Human Vision Reconstructs Time to Satisfy Causal Constraints

Abstract

The goal of perception is to infer the most plausible source of sensory stimulation. Unisensory perception of temporal order, however, appears to require no inference, since the order of events can be uniquely determined from the order in which sensory signals arrive. Here we demonstrate a novel perceptual illusion that casts doubt on this intuition: in three studies (N=607) the experienced event timings are determined by causality in real-time. Adult observers viewed a simple three-item sequence ACB, which is typically remembered as ABC (Bechlivanidis & Lagnado, 2016), in line with principles of causality. When asked to indicate the time at which events B and C occurred, points of subjective simultaneity shifted so that the assumed cause B appeared earlier and the assumed effect C later, despite full attention and repeated viewings. This first demonstration of causality reversing perceived temporal order cannot be explained by post-perceptual distortion, lapsed attention, or saccades.

Keywords: Perception; Causality; Time; Temporal Order

Statement of Relevance

There are two sources of information on the temporal order of events: the order in which we experience them, and their causal relationships, since causes precede their effects. Intuitively, direct experience of order is far more dependable than causal inference. Here, we showed participants events that looked like collisions but where the collided upon object started moving before the collision occurred. Surprisingly, participants indicated in real-time that they see events happening significantly earlier or later than they actually did, at timings compatible with causal interpretations (as if there were indeed a collision). This is evidence that perceived order is not the passive registration of the sequence of signals arriving at the observer, but an active interpretation informed by rich assumptions.

Main Text

Imagine your friend coming towards you holding a nice cup of tea. But her hands are wet and the cup slips, beginning its freefall. Even more unexpectedly, the cup shatters just *before* hitting the ground. Do you think that you'd spot such a weird succession of events? Previous research says that you probably wouldn't (Bechlivanidis & Lagnado, 2013, 2016; Tecwyn et al., 2020). Well, maybe you weren't paying attention or perhaps the scene was too weird to be remembered accurately. Now, imagine that the same scene is repeated again and again in front of your eyes, and you're asked to focus on the shattering and pinpoint the exact time when it happens. Do you think you would fare any better? Is our perception of time and temporal order a faithful reflection of what happens in the world (or at least what arrives at our retina) or can seemingly higher-level expectations, such as causality, affect the order in which we experience events occurring?

Past research shows that judgements of temporal order are not always accurate. In the prior entry effect (Titchener, 1908) attended events appear earlier due to privileged processing. Perhaps similarly, differences in luminance and contrast (Holcombe, 2015) affect the perceived order of events. In multisensory integration (Stein & Meredith, 1993), temporally separated stimulation is integrated to form unified and coherent percepts. When the timing of stimulus presentation is manipulated to occur closely before and after saccadic eye movements, their order is reversed (Kresevic, Marinovic, Johnston, & Arnold, 2016; Morrone, Ross, & Burr, 2005). More relevant to the current purposes, it has been shown that when presented with stimuli that give the impression of a recently learned (Bechlivanidis & Lagnado, 2013) or a familiar (Bechlivanidis & Lagnado, 2016) causal relationship which nevertheless violates the expected temporal order, adults and children as young as 4 (Tecwyn et al., 2020) report having seen the causal instead of the objective temporal order of events.

These prior demonstrations of order reversals, however (see Holcombe, 2015 for an extensive review), depend on split attention, stimuli that change between saccades, or integration of multimodal signals, and are usually revealed in post-hoc reports that are subject to memory distortions. Therefore, under conditions of unconstrained attention to uniform, unisensory stimuli, one would still intuit that, at the time of perception, the order of experiences will match the order of events in the world. In other words, provided that people attend closely to the events in question, use the same sensory modality, and there is no interval between their experience and its report, the perceived order will coincide with the order in which stimuli arrive at their sensory organs. Although this describes what we take to be an intuitive view, there are indeed theoretical accounts of experience, such as the brain time theory (Dennett & Kinsbourne, 1992; Holcombe, 2015) or the mirroring theory (Mellor, 1985; Phillips, 2014) that assume such a direct mapping between the temporal structure of reality and of experience. What underlies this intuition is that, whilst spatial perception, for instance, requires an inferential step to generate 3D percepts from retinal input, temporal order perception, at its most basic, does not require any inferential processes at all, since the perceptual input is itself temporally ordered. In other words, internal representations of temporal order, unless they involve cross-modal integration, switches in attention or substantial differences between stimuli, match the order of experienced external events.

Here we test this mirroring intuition by asking whether causality, which also carries temporal order information (since causes precede their effects), can affect the order in which events are perceived in real-time. We modified a paradigmatic Michottean (Michotte, 1963) causal sequence (two objects colliding), by adding a third object, to produce a domino effect collision involving three objects A, B and C. Critically, instead of the canonical order ABC (A collides with B which then collides with C), we presented a reordered version of the sequence: A moves first but at the time of its making contact with B, C starts moving, and B starts moving only 150ms later than that (i.e. ACB, Figure 1). Earlier research has demonstrated that this stimulus reliably leads people to i) report that A was the cause of B, and B the cause of C), and ii) remember having seen ABC instead of ACB (Bechlivanidis & Lagnado, 2016; Tecwyn et al., 2020). This is because, despite the objective ACB order, the event sequence best fits a causal schema ABC, where B is the presumed cause of C's motion, and thus B must have occurred before its supposed effect C. Until now, such distortions have only been demonstrated at the retrieval stage, and have been explained via the constructive nature of memory (Pedro, 2020; White, 2015).

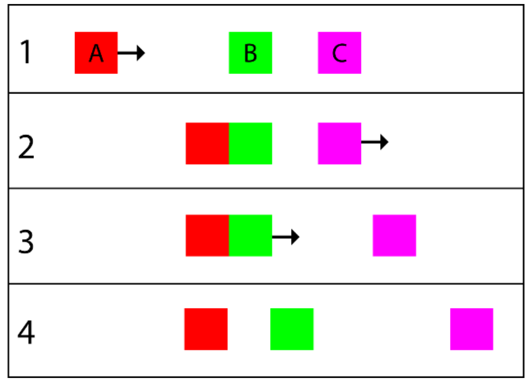

Figure 1: A reordered Michottean domino-like sequence (the arrows represent the time of motion onset). While it appears like a series of collisions, object C in fact moves (2) before its presumed cause B (3).

