Title
Twisted in Time: Interaction of Simultaneity and Causal Judgements in Multisensory Integration

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors
Chen, Tianshu
Zou, Luyi
Lagnado, David

Publication Date
2026

Twisted in Time: Interaction of Simultaneity and Causal Judgements in Multisensory Integration

Tianshu Chen, Luyi Zou, David Lagnado
Department of Experimental Psychology, University College London, 26 Bedford Way, London WC1H 0AP

Abstract

Multisensory integration tolerates limited temporal misalignment between multisensory signals, often characterised as a temporal window of integration (TWI). This study explores whether TWI size is influenced by perceived causal impressions, and whether auditory cues within the TWI can increase causal impressions when visual evidence is ambiguous. Results showed that launching and strong interaction conditions produced wider TWIs than weak conditions. Critically, this widening was directionally asymmetric: stronger causal conditions selectively expanded the post-contact integration window while leaving pre-contact tolerance unchanged, consistent with a forward causal prediction account. In the weak condition, near-synchronous sounds selectively increased causal and contact judgments. The finding of a close relationship between simultaneity, causal, and contact judgments supports a common inferential process that integrates multisensory evidence to estimate a shared cause.

Keywords: multisensory integration; temporal binding window; causal perception; simultaneity judgment

Introduction

Our experience of the world involves multiple sources of sensory information. The mechanisms of combining multisensory information have elicited extensive research in psychology. One of the major findings is that the process of integrating multiple senses can tolerate asynchrony. It is important that the brain applies some tolerance to discrepancies, due to errors incurred during physical transmission and sensory processing. For instance, the relative timing of visual and auditory signals at source is contaminated by differences in the speed that light and sound travel through air, as well as differences in neural transduction latencies (King & Palmer, 1985; Maunsell & Gibson, 1992; Recanzone, Guard& Phan, 2000). Hence the overcoming of asynchrony across models, or the tolerance of a temporal binding window, is an interesting byproduct of multisensory signal integration.

Temporal Window of Integration

Previous research has shown that the temporal binding window is strongly related to causal inference, such that prior beliefs about causal relationships shape perceived temporal proximity between events. In the field of causal inference research, temporal binding—originally termed *intentional binding* when first described by Haggard et al. (2002)—refers to the perceived temporal attraction between two causally related events. Early accounts proposed that this illusion arises primarily from predictive mechanisms associated with intentional, voluntary actions, emphasizing a motor-based explanation. As a result, intentional binding has been widely studied in the context of action perception and has often been used as an implicit measure of the sense of agency over action outcomes (Haggard & Tsakiris, 2009).

However, accumulating evidence challenges the view that temporal binding is restricted to intentional action–effect pairs. More recent research demonstrates that temporal binding also occurs between events that are causally linked without involving an intentional action (e.g., Borhani et al., 2017; Buehner, 2012, 2015). When one event is interpreted as the effect of a preceding event, the temporal interval between them is perceived as shorter than its actual duration (Klaffehn et al., 2021). These findings suggest that temporal window of integration (TWI) reflects a general tendency that the perceived time interval between two events is compressed when the two events are inferred to share a common cause. This perspective naturally aligns with Bayesian accounts of perception, in which causal structure is inferred by combining prior beliefs with sensory evidence (Körding et al., 2007). Within such frameworks, a stronger prior belief that two events are causally related increases the likelihood that they are treated as originating from a common cause, leading to stronger perceptual integration and increased TWI.

Interestingly, the majority of the studies manipulated causal judgement by associative learning, such as the Libet clock paradigm (Buehner 2015; Haggard, Clark & Kalogeras, 2002; Libet et al., 1983). In these experiments, causal beliefs were introduced by learning an association between a key press and an auditory stimulus, or between two visual stimuli (e.g., flashes of light). The dependent measure was the perceived temporal interval between the two events, compared across conditions in which the events were believed to be causally related versus not causally related. Less research has examined temporal processing in purely perceptual causal events, and few studies have directly investigated how perceived causal strength influences simultaneity judgments.

