Crossmodal Postdiction:
Conscious Perception as Revisionist History†

Noelle R. B. Stiles¹, Armand R. Tanguay Jr.², Shinsuke Shimojo³

¹*Department of Ophthalmology, University of Southern California, Los Angeles, CA, 90033, USA;* ²*Departments of Electrical and Computer Engineering, Chemical Engineering and Materials Science, Biomedical Engineering, Ophthalmology, Physics and Astronomy, and Neuroscience Graduate Program, University of Southern California, Los Angeles, CA, 90089, USA;* ³*Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, 91125, USA*
E-mail: nstiles@usc.edu

Abstract. *Postdiction occurs when later stimuli influence the perception of earlier stimuli. As the multisensory science field has grown in recent decades, the investigation of crossmodal postdictive phenomena has also expanded. Crossmodal postdiction can be considered (in its simplest form) the phenomenon in which later stimuli in one modality influence earlier stimuli in another modality (e.g., Intermodal Apparent Motion). Crossmodal postdiction can also appear in more nuanced forms, such as unimodal postdictive illusions (e.g., Apparent Motion) that are influenced by concurrent crossmodal stimuli (e.g., Crossmodal Influence on Apparent Motion), or crossmodal illusions (e.g., the Double Flash Illusion) that are influenced postdictively by a stimulus in one or the other modality (e.g., a visual stimulus in the Illusory Audiovisual Rabbit Illusion). In this review, these and other varied forms of crossmodal postdiction will be discussed. Three neuropsychological models proposed for unimodal postdiction will be adapted to the unique aspects of processing and integrating multisensory stimuli. Crossmodal postdiction opens a new window into sensory integration, and could potentially be used to identify new mechanisms of crossmodal crosstalk in the brain.* ⓒ *2022 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.Percept.Imaging.2022.5.000403]

1. INTRODUCTION

Visual flashes, auditory beeps, and tactile taps before a sensory stimulus of interest can influence the perception of that stimulus [3, 12, 43]. The influence of preceding stimuli on the perception of a stimulus is referred to as *prediction* due to the fact that preceding stimulus properties can be used to predict both sensory outcomes and behavior. Prediction has been used widely in perceptual neuroscience to disambiguate the flow of sequential perceptual processes, and to identify the subconscious steps of decision cascades [23, 46]. Until recently, this focus on prediction has derived from the assumption that the preceding stimuli and associated neural correlates generated immediately prior to a stimulus of interest can influence and therefore be used to predict the resultant cognitive awareness or behavior. However, neural processing of additional stimuli that occur *after* the initial stimuli can also potentially alter the resulting percept of the earlier stimulus of interest.

For example, later flashes, beeps, and taps can influence the perception of an earlier stimulus of interest, a phenomenon which has consequently been named *postdiction* [2, 3, 9, 47, 57]. We define postdiction herein as any perceptual phenomenon in which a stimulus presented *later* in time affects the perception of another stimulus presented *earlier* in time. While postdictive effects can also occur at the cognitive level outside of the perceptual domain, such postdictive effects likely involve different neural mechanisms than those discussed in this review.

Postdiction at its core implies a temporal paradox: How can *later* stimuli affect the perception of *earlier* ones? Perceptual postdiction provides a relatively straightforward solution to this seeming paradox, as conscious awareness progresses neither simultaneously nor synchronously with the presentation of sensory stimuli. Instead, awareness is delayed in time to allow for the integration of sensory processing from earlier stimuli. Consequently, the brain can process a series of sensory inputs as a group before an individual becomes aware of them. It has therefore been theorized that earlier and later stimuli can interact within this short temporal window before conscious awareness of a given stimulus is precipitated [10, 22, 29]. As a result, the temporal paradox described above dissolves, as later stimuli can impact the perception of earlier stimuli within this temporal window.

Another consequence of this short time window before awareness is that all of the perceptual postdictive illusions occur quickly, within a stimulus presentation window of less than approximately 400 ms in duration. This short stimulus window is a key constraint on perceptual postdictive processing. As such, the perceptual stimuli are typically presented as quick taps, beeps, or flashes in perceptual postdictive illusions (with the exception of the Flash Lag Illusion, which is generated by a moving stimulus preceding and following a quickly presented stimulus burst [36]).

Unisensory postdiction occurs in several common illusions such as Apparent Motion [24, 25, 40, 53], the Cutaneous Rabbit Illusion [14], the Flash Lag Illusion [36], Backward Masking [39], and the Line Motion Illusion [11]. In addition, most computational and psychophysical studies of postdiction have been unisensory in nature, with either visual, auditory, or tactile stimuli alone.

A new and emerging type of postdiction is crossmodal postdiction, in which not only does one or more later stimuli have a retroactive influence on an earlier stimulus, but they also bridge the sensory divide. We will define *crossmodal postdiction* as any combination of stimuli from two or more sensory modalities that generate a postdictive perception or effect.

The recent discovery of several new illusions has initiated the exploration of crossmodally postdictive phenomena, including the Illusory Audiovisual Rabbit Illusion [51], the Invisible Audiovisual Rabbit Illusion [51], and the Burst Lag Auditory–Visual Illusion [1]. In addition, while postdiction was still being formulated and investigated in the first decade of the 21st century, neuroscientists found that several previously investigated unisensory (now understood as postdictive) illusions could be extended into the crossmodal domain. It is interesting to note that these crossmodally modified illusions were not initially identified as postdictive in a number of cases, likely due to the limited application of the postdictive model available at that time. We include herein a discussion of a wide range of these crossmodal illusions and their potential implications for crossmodal postdiction.

For the purposes of this review article, we have organized crossmodal postdiction into three different types (Figures 1–3).

The first type of crossmodal postdiction that we will discuss can be called *Unimodal Postdiction with Crossmodal Influence* (as illustrated schematically in Fig. 1), and occurs when a series of stimuli that interact postdictively are presented unimodally, but a crossmodal stimulus influences the perception of the unimodal postdictive effect. For example, auditory and tactile stimuli can influence the perception of Visual Apparent Motion [13, 16, 42]. In this case, the unisensory postdictive effect (indicated with dashed arrows in Fig. 1) is the illusory motion generated between two spatially displaced and sequentially presented visual stimuli. The crossmodal feature is the fact that the auditory or tactile stimuli can influence the visual postdiction occurring among the visual stimuli (solid green arrows, Fig. 1). This type of crossmodal postdiction also includes auditory influence on the Visual Flash Lag Illusion [56], as well as auditory influence on Visual Backward Masking [8], and is the most common type of crossmodal postdiction studied to date.

The second type of crossmodal postdictive illusion occurs when stimuli of two different modalities directly interact to generate a joint postdictive illusory perception (as illustrated schematically in Fig. 2, top). This second type of crossmodal postdiction, referred to herein as *Crossmodal Postdiction with Emergent Illusory Perception*, has been shown to occur with Intermodal Apparent Motion (in which Apparent Motion occurs with two stimuli of different modalities) [21] and the Burst Lag Auditory–Visual Illusion (a crossmodal version of the Visual Flash Lag Illusion) [1]. For example, a schematic diagram of Intermodal Apparent Motion is provided in Fig. 2 (bottom). In Intermodal Apparent Motion, brief spatially separated stimuli from two modalities (such as audition and vision) are interleaved in time. As a consequence, the participant perceives smooth illusory motion among the crossmodal stimuli locations that is postdictively generated *across* the senses. The key factor in this type of postdiction is that the postdictive influence (as shown by the dashed arrows in Fig. 2) bridges from one modality to another modality (as illustrated by different colors in Fig. 2).

The third type of crossmodal postdictive illusion can be referred to as *Crossmodal Postdiction with Crossmodal Illusory Perception*, and occurs when real stimuli interact with an illusory perception generated by a crossmodal trigger (such as the illusory flash in the Double Flash Illusion [44]) to generate postdiction (Fig. 3). This type of postdiction requires perception of a crossmodally generated illusion. Examples include the Illusory Audiovisual Rabbit Illusion and the Invisible Audiovisual Rabbit Illusion. For example, the Illusory Audiovisual Rabbit Illusion stimuli in sequence are beep–flash, beep, beep–flash (Fig. 3). An illusory flash is often perceived to be paired with the lone beep. The spatial and temporal properties of the stimuli (beep–flash and then beep) are the same as in the Double Flash Illusion, and therefore the *perception* of the illusory flash is likely generated similarly. However, unlike in the Double Flash Illusion where the illusory flash is collocated with the real flash, the illusory flash in the Audiovisual Rabbit Illusion is reported to be spatially displaced and located between the first real flash (*e.g.*, on the left) and second real flash (*e.g.*, on the right). The real second flash therefore influences the *location* of the visual illusory flash postdictively. This postdictive interaction between a real stimulus and an illusory perception is characteristic of the third type of crossmodal postdiction.