To examine the possibility of the alternative and more surprising perceptual explanation, rather than asking for judgements of order, that necessarily take place post-hoc, we asked participants to indicate in real-time when they see the objects move. If reordering occurs during retrieval, the subjective timings should be accurate. If, however, the effect is already present during encoding, then to turn a non-causal ACB sequence into the causally consistent ABC, either object B will be perceived as moving earlier than it does, or C will be seen as moving later, or both (Figure 2).

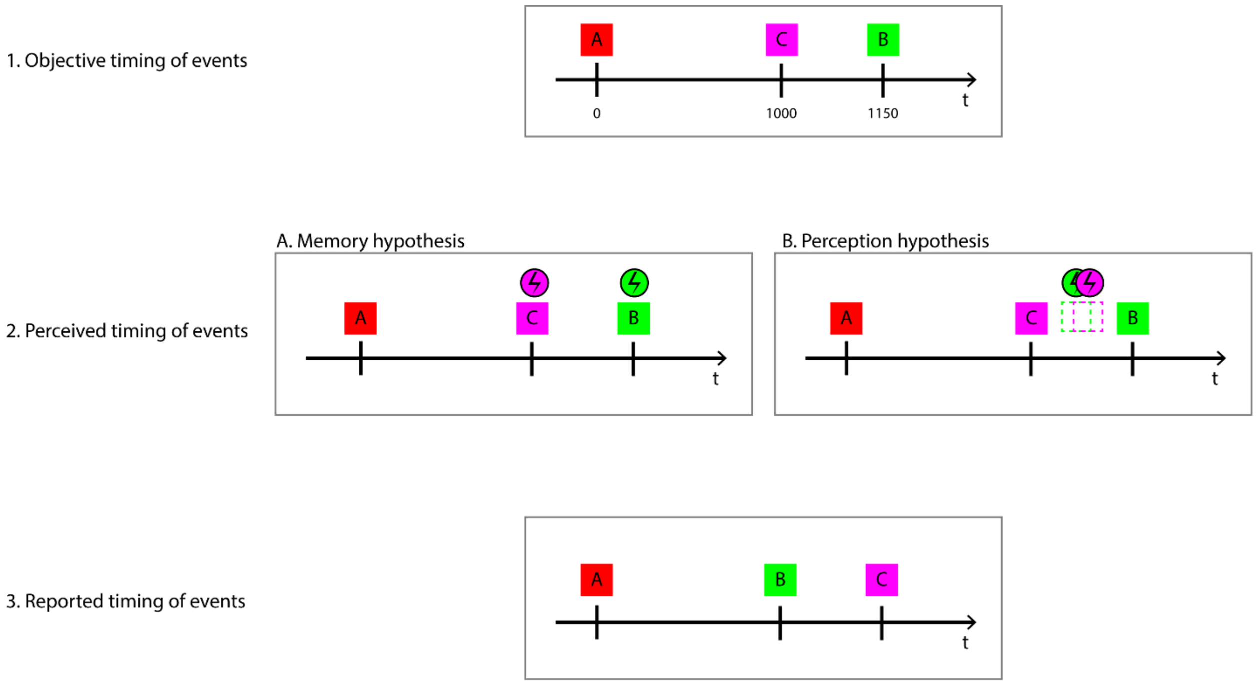

Figure 2: Given that previous research indicates that the reported ABC order (3) diverges from the objective ACB order (1), the question is whether this reordering occurs already at the time of perception or whether an initially accurate perception is distorted at a later stage, e.g. during retrieval. In the latter case, participants will be able to accurately synchronize an on-screen flash with the actual motion onsets of B and C (2A). If order perception is already distorted, however, participants will perceive B moving earlier and/or C moving later, and that would be reflected in the chosen temporal locations of the flash (2B). Although the absolute size of any temporal shift is not critical, note that the total temporal displacement required to turn a non-causal ACB sequence to the causal ABC sequence is 150ms, since that is the time that elapses between the time when C starts moving and the time when B does, in the former sequence.

Overview of experiments

In all experiments, participants saw variations of the animation depicted in Figure 1 and had to synchronize a non-localised on-screen flash with the motion onset of B or C. To that end, they were given unlimited attempts to adjust the timing of the flash via a slider, with each adjustment causing the clip to be played again using the updated flash timing. Our dependent variable was the point of subjective simultaneity (PSS - the temporal distance between motion onset of the target object and the final adjusted flash location). The critical clip was invariably the ACB sequence (Figure 3, top right panel), i.e. the clip where the temporal order does not match the apparent causal order of events. Rather than focusing on the absolute PSS, though, we were interested in comparing the PSS in the ACB clip against the PSS derived from clips where there was no tension between the temporal and causal order, either because by removing one of the objects, the appearance of any obvious causal relationship has also been removed (Experiment 1) or because the causal relation was congruent with the temporal order (Experiments 2 & 3).

