

Causal Action: A Fundamental Constraint on Perception and Inference About Body Movements



Yujia Peng¹, Steven Thurman^{1,2}, and Hongjing Lu^{1,3}

¹Department of Psychology, University of California, Los Angeles; ²U.S. Army Research Laboratory, Human Research and Engineering Directorate, Aberdeen Proving Ground, Maryland; and

³Department of Statistics, University of California, Los Angeles

Psychological Science
2017, Vol. 28(6) 798–807
© The Author(s) 2017
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797617697739
www.psychologicalscience.org/PS



Abstract

The human body navigates the environment via locomotory movements that leverage gravity and limb biomechanics to propel the body in a particular direction. This process creates a causal link between limb movements and whole-body translation. However, it is unknown whether humans use this causal relation as a constraint in perception and inference with body movements. In the present study, participants rated actions of other individuals as more natural when limb movements (as a cause) occurred before body displacements (as an effect) than when limb movements temporally lagged behind body displacements. This causal expectation for human body movements not only affected perceptual impressions regarding the naturalness of observed actions but also guided the interpretation of motion cues within a more generalized causal context. We interpret these results within a framework of causality as evidence that the constraint of causal action plays an important role in perception and inference with body movements.

Keywords

causality, biological motion, action, causal perception, open data

Received 2/17/16; Revision accepted 2/13/17

Seminal works (Heider & Simmel, 1944; Michotte, 1946/1963), coupled with contemporary developments (e.g., Scholl & Tremoulet, 2000; White, 2006), have demonstrated an intimate connection between causation and perception. A causal impression can directly arise from our perception of the world and can influence further perceptual judgments, such as event timing (Bechlivanidis & Lagnado, 2013, 2016). However, previous studies of causal perception have mostly focused on the interactions of inanimate objects in the physical world (e.g., colliding balls), which has limited generalization to more complex visual inputs, such as human actions. In contrast to dynamic events involving objects, human actions have a special status in that they afford privileged access to the experience of agency and enable discovery of causal relations in the physical and social environment through purposeful interventions (Abravanel, Levan-Goldschmidt, & Stevenson, 1976; White, 1999).

Consider one simple example of human actions. The human body navigates the environment via locomotory movements that leverage gravity and limb biomechanics

to propel the body in a particular direction. This process creates a causal link between limb movements and whole-body translation, resulting in expectations about the relation between the two motion cues (i.e., limb movements in the body-centered reference frame in relation to body displacements in the environmental reference frame). This causal linkage may help explain why the “moonwalk” dance move popularized by Michael Jackson is experienced as surprising or even thrilling. While the dancer moves his or her legs in a way that appears to simulate walking forward, the whole body glides seamlessly backward, creating a dramatic conflict with the expected relationship between limb movements and body displacements.

Recent research has revealed that humans are sensitive to the temporal binding between limb movements and body displacements, given that we commonly observe the

Corresponding Author:

Yujia Peng, Department of Psychology, University of California, Los Angeles, 6564 Franz Hall, 502 Portola Plaza, Los Angeles, CA 90095
E-mail: yjpeng@ucla.edu

two types of motion occurring in near synchrony. Disrupting the temporal congruency between the two sources of motion information curtails the perception of animacy (Thurman & Lu, 2013), the detection of social interaction between two agents (Thurman & Lu, 2014), and the discrimination of locomotion style (Masselink & Lappe, 2015; Thurman & Lu, 2016). However, it is unclear whether people show tolerance to some situations in which limb movements and body displacements are temporally misaligned but in a causally consistent way (e.g., limb movements may be shifted ahead in time but still precede body displacements in locomotion). In other words, is the degree of tolerance constrained by the directionality of the causal relation between the motion cues?

