



Temporal Binding in Multisensory and Motor-Sensory Contexts: Toward a Unified Model

Kishore Kumar Jagini*

Center for Cognitive and Brain Sciences, Indian Institute of Technology Gandhinagar, Gandhinagar, India

Our senses receive a manifold of sensory signals at any given moment in our daily lives. For a coherent and unified representation of information and precise motor control, our brain needs to temporally bind the signals emanating from a common causal event and segregate others. Traditionally, different mechanisms were proposed for the temporal binding phenomenon in multisensory and motor-sensory contexts. This paper reviews the literature on the temporal binding phenomenon in both multisensory and motor-sensory contexts and suggests future research directions for advancing the field. Moreover, by critically evaluating the recent literature, this paper suggests that common computational principles are responsible for the temporal binding in multisensory and motor-sensory contexts. These computational principles are grounded in the Bayesian framework of uncertainty reduction rooted in the Helmholtzian idea of unconscious causal inference.

Keywords: temporal binding, multisensory, motor-sensory, causal inference, Bayesian models, precision

OPEN ACCESS

Edited by:

Giuseppe Giglia,
University of Palermo, Italy

Reviewed by:

Sheila Gillard Crewther,
La Trobe University, Australia
Cristiano Cuppini,
University of Bologna, Italy

*Correspondence:

Kishore Kumar Jagini
kishore.jagini@iitgn.ac.in

Specialty section:

This article was submitted to
Sensory Neuroscience,
a section of the journal
Frontiers in Human Neuroscience

Received: 16 November 2020

Accepted: 18 February 2021

Published: 25 March 2021

Citation:

Jagini KK (2021) Temporal Binding in
Multisensory and Motor-Sensory
Contexts: Toward a Unified Model.
Front. Hum. Neurosci. 15:629437.
doi: 10.3389/fnhum.2021.629437

INTRODUCTION

We receive sensory information from the environment and the body through several distinct senses. For a coherent and unified representation of information, our brain needs to group the multisensory features emanating from an object or event (Calvert et al., 2004). For instance, imagine that you are applauding the musical performance of your friend by rhythmic hand clapping. The multiple sensory features (such as tactile, auditory, and visual) from hand-clapping are grouped and experienced as coming from a single causal event rather than separate events. Several challenges that the brain needs to overcome for grouping or often called “binding” the multisensory features of an event (Vroomen and Keetels, 2010; Vilares and Kording, 2011; Burwick, 2014; Spence and Frings, 2020). This paper focuses on two non-trivial and inter-related challenges that the brain must account for in binding the multisensory and motor-sensory features in the time domain.

The first challenge is causal determination. Our senses are bombarded with multiple sensory features that are either received passively or generated as a consequence of our motor actions. How does our brain deal with the ambiguity in matching sensory features that belong to one causal event and segregate others? Or how does our brain determine whether the sensory features are causal outcomes of our voluntary motor actions or not? The second challenge is with regard to the lack of precision in the temporal estimates of sensory features across the senses. This lack of precision in temporal estimates is assumed to be due to the noisy or uncertain sensory information and differential temporal resolution in the encoding of the temporal information across the

senses (Kersten et al., 2004; Faisal et al., 2008; Vroomen and Keetels, 2010). How does our brain account for this sensory noise and differential precision in encoding the temporal information across the senses for coherent and robust perceptual binding of sensory signals coming from a common cause? Previous studies have proposed different mechanisms for the temporal binding phenomenon in multisensory and motor-sensory contexts (Haggard et al., 2002; Chen and Vroomen, 2013). This paper reviews the recent literature on the temporal binding phenomenon in multisensory and motor-sensory contexts. Moreover, this review suggests the existence of common computational principles grounded in the Bayesian framework for temporal binding in multisensory and motor-sensory contexts. The following section briefly describes various behavioral manifestations of the temporal binding and its constraints across the multisensory and motor-sensory contexts. After describing the basic temporal binding phenomenon, the author discusses the Bayesian inference models and the extent to which these models explain the temporal binding phenomenon in the multisensory and motor-sensory contexts.

