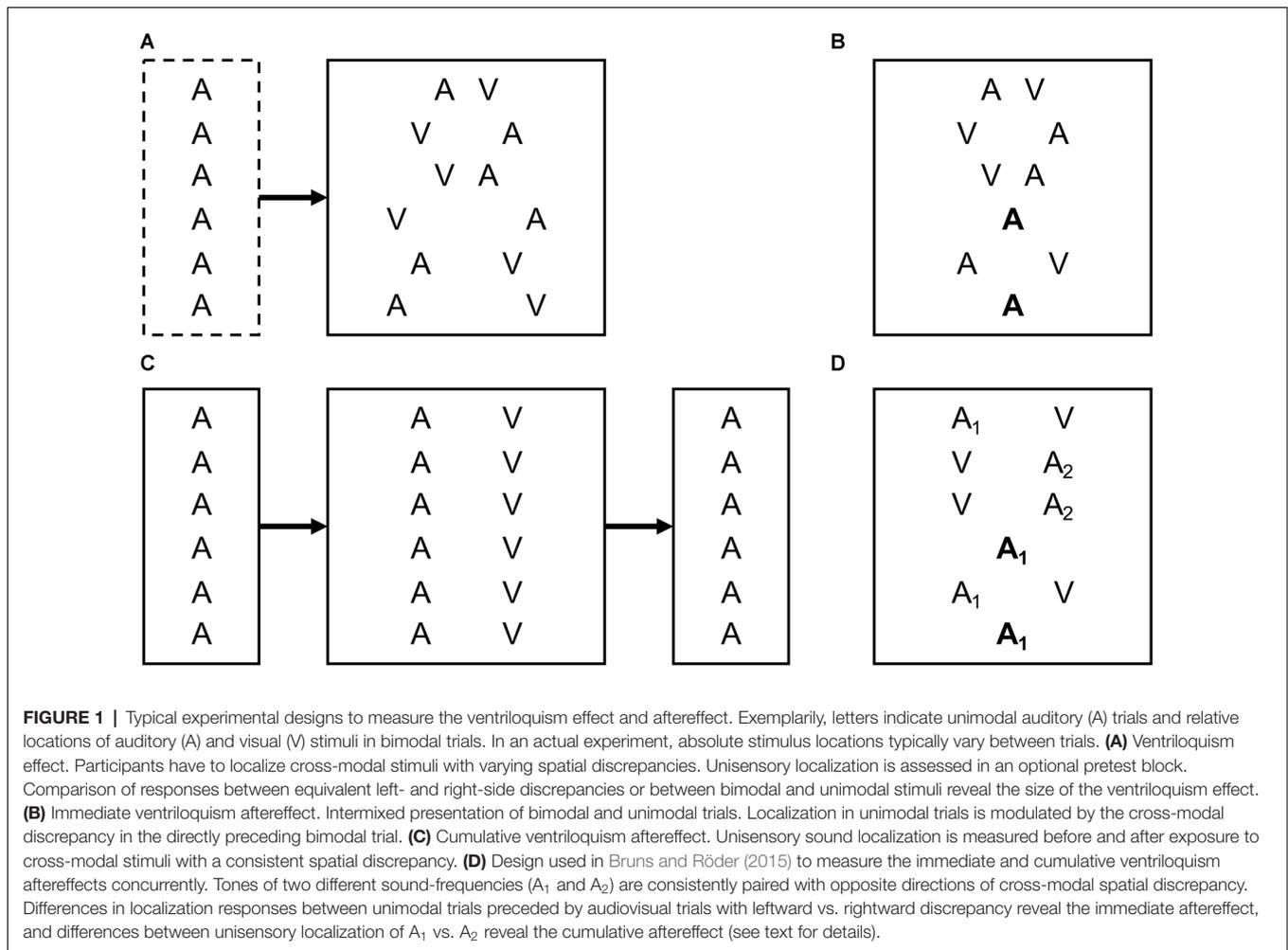
Bruns and Röder (2015) recently introduced a procedure that allows assessing both immediate and cumulative aftereffects (as well as ventriloquism effects) at the same time (see **Figure 1D**). In this paradigm, auditory-only and audiovisual trials were intermixed. Crucially, tones of two different sound frequencies were used that were paired with opposite directions of audiovisual disparity (leftward vs. rightward). Sound localization responses in auditory-only trials (averaged across tone frequencies) were modulated by the direction of audiovisual disparity in the directly preceding audiovisual trial, indicating an immediate aftereffect. Additionally, sound localization responses differed between the two tone-frequencies, indicating a frequency-specific cumulative aftereffect induced by the consistent pairing of tone-frequency and direction of audiovisual disparity (but see Frissen et al., 2003, 2005; Bruns and Röder, 2017; for a discussion of the sound frequency specificity of the cumulative aftereffect).