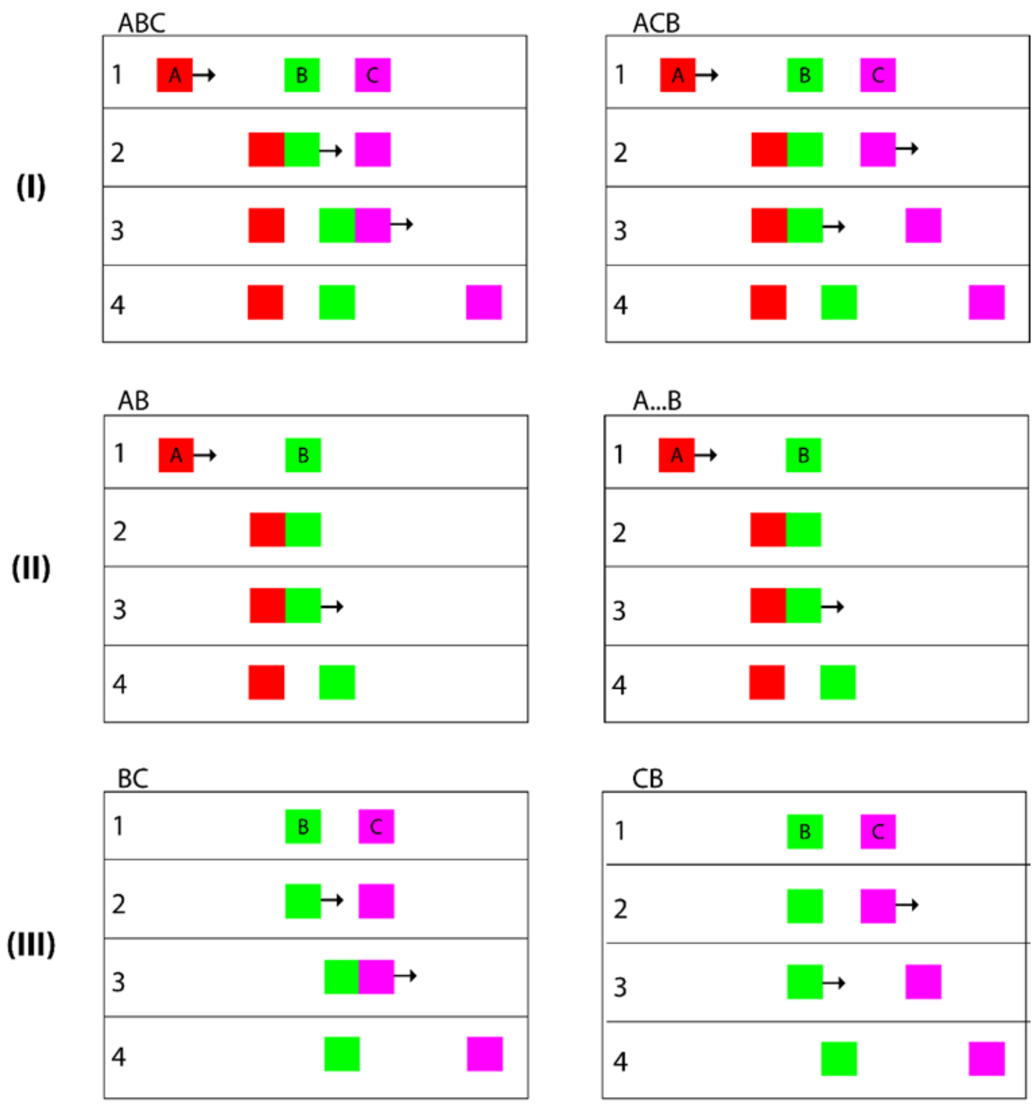

Figure 3: The target clips shown in the 3 experiments. On the left column, the order of events follows the causal direction. On the right column, object B starts moving with a delay (150ms) and after its presumed effect C (if present) has moved. On each column, the sequences shown in rows (II) and (III) are identical to the sequence shown in row (I), with a single object (A or C) removed. The clips on the right column were used in Experiment 1, the clips in rows (I) and (II) were used in Experiment 2, while Experiment 3 presented the clips shown in rows (I) and (III).

Experiment 1

Materials and Design

The experiment was approved by the UCL Research Ethics Committee (EP/2017/005), was preregistered at https://osf.io/w2qd4 and can be viewed at https://bit.ly/2BGP5rR. It was conducted within participants, with each participant seeing eight clips, four of which were non-target clips and acted as attention checks (see exclusion criteria below). All target clips followed the ACB order (Figure 3, right column) and differed in the number of objects present (two or three) and the object that participants were asked to synchronize the flash with (B or C). The two-object clips (A…B & CB) were included as comparisons to the critical ACB clip because they preserved the same temporal dynamics without implying a causal relationship. Therefore, the four target stimuli were ACB (sync with B), A…B (sync with B), ACB (sync with C) and CB (sync with C), as shown in Figure 3 (right column), resulting in a 2 x 2 within subjects design with the factors *Number of Objects* (3; 2) and *Synchronization Target* (B; C) The order of clip presentation was randomized per participant but the alternation between non-target and target clips was kept constant, starting with a non-target clip and then alternating between target and non-target clips.

All clips featured 2 or 3 red (#FF0000), green (#00FF00) and purple (#EC00F0) squares (30x30 pixels). The colors were randomly assigned in each clip, but the color of the sync target was kept constant for each participant. The squares moved at a constant speed of 0.2 pixels per frame, with the target frame rate set at 60 frames per second. The 2-object versions of the clips were identical to the 3-object ones with a single object being invisible: When the sync target was

B, object C was invisible, while when participants were asked to synchronize the flash with object C, object A was invisible[1].

The squares (2 or 3) were arranged in a row, 150 pixels from the top of the viewport (the user's visible area of a web page in the browser) and moved horizontally in the same direction, either to the left or to the right, randomly decided for each participant (in what follows we will describe only the left-to-right versions, since in the right-to-left direction clips were mirrored horizontally but otherwise identical). Square A was positioned 160 pixels from the left edge of the screen. Square B was placed 200 pixels to the right of square A and square C was placed 30 pixels to the right of square B. There was an initial period of no motion, randomly determined for each clip (1500-3400ms). This was especially important because, if the start time were fixed, the correct flash location would be identical between clips, possibly allowing transfer between trials and leading to order effects. When the clip started, square A travelled for 1000ms at 0.2 pixels/frame, stopping directly adjacent to square B. Critically, the next object to move was square C, while B moved 150ms later. Object B travelled for 30 pixels and object C for 200 pixels, both at 0.2 pixels/frame. The 2-object target clips were, as discussed, identical to 3-object ones with one of the squares being invisible. Thus, the A…B clip (Figure 3 II, right column) was the same as the ACB clip without square C, while the CB clip (Figure 3 III, right column) was the same as the ACB clip without object A.

The non-target clips featured 2 or 3 squares arranged in a vertical column, with a 30-pixel gap between them (equal to the height of each square). When the animation started, the squares

[1] The invisible object in 2-object clips was in fact still present in the animation but rendered in the same colour as the background (#FFFFFF). This was done to reduce the possibility of timing variations due to computational load. Both the intended and the actual timings of the events were recorded, showing no systematic variability.

moved horizontally in the same direction (to the left or to the right, randomly decided per participant) at 0.2 pixels/frame and came to a halt 200 pixels later. The order of motion onset was randomly determined per clip, but the relative timings were identical to those of the target clips (0ms, 1000ms, 1150ms).

At some point during the animation of each target or non-target clip, the whole viewport would flash black, i.e. the background color was set to black (#000000) for a single frame and back to white (#FFFFFF) again. The initial temporal position of the flash was randomly determined per participant to be either at the beginning of the clip (before any of the squares moved) or at the end (after all squares have reached their final location).