Causal Impression

To address this gap, the current study aims to manipulate perceived causal strength by changing observable impacts to create different strengths of causal impression. Causal

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

impression refers to the immediate phenomenology that one event directly produces another—classically exemplified by Michotte's (1946/1963) launching displays. Rather than an inference from learned covariation, this experience arises from event-level perceptual organization where observers perceive "two objects but only a single motion," as if energy is transferred from a launcher to a patient (Scholl & Tremoulet, 2000). Unlike the learned action–effect associations typically used in temporal-binding studies, causal impressions are driven by observable spatiotemporal structures in the stimulus stream.

These impressions are highly sensitive to kinematic parameters, particularly spatial and temporal contiguity at the moment of impact (Schlottmann & Shanks, 1992). This process is not a static categorization but a dynamic construction; cues presented shortly after contact can retrospectively bias whether an event is seen as a causal "launch" or a non-causal "pass" (Choi & Scholl, 2006). Such findings suggest the visual system integrates information over a brief window to establish causal organization, likely utilizing specialized perceptual routines rather than deliberate reasoning (Rolfs et al., 2013).

Theoretical accounts propose that these impressions reflect a match between incoming kinematics and stored event schemas—mental templates for familiar physical interactions (Rips, 2011; White, 2005). These schemas allow for robust causal interpretations even in physically implausible displays, provided they align with canonical interaction patterns. Crucially, causal impressions scale with the apparent "impact" on the affected object. Impressions are strongest when the patient's motion appears to be a compelled consequence of the launcher's motion, such as through pronounced acceleration or high force "explosive" changes (White, 2005). By manipulating the visual evidence of this effect, we expect to modulate strength of causal impression without associative learning during experiments.

Our Approach

We introduce a paradigm pairing multisensory signals with video clips giving different degrees of causal impression. By systematically manipulating the stimulus onset asynchrony (SOA) between an auditory cue and the moment of visual contact, we investigate how causal strength modulates the TWI. Specifically, we test whether stronger causal impressions broaden the TWI, making participants more likely to judge events as simultaneous at longer intervals (Bechlivanidis et al., 2022). Furthermore, we examine whether a simultaneous auditory cue can selectively enhance the perceived causality of weak interactions by providing additional multisensory evidence of a shared event.

Our approach targets perceptual rather than deliberate inference by employing animated launching-like events known to elicit immediate impressions of physical interaction. This allows us to assess two primary predictions: (1) that the TWI is not fixed but expands with the strength of the causal impression, and (2) that auditory signals occurring within this window reciprocally shape the perceived causal structure of the event.

The visual stimuli we used are adapted based on classic Michottean launching (Michotte, 1946/1963). Instead of having objects with straight line movement, as we also want to create the weak interaction condition as objects contact each other but show no impact on each other's movement, we have objects moving on a partial circle motion around fixed radius (Figure 1). In the launching condition, one object moves towards a stationary object and stops upon contact, the stationary object then starts moving after contact. In the no interaction separate condition, the two objects follow a half-