RECENT FINDINGS

Top-Down Influences on Cross-Modal Binding and Learning

A long-standing debate in multisensory research is the extent to which multisensory processing is influenced by top-down factors (Röder and Büchel, 2009; Talsma et al., 2010). Contrary to earlier findings suggesting that the ventriloquism effect and aftereffect reflect largely automatic processes (Bertelson et al., 2000; Vroomen et al., 2001; Passamonti et al., 2009; Odegaard et al., 2016), several recent lines of evidence have identified top-down influences on the ventriloquism effect and aftereffect.

In a study by Bruns et al. (2014), participants could earn either a high or a low monetary reward for accurate sound localization performance, which put their motivational goal of maximizing the reward in conflict with the auditory spatial bias induced by the ventriloquism effect. As compared to stimuli associated with a low reward, the ventriloquism effect was significantly reduced for high reward stimuli. A similar reduction of the ventriloquism effect was observed when emotionally salient auditory stimuli (fearful voices) were presented prior to the audiovisual test phase (Maiworm et al., 2012). In both cases, the experimental manipulations did not affect unisensory auditory localization performance, suggesting that top-down influences from the emotion and reward systems specifically reduced cross-modal binding. A similar pattern of

FIGURE 1 | Typical experimental designs to measure the ventriloquism effect and aftereffect. Exemplarily, letters indicate unimodal auditory (A) trials and relative locations of auditory (A) and visual (V) stimuli in bimodal trials. In an actual experiment, absolute stimulus locations typically vary between trials. **(A)** Ventriloquism effect. Participants have to localize cross-modal stimuli with varying spatial discrepancies. Unisensory localization is assessed in an optional pretest block. Comparison of responses between equivalent left- and right-side discrepancies or between bimodal and unimodal stimuli reveal the size of the ventriloquism effect. **(B)** Immediate ventriloquism aftereffect. Intermixed presentation of bimodal and unimodal trials. Localization in unimodal trials is modulated by the cross-modal discrepancy in the directly preceding bimodal trial. **(C)** Cumulative ventriloquism aftereffect. Unisensory sound localization is measured before and after exposure to cross-modal stimuli with a consistent spatial discrepancy. **(D)** Design used in Bruns and Röder (2015) to measure the immediate and cumulative ventriloquism aftereffects concurrently. Tones of two different sound-frequencies (A_1 and A_2) are consistently paired with opposite directions of cross-modal spatial discrepancy. Differences in localization responses between unimodal trials preceded by audiovisual trials with leftward vs. rightward discrepancy reveal the immediate aftereffect, and differences between unisensory localization of A_1 vs. A_2 reveal the cumulative aftereffect (see text for details).

results was observed in a recent study in which participants either actively initiated audiovisual stimulus presentations with a button press or were passively exposed to the same stimuli. Contrary to the intuitive assumption that self-initiation would increase the prior expectation that auditory and visual stimuli had a common cause, a reduction of the size of the ventriloquism effect was observed for self-initiated stimuli, possibly due to an increased sensitivity to cross-modal spatial discrepancies in the self-initiation condition (Zierul et al., 2019).

A second line of research investigated the effects of feedback information about the stimulus location on cross-modal recalibration. In a visuo-vestibular version of the ventriloquism aftereffect, participants received a reward for correct localization responses which was contingent either on the visual or on the vestibular cue. This manipulation resulted in a yoked recalibration of both cues in the same direction (Zaidel et al., 2013), whereas passive exposure without feedback shifted both cues independently toward each other (Zaidel et al., 2011). The importance of feedback information was substantiated in the classic audiovisual ventriloquism aftereffect. Here, asynchronous stimuli in which

the visual stimulus lagged the auditory stimulus and, thus, provided feedback about the auditory location were more effective in inducing an aftereffect than synchronous stimuli in which the visual stimulus was extinguished too quickly to provide feedback (Pages and Groh, 2013). Thus, feedback, which presumably exerts top-down influences on perception, might be an important but previously overlooked driver of cross-modal recalibration.