The present study addresses this question by examining how the cause-effect relation inherent in body movements affects the perception and inference of actions. We used a key manipulation based on a ubiquitous feature of causation: the temporal-priority principle, which holds that a cause must precede its effect (Hume, 1739/1888; Price, 1992; White, 2006; Bechlivanidis & Lagnado, 2013). In Experiment 1, we systematically manipulated the direction and magnitude of temporal offsets between limb movements and whole-body displacements. If the causal relation between the two motion cues is important, the temporal-priority principle would predict that when body displacements (the effect) temporally lag behind limb movements (the cause), observers may show greater tolerance to a deviation from close simultaneity. In this situation, the temporal relation between limb movements and body displacements remains qualitatively consistent with normal causal directionality. In contrast, when body displacement (the effect) is shifted earlier in time to occur before the supposed cause (limb movements), observers may show less tolerance because of the strong violation of the causal expectation. However, if temporal alignment per se is the critical factor (without consideration of causal directionality), then we would expect a symmetric influence of temporal offsets on perceived naturalness of actions. In Experiment 2, we varied the cover story associated with identical stimuli to examine whether inference judgments shift to conform to the causal context when different beliefs are induced. Together with a series of control experiments, results from the present study provide evidence against mere associative learning and instead support the hypothesis that causal relations in body movements play an integral role in our perception and inference with actions.

Experiment 1

Experiment 1 was designed to assess how the directionality of temporal offsets between limb movements and body displacements affects the perceived naturalness of human actions.

Method

Participants. One hundred nine online participants were recruited through Amazon's Mechanical Turk (MTurk). Each was paid \$1 for participating in the online experiment (average duration of 8 min). All experimental procedures were approved by the Committee for Protection of Human Subjects at the University of California, Los Angeles (UCLA). The sample size was determined on the basis of previous research on action recognition using MTurk (Shu, Thurman, Chen, Zhu, & Lu, 2016). Data collection for the online experiment stopped on the day when the expected sample size was reached.

Stimuli. Action stimuli were generated from the Carnegie Mellon University Motion Capture Database (<http://mocap.cs.cmu.edu>) and processed using the Biological Motion Toolbox developed by van Boxtel and Lu (2013). We selected actions in which a person walked on an uneven surface with invisible steps, and both horizontal and vertical body displacements were included in the action sequence. The stimuli used in the experiments appear in Videos S1 and S2 in the Supplemental Material available online; they can also be viewed at <http://cvi.psych.ucla.edu/causal-action-2016.html>.

Body displacements were computed as the change in the average position of the two hip joints in time, and limb movements were defined as the residual motion after subtracting the body-motion component on a frame-by-frame basis. The temporal relationship between limb movements and body displacements was manipulated by shifting the sequence of body displacements forward or backward in time relative to the sequence of limb movements, as illustrated in Figure 1a. Body displacements were manipulated to either lead or lag behind the posture change resulting from the limb movements. In the lead condition, the temporal sequence of body positions was shifted forward in time relative to limb movements (i.e., the effect preceded); in the lag condition, the temporal sequence of body positions was shifted backward in time (i.e., the cause preceded).

Procedure. Participants were presented with the following cover story:

Imagine you are viewing a walking sequence on an uneven surface with invisible steps through a slowly rotating camera. The rotation of the camera will help you perceive the 3D space. Look at the relative limb movements of the walker. Look at how the body position changes over time. Ask yourself if that could be a real person's motion in the environment.

Participants were asked to rate the naturalness of the videos on a scale from 1 (*unnatural*) to 5 (*natural*).