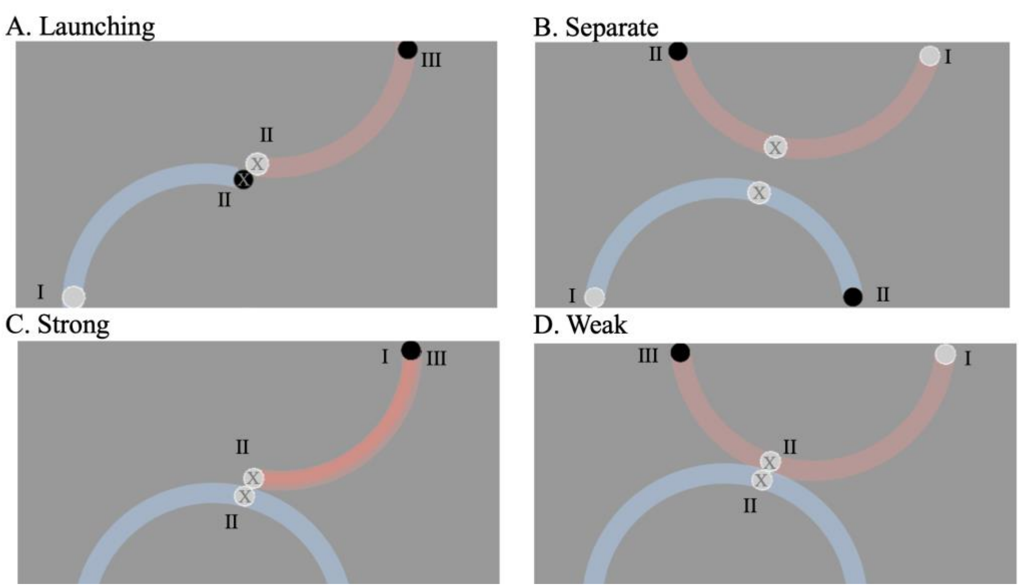

Figure 1: Four example interaction stimuli. White discs indicate the objects' starting positions; black discs indicate their final positions. Discs marked with "X" show the objects' locations at the moment the auditory cue occurred in the synchronous condition (SOA = 0). Roman numerals indicate position of each object at different time points in the clip (e.g., I = first frame of the clip). Colored arcs trace the motion paths of the two objects.

circle motion and never come into contact. In the strong causal interaction condition, both objects start moving at the beginning of the clips, upon contact one object keeps its original moving pattern but the other object changes its moving direction and moves back to its original position, showing a "bounce off" impression. In the weak causal interaction condition, the two objects follow half-circle path and come into contact for 1 frame, without changing their moving patterns. We hypothesized that causal impressions would be substantial in the launching condition, where a stationary object is set into motion, and in the strong condition, where an object abruptly changes direction, but ambiguous in the weak condition, where neither object changes state upon contact. The strong condition serves as a control for the weak condition, because in both strong and weak cases the objects are moving throughout the clip; this allows us to test whether temporal binding depends on any salient visual impact rather than specifically on a transition from rest to motion as in launching. Thus, the critical factor is not stationarity per se, but the presence of an observable effect of contact.

Additionally, both the launching and strong interactions deliberately violate Newtonian mechanics: the "launcher" object shows no reaction to the collision. This keeps physical plausibility constant across conditions and prevents differences in real-world likelihood from driving the effect. As a result, any variation in temporal binding can be attributed to perceived causal strength from visible impact, rather than to differences in how physically realistic the interactions appear[1].

We ask participants to give rating of their agreements with simultaneity, causal and contact judgements explicitly. Audiovisual simultaneity judgments provide a direct behavioral indicator of TWI, capturing the range of SOAs over which multisensory signals are bound into a single event (Dixon & Spitz, 1980; Vroomen & Keetels, 2010). Measuring simultaneity across graded delays therefore reveals how TWI varies across different causal interaction clips. To test whether such temporal binding also shapes higher-level event interpretation, we asked perceived causality judgement questions rather than asking for mere motion descriptions (Michotte, 1946/1963; Scholl & Tremoulet, 2000; White & Milne, 1997). Finally, contact judgments isolate the perception of physical touch from causal attribution, providing more low-level visual feedback against which multisensory influences can be assessed (Scholl & Nakayama, 2002). Together, these three judgments sample different levels of event representation—temporal coherence, causal impression, and physical contact—allowing us to determine whether variations in TWI across clips selectively affect simultaneity alone or extend to both causal and contact perception via shared multisensory binding mechanisms.

Method

Participant

We recruited 103 adult participants (53 female, 49 male and 1 prefer not to say; mean age 37.1). Participants were recruited via Prolific (www.prolific.com) and paid £6 per hour for their participation. The experiment received full ethical approval from the UCL Department of Experimental Psychology Ethics committee, ethical approval code 0487.

Stimuli

The stimuli were short animations depicting four types of interactions between two discs: launching, strong, weak, and separate (see Figure 1, stimuli are available via https://tinyurl.com/twistedintime). Each clip lasted 1 s at the original playback speed and was presented in a 1280 × 720 pixel frame. To minimize predictability of the contact moment, every base animation was rendered in four variants: original, horizontally mirrored, slowed to 0.8× speed, and mirrored plus slowed. This manipulation ensured that neither the exact movement direction nor the contact timing could be anticipated from completing previous trials.

An impact sound (20 ms) was paired with the visual contact in sound-present conditions. For the slowed (0.8×) variants, auditory cue was added after the speed of the video has been adjusted so that the intended SOAs relative to visual contact were preserved. Each trial began with a 500 ms central fixation cross before the animation. All stimuli were created using Adobe Animate and Adobe Premiere.