TEMPORAL BINDING AND TEMPORAL BINDING WINDOW

The term “temporal binding” refers to the subjective experience of mutual attraction between two or more events in the time domain. For example, in the audio-visual perception, the temporal aspect of a visual event, such as onset time, can be perceptually shifted and binds with a slightly asynchronous auditory event (Vroomen and Keetels, 2010; Chen and Vroomen, 2013). Similarly, in the motor-sensory contexts, the perceived onset times of self-generated motor action and its sensory outcome (e.g., visual or auditory event) are shown to be mutually attracted to each other (Haggard et al., 2002; Wolpe et al., 2013). The temporal binding phenomenon was also observed for other aspects of the time domain, such as frequency and duration. For instance, in the double-flash illusion, a single visual flash is perceived as multiple flashes when accompanied by multiple auditory beeps (Shams et al., 2000, 2005). With regard to the duration perception, studies have shown that visual events are perceived to be longer or shorter during the concurrent auditory event or motor action (Burr et al., 2009; Press et al., 2014; Anobile et al., 2019). Importantly, however, these temporal illusions are preserved over a time window known as “temporal binding window (TBW)” or “temporal integration window” (Diederich and Colonius, 2004; Wassenhove et al., 2007; Vroomen and Keetels, 2010). From the literature, it appears that there is a large variability in the extent of temporal binding windows across different combinations of paired multisensory stimuli, experimental paradigms, and stimulus (such as spatiotemporal, stimulus complexity) or cognitive factors (Andersen et al., 2004; Vroomen and Keetels, 2010; Stevenson and Wallace, 2013). Also, from the developmental perspective, studies have shown that the extent of multisensory temporal binding windows follows a U-shaped function with children and older age groups having larger binding windows compared to the young adults (Wallace et al.,

2019). The increased temporal binding windows in children, older adults, and in certain neurodevelopmental disorders (e.g., autism) lead to the disruption of various cognitive abilities and reduced behavioral performance (Barutcu et al., 2010; Downing et al., 2015; Wallace et al., 2019).

BAYESIAN INFERENCE

In recent decades, studies from neuroscientific, behavioral, and computational approaches have indicated that the brain generates various mental events by “predictive-processing” of information (Rao and Ballard, 1999; Feldman and Friston, 2010; Clark, 2013; Hohwy, 2013; Hutchinson and Barrett, 2019). The core assumption of the “predictive-processing” framework is that the brain constantly runs an internal mental model of the world and uses it to predict the causes of the sensory effect. The internal model is assumed to be continuously updated based on the discrepancy between predicted and actual sensory input which is often referred to as prediction error (Raichle, 2015). The essential role of the brain is to minimize the prediction error for the best possible causal inference of sensory information. Although formal computational models of predictive processing frameworks have been developed recently, the core assumptions have roots in the Helmholtzian idea that the brain makes unconscious perceptual inference based on prior knowledge or prior learning (Von Helmholtz, 1867). One of the worrying problems for the perceptual inference is that there is no perfect one-one mapping between cause and sensory effect. Sensory information is corrupted with noise from the external world, noise in the nervous system, and variable precision of sensory encoding across the senses (Ernst and Bühlhoff, 2004; Ernst, 2006). This variability in sensory information necessitates the brain to perform probabilistic (Bayesian) inference when computing predictions and prediction errors (Vilares and Kording, 2011). The main purpose of probabilistic processing is to update the internal models with precise prediction-error signals and ignore (or less prioritize) relatively less precise prediction-error signals. According to Bayesian probabilistic predictive processing models, perception arises from the precision-weighted probabilistic combination of prior belief or knowledge (or prior in Bayesian terms) of the world and the current sensory evidence (or likelihood in Bayesian terms). In other words, perception is determined by the trade-off between the precision of prior and likelihood.

In parallel lines, the characterization of cause-and-effect temporal relationships by Hume inspired numerous empirical studies to understand the predictive processing of the brain (Hume, 1739; Pearl, 1988, 2000; Hohwy, 2013). Hume has suggested that the inference of the relationship between cause and effect developed through statistical regularities in nature. He has proposed three fundamental cues that may support causal learning, such as temporal priority, contingency, and contiguity. Temporal priority refers to the idea that there must be an existence of cause before the sensory effect. This cause(s) and its effect(s) or causally related events are typically co-occurring together repeatedly and reliably (i.e., contingent) and co-occur

close in space and time (i.e., contiguous). Numerous studies have experimentally manipulated the rules of causal learning to understand the causal learning and predictive processing of the brain and provided empirical evidence (Alais et al., 2010; Buehner, 2014).