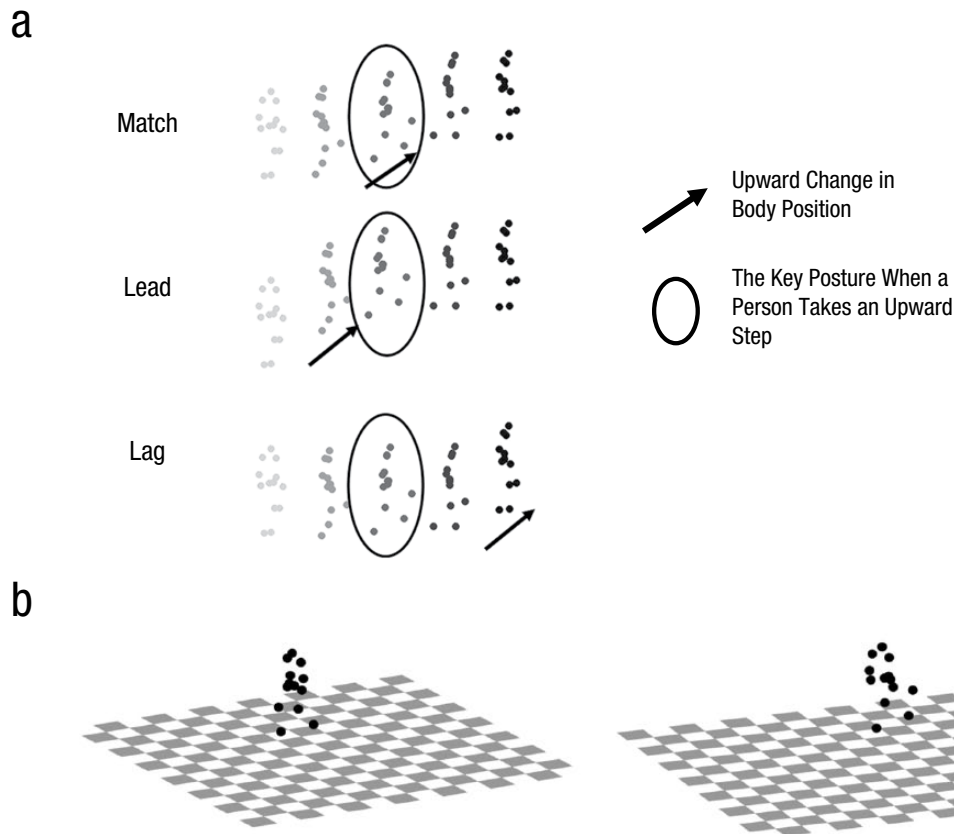


Fig. 1. Illustrations of the stimuli in Experiment 1. The dots in (a) represent point-light walkers with different temporal relationships between body displacements and limb movements. The ellipses circle the key posture when a person takes an upward step, and the black arrows indicate the upward change of body position that is associated with such a step. The point lights in the walker change from light to dark color to denote elapsed time. In the match condition, the posture and body displacements were in synchrony. In the lead condition, the body position changed before the limbs moved. In the lag condition, the body position changed after the limbs moved. The stimulus frames in (b) were taken from a sequence in which a point-light walker moved on an uneven surface that had invisible steps. The image slowly rotated in a clockwise direction. Videos S1 and S2 in the Supplemental Material show examples of dynamic stimuli. The videos are also available at <http://cvl.psych.ucla.edu/causal-action-2016.html>.

On each trial, participants saw a point-light actor walking on a checkerboard surface with invisible steps, as shown in Figure 1b. The camera rotated in a counter-clockwise direction (meaning that the checkerboard and the actor appeared to rotate clockwise) at a speed of $3^\circ/\text{s}$, which was intended to facilitate 3-D perception of the action stimulus in the environment. Each video lasted 8.33 s and consisted of 500 frames selected from the original 2,000-frame videos based on the 32-s motion-capture data.

Actions were presented with a randomly selected starting viewpoint ($\pm 45^\circ$ from a side view) to ensure that any effect did not depend on a specific viewpoint. Six temporal offsets between limb movements and body displacements were used: 0, ± 0.5 , ± 1.0 , and 8.33 s (corresponding to 0, ± 30 , ± 60 , and 500 frames at a refresh rate of 60 Hz). The conditions with no offset and a large offset

(i.e., 8.33 s) served as extreme cases to help participants anchor the two ends of the rating scale. Positive offsets constituted the lead condition (i.e., the effect of body displacements occurred before the causal cue of limb movements); negative offsets constituted the lag condition (i.e., the effect of body displacements occurred after the causal cue of limb movements).

The experimental procedure included two blocks. Each block consisted of 12 experimental trials (two starting viewpoints for each of six temporal offsets) and one randomly placed attention-check trial. The attention-check trial assigned a trivial task; participants were presented with either a walking or jumping sequence and were asked to identify the presented action. The purpose of including these two attention-check trials was to identify outlier participants who gave random responses in the online experiment.