Finally, in a third line of research, Berger and Ehrsson (2013, 2014, 2018) showed that imagining a visual stimulus at a location discrepant to an auditory stimulus had the same effect on auditory localization as actually seeing a visual stimulus at that location. Both imagery-induced ventriloquism effects (Berger and Ehrsson, 2013, 2014) and aftereffects (Berger and Ehrsson, 2018) were obtained. Explicit mental images were, thus, integrated with auditory sensory input in a similar manner as actual visual input, providing strong evidence for top-down influences on multisensory processing. A somewhat opposite approach was taken by Delong et al. (2018), who used continuous flash suppression to render an actual visual stimulus invisible. They obtained a significant ventriloquism effect with the invisible stimuli, which was, however, reduced in size compared to visible

stimuli. Taken together, these results show that the ventriloquism effect is influenced by both bottom-up and top-down processes.

Time Scales of Cross-Modal Recalibration

Cross-modal recalibration in the ventriloquism aftereffect has been described at two different time scales. Initial studies measured shifts in sound localization after exposure to several hundred audiovisual trials with a consistent spatial disparity (Radeau and Bertelson, 1974; Recanzone, 1998; Lewald, 2002), implicitly assuming that recalibration requires accumulated evidence of cross-modal mismatch. This assumption was challenged by findings demonstrating immediate effects on sound localization after a single audiovisual exposure stimulus (Wozny and Shams, 2011). Several recent studies have addressed the theoretically important question of how immediate and cumulative cross-modal recalibration are related.

A consistent finding is that the size of the ventriloquism aftereffect increases if several audiovisual exposure trials with a consistent spatial disparity precede the auditory test trials (Wozny and Shams, 2011; Bruns and Röder, 2015; Bosen et al., 2017, 2018), until the aftereffect reaches a maximum after about 180 exposure trials (Frissen et al., 2012). The last audiovisual stimulus, however, seems to have a particularly strong influence on subsequent sound localization (Mendonça et al., 2015). Theoretically, the immediate and cumulative portions of the ventriloquism aftereffect could be explained by the same underlying mechanism, a strong but rapidly decaying immediate aftereffect with a long tail that allows for accumulation across trials (Bosen et al., 2018). However, recent experimental evidence suggests dissociable mechanisms underlying immediate and cumulative recalibration (Bruns and Röder, 2015; Watson et al., 2019).

A controversial point is the longevity of the (cumulative) ventriloquism aftereffect after cessation of cross-modal discrepancy training. While some studies observed a rapid decay of the aftereffect if there was a delay between audiovisual exposure and auditory localization posttest (Bosen et al., 2017, 2018), others have found no significant decline of the aftereffect (Frissen et al., 2012). However, it was assumed that the aftereffect would last at most until new (spatially coincident) audiovisual evidence is encountered, as would naturally occur after leaving the experimental situation (Recanzone, 1998). Contrary to this assumption, a recent study showed that repeated exposure to audiovisual stimuli with a consistent spatial disparity enhanced the ventriloquism aftereffect over the course of several days, that is, aftereffects were still present after 24 h and accumulated with additional audiovisual discrepancy training (Bruns and Röder, 2019). This finding raises the possibility that cross-modal recalibration effects are context-specific (e.g., for the laboratory situation), making them more stable than previously thought.

Neural Mechanisms Underlying Cross-Modal Binding and Learning

Neuroimaging studies have shown that the ventriloquism effect is associated with a modulation of activity in space-sensitive regions of the planum temporale in auditory cortex (Bonath et al., 2007, 2014; Callan et al., 2015; Zierul et al., 2017). Behaviorally,

the ventriloquism effect is reduced if audiovisual stimuli are presented asynchronously (Slutsky and Recanzone, 2001; Wallace et al., 2004). Interestingly, Bonath et al. (2014) showed that separate (but adjacent) regions of the planum temporale coded ventriloquist illusions to synchronous and asynchronous audiovisual stimuli, which might suggest an involvement of different multisensory temporal integration windows.