Design and Procedure

The study was conducted online using the Gorilla Experiment Builder (Anwyl-Irvine et al., 2018). For each interaction type, eight SOA conditions were tested: synchronous with contact (0 ms), 100/200/300 ms before contact (B100, B200, B300), 100/200/300 ms after contact (A100, A200, A300), and a no-sound baseline. This yielded 32 trial types in total (4 interactions × 8 SOA conditions). On each trial, one of the four visual variants of the corresponding interaction was selected at random.

Trials with and without the auditory cue were divided into two blocks, with participants completing the no-sound block first. This is to measure participants' judgement of the visual interactions as a baseline response. After completing the no-sound block, participants were given an instruction page where they can familiarize themselves with the auditory cue. This is to ensure that participants' responses are not distorted by a surprise effect. Participants then continue to the block of trials with auditory cues. Trial orders within each block were randomized.

Within each trial, participants were asked to report (in the following order):

(1) Simultaneity rating: To what extent do you agree with the claim "I heard a sound at the same time as the

[1] A pilot study (N = 18) supports our design of stimuli. A one-way repeated measures ANOVA revealed a significant effect of causal condition on participant ratings, $F(3, 51) = 9.26$, $p < .001$, ges = .237. Bonferroni-corrected pairwise comparisons showed that ratings in the launching and strong conditions were significantly higher than in the separate condition ($p = .038$ and $p = .001$, respectively), and that strong condition ratings were significantly higher than weak condition ratings ($p = .007$).

objects made contact.” (A click box is next to this question by which participants can indicate “I did not hear a sound”)

(2) Causal rating: To what extent do you agree with the claim “There is an interaction between the two objects.”

(3) Contact rating: To what extent do you agree with the claim “The two objects made contact.”

For each question, the slider rating ranges from 0 (Strongly Disagree) to 100 (Strongly Agree). The exact numerical value of the slider position was hidden from the participant.

Results

Simultaneity Judgement

Simultaneity ratings were analyzed using a linear mixed-effects model fit with restricted maximum likelihood. The model included interaction condition, temporal offset, and their interaction as fixed effects, with random intercepts for participants. Degrees of freedom were estimated using Satterthwaite’s method.

The results revealed a robust main effect of different SOAs, with simultaneity ratings decreasing as the temporal separation between the auditory cue and the moment of contact increased (Figure 2A). The comparisons revealed a notable asymmetry in TWI. In the visual-leading conditions, no significant difference was observed between the synchronized condition and A100 (*Estimate* = 1.02, *SE* = 3.87, $t = 0.26$, $p = .79$), while a significant difference emerged at A200, the effect size remained small (*Estimate* = -7.96, *SE* = 3.88, $t = -2.05$, $p = .04$). In contrast, audio-leading SOAs showed significant and large differences from the synchronized condition at both B100 (*Estimate* = -23.20, *SE* = 3.87, $t = -5.99$, $p < .001$) and B200 (*Estimate* = -53.46, *SE* = 3.88, $t = -13.77$, $p < .001$).

Regarding the influence of interaction types, we observed no significant difference in simultaneity judgments between the launching and strong causal conditions (*Estimate* = 2.11, *SE* = 3.87, $t = 0.55$, $p = .58$). However, participants’ responses in the weak causal condition differed significantly from the launching condition (*Estimate* = -10.96, *SE* = 3.87, $t = -2.83$, $p = .005$). This finding is crucial to our central research question, as it shows that the TWI varies significantly across different causal interaction types, with weaker causal impressions giving a narrower window of tolerance.

Most importantly, the interaction between visual conditions and SOAs was significant, reflecting different sensitivity to temporal asynchrony across the three visual conditions. In the launching and strong conditions, simultaneity ratings did not differ significantly between small auditory delays (e.g., Strong A100–A200: *Estimate* = 6.75, *SE* = 3.87, $t = 1.74$, $p = .66$) but were significantly reduced at larger SOAs (A300, B100, B200, and B300, all *ps* < .001). In contrast, the weak condition exhibited a steeper decline, with significant reductions in simultaneity ratings emerging as early as A200 (A100–A200: *Estimate* = 21.43, *SE* = 3.88, $t = 5.52$, $p < .001$) and continuing to decrease at larger SOAs. Pairwise comparisons confirmed that temporal asynchronies disrupted simultaneity judgments more robustly and at shorter delays in the weak condition than in the launching or strong conditions. Furthermore, in the launching and strong conditions, simultaneity ratings remained reliably above the scale midpoint for small SOAs (0, A100, A200) and only dropped below 50 (i.e., disagree with claim of synchrony) at larger SOAs (B200, B300). In the weak condition, however, ratings approached the midpoint at A200 (95% CI [46.8, 59.2]) and fell reliably below 50 at A300 (95% CI [20.1, 32.6]), indicating that participants were sensitive to asynchronies at substantially smaller temporal gaps when the causal evidence was weak.