Below the clip, some of the instructions were repeated to participants (task and sync target, unrestricted number of attempts, performance-based fee) and below that there was a slider ranging from 0 to 4000ms (the actual values were not visible to participants, but the slider was labelled 'earlier' on its left and 'later' on its right edge). The position of the slider controlled the temporal position of the flash. Its initial position corresponded to the initial temporal location of the flash (extreme left = flash at 0ms, i.e. flash before animation, extreme right = flash at 4000ms, i.e. flash after animation).

Participants

We recruited 280 participants through Amazon Mechanical Turk. The sample size was decided through a power analysis based on earlier pilot studies (power: 80%, $\alpha=.05$). According to the preregistration plan we did not include in the analysis participants who, in any of the non-target clips reported a PSS exceeding 400ms in any direction (i.e. abs(PSS) > 400) or a PSS higher than 1000ms in the target clips. Eighty participants were removed based on the first criterion, while the second criterion did not lead to further exclusions. The final sample size consisted of 200 participants (mean age: 34.5, SD=10.7, 99 females) who received $0.20 for participating and an additional $0.30 if they passed the exclusion criteria.

Procedure

After providing informed consent, participants were asked for basic demographics (age, gender) and were introduced to the task: they would watch eight clips featuring moving squares. At some point during the animation the screen would flash black. A slider below the clip would allow them to adjust the temporal position of the flash. Their task was to adjust the flash position so that it occurred exactly when one of the squares started moving (the actual color of the square was mentioned but differed between participants). After each adjustment, the clip would be replayed. There was no limit in the number of adjustments allowed. Finally, it was explained to participants that their fee would depend on their performance in the task.

Participants then watched the eight clips, and for each one used the slider to adjust the temporal location of the flash. After each clip, participants were reminded that the task would remain the same for the next clip and that they had as many attempts as needed. Following the eight clips, participants were asked for any additional comment, were informed about their final fee and were thanked for participating.

Results

To reach PSS, participants made an average of 7.9 (SD=6.03) adjustments of the flash location per clip and thus watched each sequence as many times[2]. As can be seen in Figure 4, when presented with the ACB sequence and asked to indicate when object B started moving, participants positioned the flash on average 82.96ms (SD = 83.01) before it actually did and

[2] The raw data for all experiments can be found at https://osf.io/sz8yt/

when asked to indicate when object C started its motion, they positioned the flash on average 45.41ms (SD = 128.07) after C actually moved. The figure also shows a clear difference in PSS between clips where all objects were visible (orange bars) and clips where one of the objects was removed (blue bars) and thus causal impressions were weakened or not present at all (Michotte, 1963).

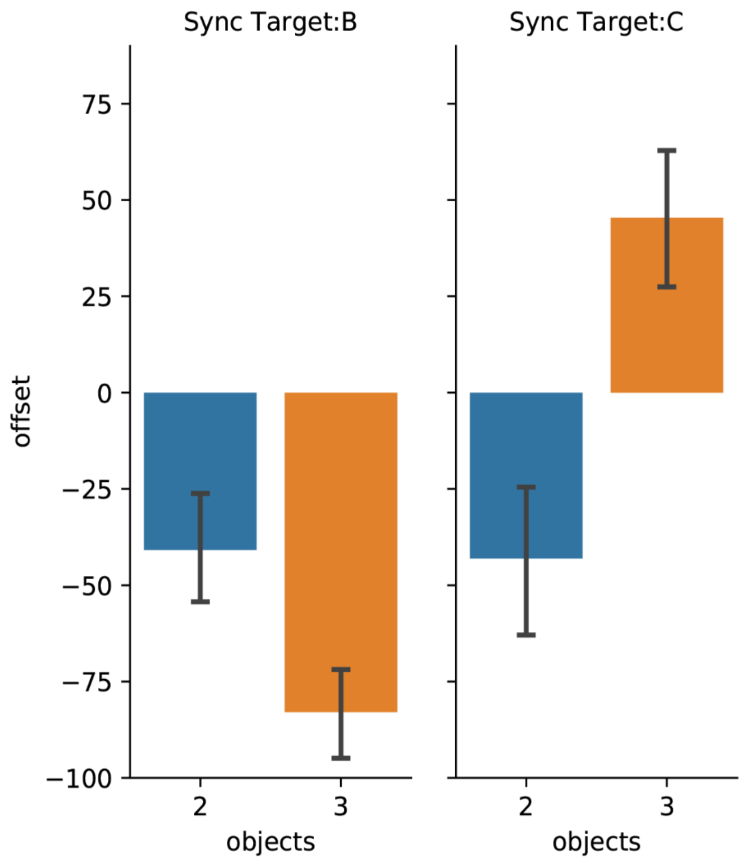

Figure 4: Average PSS per clip and synchronization target in 3-object ACB and 2-object A…B and CB sequences. Causal impressions will be weak or non-existent in the case of 2 object clips. Error bars represent 95% confidence intervals.

Formally, a repeated measures ANOVA[3] revealed significant main effects of the factors *Number of Objects* (i.e. 3- vs 2-object clips) , $F(1,199)=10.014$, $p=.002$, $\eta^2=.048$, and *Synchronization Target*, $F(1,199)=67.403$, $p<.001$, $\eta^2=.253$, and a significant interaction effect $F(1,199)=76.348$, $p<.001$, $\eta^2=.277$. When synchronizing the flash with the onset of object B's motion, the PSS was significantly lower when three objects were present (mean=-82.96, SD = 83.01) than with two objects (mean=-40.84, SD=93.59, $t(199) = 5.392$, $p<.001$, $d=.476$). When syncing with object C, the PSS was significantly higher in the three (mean=45.41, SD = 128.07) compared to the two objects clips (mean=-43.03, SD=139.43, $t(199) = 7.042$, $p<.001$, $d=.661$).