Recent studies have suggested that the human temporal perception is consistent with Bayesian inference models across different time scales and temporal aspects (Shi et al., 2013; Rhodes, 2018). For example, a well-known perceptual phenomenon in the temporal dimension called “central-tendency effect” has been demonstrated to be quantitatively predicted by Bayesian inference models (Jazayeri and Shadlen, 2010).

Bayesian Casual Inference in Multisensory Temporal Binding

Appropriate binding of multisensory features of an event and segregating others necessary for a coherent and unified perceptual representation lead to enhanced behavioral performance. For instance, previous researchers have demonstrated that the binding of multisensory information enhances the speed and accuracy of detection performance and increases the precision of sensory estimates that enhanced the discrimination performance (Ernst and Banks, 2002; Diederich and Colonius, 2004; Ernst and Bühlhoff, 2004; Ernst, 2006).

The first challenge that our brain needs to account for in binding multisensory features of an event is solving the causal inference problem—determining whether sensory signals are coming from a common causal event or different events. That is, our perceptual system needs to infer the causal structure of the world from noisy sensory data for which we do not have direct access (Körding et al., 2007; Stein, 2012). Bayesian causal inference models explain how an observer might infer the causal structure for determining the probabilistic estimation of whether sensory signals are coming from a common causal event or different events (Körding et al., 2007; Wozny et al., 2010; Noppeney, 2020). The estimation or inference of causal structure is thought to be derived by the probabilistic averaging of the common cause prior (or prior knowledge that the signals are coming from a common source) and current sensory evidence according to the Bayesian models (Ernst, 2006, 2012; Körding et al., 2007). Therefore, the extent of binding or integration of multisensory signals depends on the strength of the inferred causal structure. For instance, forced fusion might happen only if an observer infers that the multiple signals are coming from a common cause with absolute certainty, or complete segregation of signals could happen if the observer infers signals are coming from separate sources. However, due to the inherent uncertainty of the sensory data and uncertainty in causal inference, the integration of multisensory signals can arbitrate between forced fusion and segregation (Körding et al., 2007; Shams and Beierholm, 2010; Ernst, 2012).

Previous research has indicated numerous cues that are suggested to act as common cause priors (e.g., spatial and temporal mapping or correlation between sensory signals) for solving the causal inference problem (Ernst and Bühlhoff,

2004; Doehrmann and Naumer, 2008; Vroomen and Keetels, 2010; Buehner, 2014; Debats et al., 2017). For example, a typical multisensory event in the natural environment, a ball hitting a glass window, produces multiple sensory stimuli that are spatiotemporally proximal. These spatial and temporal regularities are utilized by our perceptual system to decide whether sensory cues are coming from the same or different causal events (Vroomen and Keetels, 2010; Chen and Vroomen, 2013). Moreover, the extant literature has indicated several higher-order cognitive factors such as semantic (Doehrmann and Naumer, 2008), metaphoric (Parise and Spence, 2009), or experimentally learned matching (Ernst, 2007) of paired multisensory cues involved in causal determination. This evidence also indicated that the causally related (e.g., congruent) multisensory features have a larger TBW than non-causally related (or unrelated) features. In other words, the larger TBW indicates that the causally related pairs of multisensory stimuli are more often perceived to occur together in time than the pairs of unrelated stimuli that have the same amount of asynchrony between them. Moreover, the strength of prior belief that the pair of events is causally related is shown to be positively correlated with the tendency to perceive events as co-occurring together in time (Faro et al., 2005).