Given its asymmetrical shape in the estimated marginal means, to extract theoretically interpretable parameters, split Gaussian functions were fit with the split point fixed at physical simultaneity (SOA = 0), yielding separate window widths for sounds arriving before (σ_left) and after (σ_right) contact. Conditions differed significantly on σ_left (weak vs. strong: $p = .014$), with weak showing a wider before-window than strong. Conditions also differed on σ_right (weak vs. strong: $p = .050$), with strong showing a wider after-window than weak. Critically, within-condition asymmetry tests

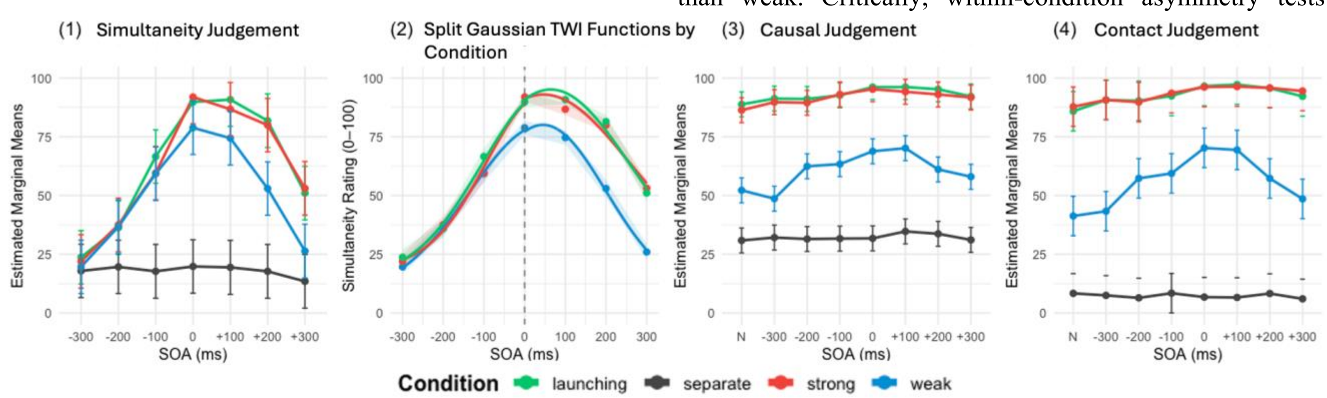

Figure 2: (1) Estimated marginal means for simultaneity judgements. (2) Fitted split-Gaussian TWI functions alongside observed mean simultaneity ratings (±1 SE) as a function of SOA (ms), with the dashed vertical line indicating simultaneity (SOA = 0). (3) Estimated marginal means for Causal Judgement and (4) Contact judgement across the full range of SOAs with the no sound baseline (N). Error bars represent 95% confidence intervals.

revealed that strong and launching conditions had significantly wider after-windows than before-windows (strong: 304ms vs. 68ms, $p < .001$; launching: 223ms vs. 106ms, $p = .033$), while weak showed no significant asymmetry (153ms vs. 135ms, $p = .743$). On the after side, weak produced a significantly narrower window than strong ($p = .006$). These results suggest that strong causal impressions may generate a forward temporal expectation that sound will follow contact, and it is this expectation that widens tolerance for post-contact asynchrony.