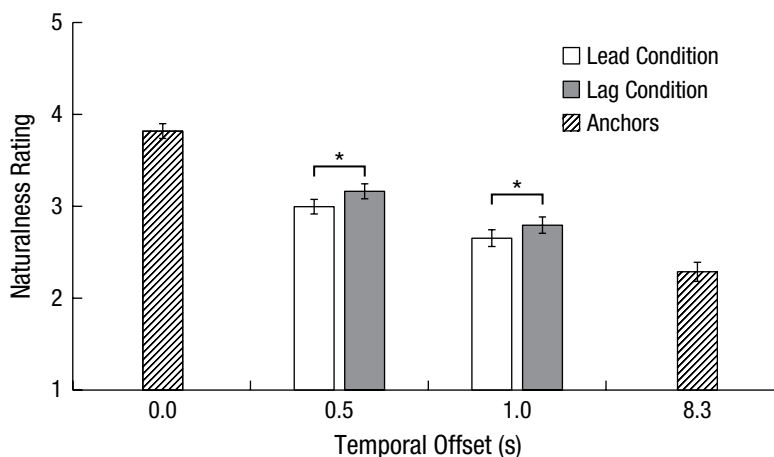


Fig. 2. Results from Experiment 1: mean naturalness rating as a function of temporal offset between limb movements and body displacements, presented separately for the lead condition and the lag condition. Also shown are the mean naturalness ratings in the conditions with offsets of 0.0 and 8.3 s, which anchored the range of ratings. Error bars represent ± 1 SEM. Asterisks indicate statistically significant differences between conditions ($*p < .05$).

Results

Fourteen of the 109 online participants were removed from the analysis because they failed to satisfy the inclusion criteria. Specifically, 9 participants were excluded because they failed to recognize the simple actions in both of the attention-check trials. Data from 5 participants were excluded because they provided the same ratings for all trials in the experiment.

As expected, naturalness ratings were highest in the zero-offset condition (i.e., perfect synchrony between limb movements and body displacements; $M = 3.82$, $SD = 0.79$); naturalness ratings were lowest in the condition with a temporal offset of 8.3 s ($M = 2.29$, $SD = 1.01$). These results demonstrate that human observers are generally sensitive to the magnitude of temporal offsets between limb movements and body displacements. The two extreme conditions (i.e., 0.0 s and 8.3 s) did not include the directional temporal shifts to generate the lead and lag offsets; consequently, ratings for these conditions were not included in the following analyses.

To examine how the directionality of temporal offsets between the two movement cues influenced naturalness ratings, we conducted repeated measures analyses of variance (ANOVAs) with two within-subjects factors, temporal offset magnitude (0.5, 1.0 s) and offset direction (lead, lag). As shown in Figure 2, the results revealed a significant main effect of temporal offset direction, $F(1, 94) = 8.95$, $p = .004$, $\eta_p^2 = .842$. This finding indicates that observers judged actions to be more natural when the temporal offset was consistent with the expected causal direction (lag condition) than when there was an equal amount of temporal offset in the lead condition (i.e., when the temporal offset was opposite the causal direction). As

expected, we found a significant main effect of offset magnitude, $F(1, 94) = 28.73$, $p < .001$, $\eta_p^2 = 1.0$, which indicates that people were sensitive to the general degree of temporal alignment between limb movements and body displacements when assessing the validity of observed actions, and larger offsets resulted in lower naturalness ratings. The two-way interaction between offset magnitude and temporal direction was not significant, $F(1, 94) = 0.10$, $p = .754$. Although the rating differences in the absolute scale may appear small, it should be emphasized that (as noted previously) participants' ratings did not span the full 5-point scale. These mean ratings indicate that participants tended to provide naturalness ratings in the middle range, as long as the observed limb movements did not obviously violate biological constraints (which is consistent with results from a previous study on action recognition; Thurman & Lu, 2013).