The next question is how the brain optimally binds, in the time domain, the causally related multisensory features which are processed at different times due to the noise in the nervous system. According to the Bayesian causal inference models, the causally related multisensory features are temporally bound together by precision-weighted probabilistic cue combination (Vilares and Kording, 2011). In other words, the less precise sensory feature is perceptually shifted closer to the more precise sensory feature to maintain temporal coherence. For example, when the audio-visual cues of an object are presented asynchronously, the visual stimulus is perceived to occur temporally closer to the auditory stimulus, called “temporal ventriloquism” (Morein-Zamir et al., 2003). Since the precision of the temporal judgment of the visual cue is lower than the auditory cue, the Bayesian sensory cue combination predicts that the auditory temporal judgment is given more weight and shifts the visual stimulus perceptually closer and bound to the auditory stimulus (Alais and Weston, 2010; Chen and Vroomen, 2013). Similarly, Ley et al. (2009) showed that auditory and vibrotactile stimuli are perceptually bound according to the Bayesian cue combination. Other studies indicated that the semantic or learned correlations (congruent) between a pair of sensory cues induced greater temporal ventriloquism compared to the non-congruent sensory pairs (Vatakis and Spence, 2007; Chen and Vroomen, 2013). Concerning the double flash illusion, since the reliability of auditory event is greater than the visual event in the temporal domain, the temporal frequency of auditory beeps perceptually dominated the temporal frequency of visual flashes (Andersen et al., 2004). Moreover, researchers have demonstrated that the double flash illusory percept follows the principles of Bayesian causal inference models by manipulating the relative reliabilities of auditory and visual stimuli (Shams et al., 2005). Similarly, duration estimates of audio-tactile and audio-visual sensory signals are found to be in accordance with

the Bayesian causal inference models (Hartcher-O'Brien et al., 2014; Ball et al., 2017). However, previous literature has indicated that the individual multisensory features are either under- or overweighted than expected by Bayesian causal inference in binding due to the inherent limitations in the models (for the detailed review, see Noppeney, 2020). Future studies are required to refine the current Bayesian models to fully account for the multisensory perception.

The Bayesian framework of the multisensory causal inference model became an influential model by systematically explaining the empirical evidence of multisensory perception literature. However, the current multisensory literature indicated that the reported multisensory binding effects are influenced by a combination of more than one factor, and it is not clear how they independently and interactively modulate the multisensory temporal perception. For instance, factors such as spatial and temporal proximity, semantic (or learned) congruency between pairs of cues, and attentional allocation are all known to influence temporal perception (Oever et al., 2016). Future studies are required to orthogonally manipulate these factors within an experimental paradigm in order to understand their independent and interactive roles in the temporal binding phenomenon.

Bayesian Casual Inference in Motor-Sensory Temporal Binding

The last few decades of research have focused on the temporal processing of multisensory features that are passively received by the study participants. However, in the real-world, multisensory features can also occur because of our interactions with the environment. The broader question is whether the process by which motor-sensory cues generated by voluntary action are bound differs from the passively received sensory cues. Previous literature indicated temporal binding between voluntary motor action and its causal sensory outcome (Haggard et al., 2002; Hughes et al., 2013). For instance, Haggard et al. (2002) indicated the perceived temporal attraction between voluntary action onset (keypress) and its predictable sensory outcome, such as a brief tone (Haggard et al., 2002). In their study, participants were asked to watch a clock face and report when an action was performed and when the sensory outcome was presented in two conditions. In baseline conditions (single event conditions), the study participants reported the onset times of keypress (voluntary action), time of muscle twitch produced by Trans-cranial Magnetic stimulation (TMS condition) on the motor cortex, time of audible sound created by TMS without muscle twitch (TMS sham condition), and time of a tone onset (tone condition) in separate trials. An audible tone appeared in operant conditions after 250 milliseconds of each voluntary keypress condition, TMS, and sham TMS conditions. The task of the subjects was to report the time of both events in operant condition at the end of each trial. Their study results indicated the perceived temporal attraction between action and its outcome (tone) when participants intentionally performed an action rather than TMS-induced involuntary action (Haggard et al., 2002). In other words, action and outcome are bound together by shifting the perceived temporal onsets toward each other when participants