Causal Judgement

Participants' causal ratings were analyzed using a linear mixed-effects model with condition, SOA, and their interaction as fixed effects, alongside random intercepts for participants. There was a strong main effect of condition, with significantly lower causal ratings in the weak (*Estimate* = -27.30, *SE* = 3.39, $t = -8.07$, $p < .001$) and separate (*Estimate* = -64.36, *SE* = 3.39, $t = -19.01$, $p < .001$) conditions relative to the baseline launching condition. Causal ratings for the strong condition did not differ significantly from the launching condition ($p = .79$). Follow-up analyses examined the effect of SOA within each condition using estimated marginal means with Holm-corrected comparisons against the no-sound baseline (Figure 2B). In the weak condition, causal ratings varied significantly across SOAs, with several sound-present trials (0, A100, A200, B100, B200) differing reliably from the no-sound baseline (all *ps* < .05). Notably, in the weak condition, ratings for the no-sound baseline and the B300 SOA were statistically indistinguishable from the scale midpoint, as indicated by confidence intervals spanning 50 (e.g., no-sound baseline: 95% CI [46.9, 57.5]). This suggests that participants were uncertain about the causal impression of the visual stimuli when the auditory cue was absent or occurred outside the TWI. In contrast, ratings for all other SOAs in the weak condition were reliably above the midpoint (e.g., simultaneous: 95% CI [63.5, 74.1]), indicating that the presence of an auditory cue within the TWI successfully increased causal impression.

This pattern suggests that extreme sound–contact misalignment selectively disrupts causal judgments specifically when visual causal evidence is weak. Meanwhile, no SOA comparisons reached significance in the launching, strong, or separate conditions after correction. This indicates that multisensory integration within a tolerable temporal binding window significantly modulates causal judgments when visual evidence is ambiguous but is largely discounted when the causal structure is either clear or entirely absent.

Contact Judgement

Participants' agreement regarding whether the two objects made contact was analyzed using a linear mixed-effects model with condition, SOA, and their interaction as fixed effects, alongside random intercepts for participants. Similar to the simultaneity judgments, we observed no significant difference between the launching and strong causal conditions (*Estimate* = -0.43, *SE* = 3.16, $t = -0.14$, $p = .89$), with both conditions receiving contact ratings significantly higher than the other two interaction types (Figure 2C). As expected, participants consistently reported an absence of contact in the separate condition. In the weak causal condition, however, contact judgments were significantly modulated by the SOA, despite the objects making physical contact in all corresponding visual stimuli.

Follow-up analyses examined the effect of SOA on contact judgments within each condition using estimated marginal means with Holm-corrected comparisons against the no-sound baseline. A graded pattern was observed in the weak condition: ratings were reliably above the scale midpoint for synchronous and small-SOA trials (0, A100, A200, B100, B200), indicating strong agreement that contact occurred when temporal alignment was close. In contrast, contact ratings for larger SOAs (A300, B300) and the no-sound baseline fell below the midpoint (e.g., B300: 95% CI [38.7, 47.9]), suggesting that participants became uncertain about the occurrence of contact. This pattern indicates that additional auditory cues can influence the judgment of basic physical events, such as contact, particularly when visual evidence is ambiguous.

Discussion

This study investigated whether multisensory integration influences simultaneity, causal, and contact judgments during visual object interactions. Using an adapted Michottean launching paradigm with different levels of causal evidence, we tested whether the temporal window of integration depends on perceived causal impression. Where causal impression is obvious in the launching and strong causal conditions, simultaneity judgments did not differ from synchronous trials when the auditory cue followed object contact by up to 200 ms. In contrast, simultaneity judgments in the weak causal condition declined at the same temporal offset. Causal ratings showed a parallel pattern: synchronous and near-synchronous sounds increased causal impressions in the weak condition. The same temporal sensitivity was observed in contact judgments.

By comparing simultaneity judgments across seven stimulus onset asynchronies (SOAs; 0 ± 300 ms), the paradigm successfully captured the full extent of the temporal window of integration. In conditions where visual stimuli show strong causal impression, participants showed high agreement that the events were simultaneous when SOAs fell within approximately ±200 ms, whereas simultaneity judgments dropped sharply at ±300 ms, particularly when the auditory cue preceded the visual contact. This pattern indicates that the TWI in the present task is unlikely to extend beyond ~300 ms. The chosen SOA range therefore spans both the region of maximal multisensory binding and the points at which integration breaks down, suggesting that the design provides sufficient temporal coverage to characterize the complete integration window.