In this review article, we will also discuss the impact of crossmodal interactions on the modeling of postdictive sensory processing. In order to accommodate the unique aspects of crossmodal postdiction, unimodal postdiction models must be adapted to reflect cross-sensory differences in sensory transmission and transduction, cross-sensory differences in cortical sensory processing rates, and the addition of sensory processing in multisensory cortical regions.
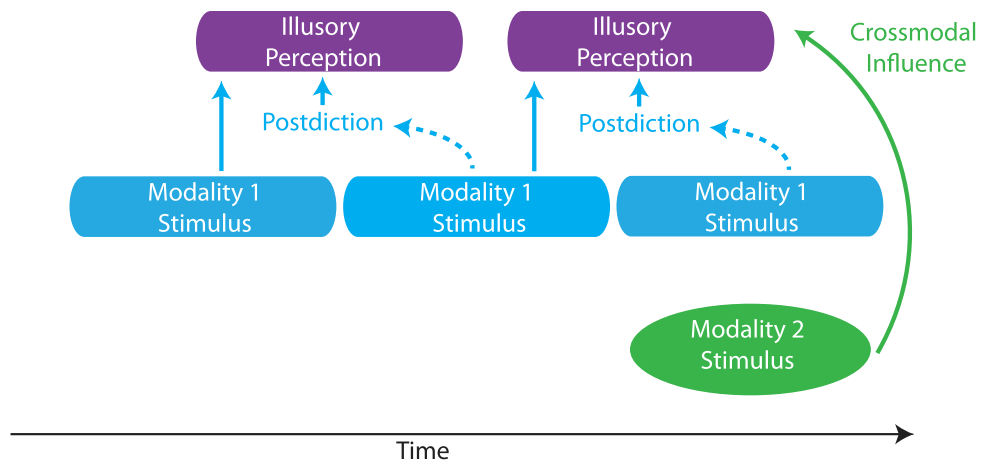
Each of the three types of crossmodal postdiction are treated in subsequent sections. Unimodal Postdiction with Crossmodal Influence is discussed in Section 2, Crossmodal Postdiction with Emergent Illusory Perception is discussed in Section 3, and Crossmodal Postdiction with Crossmodal Illusory Perception is discussed in Section 4. In Section 5, we describe several of the previous unisensory models for postdiction, and the modifications and additions to these
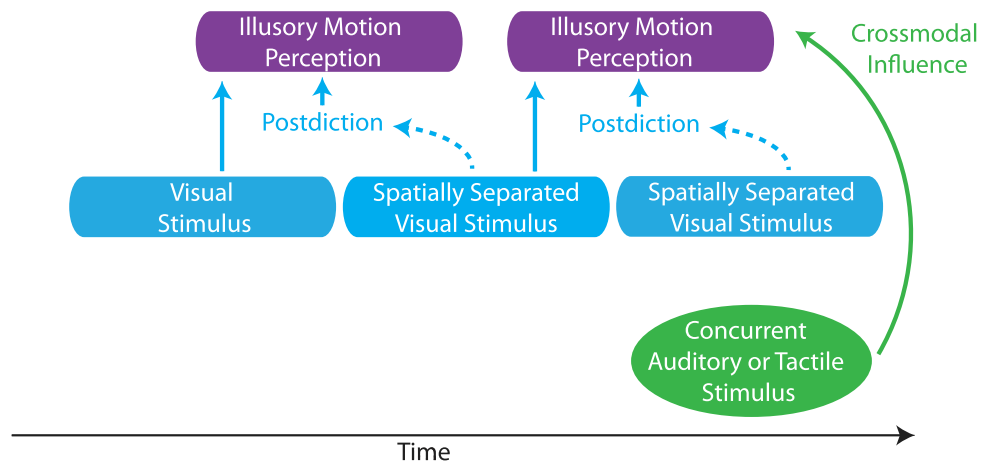
Unimodal Postdiction with Crossmodal Influence

Unimodal Postdiction
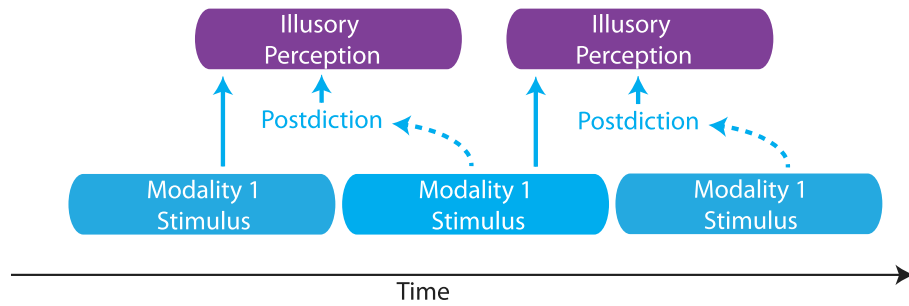
Unimodal Postdiction with Crossmodal Influence

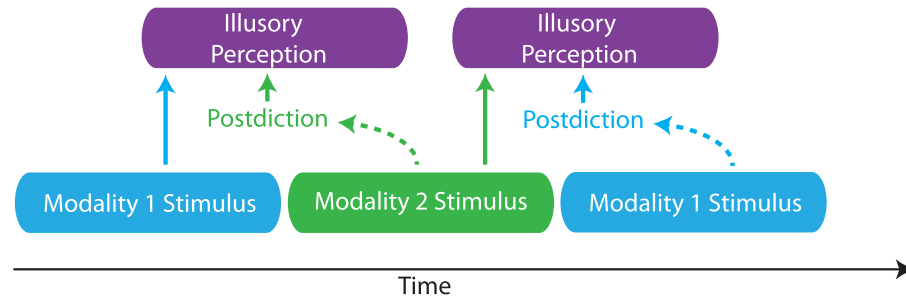Example: Auditory and Tactile Influence on Visual Apparent Motion

Figure 1. Unimodal postdiction with crossmodal influence. These schematic diagrams illustrate unimodal postdiction dynamics, and how crossmodal stimuli can strengthen or weaken otherwise unimodal postdiction.

models that are necessary to fit the crossmodal postdiction case. Avenues of opportunity for future research on crossmodal postdiction, as well as current and future applications of crossmodal postdiction in emerging technologies, are discussed in Section 6.

Over the past few decades, key discoveries in perceptual neuroscience have demonstrated that the senses are densely interconnected with extensive crosstalk and feedback at every stage of sensory processing. Postdictive sensory integration is no exception. The extension of postdiction

Crossmodal Postdiction with Emergent Illusory Perception
General Diagram

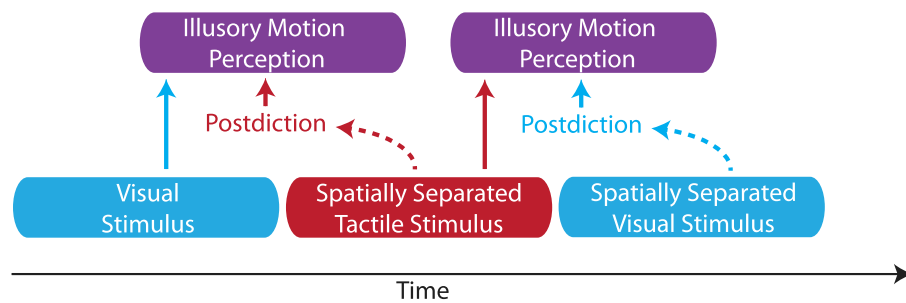Example: Intermodal Apparent Motion

Figure 2. Crossmodal postdiction with emergent illusory perception. These schematic diagrams illustrate the temporal sequence of crossmodal illusions that are characterized by a postdictive illusion that derives from crossmodal stimuli.

into the multimodal domain was initiated during this same time frame, and now continues to expand with numerous new investigations of cross-sensory interactions. In this review article, we will explore the implications of these crossmodal postdictive illusions and interactions on perceptual processing in the brain.

2. UNIMODAL POSTDICTION WITH CROSSMODAL INFLUENCE

2.1 Overview
In this section of the article, we will discuss traditional postdictive effects in one modality that are influenced by the presence of a stimulus in another modality. In effect, these illusions involve unimodal postdiction with a crossmodal influence on the perception of the unimodal postdictive effect. Therefore, we will first introduce the unimodal postdictive effect for each of these crossmodal postdictive illusions, and then discuss how this unimodal effect was modified by the addition of a stimulus from another modality.

2.2 Auditory and Tactile Influence on Visual Apparent Motion
Visual Apparent Motion occurs when two spatially separated visual flashes are presented sequentially [24, 25, 40, 53]. Rather than just perceiving two individual flashes in two locations across the visual field, participants report perceiving smooth continuous motion between the individual flashes.

The Visual Apparent Motion Illusion occurs with equal strength if the direction of the flashes is randomized across trials, thereby preventing prediction of the flash locations. Therefore, this observation leads to the conclusion that the perception of the illusory motion between the flashes is postdictive, with the second flash triggering the illusory perception of motion between the two flashes. Neuroimaging studies have also indicated that perceived illusory motion between flashes generates corresponding activation along the perceived path in early visual cortex, which is likely due to feedback from higher regions such as MT or V5 [35]. Apparent Motion has also been shown to occur in the auditory domain (a sequence of beeps) [52], as well as in the tactile domain (a sequence of taps) [7, 27, 28].

Crossmodal (*e.g.*, auditory) stimuli have been added before, concurrent with, and after visual flashes designed to induce apparent motion, and have been shown to influence the perceived illusory motion [16]. We should note that the original publications on auditory and tactile influence on visual apparent motion did not explicitly identify these crossmodal illusions as incorporating postdiction. In a broad sense, though, these illusions do incorporate postdictive processing that is strengthened or weakened by crossmodal stimuli, and therefore represent additions to the crossmodal postdictive illusion repertoire.