Adjustments of auditory spatial processing in the ventriloquism effect have been linked to feedback influences on auditory cortex activity (Bonath et al., 2007; Bruns and Röder, 2010). Recent EEG and functional magnetic resonance imaging (fMRI) evidence has indeed implicated multisensory association areas of the intraparietal sulcus in the generation of the ventriloquism effect. While primary sensory areas initially encoded the unisensory location estimates, posterior intraparietal sulcus activity reflected the integrated estimate which depends on the relative reliabilities of the auditory and visual estimates (Rohe and Noppeney, 2015a). The brain needs to weigh the unisensory estimate against the integrated estimate due to the inherent uncertainty about the true causal structure (Körding et al., 2007), and this weighing was reflected in anterior intraparietal sulcus activity emerging from 200 ms poststimulus onwards (Rohe and Noppeney, 2015a; Aller and Noppeney, 2019). Parietal representations were found to mediate both multisensory integration and the immediate recalibration of unisensory perception in the subsequent auditory trial (Park and Kayser, 2019). In a re-analysis of their data, Rohe and Noppeney (2016) further showed that parietal areas take into account top-down task relevance (i.e., which modality had to be reported), which might suggest a neural basis for other top-down influences discussed in the subsection "Top-Down Influences on Cross-Modal Binding and Learning." EEG and MEG studies have revealed a crucial role of neural oscillations in orchestrating the interplay between stimulus-driven and top-down effects in multisensory processing (Senkowski et al., 2008; Keil and Senkowski, 2018). Based on the available evidence, neural network models of the ventriloquism effect have been developed (Magosso et al., 2012; Cuppini et al., 2017).

While the neural computations underlying multisensory spatial integration and immediate recalibration might critically depend on parietal areas, cross-modal recalibration in the cumulative ventriloquism aftereffect was found to result in an enduring change of spatial representations in the planum temporale and an increase of connectivity between the planum temporale and parietal areas (Zierul et al., 2017). This suggests that sustained changes in unisensory sound localization reflect altered bottom-up processing along the auditory "where" pathway (Bruns et al., 2011a).

FUTURE DIRECTIONS

The ventriloquism effect and aftereffect have generated an abundance of new insights into the mechanisms of multisensory processing in recent years. Future challenges include translating these new findings into a more general theoretical framework

of multisensory processing in naturalistic environments as well as clarifying the developmental trajectory of multisensory spatial integration and learning.

In real-world scenarios, cross-modal stimuli are usually accompanied by a myriad of other continuously changing stimuli. This sensory context inevitably modulates how a particular stimulus is processed (Bruns and Röder, 2017; Bruns and Watanabe, 2019) and shapes priors for processing that stimulus during future encounters (Habets et al., 2017; Odegaard et al., 2017). In addition, the sensory evidence itself might be corrupted by varying amounts of noise. Interestingly, in a phenomenon referred to as cross-modal stochastic resonance, it has been found that intermediate levels of noise in one sensory modality can enhance (rather than impair) responses to weak stimuli in another sensory modality (Manjarrez et al., 2007; Mendez-Balbuena et al., 2018). Future studies should address how learned priors and sensory context interact with bottom-up sensory evidence in the brain. To address these questions, emerging technologies like augmented and virtual reality might help bringing the ventriloquism effect and aftereffect paradigm closer to more complex real-world scenarios (Sarlat et al., 2006; Kytö et al., 2015).

Multisensory spatial processing appears relatively stable over time during adulthood (Odegaard and Shams, 2016), but surprisingly few studies have tested its ontogenetic development in humans. Non-human animal studies have typically investigated visual calibration of auditory spatial representations over rather long time scales of weeks to months (King, 2009), but the developmental trajectory of short-term recalibration effects (as observed in the ventriloquism aftereffect) and its relation to optimal cross-modal integration (as measured in the ventriloquism effect) remains unknown. To assess developmental influences on multisensory spatial functions, retrospective studies in which the impact of sensory deprivation during sensitive periods of development (e.g., due to blindness) is tested in adult individuals are needed as well (Occelli et al., 2012).