The temporal window of integration showed a clear asymmetry between auditory-leading and visual-leading conditions. Consistent with previous findings, auditory-leading asynchronies produced narrower integration windows than visual-leading asynchronies (Cecere et al., 2016; Dixon & Spitz, 1980; Hillock-Dunn & Wallace, 2012). Another interpretation of this asymmetry is that the auditory cue is perceived as an effect of object contact. Under this interpretation, visual-leading asynchronies follow a plausible cause–effect order, whereas auditory-leading asynchronies violate this order, making them less tolerable and more difficult to be integrated across sensory modalities. Split Gaussian modelling confirmed that stronger visual causal impression selectively widened post-contact integration windows, while weak causal impression showed no such asymmetry. This pattern is consistent with a forward causal prediction account: when visual stimuli strongly imply a causal interaction, the perceptual system generates a temporal expectation that sound — as the effect — will follow contact, selectively widening tolerance for post-contact asynchrony. The absence of this asymmetry in the weak condition suggests that forward temporal prediction is contingent on the strength of the causal impression rather than being a general feature of audiovisual processing. Future studies could test this interpretation by using auditory cues that are clearly non-impact-related, or by introducing paradigms imitating auditory-leading causal interactions, such as visualization of volume monitor.

The significant conditional differences observed across launching, strong, and weak interactions validate our experimental design, confirming that manipulating the magnitude of visual impact effectively elicits distinct levels of causal impression. Our manipulation strategy—varying visible impact while holding overall scene structure constant—aligns with prior evidence that causal impression is not tightly constrained by physical realism. Classic work on visual causality shows that observers experience compelling causal interactions in highly simplified displays, and that these impressions are tuned primarily by spatiotemporal and kinematic cues (e.g., contact timing, contiguity, relative motion) rather than by explicit evaluation of Newtonian physics (Scholl & Tremoulet, 2000; Rips, 2011). More directly, recent experiments showed "causation without realism": strong causal impressions can arise upon first encounter even for novel or physically atypical interactions, challenging accounts that require realistic, frequently experienced physical sequences for causal perception (Bechlivanidis, Schlottmann, & Lagnado, 2019). This fits schema-based views in which incoming motion is matched against event schemas (e.g., collision → transfer → displacement), yielding robust causal interpretations so long as the display provides sufficient evidence of an effect—even if other physical details (such as equal-and-opposite reaction) are missing or inconsistent (Rips, 2011). At the same time, work on adaptation indicates that launching-like causality engages specialized perceptual routines that operate at early, retinotopic stages of processing, supporting the idea that these judgments are not purely deliberative or "rational physics" computations (Rolfs, Dambacher, & Cavanagh, 2013). Taking together, these findings support our design rationale: observers rely heavily on perceptually available impact cues when constructing causal interpretations—consistent with our claim that causal strength can be manipulated via observable impact without depending on physical plausibility. More importantly, using these perceptual displays revealed that the temporal window of integration—central to multisensory processing (Vroomen & Keetels, 2010)—modulates not only simultaneity judgments but also explicit reports of causality and contact. While previous work has often inferred causality indirectly from behavioral adaptation (e.g., Buehner, 2012), our perceivable causal impression experiment design highlights a bidirectional relationship: causal perception shapes temporal binding, yet multisensory integration dynamics reciprocally shape causal and contact judgments.

This pattern aligns with a Bayesian causal inference account (Körding et al., 2007), wherein the perceptual system integrates all available sensory evidence to estimate the probability of a common source. Specifically, when visual kinematics strongly imply a launching interaction, the prior for a common cause is high; this elevates the tolerance for temporal discrepancy, allowing auditory and visual signals to be bound despite moderate offsets. On the other hand, in weaker causal contexts where this prior is weak, the same audiovisual delay more readily favors an independent-cause interpretation unless near-synchronous sound provides sufficient likelihood evidence for a shared cause. On this account, a causal impression arises from combining information across the senses. Visual and auditory cues are integrated to estimate how likely they come from the same event, and judgments of simultaneity and contact co-varied because they are different expressions of this single underlying inference process. Whether this influence reflects a fundamental, automatic perceptual mechanism or is specific to explicit reporting remains a question for future research using implicit measures.