intentionally performed an action. Hence, it has been called the "intentional binding" (IB) effect. Further, their study indicated the increased IB effect when the outcome was short delayed after the action and temporally predictable. However, as the delay increased between action and outcome, and the outcome temporally became unpredictable, the IB effect was reduced. This evidence indicates the importance of spatiotemporal factors for causal determination and temporal binding of action and its sensory outcome. The IB effect was attributed to the motor-based predictive mechanisms since IB appeared for voluntary (intentional) and not for involuntary (TMS-induced) actions (Haggard et al., 2002; Hughes et al., 2013). Waszak et al. (2012) proposed a pre-activation account that explains how the sensory action–outcome binds to the action (Waszak et al., 2012). According to the pre-activation account, predicted action–outcomes are pre-activated and increase their baseline neural-activity before the outcome occurs. Since the neural units of predicted outcomes are already activated to some baseline level by the motor-based predictive mechanisms, less strength of the signal is required for reaching the detection threshold. Thus, the action–outcome reaches threshold awareness faster and is perceived temporally closer to the action.

Contrastingly, studies also indicated that IB-like effects appeared even for non-intentional (passive) actions (Buehner, 2015; Borhani et al., 2017; Suzuki et al., 2019), machine-made action and its causal outcome (Buehner, 2012), or observation of other's action and its causal outcome (Poonian et al., 2015). This evidence casts severe doubts on the role of motor-based (forward model) predictive mechanisms on IB and suggests a general predictive mechanism responsible for the temporal binding between action and its sensory outcome (Dogge et al., 2019; Press et al., 2019).

A number of recent studies have begun to investigate IB mechanisms from the perspective of Bayesian cue integration (Moore and Obhi, 2012; Wolpe et al., 2013; Lush et al., 2019). Considering the action and its sensory outcome are causally related, and the temporal judgments of action timing and its outcome are prone to inaccuracies due to the noise, one can model the IB in terms of the Bayesian cue integration framework. For example, Wolpe et al. (2013) manipulated the action outcome's (a brief tone) temporal precision or reliability (inverse of the variance) by adding white noise. They found that the perceived onset time of auditory outcome attracted more to the action when the reliability of the tone was weak (e.g., with added noise) compared to the high-reliability tone (e.g., with no added noise). In another study by Lush et al. (2019), the participants were divided into two groups based on their reliability of time judgments of intentional action (low and high-reliability groups) and measured the perceived temporal attractions between action and its outcome. Their study indicated that the perceived time of action attracted more toward the outcome in the low-reliability group than in the high-reliability group. Legaspi and Toyozumi (2019) explicitly compared the results of observed IB effects in the studies of Haggard et al. (2002) and Wolpe et al. (2013) with the predictions of the Bayesian cue combination model (Legaspi and Toyozumi, 2019). Interestingly, their model reliably predicted the intentional binding effects observed in the studies of Haggard

et al. (2002) and Wolpe et al. (2013). Concerning the duration aspect of the time dimension, auditory or visual perceived durations are modulated in action contexts (Press et al., 2014; Anobile et al., 2019). However, there is a lack of studies assessing the Bayesian integration of duration estimates in motor-sensory contexts. The abovementioned studies indicated that temporal binding between motor action and its sensory outcome follow general rules of Bayesian cue integration common to the multisensory perceptual phenomenon and not necessarily restricted to the motor-based predictive mechanisms. Future studies are required to evaluate the Bayesian cue integration model to understand how action modulates the temporal binding of multisensory outcomes having differential temporal precisions. This leads to a more naturalistic understanding of the role of action on perception since our actions often produce multiple sensory stimuli.

CONCLUSIONS

This review explored the temporal binding mechanisms in multisensory and motor-sensory contexts. By critically evaluating the recent empirical evidence, this paper suggests

that the common computational mechanisms grounded in Bayesian causal inference models are responsible for the temporal binding in multisensory and motor-sensory contexts. Moreover, the extent of temporal binding depends on the strength of prior and the precision of sensory likelihoods. Future studies are required to understand the independent and interactive roles of multiple priors and sensory likelihoods on temporal binding across the multisensory and motor-sensory features.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGMENTS

Author want to thank Meera Mary Sunny, Ishita Arun, and Midhula Chandran for helpful feedback on the earlier versions of the manuscript. I also want to thank Monal Desai for her helpful comments on English language issues in the manuscript.