Experiment 2a

In Experiment 2, we aimed to gauge the inferential aspects of causality in action perception. We drew on a design typically used in studies of causal inference to explicitly separate the causal cue and its effect. We created a reasoning task in which the two movement cues were presented by distinct visual entities in the display. This new reasoning task assessed whether people used the default causal relation to infer the binding between the two types of movements.

Method

Participants. Twenty UCLA undergraduate students (mean age = 20.8 years; 14 female) participated in the

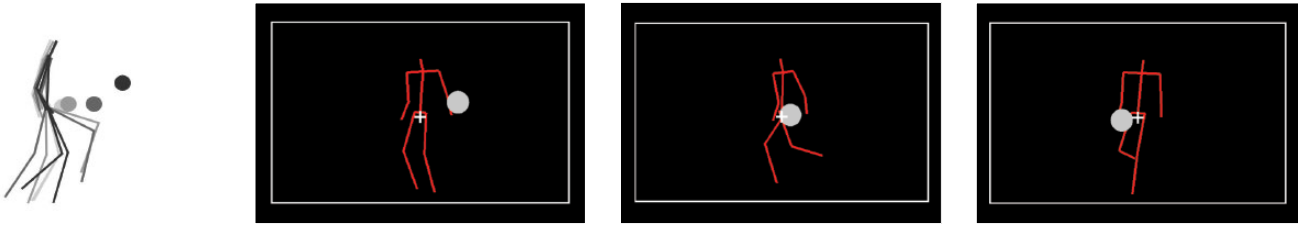


Fig. 3. Illustrations of the stimuli in Experiments 2a and 2b. The illustration on the left shows several possible limb movements for a stick-figure walker resulting from posture changes over time. The walker remained in a stationary location; the dot depicts the change in body position that results from body displacements. The sticks in the walker and the dot change from light to dark color to denote elapsed time. Three sample frames from an experimental trial are shown on the right to demonstrate how a dot (represented as a GPS dot in Experiment 2a and a laser spot in Experiment 2b) moved according to the assigned trajectory of body displacements with a particular temporal offset from limb movements. Videos S3 and S4 in the Supplemental Material show examples of the dynamic stimuli used in these experiments (they are also available at <http://cvl.psych.ucla.edu/causal-action-2016.html>).

experiment for course credit. All participants had normal or corrected-to-normal vision. The sample size was estimated on the basis of sample sizes in previous studies of causal perception in a laboratory setup ($N = 14$ in Scholl & Nakayama, 2002). Data collection for this experiment stopped in the week when the expected sample size was reached.

Stimuli. Four action sequences of a person walking on an uneven surface were displayed from two viewing directions with orthogonal projection. The size of the walker was a maximum of 3.2° wide by 5° high. The walker was displayed as a red stick figure (3.5 cd/m^2) on a black background (0 cd/m^2); the size of the frame was 22° by 12° . The walker appeared to walk on a treadmill by maintaining a stationary position for the average location of two hip joints at the center fixation point. A gray dot (75.9 cd/m^2 ; diameter = 1°), tracking the position change of the body over time (as a GPS navigation system shows the location of a vehicle), moved separately according to the trajectory of body displacements. Figure 3 provides a schematic illustration of an example stimulus. A white fixation cross (146.5 cd/m^2) was always shown at the center of the screen. Participants used a chin rest to maintain a fixed viewing distance of 35 cm.

On each trial, action stimuli were presented for 6.67 s. The first 100 frames (i.e., 1.67 s) presented only the walker to encourage participants to maintain fixation on the walking action. The GPS dot then appeared at the center of the screen in the 101st frame and subsequently moved according to the assigned trajectory of body displacements. In the experiment, the stimuli were shown from one of four viewpoints (45° , 135° , 225° , and 315°). The order of conditions was randomized.

Procedure. Participants were given a cover story.

Imagine that you work for a specialized video analysis company and are given two sources of

information: (1) A posture-change video from a motion tracking system, which records a person's posture change over time and keeps the figure always at the center, and (2) A dot-motion video from a GPS system, which tracks the location of the person.