As an example in the auditory–visual domain, the addition of a clicking sound presented at the same time as each of two sequential, spatially separated visual flashes

Crossmodal Postdiction with Crossmodal Illusory Perception
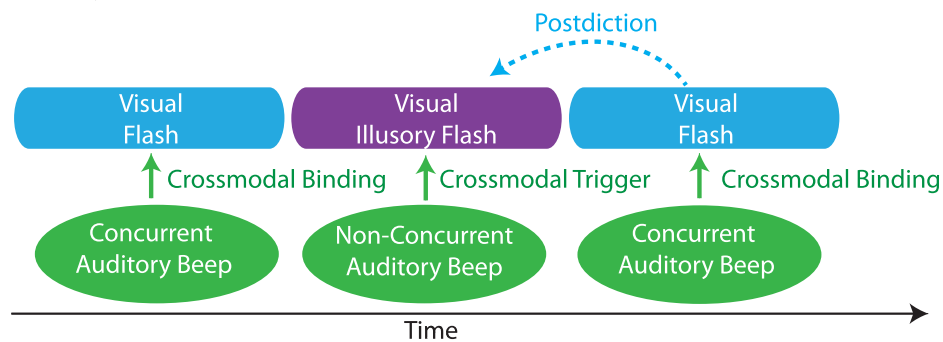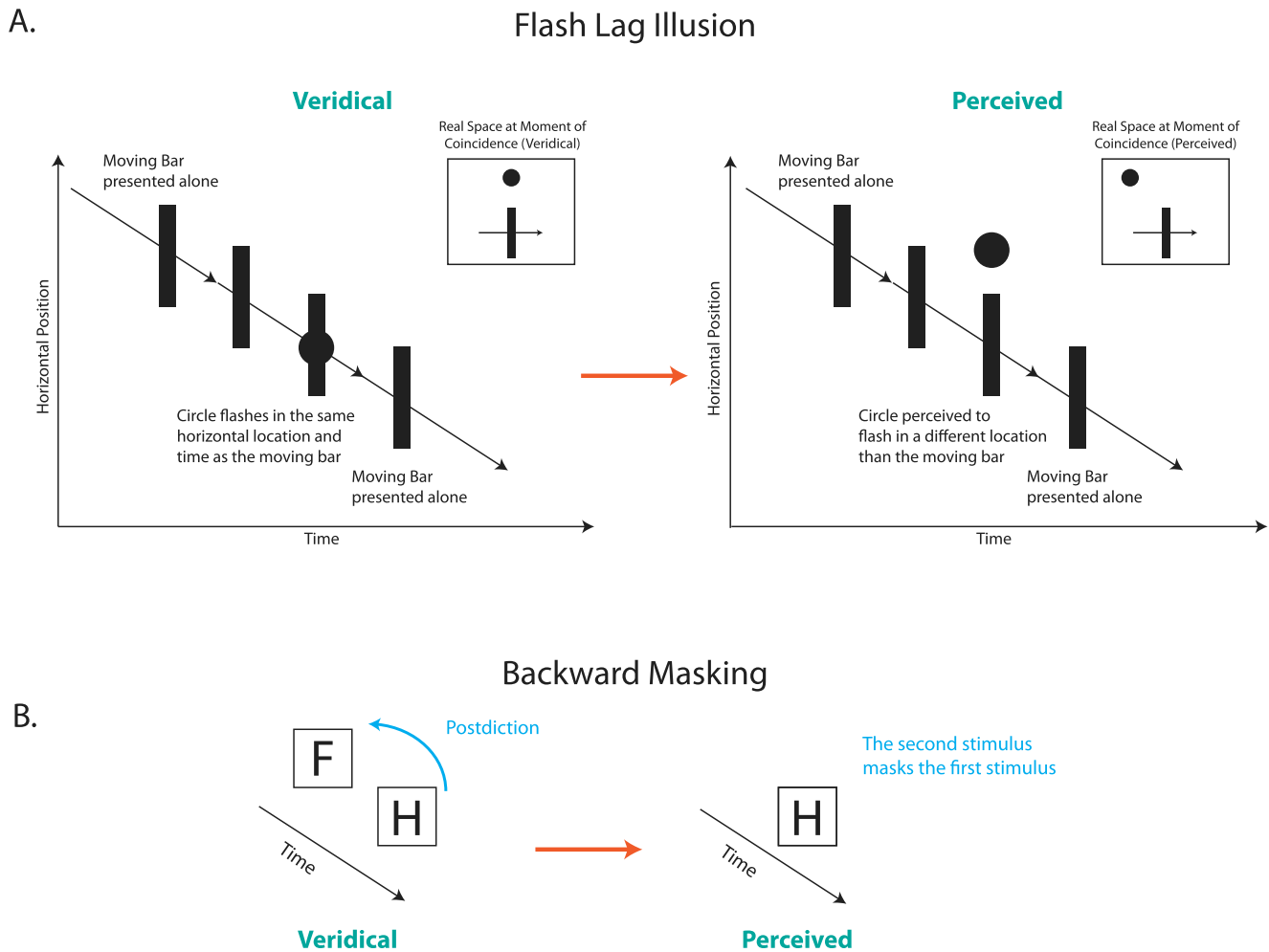General Diagram
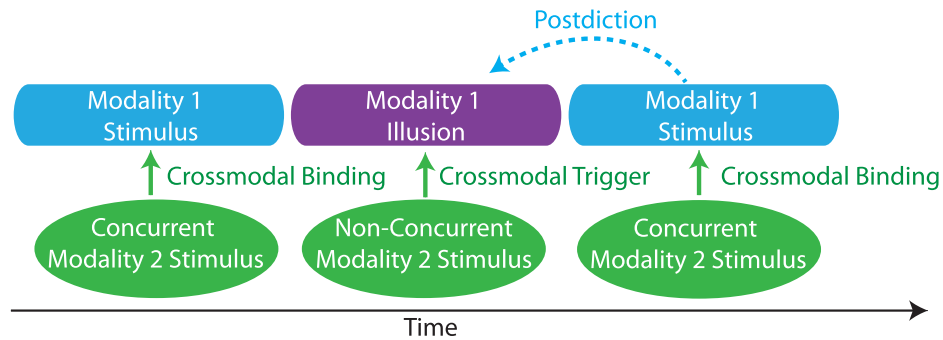
Example: Illusory Audiovisual Rabbit

Figure 3. Crossmodal postdiction with crossmodal illusory perception. These schematic diagrams illustrate crossmodal postdiction that is characterized by an interplay between real and illusory stimuli.

weakened the perception of visual apparent motion relative to that perceived with visual flashes presented alone [16]. By way of contrast, a click presented in between the two successive flashes strengthened the visual perceived apparent motion [16]. Therefore, the additional auditory information appears to be integrated with the visual postdictive processing to strengthen or weaken postdictive illusory motion by reinforcing the perception of either discrete flashes or continuous motion.

Soto-Faraco *et al.* and Sanabria *et al.* performed auditory apparent motion experiments with sequential, spatially separated auditory beeps that generate the perception of motion [42, 50]. In their experiments, flashes or taps were presented simultaneously with the auditory beeps, either in the same or opposite spatial order. They observed that the percentage of correct identification of the auditory apparent motion direction was reduced when spatially incongruent crossmodal stimuli were presented relative to when spatially congruent crossmodal stimuli were presented. This effect is likely due to the influence of visual stimulus location on the perception of auditory stimulus location (the Ventriloquist Effect). When the auditory locations are shifted closer together spatially by the Ventriloquist Effect, it becomes more difficult to identify the spatial shift in beep location, and therefore (postdictive) apparent motion perception is diminished.

Freeman and Driver modified the Visual Apparent Motion Effect by presenting three visual flashes, the first on the left, the second on the right, and the third on the left (at the same location as the first flash) [13]. Von Grünau had previously shown that when the time interval between the flashes is adjusted such that there is a shorter time interval between the first and second flashes than the second and third flashes, the left-to-right apparent motion perception dominates over the right-to-left apparent motion perception [20]. Freeman and Driver presented the three flashes with equal intervals, but added auditory beeps that were lagging or leading the visual stimuli in time (all of the beeps were centrally located). Through the influence of the Temporal Ventriloquism Effect [56] (auditory stimuli influencing the perceived timing of visual stimuli), the auditory beeps caused the visual flashes to appear closer together or farther apart in time, and therefore strengthened the left-to-right or right-to-left movement (depending on the beep timing). Therefore, auditory stimuli influenced the perceived timing of the visual stimuli, which strengthened or weakened the postdictive influences of the second and third flashes. This study also showed that visual stimuli with substantial spatial separations (up to 14 degrees in visual angle) and long temporal delays (over 300 ms) can still generate illusory apparent motion by means of postdiction. Furthermore, auditory manipulation of the

A.

Flash Lag Illusion

Veridical

Moving Bar presented alone

Circle flashes in the same horizontal location and time as the moving bar

Moving Bar presented alone

Horizontal Position

Time

Real Space at Moment of Coincidence (Veridical)

Perceived

Moving Bar presented alone

Circle perceived to flash in a different location than the moving bar

Moving Bar presented alone

Horizontal Position

Time

Real Space at Moment of Coincidence (Perceived)

Backward Masking

B.

F

H

Postdiction

Time

The second stimulus masks the first stimulus

H

Time

Veridical

Perceived

Figure 4. The flash lag illusion and backward masking. These two sets of diagrams highlight the spatial and temporal features of the Flash Lag Illusion and Visual Backward Masking. The left stimulus representation for each illusion shows the stimuli as presented in the external environment (*i.e.*, veridical). The right stimulus representation shows the stimuli as they are perceived by the participant due to postdictive processing (*i.e.*, as perceived).

visual postdiction effect was also shown to occur over this relatively wide range of spatial and temporal scales.

2.3 *Auditory Influence on the Visual Flash Lag Illusion*

The Visual Flash Lag Illusion occurs when a visual stimulus moves smoothly across the visual field, and then a single flash is presented that is vertically aligned with the moving stimulus (Figure 4A) [31, 36]. Even though the flashed and moving stimuli are aligned when the flash is presented, the observer perceives the visual moving stimulus to be ahead of the single flash at the moment of coincidence. This visual phenomenon has been named the Flash Lag Illusion, as the location of the flash is perceived to lag the location of the smoothly moving object.

The smoothly moving object has been hypothesized to be projected forward in perception to compensate for the delay in perceptual processing (a form of prediction) [36]. By way of contrast, the flashed stimulus does not have a predicted perceptual evolution (path), and therefore lags the smoothly moving object, which does.

However, this predictive model of the Visual Flash Lag Illusion is not supported by the two cases of flash termination and flash initiation of movement. If the smoothly moving object abruptly stops when the coincident flash is presented (*i.e.*, the flash terminated case), no flash lag is perceived. But if the smoothly moving object instead initiates movement when the coincident flash is presented (*i.e.*, the flash initiated case), flash lag *is* perceived. Therefore, the after-the-fact smooth movement of the visual stimulus influences the relative perceived positions of the stationary and moving visual stimuli backward in time [10].