With their long history, the ventriloquism effect and aftereffect are timeless experimental paradigms and invaluable tools for the field of multisensory research. Hopefully, this review article will stimulate further discoveries in the years to come.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

REFERENCES

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262. doi: 10.1016/j.cub.2004.01.029

Aller, M., and Noppeney, U. (2019). To integrate or not to integrate: temporal dynamics of hierarchical Bayesian causal inference. *PLoS Biol.* 17:e3000210. doi: 10.1371/journal.pbio.3000210

Arnold, D. H., Petrie, K., Murray, C., and Johnston, A. (2019). Suboptimal human multisensory cue combination. *Sci. Rep.* 9:5155. doi: 10.1038/s41598-018-37888-7

Berger, C. C., and Ehrsson, H. H. (2013). Mental imagery changes multisensory perception. *Curr. Biol.* 23, 1367–1372. doi: 10.1016/j.cub.2013.06.012

Berger, C. C., and Ehrsson, H. H. (2014). The fusion of mental imagery and sensation in the temporal association cortex. *J. Neurosci.* 34, 13684–13692. doi: 10.1523/JNEUROSCI.0943-14.2014

Berger, C. C., and Ehrsson, H. H. (2018). Mental imagery induces cross-modal sensory plasticity and changes future auditory perception. *Psychol. Sci.* 29, 926–935. doi: 10.1177/0956797617748959

Bertelson, P., and Aschersleben, G. (1998). Automatic visual bias of perceived auditory location. *Psychon. Bull. Rev.* 5, 482–489. doi: 10.3758/bf03208826

Bertelson, P., and de Gelder, B. (2004). "The psychology of multimodal perception," in *Crossmodal Space and Crossmodal Attention*, eds C. Spence and J. Driver (Oxford, England: Oxford University Press), 141–177.

Bertelson, P., and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept. Psychophys.* 29, 578–584. doi: 10.3758/bf03207374

Bertelson, P., Vroomen, J., de Gelder, B., and Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Percept. Psychophys.* 62, 321–332. doi: 10.3758/bf03205552

Bonath, B., Noesselt, T., Krauel, K., Tyll, S., Tempelmann, C., and Hillyard, S. A. (2014). Audio-visual synchrony modulates the ventriloquist illusion and its neural/spatial representation in the auditory cortex. *Neuroimage* 98, 425–434. doi: 10.1016/j.neuroimage.2014.04.077

Bonath, B., Noesselt, T., Martinez, A., Mishra, J., Schwiecker, K., Heinze, H.-J., et al. (2007). Neural basis of the ventriloquist illusion. *Curr. Biol.* 17, 1697–1703. doi: 10.1016/j.cub.2007.08.050

Bosen, A. K., Fleming, J. T., Allen, P. D., O'Neill, W. E., and Paige, G. D. (2017). Accumulation and decay of visual capture and the ventriloquism aftereffect caused by brief audio-visual disparities. *Exp. Brain Res.* 235, 585–595. doi: 10.1007/s00221-016-4820-4

Bosen, A. K., Fleming, J. T., Allen, P. D., O'Neill, W. E., and Paige, G. D. (2018). Multiple time scales of the ventriloquism aftereffect. *PLoS One* 13:e0200930. doi: 10.1371/journal.pone.0200930

Bosen, A. K., Fleming, J. T., Brown, S. E., Allen, P. D., O'Neill, W. E., and Paige, G. D. (2016). Comparison of congruence judgment and auditory localization tasks for assessing the spatial limits of visual capture. *Biol. Cybern.* 110, 455–471. doi: 10.1007/s00422-016-0706-6

Bruns, P., Liebnau, R., and Röder, B. (2011a). Cross-modal training induces changes in spatial representations early in the auditory processing pathway. *Psychol. Sci.* 22, 1120–1126. doi: 10.1177/0956797611416254

Bruns, P., Spence, C., and Röder, B. (2011b). Tactile recalibration of auditory spatial representations. *Exp. Brain Res.* 209, 333–344. doi: 10.1007/s00221-011-2543-0

Bruns, P., Maiworm, M., and Röder, B. (2014). Reward expectation influences audiovisual spatial integration. *Atten. Percept. Psychophys.* 76, 1815–1827. doi: 10.3758/s13414-014-0699-y

Bruns, P., and Röder, B. (2010). Tactile capture of auditory localization: an event-related potential study. *Eur. J. Neurosci.* 31, 1844–1857. doi: 10.1111/j.1460-9568.2010.07232.x

Bruns, P., and Röder, B. (2015). Sensory recalibration integrates information from the immediate and the cumulative past. *Sci. Rep.* 5:12739. doi: 10.1038/srep12739

Bruns, P., and Röder, B. (2017). Spatial and frequency specificity of the ventriloquism aftereffect revisited. *Psychol. Res.* doi: 10.1007/s00426-017-0965-4 [Epub ahead of print].