For each participant, we calculated the total PSS when syncing with B and when syncing with C and compared it against the objectively correct 0ms and against the minimum 150ms total deviation required for a causally plausible sequence. As shown in Figure 5, for the 3-object ACB clip, the total deviation (mean=128.36, SD=138.51) differed strongly from 0 ($t(199)=13.246$, $p<.001$, $d=.927$) but less so compared to 150 ($t(199)=2.210$, $p=.028$, $d=.156$). In contrast, for the 2-object clips (A…B and BC) the total deviation (mean=-2.19, SD=163.58) was not significantly different from 0 ($t(199) = -.189$, $p=.850$, $d=.013$) but was clearly lower than 150ms ($t(199)=3.157$, $p=.0018$, $d=.930$). The total PSS in 3-object clips was significantly higher than the total PSS in 2-object clips ($t(199)=8.738$, $p<.001$, $d=.861$). Finally, a McNemar test comparing the proportion of participants whose total deviation exceeded 150ms in the 3-object clip (45.5%) and 2-object clip (13.5%) was also significant ($\chi^2=91.0$, $p<.001$).

---

[3] Note that in most cases normality assumptions were violated (Shapiro-Wilk tests) leading us to conduct additional non-parametric equivalents for all tests reported here (e.g. Friedman chi square, Wilcoxon signed-rank). Given that we did not find any noteworthy differences between the parametric and the non-parametric tests, we opted to report the former which were preregistered, arguably robust to normality violations and more familiar to readers.

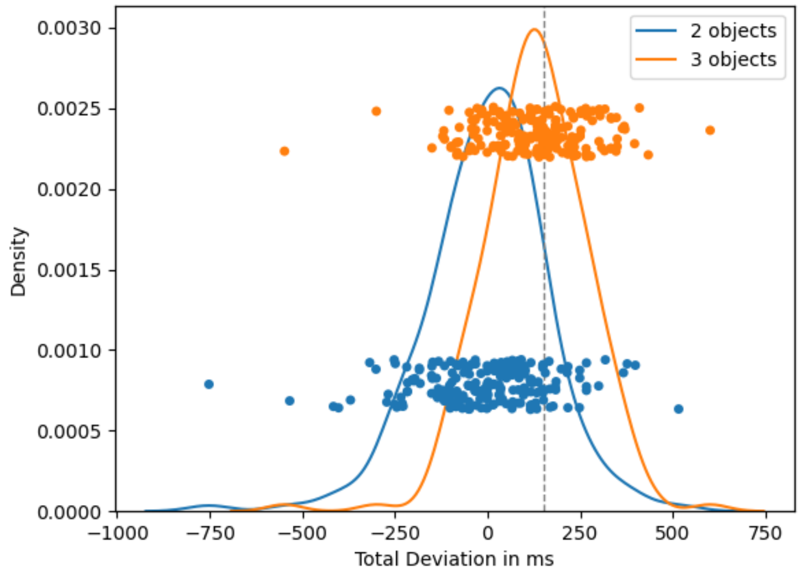

Figure 5: Total shift per participant in 3 and 2 object clips. The curves were computed through kernel density estimation (Gaussian kernel, Scott bandwidth). The vertical line marks the critical 150ms, the minimum total shift required to convert the ACB sequence to the causally canonical ABC sequence.

Discussion

The results support the perceptual basis of the effect. When watching the reordered ACB sequence, participants actually perceive B happening earlier and C happening later, at timings that in total approach the temporal displacement necessary to turn the ACB sequence into the causal ABC one. Displacements of such magnitude were not observed when one of the objects was hidden. It is thus the illusory causal context that produces the online reversal of temporal order: The insertion of C into A…B to produce ACB, shifts perception of B earlier in time to yield a causally meaningful ABC percept, while the addition of A to CB (to also produce ACB) pushes perception of C later in time to yield the ABC percept.

It may be argued that the difference between the 3-object sequence and its 2-object counterpart lies not only in the resulting causal impressions but also in varying perceptual loads.

If participants attempt to keep track of all objects present, an extra object might increase perceptual load which may explain the observed inaccuracies. Experiments 2 and 3 compare performance in the ACB clip against a 3-object ABC sequence where the causal and temporal orders coincide and thus no deviations are expected.

Experiment 2

Materials and Design

This experiment was preregistered at https://osf.io/64rjb and can be viewed at https://bit.ly/2BGP5rR. The design was adapted from Experiment 1, such that two of the target sequences were replaced by sequences where the causal and temporal order were congruent, and participants were asked to synchronize the flash only with the onset of object B in all clips. Specifically, to the 3-object ACB sequence and its 2 object A…B counterpart, we added the canonical ABC sequence and the 2-object AB sequence that results after removing C (Figure 3, I and II). Consequently, the 3-object sequences, ACB and ABC differ only with respect to the object that moves after A stops, and the 2-object sequences differ on whether B moves immediately (AB) or with a 150ms delay (A…B) after A stops moving. Thus, the design crossed the factors *Congruency* (Congruent [i.e. causal = temporal order]; Incongruent) and *Number of Objects* (3; 2). The non-target clips and the order of presentation remained the same.

Participants

A different sample of 280 participants was recruited through Mechanical Turk. The exclusion criteria were the same as before but instead of 400ms we used a stricter threshold of 300ms (pre-registered) for the non-target clips. As a result, 74 participants were excluded for an absolute PSS over 300ms in at least one of the non-target clips and another 2 participants for an absolute PSS exceeding 1000ms in at least one of the target clips. The resulting sample consisted of 204 participants (mean age: 35.5, SD=10.4, 94 females) who received $0.20 for participating and an additional $0.30 if they passed the exclusion criteria.

Procedure

The procedure was identical to Experiment 1 but this time participants were only asked to synchronize the flash with the onset of movement of object B in all clips.

Results

As before, participants changed the position of the flash and thus watched each clip 7.48 (SD=5.05) times on average. In the left panel of Figure 6, we can see that the results closely replicate the findings of Experiment 1. In the ACB clip participants placed the flash on average 83.97ms (SD=108.01) before B actually started moving and in the A…B clip 50.31ms (SD=118.29) earlier. However, when the temporal order matched the causal order of events (Figure 6, right panel), there was actually a small positive offset both when 3 objects (mean=17.63, SD=88.29) and when 2 objects (mean=14.54, SD=67.23) were present.