Participants were then presented with a few videos demonstrating how the posture changes were separated from the change in body position over time, based on the original motion-capture video. They were informed that

It turns out that in preparing the combined videos, mistakes are sometimes made. Sometimes the posture-change video is correctly linked to the dot-motion video. However, in other cases, the posture-change and dot-motion videos were mixed up so that the two videos shown together do not match.

Participants were asked to judge whether the posture-change video and dot-motion video matched by pressing one of the two response buttons.

Participants were given two practice blocks with feedback. Each practice block consisted of 12 trials, 6 trials of matched stimuli (i.e., temporal offset of zero) and 6 trials of unmatched stimuli with obvious temporal misalignment (i.e., body motion was 8.33 s ahead of limb movements). Matched and unmatched trials were randomly interleaved. Feedback was provided after each practice trial. For correct responses, participants were provided with a beep sound and the word "Correct" on the screen. For wrong responses, the screen displayed the word "Incorrect" at the end of the practice trial.

In the subsequent test block, 96 trials were presented to participants. The trials included eight levels of temporal offsets ($\pm 0.02 \text{ s}$, $\pm 0.5 \text{ s}$, $\pm 1.0 \text{ s}$, and $\pm 1.5 \text{ s}$) between body displacements and the posture change resulting from limb movements. The experiment included 10 filler trials that came from the practice block. The 10 filler trials were

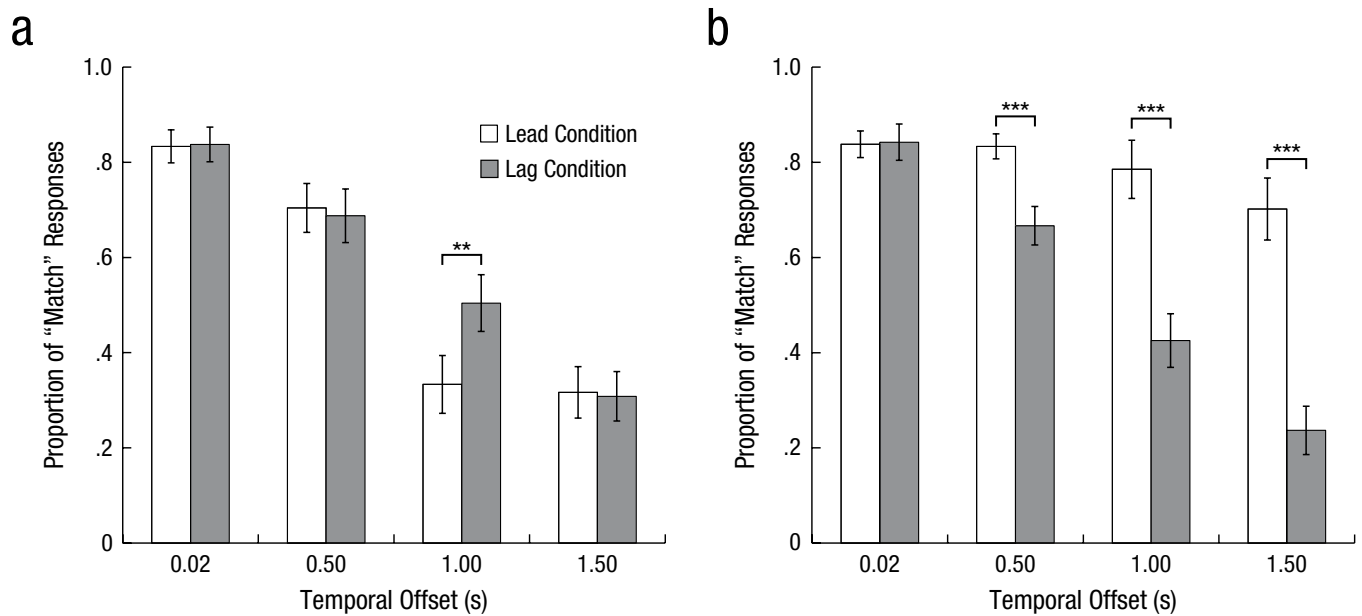


Fig. 4. Results from (a) Experiment 2a and (b) Experiment 2b: mean proportion of “match” responses as a function of temporal offset, presented separately for the lead condition and the lag condition. In Experiment 2a, the temporal offset was between relative limb movements and body displacements. In Experiment 2b, the temporal offset was between relative limb movements (effect) and the motion of the laser dot (cause). Error bars represent ± 1 SEM. Asterisks indicate significant differences between conditions ($**p < .01$, $***p < .001$).

randomly inserted into the experiment as attention-check trials. The entire experiment lasted for about 20 min.