Conclusion

To sum up, the temporal window of multisensory integration is not fixed but flexibly shaped by perceived causal impressions. Strong causal evidence widened the integration window, and this widening was directionally asymmetric — selectively expanding post-contact tolerance while leaving pre-contact windows unchanged, consistent with a forward causal prediction account. Meanwhile, synchronous and near synchronous auditory cue would enhance the causal impression of a visual contact. The intertwining relationship between simultaneity, contact, and explicit causal judgments indicates that these percepts reflect a shared inferential process rather than independent pathways. Together, the results support a Bayesian view in which human integrates evidence across senses to estimate common causation, and causal impression emerges when the accumulated multisensory evidence crosses a threshold for a unified event interpretation.

Acknowledgement

This work is part of Tianshu Chen's PhD project and Luyi Zou's MSc research project with Prof David Lagnado as supervisor.
We are grateful to anonymous reviewers and Dr Christos Bechlivanidis for the Gaussian curve fitting suggestion.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407.

Bechlivanidis, C., Buehner, M. J., Tecwyn, E. C., Lagnado, D. A., Hoerl, C., & McCormack, T. (2022). Human vision reconstructs time to satisfy causal constraints. *Psychological Science, 33*(2), 224–235.

Bechlivanidis, C., Schlottmann, A., & Lagnado, D. A. (2019). Causation without realism. *Journal of Experimental Psychology: General, 148*(5), 785–803.

Borhani, K., Beck, B., & Haggard, P. (2017). Choosing, doing, and controlling: Implicit sense of agency over somatosensory events. *Psychological Science, 28*(7), 882–893.

Buehner, M. J. (2012). Understanding the past, predicting the future: Causation, not intentional action, is the root of temporal binding. *Psychological Science, 23*(12), 1490–1497.

Buehner, M. J. (2015). Awareness of voluntary and involuntary causal actions and their outcomes. *Psychology of Consciousness: Theory, Research, and Practice, 2*(3), 237–248.

Cecere, R., Gross, J., & Thut, G. (2016). Behavioural evidence for separate mechanisms of audiovisual temporal binding as a function of leading sensory modality. *European Journal of Neuroscience, 43*(12), 1561–1568.

Choi, H., & Scholl, B. J. (2006). Measuring causal perception: Connections to representational momentum? *Acta Psychologica, 123*(1–2), 91–111.

Dixon, N. F., & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception, 9*(6), 719–721.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5*(4), 382–385.

Haggard, P., & Tsakiris, M. (2009). The experience of agency: Feelings, judgments, and responsibility. *Current Directions in Psychological Science, 18*(4), 242–246.

Hillock-Dunn, A., & Wallace, M. T. (2012). Developmental changes in the multisensory temporal binding window persist into adolescence. *Developmental Science, 15*(5), 688–696.

King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Experimental Brain Research, 60*(3), 492–500.

Klaffehn, A. L., Sellmann, F. B., Kirsch, W., Kunde, W., & Pfister, R. (2021). Temporal binding as multisensory integration: Manipulating perceptual certainty of actions and their effects. *Attention, Perception, & Psychophysics, 83*(8), 3135–3145.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE, 2*(9), e943.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain, 106*(3), 623–642.

Maunsell, J. H. R., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology, 68*(4), 1332–1344.

Michotte, A. (1963). *The perception of causality* (T. R. Miles & E. Miles, Trans.). Methuen. (Original work published 1946)

Prolific. (2026). *Prolific*. https://www.prolific.com

Recanzone, G. H., Guard, D. C., & Phan, M. L. (2000). Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *Journal of Neurophysiology, 83*(4), 2315–2331.

Rips, L. J. (2011). Causation from perception. *Perspectives on Psychological Science, 6*(1), 77–97.

Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology, 23*(3), 250–254.

Schlottmann, A., & Shanks, D. R. (1992). Evidence for a distinction between judged and perceived causality. *The Quarterly Journal of Experimental Psychology, 44*(2), 321–342.

Scholl, B. J., & Nakayama, K. (2002). Causal capture: Contextual effects on the perception of collision events. *Psychological Science, 13*(6), 493–498.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*(8), 299–309.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics, 72*(4), 871–884.

White, P. A. (2005). Visual causal impressions in the perception of several moving objects. *Visual Cognition, 12*(2), 395–404.

White, P. A., & Milne, A. (1997). Phenomenal causality: Impressions of pulling in the visual perception of objects in motion. *The American Journal of Psychology, 110*(4), 573–602.