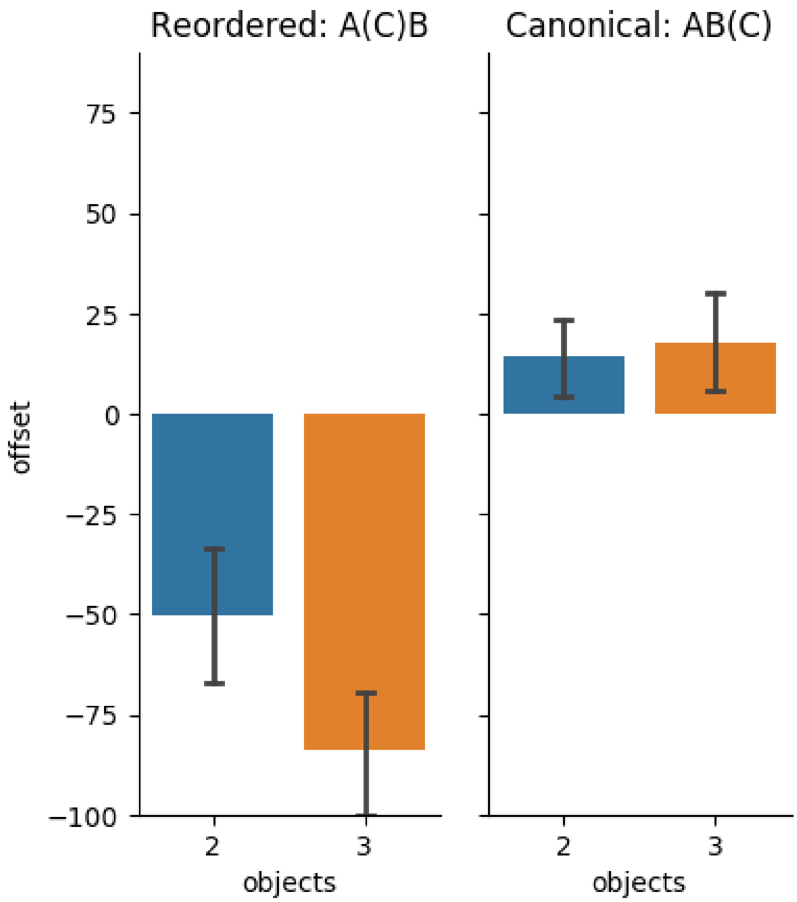

Figure 6: Average PSS per clip type (order of events) and number of objects present in Experiment 2 (synchronization was only to object B in this experiment). Error bars represent 95% confidence intervals.

A repeated measures ANOVA shows significant effects of *Congruency* ($F(1,203)=177.103$, $p<.001$, $\eta^2=.466$) and *Number of Objects* ($F(1,203)=6.029$, $p=.015$, $\eta^2=.029$) and, critically, their interaction ($F(1,203)=6.864$, $p=.009$, $\eta^2=.033$). Post-hoc paired t-tests show, as before, a significant difference between the PSS generated from 3-object ACB and the 2-object A…B ($t(203)=3.049$, $p=.003$, $d=.213$) but no significant difference between the 3-object ABC and its AB counterpart ($t(203)=0.420$, $p=.675$, $d=.029$). Finally, there was a significant difference between the 3-object ACB and the 3-object ABC clips ($t(203)=11.52$, $p<.001$, $d=.807$).

Discussion

Experiment 2 replicated the significantly negative offset of the perceived temporal location of B when the objective temporal order of events did not follow the causal order. Conversely, when the temporal order was congruent with the causal order, there was a small positive offset. This suggests that rather than the number of objects and the associated perceptual load, it is indeed causality that affects the perceived timing of B's onset of motion. Experiment 3 follows the same methodology, while asking participants to synchronize the flash with the onset of C.

Experiment 3

Materials and Design

This experiment was preregistered at https://osf.io/kcw4v and can be viewed at https://bit.ly/2BGP5rR. We used the same experimental design as in the first two experiments, but this time asked participants to synchronize the flash only with the onset of C. To that end, although the 3-object clips were the same as in Experiment 2 (reordered ACB and canonical ABC), the 2-objects clips were modified by rendering object A invisible, to generate the CB and BC clips (Figure 3, III). The design thus, as in Experiment 2, crossed the within-subjects factors *Congruency* (Congruent [i.e. causal = temporal order]; Incongruent) and *Number of Objects* (3; 2).

Participants

We once again recruited 280 participants and applied the same exclusion criteria as in Experiment 2, resulting in the exclusion of 74 participants for deviations exceeding 300ms in the non-target clips and 3 participants for offsets over 1000ms in one of the target clips. The resulting sample consisted of 203 participants with mean age 33.46 (SD=10.53), of which 97 were female. The participants received the same compensation as before.

Procedure

There were no changes in procedure compared to the other two experiments.

Results and Discussion

Participants required a similar number of 7.85 (SD=5.73) adjustments to reach PSS for each clip. As shown in Figure 7, we recorded the same effect as in Experiment 1, with a positive temporal displacement of the onset of C in the reordered ACB clip (mean=49.70, SD=106.37). When A is hidden and thus the causal impression is not present, the PSS for C turns negative (mean=-41.48, SD=129.19) as observed earlier. Both the 2-object and the 3-object canonical clips produced small offsets, negative in the former case (mean=-2.03, SD=124.36) and positive in the latter (mean=18.55, SD=89.97).