Vroomen and de Gelder also investigated the Visual Flash Lag Illusion, and found that they could reduce the magnitude of the effect by adding a brief auditory tone (beep) coincident with the visual flashed stimulus [56]. In their analysis, the auditory burst acted in part to heighten attention, thereby speeding up the processing of the flash and causing it to appear slightly earlier than it appeared without the auditory burst, thereby decreasing the difference between the relative perceived positions of the stationary and moving flashes [48]. In addition, when the auditory burst was

presented *before* the flash, the Temporal Ventriloquism Effect likely caused the flash to appear even earlier (closer in time to that of the beep), which also decreased the reported flash lag (*i.e.*, the moving stimulus appears shifted less relative to the stationary flash because the flash was perceived earlier rather than later in time). When the sound burst was presented *after* the flash, temporal ventriloquism likely caused the flash to appear later, thereby increasing the flash lag effect. In their experiments, these two effects were borne out as hypothesized. Unfortunately, the flash initiated and flash terminated cases have not yet been investigated, which limits the implications of this study for crossmodal postdiction.

2.4 *Auditory Influence on Backward Masking*
The Visual Backward Masking Effect occurs when a visual masking stimulus is presented after a visual target stimulus, and thereby suppresses the visibility of the earlier visual stimulus (as shown schematically in Fig. 4B using the letter masking paradigm) [8]. Visual backward masking can be altered by the addition of an auditory stimulus that is coincident with the visual target stimulus, which has the effect of dissociating the two visual stimuli into two separate visual events rather than one combined visual percept [8]. As such, the inherently unimodal postdictive backward masking effect is reduced, as perception of the target is restored and not suppressed. Similarly, an auditory burst coincident with the mask also dissociates the two visual stimuli, once again reducing the backward masking effect. Auditory modulation of the Visual Backward Masking Effect thus indicates a modulation of the postdictive influence of the mask on the target by means of the auditory stimulus, an example of Unimodal Postdiction with Crossmodal Influence (Fig. 1).

3. CROSSMODAL POSTDICTION WITH EMERGENT ILLUSORY PERCEPTION
3.1 *Overview*
Postdictive illusions can also occur when two or more stimuli are presented in different modalities, but are bound together by common properties such as spatial location and coincidence (or sequence) in time. In the two examples below, postdictive apparent (illusory) motion and temporal lag are perceived across the senses (between somatosensation and vision, or between audition and vision, respectively) when the crossmodal stimuli are either collocated in space or coincident in time (including rapidly sequenced), and are thereby perceptually bound.

3.2 *Intermodal Apparent Motion*
Visual Apparent Motion was described in the previous section, and occurs unimodally (for example) when multiple flashed stimuli are presented sequentially in displaced locations. The resulting postdictive illusory perception is one of smooth visual motion between each of the discrete flashes.

Apparent motion can also occur crossmodally when the sequential stimuli are presented (for example) as interleaved flashes and taps [21]. Harrar and colleagues attached light emitting diodes (LEDs) and tactile stimulators to two of

a participant's fingertips (one on each hand) that were spatially separated by distances ranging from 2 cm to 56 cm, corresponding to a range of visual angles between 2.3° and 68.1°, respectively. They found that when participants fixed their gaze between their fingertips, visual–visual apparent motion was perceived when the LEDs were flashed sequentially with time delays ranging from 40 ms to 300 ms, and tactile–tactile apparent motion was perceived when the tactile stimulators were pulsed sequentially over the same range of time delays. In both cases, the preferred time delay between the stimuli (*i.e.*, the time delay that elicited the strongest perceived apparent motion effect) varied with finger (hand) separation. Visual–tactile apparent motion was also perceived between the two fingers, but intriguingly did not vary in preferred time delay over the range of finger separations tested. Harrar *et al.* consequently hypothesized that the mechanism for Intermodal Apparent Motion (visual–tactile, as shown schematically in Fig. 2, bottom) was different than the mechanism for Intramodal Apparent Motion, at least for stimulus onset asynchronies less than 200 ms or so.

Intermodal Apparent Motion is thus an example of Crossmodal Postdiction with Emergent Illusory Perception, with spatially separated stimuli in two different modalities that are interleaved in sequence as illustrated in Fig. 2. In the specific example shown in Fig. 2 (bottom), the LED flash is assumed to be Modality 1, and the tactile pulse is assumed to be Modality 2. Consider, for example, the first two stimuli. The illusion of apparent motion emerges from the postdictive influence of the spatially separated but temporally sequenced tactile pulse on the perception of the visual flash. The roles of the LED flash and tactile pulse can also be reversed to generate an equivalent example of Intermodal Apparent Motion, as illustrated by the second and third stimuli in the sequence.

3.3 *Burst Lag Auditory–Visual Illusion*
In the Visual Flash Lag Illusion (as shown schematically in Fig. 4A), a delay or lag in the timing of a flashed visual stimulus is perceived relative to that of a smoothly moving visual stimulus, leading to the perception of a spatial misalignment of the two stimuli at the moment determined by the flash (as discussed in detail in Section 2.3 above).

The Visual Flash Lag Effect has also been shown to occur crossmodally (the Burst Lag Auditory–Visual Effect) when the smooth moving stimulus is visual and the brief stimulus is auditory, and *vice versa* [1]. Similar to the case of the visual stimuli in the Visual Flash Lag Effect described in Section 2.3 above, the auditory stimuli are spatially localized and designed to either sweep across space, or to be presented momentarily as short beeps that are spatially co-located with moving visual objects. Alais and Burr found that crossmodal burst lags (auditory–visual and visual–auditory) were smaller in lag magnitude than auditory–auditory burst lags, but larger in magnitude than visual–visual flash lags. Similar to the original visual–visual flash lags, the crossmodal burst lags were also perceived to have no spatial shift (*i.e.*,

no temporal lag) in the burst terminated case, and to have a spatial shift (*i.e.*, a temporal lag) in the burst initiated case.

These crossmodal burst lag results support the interpretation that they involve crossmodal postdiction, although other models for the crossmodal burst lag effect have not yet been fully ruled out, including both the temporal averaging and positional sampling models [1]. Both the temporal averaging and positional sampling models rely on longer integration periods for audition relative to vision in order to enable later visual stimuli to influence earlier auditory stimuli. These two models in our view would also qualify as postdictive under the general definition used in this article, although they were not originally labeled as such.

## 4. CROSSMODAL POSTDICTION WITH CROSSMODAL ILLUSORY PERCEPTION

4.1 *Overview*

The combination of a crossmodal illusion, such as the Double Flash Illusion [44], with postdictive processing was discovered quite recently [51], and indicates that postdictive processing can act not only on real stimuli but also on *illusory* perceptions. For example, in the Illusory Audiovisual Rabbit Illusion (as shown schematically in Figure 5), the perceived *location* of an *illusory* flash induced crossmodally earlier in time can be postdictively influenced by a later *real* beep–flash pair. Additionally, in the Invisible Audiovisual Rabbit Illusion, a real visual flash can be crossmodally and postdictively suppressed by a following beep–flash pair, thereby creating the illusion of only two instead of three perceived flashes. Not only do these illusions indicate the versatility of postdictive processing, but they have also begun to suggest that postdiction may be effective over a longer than previously expected temporal window of influence.

4.2 *Illusory Audiovisual Rabbit Illusion*

In the Illusory Audiovisual Rabbit Illusion, a beep–flash pair is first presented, followed by a single beep and subsequently by a second beep–flash pair (Fig. 5). The two real flashes are presented peripherally as well as spatially separated, with the first flash in the center and the second flash on the right (or the first flash in the center and the second flash on the left) below a central fixation cross. The auditory tones (beeps) are all centrally located. An illusory flash is reported by most participants to be perceived between the two real flashes in time. The *perception* of the illusory flash is triggered by the second beep crossmodally, as in the Double Flash Illusion. In addition, the illusory flash is perceived by most participants to be spatially located between the first real flash and the second (final) real flash. Furthermore, the shifted position of the illusory flash occurs even when the direction of the visual flash presentation is randomized between left-to-right and right-to-left across trials. Evidently the *location* of the illusory flash is influenced by the position of the final visual flash postdictively [51].