REFERENCES

- Alais, D., Newell, F., and Mamassian, P. (2010). Multisensory processing in review: from physiology to behaviour. *See. Perceiv.* 23, 3–38. doi: 10.1163/187847510X488603
- Alais, D., and Weston, E. (2010). Temporal ventriloquism: perceptual shifts in temporal position and improved audiovisual precision predicted by maximum likelihood estimation. *J. Vis.* 6, 171–171. doi: 10.1167/6.6.171
- Andersen, T. S., Tiippana, K., and Sams, M. (2004). Factors influencing audiovisual fission and fusion illusions. *Cogn. Brain Res.* 21, 301–308. doi: 10.1016/j.cogbrainres.2004.06.004
- Anobile, G., Domenici, N., Togoli, I., Burr, D., and Arrighi, R. (2019). Distortions of visual time induced by motor adaptation. *J. Exp. Psychol. Gen.* 149, 1333–1343. doi: 10.1037/xge0000709
- Ball, D. M., Arnold, D. H., and Yarrow, K. (2017). Weighted integration suggests that visual and tactile signals provide independent estimates about duration. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 868–880. doi: 10.1037/xhp0000368
- Barutchu, A., Danaher, J., Crewther, S. G., Innes-Brown, H., Shivdasani, M. N., and Paolini, A. G. (2010). Audiovisual integration in noise by children and adults. *J. Exp. Child Psychol.* 105, 38–50. doi: 10.1016/j.jecp.2009.08.005
- Borhani, K., Beck, B., and Haggard, P. (2017). Choosing, doing, and controlling: implicit sense of agency over somatosensory events. *Psychol. Sci.* 28, 882–893. doi: 10.1177/0956797617697693
- Buehner, M. J. (2012). Understanding the past, predicting the future. *Psychol. Sci.* 23, 1490–1497. doi: 10.1177/0956797612444612
- Buehner, M. J. (2014). Time and causality: editorial. *Front. Psychol.* 5:228. doi: 10.3389/fpsyg.2014.00228
- Buehner, M. J. (2015). Awareness of voluntary and involuntary causal actions and their outcomes. *Psychol. Conscious Theory Res. Pract.* 2, 237–252. doi: 10.1037/cns0000068
- Burr, D., Banks, M. S., and Morrone, M. C. (2009). Auditory dominance over vision in the perception of interval duration. *Exp. Brain Res.* 198:49. doi: 10.1007/s00221-009-1933-z
- Burwick, T. (2014). The binding problem. *Wiley Interdiscip. Rev. Cogn. Sci.* 5, 305–315. doi: 10.1002/wcs.1279
- Calvert, G. A., Spence, C., and Stein, B. E. (2004). *The Handbook of Multisensory Processes*. Cambridge, MA: MIT Press.
- 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
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477
- Debats, N. B., Ernst, M. O., and Heuer, H. (2017). Kinematic cross-correlation induces sensory integration across separate objects. *Eur. J. Neurosci.* 46, 2826–2834. doi: 10.1111/ejn.13758
- Diederich, A., and Colonius, H. (2004). Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Percept. Psychophys.* 66, 1388–1404. doi: 10.3758/BF03195006
- Doehrmann, O., and Naumer, M. J. (2008). Semantics and the multisensory brain: how meaning modulates processes of audio-visual integration. *Brain Res.* 1242, 136–150. doi: 10.1016/j.brainres.2008.03.071
- Dogge, M., Custers, R., and Aarts, H. (2019). Moving forward: on the limits of motor-based forward models. *Trends Cogn. Sci.* 23, 743–753. doi: 10.1016/j.tics.2019.06.008
- Downing, H. C., Barutchu, A., and Crewther, S. G. (2015). Developmental trends in the facilitation of multisensory objects with distractors. *Front. Psychol.* 5:1559. doi: 10.3389/fpsyg.2014.01559
- Ernst, M. O. (2006). “A Bayesian view on multimodal cue integration,” in *Human Body Perception From the Inside Out*, eds G. Knoblich, I. Thornton, M. Grosjean, and M. Shiffrar (New York, NY: Oxford University Press), 105–131.
- Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *J. Vis.* 7:7. doi: 10.1167/7.5.7
- Ernst, M. O. (2012). “Optimal multisensory integration: assumptions and limits,” in *The New Handbook of Multisensory Processes*, ed B. E. Stein (Cambridge, MA: MIT Press), 1084–1124.
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433. doi: 10.1038/415429a
- Ernst, M. O., and Bühlhoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169. doi: 10.1016/j.tics.2004.02.002
- Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nat. Rev. Neurosci.* 9, 292–303. doi: 10.1038/nrn2258
- Faro, D., Leclerc, F., and Hastie, R. (2005). Perceived causality as a cue to temporal distance. *Psychol. Sci.* 16, 673–677. doi: 10.1111/j.1467-9280.2005.01594.x