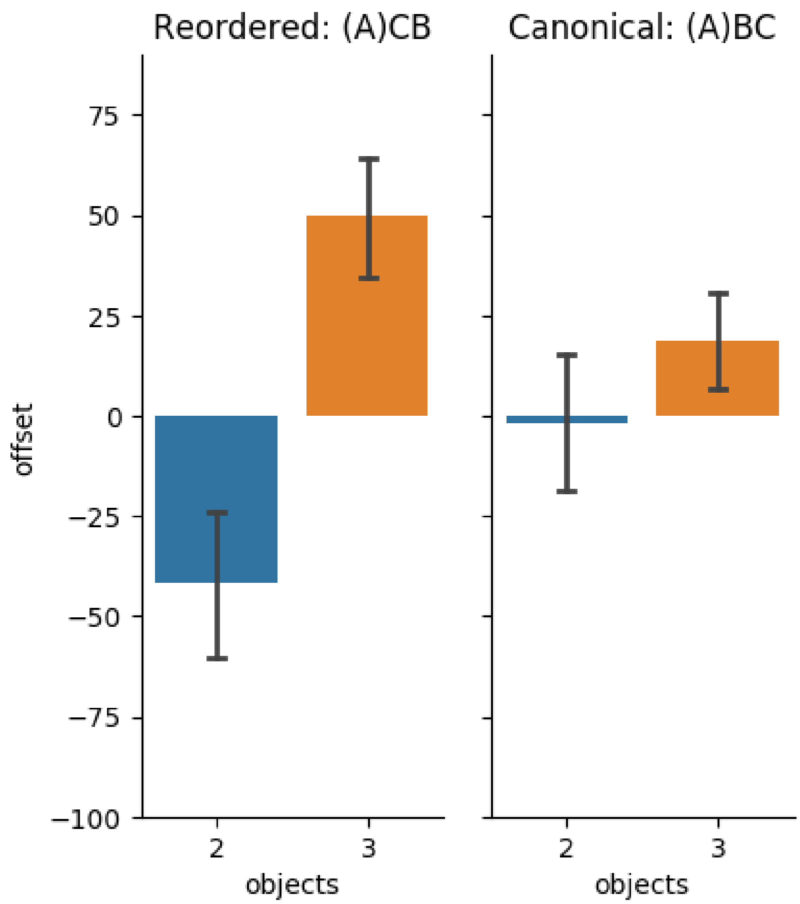

Figure 7: Average PSS per clip type (order of events) and number of objects present in Experiment 3 (synchronization was only to object C in this experiment). Error bars represent 95% confidence intervals.

A repeated measures ANOVA only showed a significant main effect of the *Number of Objects* (F(1,202)=48.130, p<.001, η^2=.192), but not a significant effect of *Congruency* (F(1,202)=0.306, p=.581, η^2=.002). Crucially, the interaction was significant, F(1,202)=23.868, p<.001, η^2=.106. Planned post hoc t-tests showed that the offsets in the ACB clip were significantly higher compared both to the offsets in the CB counterpart (t(202)=7.992, p<.001, d=.561) and the canonical ABC clip (t(202)=3.696, p<.001, d=.259).

Experiment 3 thus corroborated the findings in Experiments 1 and 2, showing that it is not the number of objects that produces the perceived temporal displacement of events. Only in the presence of a causal expectation or a causal impression and only when that is incongruent with the objective temporal order of events, does the perceptual system shift the timing of events to match a causal interpretation.

General Discussion

Collectively, our findings constitute the first demonstration of a unisensory perceptual illusion of temporal order induced by causal impressions, indicating that the visual system generates the experienced order through a process of interpretation (Grush, 2016; Holcombe, 2015). Participants were given precise instructions and sufficient time to repeatedly view the sequences, they attended to the critical events using the same modality, and they synchronized object motion with a non-localized flash. We can thus confidently rule out alternative explanations based on inattentional blindness, multimodal integration, flash lag and motion-after effects. Because stimulus presentation was free and unconstrained relative to the time of saccades, our results cannot be accounted for by transient perisaccadic mislocalisation either (Kresevic et al., 2016; Morrone et al., 2005). Although in this case we examined the effect only with an adult population recruited from a crowdsourcing platform, previous research suggests that children as young as four are also susceptible to causal reordering, at least when using post-hoc reports (Tecwyn et al 2020). More research needs to be carried out to study the degree of perceptual shift and, more broadly, the generalizability of the current results.

One potential limitation of the work reported here is the possibility[4] that the perceptual signals of temporal order in our stimuli could have been ambiguous, i.e. within the range of a just noticeable difference (JND). Specifically, it is possible that when the onset of motion and the

[4] We'd like to thank one of our reviewers, Alex Holcombe, for suggesting this.

flash are less than 100ms apart, they fall within the same simultaneity window (Holcombe, 2015) and, therefore, causal impressions are guiding temporal order judgment in the presence of a completely uninformative temporal signal. We are unaware of any prior research examining JNDs of temporal order judgments in perceptual causality stimuli. However, more generally, JNDs in temporal order judgements tend to be much smaller than the differences we report, e.g. 5ms (Sweet, 1953) up to 40ms (Tadin, Lappin, Blake, & Glasser, 2010), depending on the nature of the stimuli. However, more generally, JNDs in temporal order judgements tend to be much smaller than the differences we report, e.g. 5ms (Sweet, 1953) up to 40ms (Tadin, Lappin, Blake, & Glasser, 2010) depending on the nature of the stimuli. One preparation, aimed at deliberately interfering with temporal order judgments by surrounding each critical target event with ten extraneous flickering discs saw JNDs deteriorate as far as 110ms (Cass & Van Der Berg, 2014). Because we observed offsets as large as 80ms and thus outside the range of typically observed JNDs, we are reasonably confident that our results go beyond mere cognitive bias. However, the extent with which preceding and/or subsequent motion of the other object(s) and the possibility of the flash masking critical events may have interfered with temporal order judgment is unknown, and thus we cannot rule out the possibility that causal impressions served as a cognitive factor biasing perceptual responses in light of ambiguous perceptual signals.

It is interesting to note that although our interpretation of the current findings hinges on the presence of causal impressions and participants in past research are indeed reporting such strong impressions (Bechlivanidis & Lagnado, 2016; Tecwyn et al 2020), the critical ACB sequence is objectively not causal. How can a causal impression, strong enough to undermine temporal information, be generated from non-causal stimuli? This is a recurring question in the causal perception literature. The almost universal causal impressions resulting from a prototypical Michottean launching stimulus (Michotte, 1963) are often described as illusions of causality (White, 2006) since the stimulus consists of a highly improbable frictionless, perfectly elastic collision (Runeson, 1983; White, 1988). The explanations offered in that case and which may also apply to our stimuli, refer either to the similarity of the stimulus with a stored schema (Weir, 1978; White, 2006) or to the inadvertent activation of a low-level causal detector (Michotte, 1963; Scholl & Tremoulet, 2000). Thus, one possibility is that the ACB sequence is, despite its inconsistencies, similar enough to a series of collisions, that a causal schema of a domino-like effect remains the most plausible account of what transpired. Alternatively, the speculated low-level causal detector might be activated, since many of the cues to causality (spatiotemporal contiguity, property transmission) are present. The only difference between the causal ABC and the non-causal ACB sequence is that the identity of the object that moves after contact does not match the identity of the object that was interacted upon, and this might not suffice to preclude a causal impression.