Bruns, P., and Röder, B. (2019). Repeated but not incremental training enhances cross-modal recalibration. *J. Exp. Psychol. Hum. Percept. Perform.* 45, 435–440. doi: 10.1037/xhp0000642

Bruns, P., and Watanabe, T. (2019). Perceptual learning of task-irrelevant features depends on the sensory context. *Sci. Rep.* 9:1666. doi: 10.1038/s41598-019-38586-8

Caclin, A., Soto-Faraco, S., Kingstone, A., and Spence, C. (2002). Tactile "capture" of audition. *Percept. Psychophys.* 64, 616–630. doi: 10.3758/bf03194730

Callan, A., Callan, D., and Ando, H. (2015). An fMRI study of the ventriloquism effect. *Cereb. Cortex* 25, 4248–4258. doi: 10.1093/cercor/bhu306

Canon, L. K. (1970). Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *J. Exp. Psychol.* 84, 141–147. doi: 10.1037/h0028925

Chen, Y.-C., and Spence, C. (2017). Assessing the role of the 'unity assumption' on multisensory integration: a review. *Front. Psychol.* 8:445. doi: 10.3389/fpsyg.2017.00445

Chen, L., and Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Atten. Percept. Psychophys.* 75, 790–811. doi: 10.3758/s13414-013-0475-4

Connor, S. (2000). *Dumbstruck: A Cultural History of Ventriloquism.* Oxford, England: Oxford University Press.

Cuppini, C., Shams, L., Magosso, E., and Ursino, M. (2017). A biologically inspired neurocomputational model for audiovisual integration and causal inference. *Eur. J. Neurosci.* 46, 2481–2498. doi: 10.1111/ejn.13725

Delong, P., Aller, M., Giani, A. S., Rohe, T., Conrad, V., Watanabe, M., et al. (2018). Invisible flashes alter perceived sound location. *Sci. Rep.* 8:12376. doi: 10.1038/s41598-018-30773-3

Frissen, I., Vroomen, J., and de Gelder, B. (2012). The aftereffects of ventriloquism: the time course of the visual recalibration of auditory localization. *Seeing Perceiving* 25, 1–14. doi: 10.1163/187847611x620883

Frissen, I., Vroomen, J., de Gelder, B., and Bertelson, P. (2003). The aftereffects of ventriloquism: are they sound-frequency specific? *Acta Psychol.* 113, 315–327. doi: 10.1016/s0001-6918(03)00043-x

Frissen, I., Vroomen, J., de Gelder, B., and Bertelson, P. (2005). The aftereffects of ventriloquism: generalization across sound-frequencies. *Acta Psychol.* 118, 93–100. doi: 10.1016/j.actpsy.2004.10.004

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

Hairston, W. D., Wallace, M. T., Vaughan, J. W., Stein, B. E., Norris, J. L., and Schirillo, J. A. (2003). Visual localization ability influences cross-modal bias. *J. Cogn. Neurosci.* 15, 20–29. doi: 10.1162/089892903321107792

Howard, I. P., and Templeton, W. B. (1966). *Human Spatial Orientation.* Oxford, England: Wiley.

Jackson, C. V. (1953). Visual factors in auditory localization. *Q. J. Exp. Psychol.* 5, 52–65. doi: 10.1080/17470215308416626

Keil, J., and Senkowski, D. (2018). Neural oscillations orchestrate multisensory processing. *Neuroscientist* 24, 609–626. doi: 10.1177/1073858418755352

King, A. J. (2009). Visual influences on auditory spatial learning. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 331–339. doi: 10.1098/rstb.2008.0230

Klemm, O. (1909). Lokalisation von Sinneseindrücken bei disparaten Nebenreizen [Localization of sensory impressions with disparate distracters]. *Psychol. Stud.* 5, 73–161.

Kopco, N., Lin, I.-F., Shinn-Cunningham, B. G., and Groh, J. M. (2009). Reference frame of the ventriloquism aftereffect. *J. Neurosci.* 29, 13809–13814. doi: 10.1523/JNEUROSCI.2783-09.2009

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., and Shams, L. (2007). Causal inference in multisensory perception. *PLoS One* 2:e943. doi: 10.1371/journal.pone.0000943

Kytö, M., Kusumoto, K., and Oittinen, P. (2015). The ventriloquist effect in augmented reality. *Proc. IEEE Int. Symp. Mix. Augment. Real.* 2015, 49–53. doi: 10.1109/ismar.2015.18