The Illusory Audiovisual Rabbit Illusion is named after and has similarities in structure to the Cutaneous Rabbit Illusion, a frequently researched unimodal postdictive
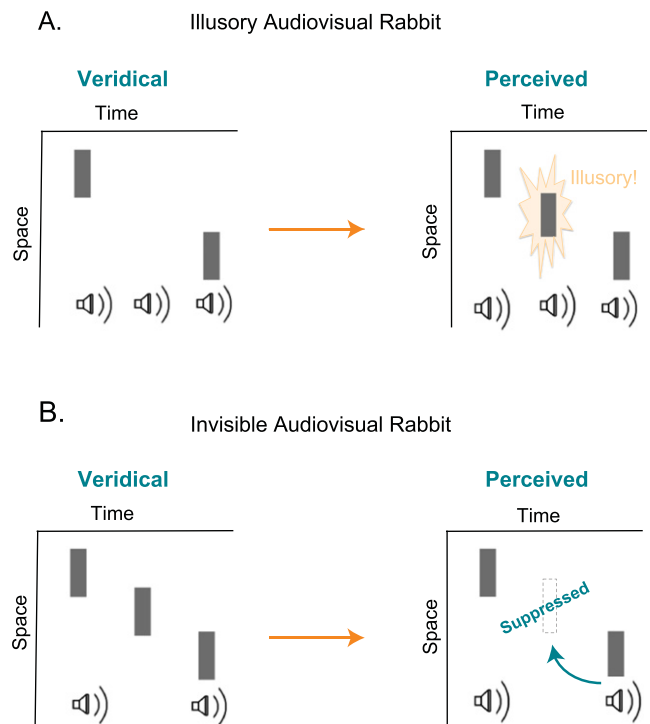
Figure 5. The Illusory Audiovisual Rabbit and Invisible Audiovisual Rabbit Illusions [51]. These schematic diagrams show plots of time and space, with flashes represented as vertical gray bars and beeps represented as speaker symbols. The diagrams on the left (Veridical) indicate the real stimuli presented to the participant. The diagrams on the right (Perceived) indicate the illusion as perceived by the majority of participants.

illusion that was discovered by Geldard and Sherrick in 1972 [14]. The Cutaneous Rabbit Illusion consists of several quick taps on the forearm, with the first and second taps in the same location and the third tap in a laterally shifted location. Participants report perceiving three taps, with the second tap not collocated with the first, but rather shifted toward the location of the third tap. In addition to the tactile alone illusion, both a unisensory visual version and a unisensory auditory version of the Cutaneous Rabbit Illusion have been shown to induce this spatial shift of perceived stimulus location [15, 49]. The multisensory Illusory Audiovisual Rabbit Illusion exhibits a similar hopping or saltatory dynamic, like the taps of the Cutaneous Rabbit Illusion, but instead is based on a sequence of visual flashes that appear to move in steps across the visual field when accompanied by auditory tones. Both the Cutaneous Rabbit and the Audiovisual Rabbit illusions have opened up new possibilities for a deeper understanding of the role of postdiction in sensory perception. For example, the unimodal models developed for the Cutaneous Rabbit (such as the low-speed prior model by Goldreich and Tong [17]) potentially provide the basis for multimodal models of the Audiovisual Rabbit Illusion, as described in Ref. [51].

The Illusory Audiovisual Rabbit Illusion is the first crossmodal illusion to demonstrate crossmodal triggering of an illusory flash that is postdictively modified by a later

real flash. In effect, the Illusory Audiovisual Rabbit Illusion includes a postdictive illusion that combines with or affects a crossmodal illusion. Therefore, the postdictive processing by necessity must either follow or be interleaved with the crossmodal interaction that generates the illusory flash. This sequential or interleaved combination of postdictive and crossmodal illusions hints at a potentially longer time window for postdictive processing than previously thought, as will be discussed further in Section 4.4 below.

4.3 *Invisible Audiovisual Rabbit Illusion*
The Invisible Audiovisual Rabbit Illusion is similar to the Illusory Audiovisual Rabbit Illusion; however, instead of a lone central beep, the Invisible Rabbit involves a lone central flash (Fig. 5) [51]. In this case, the real central flash is perceived to be *suppressed* by the crossmodal stimuli (beeps and flashes) that precede and follow it. By comparing the primary illusion with multiple control stimuli, the suppression of the second real flash was demonstrated to be postdictively generated in part by the final beep–flash pair (following the second real flash) [51]. In this case, it is likely that the crossmodal binding of the beeps and flashes (*i.e.*, crossmodal interaction) either precedes or interleaves with the postdictive suppression of the second real flash in the pipeline of sensory processing.

4.4 *The Extended Illusory Audiovisual Rabbit Illusion*
The time dynamics of the Illusory Audiovisual Rabbit Illusion have been further investigated to determine the relative influences of prediction and postdiction on the central illusory flash location [54]. In this follow-up study, the temporal delay between the lone (second) beep and the final beep–flash pair was extended from 58 ms in the original implementation of the illusion to 100, 300, 500, 700, and 900 ms. Surprisingly, the illusory flash was observed to be significantly shifted toward the final beep–flash pair even with delays between the stimuli up to 500 ms. This research suggests that the subconscious window within which perceptual postdiction can occur is able to extend up to 500 ms or more.

5. MODELS OF CROSSMODAL POSTDICTION
5.1 *Overview*
Several types of models have been proposed for unimodal postdiction, including Bayesian models [17], neuropsychological models [3, 47], and consciousness based models [10, 22, 29]. Of course, these three types of models are not mutually exclusive. In this section, we will explore the adaptation of several neuropsychological models to crossmodal postdiction, with a particular emphasis on differences in early sensory processing and transduction among the senses, as well as the addition of multisensory regions to the sensory processing pipeline.

In the following, we describe three conceptual neuropsychological models that permit later stimuli to be integrated with earlier stimuli. These three models are the Catch Up Model, the Reentry Model, and the Different Pathways

Model [47]. We have adapted each of these models to multimodal postdiction (as shown schematically in Figure 6). We will first describe the unimodal versions of the models, and then discuss the adaptations made to the models in order to allow them to be applied to crossmodal postdiction.

In the neuropsychological models discussed below, two different time domains will be evaluated. Brain Time is the time course of processing within the brain itself, and is defined by when a stimulus is processed in a given brain region and for how long relative to the timing of other events within the brain. External Time refers to the sequence of events in the environment, where stimuli are presented and have given onset times and durations as measurable by external instruments. The interplay between these two temporal domains represents an important element in the models discussed below.

5.2 *The Catch Up Model*
The Catch Up Model postulates that a second stimulus can "catch up" to a first stimulus if the first stimulus is processed in a given brain region long enough that the second stimulus arrives during the processing. In effect, the Catch Up Model takes into account the time required to process sensory stimuli before conscious awareness (due to both feedforward and local recurrent processes), which permits a temporal overlap in the stimulus processing periods. This temporal overlap could thereby allow a later stimulus to impact the perception of an earlier stimulus. In the Catch Up Model, Brain Time, which represents the period of stimulus processing in the brain, is differentiated from External Time, which is the time of presentation of stimuli to the participant. Even though the overlap of the stimuli does not occur in real External Time, it does occur during the sensory processing period in Brain Time. We should also note that this model is entirely feedforward in the brain (with the possible exception of local recurrent processes as noted above), and requires that the stimuli are presented in External Time such that the neural processing of the first is still incomplete before the second stimulus reaches the same brain region, enabling a temporal overlap in neural processing.

The Catch Up Model (as shown schematically in Fig. 6, top) can be applied to crossmodal stimuli in which one of the stimuli is in Modality 1 and one of the stimuli is in Modality 2 (as in the Crossmodal Postdiction with Emergent Illusory Perception category that was discussed in Section 3 above).

If this is the case, then the two modalities are most likely to be integrated or interact in a multisensory brain region such as superior temporal sulcus (STS), among others [4, 32]. The stimuli from two different modalities might also be processed in one or more multisensory regions for unequal amounts of time, such that these differences in sensory processing duration increase the likelihood for processing overlap to occur.
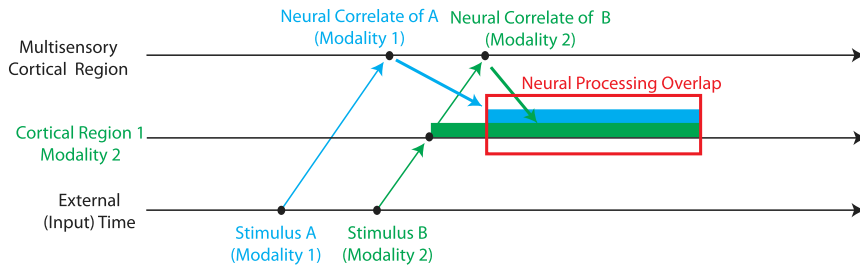
5.3 *The Reentry Model*
The Reentry Model takes advantage of the extensive feedback that exists among cortical regions to enable postdiction
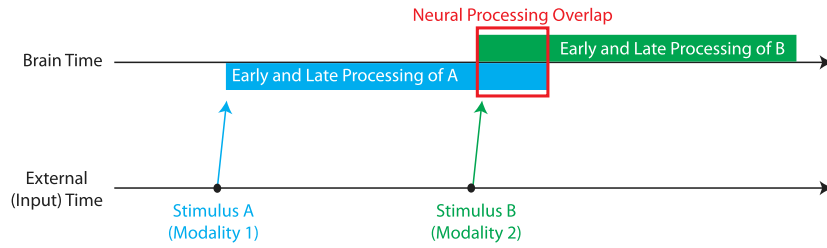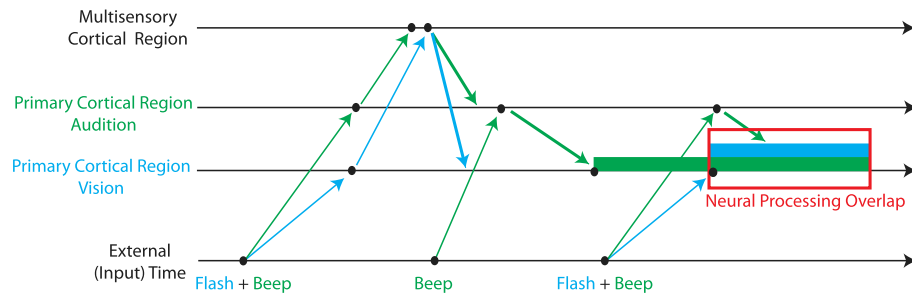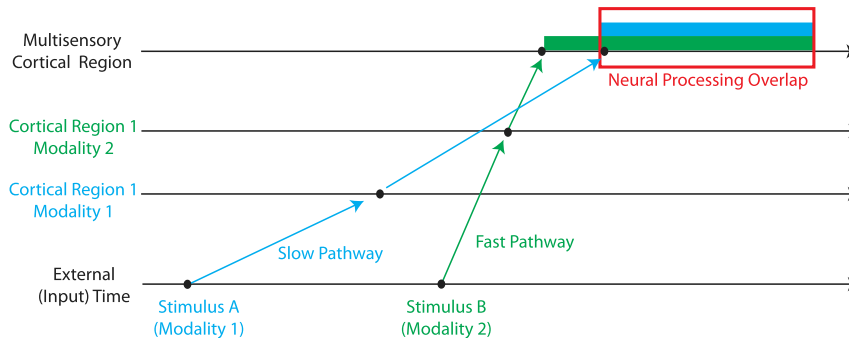
Figure 6. Crossmodal postdiction models. This figure presents schematic diagrams of three crossmodal postdiction neural models, adapted from unisensory postdiction models [47] to apply to crossmodal postdiction.