- Feldman, H., and Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* 4:215. doi: 10.3389/fnhum.2010.00215
- Haggard, P., Clark, S., and Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nat. Neurosci.* 5, 382–385. doi: 10.1038/nn827
- Hartcher-O'Brien, J., Luca, M. D., and Ernst, M. O. (2014). The duration of uncertain times: audiovisual information about intervals is integrated in a statistically optimal fashion. *PLoS ONE* 9:e89339. doi: 10.1371/journal.pone.0089339
- Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.
- Hughes, G., Desantis, A., and Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: the role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychol. Bull.* 139, 133–151. doi: 10.1037/a0028566
- Hume, D. (1739). *A Treatise of Human Nature*. Oxford: Clarendon Press.
- Hutchinson, J. B., and Barrett, L. F. (2019). The power of predictions: an emerging paradigm for psychological research. *Curr. Dir. Psychol. Sci.* 28, 280–291. doi: 10.1177/0963721419831992
- Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13, 1020–1026. doi: 10.1038/nn.2590
- Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as Bayesian inference. *Annu. Rev. Psychol.* 55, 271–304. doi: 10.1146/annurev.psych.55.090902.142005
- 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
- Legaspi, R., and Toyozumi, T. (2019). A Bayesian psychophysics model of sense of agency. *Nat. Commun.* 10:4250. doi: 10.1038/s41467-019-12170-0
- Ley, I., Haggard, P., and Yarrow, K. (2009). Optimal integration of auditory and vibrotactile information for judgments of temporal order. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 1005–1019. doi: 10.1037/a0015021
- Lush, P., Roseboom, W., Cleeremans, A., Scott, R. B., Seth, A. K., and Dienes, Z. (2019). Intentional binding as Bayesian cue combination: testing predictions with trait individual differences. *J. Exp. Psychol. Hum. Percept. Perform.* 45, 1206–1217. doi: 10.1037/xhp0000661
- Moore, J. W., and Obhi, S. S. (2012). Intentional binding and the sense of agency: a review. *Conscious. Cogn.* 21, 546–561. doi: 10.1016/j.concog.2011.12.002
- Morein-Zamir, S., Soto-Faraco, S., and Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cogn. Brain Res.* 17, 154–163. doi: 10.1016/S0926-6410(03)00089-2
- Noppeney, U. (2020). “Multisensory perception: behavior, computations and neural mechanisms,” in *The cognitive neurosciences, 6th Edn*, eds D. Poeppel, G. R. Mangun, and M. S. Gazzaniga (Cambridge, MA: MIT Press), 141–150
- Oever, S., ten, Romei, V., Atteveldt, N., van, Soto-Faraco, S., Murray, M. M., and Matusz, P. J. (2016). The COGs (context, object, and goals) in multisensory processing. *Exp. Brain Res.* 234, 1307–1323. doi: 10.1007/s00221-016-4590-z
- Parise, C. V., and Spence, C. (2009). ‘When birds of a feather flock together’: synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS ONE* 4:e5664. doi: 10.1371/journal.pone.0005664
- Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, CA: Morgan Kaufmann Publishers.
- Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
- Poonian, S. K., McFadyen, J., Ogden, J., and Cunnington, R. (2015). Implicit agency in observed actions: evidence for N1 suppression of tones caused by self-made and observed actions. *J. Cogn. Neurosci.* 27, 752–764. doi: 10.1162/jocn_a_00745
- Press, C., Berlot, E., Bird, G., Ivry, R., and Cook, R. (2014). Moving time: the influence of action on duration perception. *J. Exp. Psychol. Gen.* 143, 1787–1793. doi: 10.1037/a0037650