Lewald, J. (2002). Rapid adaptation to auditory-visual spatial disparity. *Learn. Mem.* 9, 268–278. doi: 10.1101/lm.51402

Magosso, E., Cuppini, C., and Ursino, M. (2012). A neural network model of ventriloquism effect and aftereffect. *PLoS One* 7:e42503. doi: 10.1371/journal.pone.0042503

Maiworm, M., Bellantoni, M., Spence, C., and Röder, B. (2012). When emotional valence modulates audiovisual integration. *Atten. Percept. Psychophys.* 74, 1302–1311. doi: 10.3758/s13414-012-0310-3

Manjarrez, E., Mendez, I., Martinez, L., Flores, A., and Mirasso, C. R. (2007). Effects of auditory noise on the psychophysical detection of visual signals: cross-modal stochastic resonance. *Neurosci. Lett.* 415, 231–236. doi: 10.1016/j.neulet.2007.01.030

Meijer, D., Veselič, S., Calafiore, C., and Noppeney, U. (2019). Integration of audiovisual spatial signals is not consistent with maximum likelihood estimation. *Cortex* 119, 74–88. doi: 10.1016/j.cortex.2019.03.026

Mendez-Balbuena, I., Arrieta, P., Huidobro, N., Flores, A., Lemuz-Lopez, R., Trenado, C., et al. (2018). Augmenting EEG-global-coherence with auditory and visual noise: multisensory internal stochastic resonance. *Medicine* 97:e12008. doi: 10.1097/md.0000000000012008

Mendonça, C., Escher, A., van de Par, S., and Colonius, H. (2015). Predicting auditory space calibration from recent multisensory experience. *Exp. Brain Res.* 233, 1983–1991. doi: 10.1007/s00221-015-4259-z

Occelli, V., Bruns, P., Zampini, M., and Röder, B. (2012). Audiotactile integration is reduced in congenital blindness in a spatial ventriloquism task. *Neuropsychologia* 50, 36–43. doi: 10.1016/j.neuropsychologia.2011.10.019

Odegaard, B., and Shams, L. (2016). The brain's tendency to bind audiovisual signals is stable but not general. *Psychol. Sci.* 27, 583–591. doi: 10.1177/0956797616628860

Odegaard, B., Wozny, D. R., and Shams, L. (2016). The effects of selective and divided attention on sensory precision and integration. *Neurosci. Lett.* 614, 24–28. doi: 10.1016/j.neulet.2015.12.039

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

Pages, D. S., and Groh, J. M. (2013). Looking at the ventriloquist: visual outcome of eye movements calibrates sound localization. *PLoS One* 8:e72562. doi: 10.1371/journal.pone.0072562

Park, H., and Kayser, C. (2019). Shared neural underpinnings of multisensory integration and trial-by-trial perceptual recalibration in humans. *Elife* 8:e47001. doi: 10.7554/eLife.47001

Passamonti, C., Frissen, I., and Làdavas, E. (2009). Visual recalibration of auditory spatial perception: two separate neural circuits for perceptual learning. *Eur. J. Neurosci.* 30, 1141–1150. doi: 10.1111/j.1460-9568.2009.06910.x

Pick, H. L. Jr., Warren, D. H., and Hay, J. C. (1969). Sensory conflict in judgments of spatial direction. *Percept. Psychophys.* 6, 203–205. doi: 10.3758/bf03207017

Radeau, M., and Bertelson, P. (1974). The after-effects of ventriloquism. *Q. J. Exp. Psychol.* 26, 63–71. doi: 10.1080/14640747408400388

Recanzone, G. H. (1998). Rapidly induced auditory plasticity: the ventriloquism aftereffect. *Proc. Natl. Acad. Sci. U S A* 95, 869–875. doi: 10.1073/pnas.95.3.869

Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time. *Hear. Res.* 258, 89–99. doi: 10.1016/j.heares.2009.04.009

Röder, B., and Büchel, C. (2009). Multisensory interactions within and outside the focus of visual spatial attention (Commentary on Fairhall and Macaluso). *Eur. J. Neurosci.* 29, 1245–1246. doi: 10.1111/j.1460-9568.2009.06715.x

Rohe, T., and Noppeney, U. (2015a). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biol.* 13:e1002073. doi: 10.1371/journal.pbio.1002073

Rohe, T., and Noppeney, U. (2015b). Sensory reliability shapes perceptual inference *via* two mechanisms. *J. Vis.* 15:22. doi: 10.1167/15.5.22