(Fig. 6, middle). In the unisensory version of the Reentry Model, the first stimulus (Stimulus A) may progress through the sensory processing hierarchy and then be fed back to an earlier (lower) cortical region in the hierarchy. By the time that Stimulus A feedback arrives in this earlier sensory region, Stimulus B is assumed to have already started its processing, allowing for both stimuli to be integrated at least in part and progress to subsequent processing regions together.

In the multisensory cortical network, feedback plays a critical role in crossmodal integration, with multisensory regions often feeding signals back to unisensory cortical regions [33]. This feedback process can provide integrated sensory information (such as processed and bound auditory–visual signals) to unisensory regions. A key feature of this model is that feedback typically requires more time than feedforward processing, and given the time constraints on postdictive processing, feedback from multisensory regions to primary sensory regions must occur relatively rapidly.

For example, a multisensory Reentry Model (Fig. 6, middle) could have Stimulus A in Modality 1 progressing to a multisensory region and then being fed back to the early cortical region of Modality 2. The processing of Stimulus B in Modality 2 could also be ongoing at the same time in the early cortical region of Modality 2. In this way, the fed back component of the multisensory-processed neural correlate of Stimulus A could be processed along with both the unimodally processed neural correlate of Stimulus B and the fed back component of the multisensory-processed neural correlate of Stimulus B, such that an integrated multimodal signal then progresses up to a multisensory cortical region for further processing.

The Double Flash Illusion is an interesting case of rapid crossmodal integration that could involve feedback. The Double Flash Illusion presents one beep–flash pair followed by another (lone) beep. Participants often perceive two flashes, and the perception of the illusory flash is known to be correlated with activation in early visual cortex [33, 45]. It has been previously hypothesized that the early visual cortical activation generating the perception of the illusory flash is triggered by a direct connection between primary auditory cortex and early visual cortex [33].

As a consequence, we further hypothesize that the first beep–flash pair is processed and bound in multisensory regions, and then this information is fed back in part to both primary auditory cortex and early visual cortex in order to prime these regions for future multisensory stimuli. When the lone beep is then presented, primary auditory cortex is primed to trigger activation in early visual cortex rapidly.

This cascade of feedforward and feedback interactions also likely occurs in the Illusory Audiovisual Rabbit Illusion, which similarly involves an illusory flash that is triggered by an auditory beep, as in the Double Flash Illusion. If so, this feedforward and feedback cascade for the Audiovisual Rabbit's crossmodal interaction may slow the processing of the illusory flash enough that the illusory flash and final beep–flash pair can overlap in Brain Time. As a consequence, the illusory flash and final beep–flash pair could be integrated and modified in early visual regions and subsequent multisensory regions to generate the perception of an illusory flash that is shifted toward the location of the final flash.

This feedforward and feedback dynamic of the Illusory Audiovisual Rabbit Illusion could, therefore, be an example of the Reentry Model of postdictive processing (as shown schematically in Fig. 6, middle). In Fig. 6, the first beep–flash pair initially stimulates their respective primary cortical regions and then a multisensory cortical region, with feedback thereafter priming both primary cortical regions. The illusory flash that is then triggered by the lone auditory beep is delayed both by this priming from the first beep–flash pair and also by initial processing of the beep in auditory cortex. Therefore, when the lone beep triggers the perception of the illusory flash in primary visual cortex (labeled Primary Cortical Region, Vision in Fig. 6, middle), it may have been delayed long enough to overlap in processing with, and therefore be modified by, the second beep–flash pair in visual cortex, and perhaps further in a subsequent multisensory region.

It is of considerable interest to postulate how the spatiotemporal dynamics of the illusory flash evolves in visual cortex. The illusory flash, generated by the first beep–flash pair and following lone beep, could initially be collocated within the receptive field of the first flash, and then be modified in perceived location by the second beep–flash pair. Alternatively, the illusory flash could be initiated within a much wider receptive field that is coalesced to its final perceptual location by the final beep–flash pair. Resolution of these alternatives, among others, awaits further study.

Overall, the Reentry Model can take many forms, as the sensory neural network is capable of many combinations of feedforward and feedback cascades that are based on the type of multimodal processing involved. The key overarching aspect of the Reentry Model is the presence of multisensory cortical feedback in the multimodal processing pipeline, which allows for delays in processing that thereby permit postdiction to occur across the senses.

5.4 The Different Pathways Model

The Different Pathways Model assumes that one sensory pathway processes sensory information significantly faster than another sensory pathway [19]. If Stimulus A in Modality 1 is presented first, but is processed in part by a slower pathway, it could cross paths and be integrated with the later Stimulus B in Modality 2, if Stimulus B is processed in part by a faster pathway (as shown schematically in Fig. 6, bottom).

This model suggests many possibilities for crossmodal postdiction that have Stimulus A and Stimulus B processed in different modalities. As in the case of thunder and lightning, each modality is likely to have a different latency for transmission of the auditory sound or visual signal from the causal phenomenon in the environment to the sensors (ears and eyes, respectively) in the head. Furthermore, the

act of physiological transduction in which the environmental signal (pressure waves or photons, respectively) is converted into a neural signal is different for each modality. Finally, the pathways in the brain for both senses are also different, and also have unequal processing speeds (*e.g.*, vision tends to be slower than audition or somatosensation). Therefore, the Different Pathways Model of postdiction can be applied to crossmodal postdiction easily by simply taking into account the different routes of environmental, sensory, and neural processing for the different senses.

Of course, the Different Pathways Model in combination with sensory differences can only play a role when the first stimulus presented (and therefore the first modality used) is processed more slowly than the second modality presented. If such stimuli are presented in the opposite order, then the differences in sensory processing will work *against* the postdictive processing of crossmodal stimuli, and in favor of perceptual segregation. Since at least a few of the crossmodal postdiction illusions (such as Intermodal Apparent Motion) occur with presentation of either modality first, the Different Pathways model cannot explain all instances of crossmodal postdiction.

5.5 *Higher Order Cortical Processes*
The models of postdictive interactions described in the previous sections of this review apply in particular to perception over relatively short time scales of a few ms up to approximately 500 ms. As such, they provide possible mechanisms for most common multisensory illusions.

Postdictive interactions, however, including these multisensory illusions, can be affected to a certain extent by a number of higher order cortical processes, such as prior brain state (including oscillations), emotional state, previous experiences with multisensory illusions, instructions received prior to viewing the illusions, directed attention, and participant expectations of the possible outcomes that operate primarily at the subconscious level [26, 47].

In addition, participant reporting of trial outcomes in multisensory illusions is inherently both cognitive and conscious. For example, in the Illusory Audiovisual Rabbit Illusion, participants are asked to report the number of flashes that they perceived, as well as the location of the second flash if three flashes were perceived, after each trial. This reporting process clearly involves memory, as well as the determination of the sequence of a set of multisensory stimuli. As has been demonstrated previously, sequential (temporal order) determination does not derive solely from delays in reaction times to each sense individually in conjunction with the temporal offset, but also appears to include a higher order process that introduces its own temporal asynchrony [41]. To a certain extent, this additional temporal asynchrony is compensated for by optimization of the order and timing of presentation of the individual multisensory stimuli with respect to the particular illusion of interest. In both the classic Double Flash Illusion and the Illusory Audiovisual Rabbit Illusion, the first auditory beep typically precedes the first real flash by 23 ms, even though

the reaction time for audition is typically shorter than the reaction time for vision.

As a consequence, the reporting of interrogated aspects of a multisensory illusion directly involves both subconscious and conscious cognitive processes. Considering once again the example of the Illusory Audiovisual Rabbit Illusion, in order for participants to respond to queries several seconds after the rapid sequence that generates the illusion, memory must be accessed in order to consciously report the number of flashes perceived, the order in which they were perceived, and the location of the illusory flash. In a way, the questions could be stated as "*What* are participants conscious of that just happened?", "*When* are the participants first conscious of the *what* that just happened?", and finally "What do participants' conscious perceptions reveal about *when* the *what* happened?". Resolution of these questions is beyond the scope of this review, but provides a key avenue of opportunity for further exploration and refinement of crossmodal postdiction models.

Longer term postdictive effects are also possible at the cognitive though perhaps still subconscious level, such as cognitive reorganization, hindsight bias, and the witness effect [47], in which the brain attempts to form the best possible interpretation of ambiguous events or sequences. These longer term postdictive effects are also outside the scope of this review.

6. DISCUSSION AND NEXT STEPS
6.1 *Overview*
In this review article, we have discussed the range of postdictive phenomena that are inherently crossmodal in that they include stimuli from two or more modalities. We explored several instances in which postdiction that occurs between two stimuli in one modality was influenced by the presentation of a stimulus in another modality (Unimodal Postdiction with Crossmodal Influence). Postdiction was also observed to occur in Intermodal Apparent Motion with two stimuli of different modalities (Crossmodal Postdiction with Emergent Illusory Perception), and in the Burst Lag Auditory–Visual Illusion with an auditory burst and a moving visual stimulus (or *vice versa*). In addition, a real visual stimulus was found to affect an illusory visual perception postdictively when the illusory visual perception was triggered crossmodally (*e.g.*, by an auditory beep, as in the Illusory Audiovisual Rabbit Illusion; Crossmodal Postdiction with Crossmodal Illusory Perception). Finally, three neuropsychological models for crossmodal postdiction were also outlined and discussed (the Catch Up Model, the Reentry Model, and the Different Pathways Model). Overall, crossmodal postdiction is an emerging type of crossmodal processing that adds a new twist to the expanding field of multisensory science, and highlights new models for multisensory integration.