- Press, C., Kok, P., and Yon, D. (2019). The perceptual prediction paradox. *Trends Cogn. Sci.* 24, 13–24. doi: 10.1016/j.tics.2019.11.003
- Raichle, M. E. (2015). The restless brain: how intrinsic activity organizes brain function. *Philos. Trans. R. Soc. B Biol. Sci.* 370:20140172. doi: 10.1098/rstb.2014.0172
- Rao, R. P. N., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580
- Rhodes, D. (2018). On the distinction between perceived duration and event timing: towards a unified model of time perception. *Timing Time Percept.* 6, 90–123. doi: 10.1163/22134468-20181132
- Shams, L., and Beierholm, U. R. (2010). Causal inference in perception. *Trends Cogn. Sci.* 14, 425–432. doi: 10.1016/j.tics.2010.07.001
- Shams, L., Kamitani, Y., and Shimojo, S. (2000). What you see is what you hear. *Nature* 408, 788–788. doi: 10.1038/35048669
- Shams, L., Ma, W. J., and Beierholm, U. (2005). Sound-induced flash illusion as an optimal percept. *Neuroreport* 16, 1923–1927. doi: 10.1097/01.wnr.0000187634.68504.bb
- Shi, Z., Church, R. M., and Meck, W. H. (2013). Bayesian optimization of time perception. *Trends Cogn. Sci.* 17, 556–564. doi: 10.1016/j.tics.2013.09.009
- Spence, C., and Frings, C. (2020). Multisensory feature integration in (and out) of the focus of spatial attention. *Atten. Percept. Psychophys.* 82, 363–376. doi: 10.3758/s13414-019-01813-5
- Stein, B. E. (ed.). (2012). *The New Handbook of Multisensory Processing*. Cambridge, MA: MIT Press.
- Stevenson, R. A., and Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Exp. Brain Res.* 227, 249–261. doi: 10.1007/s00221-013-3507-3
- Suzuki, K., Lush, P., Seth, A. K., and Roseboom, W. (2019). Intentional binding without intentional action. *Psychol. Sci.* 30, 842–853. doi: 10.1177/0956797619842191
- Vatakis, A., and Spence, C. (2007). Crossmodal binding: evaluating the “unity assumption” using audiovisual speech stimuli. *Percept. Psychophys.* 69, 744–756. doi: 10.3758/BF03193776
- Vilares, I., and Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Ann. N. Y. Acad. Sci.* 1224, 22–39. doi: 10.1111/j.1749-6632.2011.05965.x
- Von Helmholtz, H. (1867). *Handbuch der Physiologischen Optik*. Leipzig: Leopold Voss.
- Vroomen, J., and Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Atten. Percept. Psychophys.* 72, 871–884. doi: 10.3758/APP.72.4.871
- Wallace, M. T., Woynaroski, T. G., and Stevenson, R. A. (2019). Multisensory integration as a window into orderly and disrupted cognition and communication. *Annu. Rev. Psychol.* 71, 1–27. doi: 10.1146/annurev-psych-010419-051112
- Wassenhove, V., van, Grant, K. W., and Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45, 598–607. doi: 10.1016/j.neuropsychologia.2006.01.001
- Waszak, F., Cardoso-Leite, P., and Hughes, G. (2012). Action effect anticipation: neurophysiological basis and functional consequences. *Neurosci. Biobehav. Rev.* 36, 943–959. doi: 10.1016/j.neubiorev.2011.11.004
- Wolpe, N., Haggard, P., Siebner, H. R., and Rowe, J. B. (2013). Cue integration and the perception of action in intentional binding. *Exp. Brain Res.* 229, 467–474. doi: 10.1007/s00221-013-3419-2
- Wozny, D. R., Beierholm, U. R., and Shams, L. (2010). Probability matching as a computational strategy used in perception. *PLoS Comput. Biol.* 6:e1000871. doi: 10.1371/journal.pcbi.1000871

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Jagini. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.