The influence of causality on time perception is also apparent in multisensory integration (Stein & Meredith, 1993). However, in that case the temporal distortions are usually explained as the attempt of the perceptual system to account for the different transmission media and transduction speeds between, for example, visual and auditory signals with a common source. Our results show that the assumed causal structure of the incoming signals (common cause in multisensory integration or causal chains here) affects the experienced timing of those signals, even in the absence of variable transmission/transduction speeds. A general principle emerges, according to which the relative timing of signal arrival is superseded by inferences regarding the timing of transmission, irrespective of the nature of those signals.

Regarding the process-basis of the reordering effect, we discern two possibilities: Based on predictive coding (Hosoya, Baccus, & Meister, 2005) or integration of sensory evidence with prior experience (Eagleman & Holcombe, 2002), strong causal expectations overpower the information from the incoming visual signal. Alternatively, if some causal impressions are, as has been argued (Bechlivanidis, Schlottmann, & Lagnado, 2019; Schlottmann, 2000; Scholl & Tremoulet, 2000), the result of low-level perceptual processes, our stimulus is generating two contradicting sensory signals, one due to the objective temporal order and one due to the implied causal order of events, with the latter being weighted more heavily. As in the checkershadow illusion (Adelson, 1995) where color perception is shown to incorporate assumptions about shadows, temporal order perception is shown here to account for assumptions about causality. And as the recipient of two letters does not rely solely on their order of arrival to infer the order of posting (Dennett & Kinsbourne, 1992), the human visual system uses causation as a postmark to determine the most plausible order of events in the world.

References

Adelson, E. H. (1995). Checkershadow illusion. Retrieved from http://persci.mit.edu/gallery/checkershadow

Bechlivanidis, C., & Lagnado, D. A. (2013). Does the “why” tell us the “when”? *Psychological Science*, *24*(8), 1563–1572. https://doi.org/10.1177/0956797613476046

Bechlivanidis, C., & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order. *Cognition*, *146*, 58–66. https://doi.org/10.1016/j.cognition.2015.09.001

Bechlivanidis, C., Schlottmann, A., & Lagnado, D. A. (2019). Causation without realism.

Journal of Experimental Psychology: General, *148*(5), 785–804. https://doi.org/10.1037/xge0000602

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. https://doi.org/10.1177/0956797612444612

Cass, J., & Van der Burg, E. (2014). Remote temporal camouflage: Contextual flicker disrupts perceived visual temporal order. Vision Research, 103(5), 92–100. https://doi.org/10.1016/j.visres.2014.08.008

Dennett, D. C., & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences*, *15*, 183–247. https://doi.org/10.1017/S0140525X12000477

Eagleman, D. M., & Holcombe, A. O. (2002). Causality and the perception of time. *Trends in Cognitive Sciences*, *6*(8), 323–325. https://doi.org/10.1016/S1364-6613(02)01945-9

Grush, R. (2016). On the temporal character of temporal experience, its scale non-invariance, and its small scale structure.

Holcombe, A. O. (2015). The temporal organization of perception. In J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199686858.013.057

Hosoya, T., Baccus, S. A., & Meister, M. (2005). Dynamic predictive coding by the retina. *Nature*, *436*(7047), 71–77. https://doi.org/10.1038/nature03689

Hubbard, T. L. (2012). Phenomenal causality II: integration and implication. Axiomathes, 1 40. https://doi.org/DOI 10.1007/s10516-012-9200-5

Hubbard, T. L. (2013). Phenomenal Causality I: Varieties and Variables. Axiomathes, 1 42. https://doi.org/DOI 10.1007/s10516-012-9198-8

Humphreys, G. R., & Buehner, M. J. (2010). Temporal binding of action and effect in interval reproduction. *Experimental Brain Research*, *203*(2), 465–470. https://doi.org/10.1007/s00221-010-2199-1

Kresevic, J. L., Marinovic, W., Johnston, A., & Arnold, D. H. (2016). Time order reversals and saccades. *Vision Research*, *125*, 23–29. https://doi.org/10.1016/j.visres.2016.04.005

Mellor, D. H. (1985). *Real Time*. Cambridge: Cambridge University Press.

Michotte, A. (1963). *The Perception of Causality*. London: Methuen & Co Ltd.

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience*, *8*(7), 950–954. https://doi.org/10.1038/nn1488

Pedro, T. (2020). Causal Beliefs and Perception of Temporal Order. The ‘Reordering Effect.’ *Timing & Time Perception*, *8*(2), 137–161. https://doi.org/https://doi.org/10.1163/22134468-20191155

Phillips, I. B. (2014). Experience of and in time. *Philosophy Compass*, *9*(2), 131–144. https://doi.org/10.1111/phc3.12107

Runeson, S. (1983). *On visual perception of dynamic events*. *Acta Universitatis Upsaliensis: Studia Psychologica Upsaliensia*.

Schlottmann, A. (2000). Is perception of causality modular? *Trends in Cognitive Sciences*, *4*(12), 441–442.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive*

Sciences, *4*(8), 299–309. https://doi.org/10.1016/S1364-6613(00)01506-0

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. MIT Press.

Sweet, A. L. (1953). Temporal discrimination by the human eye. The American journal of psychology, 66(2), 185-198.

Tadin, D., Lappin, J. S., Blake, R., & Glasser, D. M. (2010). High temporal precision for perceiving event offsets. Vision research, 50(19), 1966-1971.

Tecwyn, E. C., Bechlivanidis, C., Lagnado, D. A., Hoerl, C., Lorimer, S., Blakey, E., … Buehner, M. J. (2020). Causality Influences Children's and Adults' Experience of Temporal Order. *Developmental Psychology*, *56*(4), 739–755. https://doi.org/10.1037/dev0000889

Titchener, E. B. (1908). *Lectures on the elementary psychology of feeling and attention*. New York: Macmillan.

Weir, S. (1978). The perception of motion: Michotte revisited. *Perception*, *7*(3), 247–260. https://doi.org/10.1068/p070247

White, P. A. (1988). Causal processing: Origins and development. *Psychological Bulletin*, *104*(1), 36–52. https://doi.org/10.1037//0033-2909.104.1.36

White, P. A. (2006). The role of activity in visual impressions of causality. *Acta Psychologica*, *123*(1–2), 166–185. https://doi.org/10.1016/j.actpsy.2006.05.002

White, P. A. (2015). Visual impressions of generative transmission. *Visual Cognition*, *6285*(March), 1–37. https://doi.org/10.1080/13506285.2016.1149533