In the next subsections, we discuss the use of neuroimaging techniques for analyzing multisensory integration with postdiction, the implications of crossmodal postdiction for sensory processing, and remaining questions on crossmodal

postdiction. Finally, we describe several key emerging applications that have the potential for incorporating crossmodal postdiction, and therefore may be more optimally interpreted and potentially designed by taking crossmodal postdiction into account.

6.2 *Neuroimaging and Crossmodal Postdiction*

Neuroimaging (*e.g.*, functional Magnetic Resonance Imaging (fMRI) and Electroencephalography (EEG)) have become effective tools for understanding crossmodal interactions and their processing in the sensory neural network. For example, EEG has been employed to show that the illusory flash triggered by an auditory beep in the Double Flash Illusion is due to visual activation in early visual regions driven by activation in auditory regions [33, 44]. If a similar experimental paradigm was used to investigate Intermodal Apparent Motion, the Burst Lag Illusion, or the Audiovisual Rabbit Illusions, the interactions among primary sensory regions, and between primary sensory regions and multimodal regions, could be further explored. In addition, neuroimaging would be useful for studying the role of pre-experiment connectivity among primary sensory regions on the perception of crossmodal postdiction. While it would be difficult to pin down the exact neuropsychological model that generates a given illusion with neuroimaging, neuroimaging could nonetheless provide useful limitations and boundary conditions to constrain current and future models of crossmodal postdiction.

6.3 *Implications of Crossmodal Postdiction*

It can be argued that postdiction, as well as prediction, forms a critical building block for the generation of sensory perception. Given the importance of multisensory integration to the holistic perception of the environment, the mounting evidence for the extension of unimodal postdiction into the multimodal domain further fortifies the argument that postdiction is critical to sensory perception.

In addition, research exploring an extended body schema with both natural and virtual objects has shown that the Cutaneous Rabbit Illusion, and therefore postdiction, occurs in this domain as well [5, 18, 34]. In particular, in the papers by Miyazaki *et al.* as well as by Berger and Gonzalez-Franco, a version of the tactile Rabbit illusion was studied in which two taps were made on one hand and a third tap on the other hand [18, 34]. The second tap was "perceived" to occur on either a real stick placed between the two hands [34], or on a virtual stick placed between two *virtual* hands (presented by means of a virtual reality headset in combination with two vibrotactile stimulators) [18, 34]. These studies show that postdiction can actually cause stimuli to be "perceived" in an inanimate object that is either real or virtual.

Postdiction has also been found to occur in visuomotor perception [37] with a modified version of the Visual Flash Lag Illusion, in which the participant's arm movement generates the movement of the visual stimulus. Therefore, in this case, visual perception is integrated with self-initiated movement to generate a postdictive perception of the environment.

The evidence of unimodal postdiction and crossmodal postdiction, as well as of motor and extended body schema postdiction (as highlighted above), all broaden the previously understood purview of postdictive perception. Furthermore, all of these areas of research and the associated illusions studied therein combine to support the emerging understanding that postdiction is likely an essential element of perceptual processing. In addition, we propose that additional research will continue to expand the impact of *postdiction* on perception, and will eventually show that postdictive effects pervade most forms of early sensory processing, similar to the better known case of *prediction*.

6.4 *Crossmodal Postdiction and the Metamodal Organization of the Brain*

Crossmodal postdiction also potentially has implications for the theory of the metamodal organization of the brain. The metamodal structure of the brain as theorized by Pascual-Leone and Hamilton proposes that cortical regions should not be segregated and organized only by the primary sensory modality that they process (such as vision or audition), but rather by the computational processes that they primarily support (such as spatial processing, temporal processing, or shape processing) [38]. If it can be shown that crossmodal postdiction (as another form of multimodal integration) occurs even in primary sensory regions (as proposed by the neural models described earlier), this will further support the concept that each early cortical region performs particular computational functions rather than being solely dedicated to the processing of a single sensory modality. Future neuroimaging studies with crossmodal postdictive illusions may shed further light on this theory of the sensory organization of the brain.

6.5 *Key Remaining Questions Regarding Crossmodal Postdiction*

This review highlights several perceptual effects and illusions in which both crossmodal prediction and crossmodal postdiction have been shown to occur. A key remaining question for the psychophysical research field is how these two types of processing interact. For example, preliminary research on the Extended Illusory Audiovisual Rabbit Illusion has indicated that the influence of prediction on the perceived location of the illusory flash may be somewhat stronger than the influence of postdiction [49]. In this particular illusion, this means that the illusory flash was shifted more toward the first real flash than the last real flash. (Note: This comparison is derived from two related experiments with the same participants; a within experiment comparison would strengthen this result.)

Additional research is required to determine whether the influence of prediction is stronger or weaker than the influence of postdiction on perception when stimuli are presented with temporally symmetric timing (*i.e.*, when preceding and following stimuli have the same delay relative

to the target). Furthermore, there may be brain states (as measured by MRI or EEG, such as the measurement of the fluctuating connectivity between sensory regions) that predispose either prediction or postdiction to be a stronger influence in a given multisensory process. These questions will hopefully be further explored and resolved in future research studies.

In addition, it is currently unknown whether differences in the sensory processing pipelines of unimodal and crossmodal postdiction could cause crossmodal postdiction to have a different time window of integration than unimodal postdiction. Comparisons between the temporal windows of illusion perception for both unisensory and crossmodal postdiction will better elucidate key differences in the sensory processing and neural models that best describe these modes of perception.

6.6 *Applications of Crossmodal Postdiction*

Multisensory interactions are commonplace in the natural environment and in daily life, ranging from the immediacy of hitting a nail with a hammer (combining visual, tactile, and auditory inputs) to the more temporally extended thunder and lightning example mentioned earlier (combining auditory and visual inputs). Nonetheless, neither unisensory nor multisensory (crossmodal) postdiction are likely to occur frequently in the natural environment in such a manner as to be self-evident. In particular, it would seem at first glance that the types of multisensory postdictive illusions described herein are both structured and parametrically optimized to generate the largest possible effects in order to allow for scientific research and analysis. Emerging technological applications, however, may involve digitally generated multisensory interactions with significant postdictive consequences for perception.

As an example of these emerging applications, consider the auditory, visual, and tactile multisensory environments presented by Extended Reality (XR), Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). In these applications, digitally generated sounds, images, and touches are presented over time scales characteristic of numerous multisensory postdictive illusions (typically 100 to 200 ms), based on typical video frame rates of 30 to 60 and even 120 fps (frames per second, with frame times of 33.3 ms, 16.7 ms, and 8.3 ms, respectively).

One such application of Augmented Reality involves the use of both head mounted and dashboard mounted head-up displays (HUDs) in aircraft and automobiles to provide functional and navigational information in real time to the pilot or driver. Another such emerging application involves the use of Augmented Reality and Mixed Reality in manufacturing to guide complicated assembly tasks. Whereas in most circumstances the rate of change of the information displayed may be modest, in other circumstances the information rate of change may be extreme, as in the takeoff and landing of an aircraft, and in air-to-air combat. In these circumstances, both desired and undesired crossmodal postdictive illusions may occur, and thereby impact design constraints for spatial proximity and timing in the presentation of digitally generated information.

Another such application of either Augmented Reality or Mixed Reality involves the development of assistive devices for the blind and those with low vision [6, 30]. Consider the case of Age-Related Macular Degeneration, for example, in which the region of central vision is compromised, but often the region of peripheral vision is intact. An assistive aid could be envisioned to incorporate both a head mounted scene camera and an Augmented Reality or Mixed Reality display, which when combined with an eye-tracker could provide essential task-related and navigational information from the obscured region of central vision to the functional periphery. As in the previous case described above, careful attention to both spatial and temporal design will be important to optimize the visibility and interpretation of key information, perhaps using crossmodal postdiction to advantage, as well as to prevent the occurrence of potentially misinforming or disorienting crossmodal postdictive illusions.

These considerations point to the need for thoughtful design principles, and will be especially important in high noise (low contrast) environments, which tend to enhance multisensory ambiguity and hence the perception of illusions. For example, crossmodal postdiction could be used to advantage in emphasizing irregular and often missed signals in one modality with supporting signals in another modality (as suggested by the Illusory Audiovisual Rabbit Illusion), and in suppressing spurious signals in one modality with compensating signals in another modality (as suggested by the Invisible Audiovisual Rabbit Illusion).

The general principle might be that when problems are encountered in the presentation of information with different temporal resolutions or phases across modalities, attention should be paid to both predictive and postdictive compensatory mechanisms.

Finally, we note that even though neither unisensory nor multisensory (crossmodal) postdiction are likely to be self-evident in the natural environment, they may nonetheless be commonplace in everyday perception, as conscious percepts evolving from ambiguous or noisy environmental stimuli are subconsciously formed from both lower order and higher order cortical processes with extensive feedback and reorganization. As such, clinical applications for crossmodal postdiction may reveal multisensory processing behavior that is not only altered in abnormal eye and brain conditions, but also may anticipate useful multisensory training and rehabilitation paradigms [55].

REFERENCES

[1] D. Alais and D. Burr, "The 'flash-lag' effect occurs in audition and cross-modally," Curr. Biol. **13**, 59–63 (2003).

[2] V. Arstila, "Keeping postdiction simple," Conscious. Cogn. **38**, 205–216 (2015).

[3] T. Bachmann, "Neurobiological mechanisms behind the spatiotemporal illusions of awareness used for advocating prediction or postdiction," Frontiers Psychol. **3**, 593 (2013).