Rohe, T., and Noppeney, U. (2016). Distinct computational principles govern multisensory integration in primary sensory and association cortices. *Curr. Biol.* 26, 509–514. doi: 10.1016/j.cub.2015.12.056

Samad, M., and Shams, L. (2016). Visual-somatotopic interactions in spatial perception. *Neuroreport* 27, 180–185. doi: 10.1097/wnr.0000000000000521

Samad, M., and Shams, L. (2018). Recalibrating the body: visuotactile ventriloquism aftereffect. *PeerJ* 6:e4504. doi: 10.7717/peerj.4504

Sarlat, L., Warusfel, O., and Viaud-Delmon, I. (2006). Ventriloquism aftereffects occur in the rear hemisphere. *Neurosci. Lett.* 404, 324–329. doi: 10.1016/j.neulet.2006.06.007

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

Slutsky, D. A., and Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport* 12, 7–10. doi: 10.1097/00001756-200101220-00009

Stratton, G. M. (1897). Vision without inversion of the retinal image. *Psychol. Rev.* 4, 341–360. doi: 10.1037/h0075482

Talsma, D., Senkowski, D., Soto-Faraco, S., and Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* 14, 400–410. doi: 10.1016/j.tics.2010.06.008

Thomas, G. J. (1941). Experimental study of the influence of vision on sound localization. *J. Exp. Psychol.* 28, 163–177. doi: 10.1037/h0055183

Thurlow, W. R., and Jack, C. E. (1973). Certain determinants of the "ventriloquism effect". *Percept. Mot. Skills* 36, 1171–1184. doi: 10.2466/pms.1973.36.3c.1171

Van Wanrooij, M. M., Bremen, P., and Van Opstal, A. J. (2010). Acquired prior knowledge modulates audiovisual integration. *Eur. J. Neurosci.* 31, 1763–1771. doi: 10.1111/j.1460-9568.2010.07198.x

Vroomen, J., Bertelson, P., and de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Percept. Psychophys.* 63, 651–659. doi: 10.3758/bf03194427

Vroomen, J., and Stekelenburg, J. J. (2014). A bias-free two-alternative forced choice procedure to examine intersensory illusions applied to the ventriloquist effect by flashes and averted eye-gazes. *Eur. J. Neurosci.* 39, 1491–1498. doi: 10.1111/ejn.12525

Wallace, M. T., Roberson, G. E., Hairston, W. D., Stein, B. E., Vaughan, J. W., and Schirillo, J. A. (2004). Unifying multisensory signals across time and space. *Exp. Brain Res.* 158, 252–258. doi: 10.1007/s00221-004-1899-9

Watson, D. M., Akeroyd, M. A., Roach, N. W., and Webb, B. S. (2019). Distinct mechanisms govern recalibration to audio-visual discrepancies in remote and recent history. *Sci. Rep.* 9:8513. doi: 10.1038/s41598-019-44984-9

Woods, T. M., and Recanzone, G. H. (2004). "Cross-modal interactions evidenced by the ventriloquism effect in humans and monkeys," in *The Handbook of Multisensory Processes*, eds G. Calvert, C. Spence and B. E. Stein (Cambridge, MA: MIT Press), 35–48.

Wozny, D. R., and Shams, L. (2011). Recalibration of auditory space following milliseconds of cross-modal discrepancy. *J. Neurosci.* 31, 4607–4612. doi: 10.1523/jneurosci.6079-10.2011

Zaidel, A., Ma, W. J., and Angelaki, D. E. (2013). Supervised calibration relies on the multisensory percept. *Neuron* 80, 1544–1557. doi: 10.1016/j.neuron.2013.09.026

Zaidel, A., Turner, A. H., and Angelaki, D. E. (2011). Multisensory calibration is independent of cue reliability. *J. Neurosci.* 31, 13949–13962. doi: 10.1523/jneurosci.2732-11.2011

Zierul, B., Röder, B., Tempelmann, C., Bruns, P., and Noesselt, T. (2017). The role of auditory cortex in the spatial ventriloquism aftereffect. *Neuroimage* 162, 257–268. doi: 10.1016/j.neuroimage.2017.09.002

Zierul, B., Tong, J., Bruns, P., and Röder, B. (2019). Reduced multisensory integration of self-initiated stimuli. *Cognition* 182, 349–359. doi: 10.1016/j.cognition.2018.10.019