[4] M. S. Beauchamp, B. Argall, J. Bodurka, J. Duyn, and A. Martin, "Unraveling multisensory integration: Patchy organization within human STS multisensory cortex," Nat. Neurosci. **7**, 1190–1192 (2004).

[5] C. C. Berger and M. Gonzalez-Franco, "Expanding the sense of touch outside the body," *Proc. 15th ACM Symposium on Applied Perception* (Association for Computing Machinery, Vancouver, BC, Canada, 2018).

[6] N. Binetti, T. Cheng, I. Mareschal, D. Brumby, S. Julier, and N. Bianchi-Berthouze, "Assumptions about the positioning of virtual stimuli affect gaze direction estimates during Augmented Reality based interactions," Sci. Rep. **9**, 2566 (2019).

[7] O. Carter, T. Konkle, Q. Wang, V. Hayward, and C. Moore, "Tactile rivalry demonstrated with an ambiguous apparent-motion quartet," Curr. Biol. **18**, 1050–1054 (2008).

[8] Y.-C. Chen and C. Spence, "The crossmodal facilitation of visual object representations by sound: Evidence from the backward masking paradigm," J. Exp. Psychol. Hum. Percept. Perform. **37**, 1784 (2011).

[9] H. Choi and B. J. Scholl, "Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception," Perception **35**, 385–399 (2006).

[10] D. M. Eagleman and T. J. Sejnowski, "Motion integration and postdiction in visual awareness," Science **287**, 2036–2038 (2000).

[11] D. M. Eagleman and T. J. Sejnowski, "The line-motion illusion can be reversed by motion signals after the line disappears," Perception **32**, 963–968 (2003).

[12] R. D. Easton, A. J. Greene, and K. Srinivas, "Transfer between vision and haptics: Memory for 2-D patterns and 3-D objects," Psychonomic Bull. Rev. **4**, 403–410 (1997).

[13] E. Freeman and J. Driver, "Direction of visual apparent motion driven solely by timing of a static sound," Curr. Biol. **18**, 1262–1266 (2008).

[14] F. A. Geldard and C. E. Sherrick, "The cutaneous 'rabbit': A perceptual illusion," Science **178**, 178–179 (1972).

[15] F. A. Geldard, "The saltatory effect in vision," Sens Processes. **1**, 77–86 (1976).

[16] S. Getzmann, "The effect of brief auditory stimuli on visual apparent motion," Perception **36**, 1089–1103 (2007).

[17] D. Goldreich and J. Tong, "Prediction, postdiction, and perceptual length contraction: A Bayesian low-speed prior captures the cutaneous rabbit and related illusions," Frontiers Psychol. **4**, 221 (2013).

[18] M. Gonzalez-Franco and C. C. Berger, "Avatar embodiment enhances haptic confidence on the out-of-body touch illusion," IEEE Trans. Haptics **12**, 319–326 (2019).

[19] M. A. Goodale and A. D. Milner, "Separate visual pathways for perception and action," Trends Neurosci. **15**, 20–25 (1992).

[20] M. W. von Grünau, "A motion aftereffect for long-range stroboscopic apparent motion," Percept. Psychophys. **40**, 31–38 (1986).

[21] V. Harrar, R. Winter, and L. R. Harris, "Visuotactile apparent motion," Percept. Psychophys. **70**, 807–817 (2008).

[22] M. H. Herzog, L. Drissi-Daoudi, and A. Doerig, "All in good time: Long-lasting postdictive effects reveal discrete perception," Trends Cogn. Sci. **24**, 826–837 (2020) Epub 2020 Sep 3.

[23] M. Kaiser, D. Senkowski, N. A. Busch, J. Balz, and J. Keil, "Single trial prestimulus oscillations predict perception of the sound-induced flash illusion," Sci. Rep. **9**, 1–8 (2019).

[24] T. Kawabe, "Nonretinotopic processing is related to postdictive size modulation in apparent motion," Atten. Percept. Psychophys. **73**, 1522–1531 (2011).

[25] P. A. Kolers and M. von Grünau, "Shape and color in apparent motion," Vis. Res. **16**, 329–335 (1976).

[26] J. Keil, "Double flash illusions: Current findings and future directions," Frontiers Neurosci. **14**, 298 (2020).

[27] S. Lakatos and R. N. Shepard, "Constraints common to apparent motion in visual, tactile, and auditory space," J. Exp. Psychol.: Hum. Percept. Perform. **23**, 1050 (1997).

[28] E. Liaci, M. Bach, L. Tebartz van Elst, S. P. Heinrich, and J. Kornmeier, "Ambiguity in tactile apparent motion perception," PLoS ONE **11**, e0152736 (2016).

[29] B. Libet, *Mind Time: The Temporal Factor in Consciousness* (Harvard University Press, Cambridge, MA, 2009).

[30] Y. Liu, N. R. B. Stiles, and M. Meister, "Augmented reality powers a cognitive assistant for the blind," eLife **7**, e37841 (2018).

[31] D. M. Mackay, "Perceptual stability of a stroboscopically lit visual field containing self-luminous objects," Nature **181**, 507–508 (1958).

[32] J. L. Marchant, C. C. Ruff, and J. Driver, "Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities," Human Brain Mapping **33**, 1212–1224 (2012).

[33] J. Mishra, A. Martinez, T. J. Sejnowski, and S. A. Hillyard, "Early crossmodal interactions in auditory and visual cortex underlie a sound-induced visual illusion," J. Neurosci. **27**, 4120–4131 (2007).

[34] M. Miyazaki, M. Hirashima, and D. Nozaki, "The 'Cutaneous Rabbit' hopping out of the body," J. Neurosci. **30**, 1856–1860 (2010).

[35] L. Muckli, A. Kohler, N. Kriegeskorte, and W. Singer, "Primary visual cortex activity along the apparent-motion trace reflects illusory perception," PLoS Biol. **3**, e265 (2005).

[36] R. Nijhawan, "Motion extrapolation in catching," Nature **370**, 256–257 (1994).

[37] R. Nijhawan and K. Kirschfeld, "Analogous mechanisms compensate for neural delays in the sensory and the motor pathways: Evidence from motor flash-lag," Curr. Biol. **13**, 749–753 (2003).

[38] A. Pascual-Leone and R. Hamilton, "The metamodal organization of the brain," Progress Brain Res. **134**, 427–445 (2001).

[39] D. H. Raab, "Backward masking," Psychol. Bull. **60**, 118 (1963).

[40] V. S. Ramachandran and S. M. Anstis, "The perception of apparent motion," Sci. Am. **254**, 102–109 (1986).

[41] J. Rutschmann and R. Link, "Perception of temporal order of stimuli differing in sense mode and simple reaction time," Perceptual and Motor Skills **18**, 345–352 (1964).

[42] D. Sanabria, S. Soto-Faraco, and C. Spence, "Assessing the effect of visual and tactile distractors on the perception of auditory apparent motion," Exp. Brain Res. **166**, 548–558 (2005).

[43] D. L. Schacter and R. L. Buckner, "Priming and the brain," Neuron **20**, 185–195 (1998).

[44] L. Shams, Y. Kamitani, and S. Shimojo, "What you see is what you hear," Nature **408**, 788 (2000).

[45] L. Shams, Y. Kamitani, S. Thompson, and S. Shimojo, "Sound alters visual evoked potentials in humans," Neuroreport **12**, 3849–3852 (2001).

[46] S. Shimojo, C. Simion, E. Shimojo, and C. Scheier, "Gaze bias both reflects and influences preference," Nat. Neurosci. **6**, 1317–1322 (2003).

[47] S. Shimojo, "Postdiction: Its implications on visual awareness, hindsight, and sense of agency," Frontiers Psychol. **5**, 196 (2014).

[48] S. Shimojo, S. Miyauchi, and O. Hikosaka, "Visual motion sensation yielded by non-visually driven attention," Vis. Res. **37**, 1575–1580 (1997).

[49] D. I. Shore, S. E. Hall, and R. M. Klein, "Auditory saltation: A new measure for an old illusion," J. Acoust. Soc. Am. **103**, 3730–3733 (1998).

[50] S. Soto-Faraco, C. Spence, and A. Kingstone, "Cross-modal dynamic capture: Congruency effects in the perception of motion across sensory modalities," J. Exp. Psychol.: Hum. Percept. Perform. **30**, 330 (2004).

[51] N. R. B. Stiles, M. Li, C. A. Levitan, Y. Kamitani, and S. Shimojo, "What you saw is what you will hear: Two new illusions with audiovisual postdictive effects," PLoS ONE **13**, e0204217 (2018).

52 T. Z. Strybel, C. L. Manligas, O. Chan, and D. R. Perrott, "A comparison of the effects of spatial separation on apparent motion in the auditory and visual modalities," Percept. Psychophys. **47**, 439–448 (1990).

53 L. Sun, S. M. Frank, K. C. Hartstein, W. Hassan, and P. U. Tse, "Back from the future: Volitional postdiction of perceived apparent motion direction," Vis. Res. **140**, 133–139 (2017).

54 A. R. Tanguay Jr., N. R. B. Stiles, I. Ganguly, and S. Shimojo, "Time dependence of predictive and postdictive auditory–visual processing: The temporally extended audiovisual rabbit illusion," J. Vis. **19**, 19b (2019).

55 A. R. Tanguay Jr., N. R. B. Stiles, I. Ganguly, and S. Shimojo, "Mapping audio-visual crossmodal interactions in the visually impaired," J. Vis. **20**, 1768 (2020).

56 J. Vroomen and B. de Gelder, "Temporal ventriloquism: Sound modulates the flash-lag effect," J. Exp. Psychol.: Hum. Percept. Perform. **30**, 513 (2004).

57 Y. Yamada, T. Kawabe, and M. Miyazaki, "Awareness shaping or shaped by prediction and postdiction," Frontiers Psychol. **6**, 166 (2015).