Results

The results of Experiment 2a are shown in Figure 4a. A repeated measures ANOVA with two within-subjects factors (condition: lead, lag; temporal offset: ± 0.02 s, ± 0.5 s, ± 1.0 s, ± 1.5 s) revealed a significant main effect of offset magnitude, $F(3, 17) = 21.99$, $p < .001$, $\eta_p^2 = 1.0$, which indicates that participants were sensitive to the temporal misalignment between the two motion cues in this binding task. The interaction of offset magnitude and temporal direction of the offset was marginally significant, $F(3, 17) = 3.14$, $p = .053$, $\eta_p^2 = .622$, which suggests that the influence of temporal direction between the two movement cues on the binding judgment depended on the magnitude of temporal offsets. We found that with a 1-s offset, there were a significantly higher proportion of matched responses in the lag condition ($M = .50$, $SD = .27$) than in the lead condition ($M = .33$, $SD = .27$), $F(1, 19) = 9.34$, $p = .006$, $\eta_p^2 = .826$. The difference remained significant after adjusting for multiple comparisons using the Bonferroni correction procedure. A causal-asymmetry effect was thus observed within a middle range of the temporal window when the two motion cues could be interpreted as originating from a single walker but with a noticeable temporal delay between the two motion signals.

A temporal-asymmetry effect was not observed for the temporal offsets in the extreme conditions. When the temporal offset was very small (e.g., 0.02 s), observers might not have detected the temporal differences between the lead and lag conditions. However, when the temporal offset was very large (e.g., 1.5 s), observers might infer that the two videos were generated from different sources; therefore, they judged the two signals to be mismatched, regardless of the temporal direction of offset. Two post hoc control experiments were conducted to test these predictions.

In one control experiment, we showed two displays side by side, each with a temporal offset of the same magnitude, but one was positive and the other negative. Eight observers were asked to judge whether the two displays were the same or different. We found that observers judged the two displays with temporal offsets of 0.02 s and -0.02 s to be the same on a high proportion ($M = .92$, $SD = .08$) of trials. In contrast, for each of the other three magnitudes of temporal offsets (0.5, 1.0, and 1.5 s), people judged them to be the same much less often (M s = .22, .13, and .13, respectively). These findings support the hypothesis that people can barely detect the difference between temporal offsets of 0.02 s and -0.02 s but are sensitive to the difference in temporal direction when the magnitude of temporal offsets was 0.5 s or more.

In a second control experiment, we showed the same visual stimuli that we used in Experiment 2a. In the cover

story, we introduced the participants ($N = 21$) to a one-person situation, in which the two sources of motion (i.e., the limb movements and the dot motions) came from the same walker, and a two-person situation, in which the each of the two sources of motion came from a different walker. The participants' task was to make a two-alternative forced choice about whether the two sources of motion came from one person or from two people. Results showed that the proportion who chose the two-person situation increased with the temporal offset (0.02-s offset: $M = .22$; 0.5-s offset: $M = .38$; 1.0-s offset: $M = .56$; and 1.5-s offset: $M = .63$). Only for the longest temporal offset (1.5 s) did the proportion of "two-person" choices significantly surpass the chance level of .50 ($M = .63$, $SD = .18$, $p = .004$). This finding indicates that when the display shows a longer temporal offset, people are likely to attribute the two motion cues to two different sources, which removes the dependence of judgments on temporal directionality.

In summary, Experiment 2a used a binding task to provide evidence that observers are sensitive to the directionality of temporal offset between limb movements and body displacements in reasoning about the relation between the two motion sources. Specifically, starting at around 1 s of temporal offset, the two motion cues were more likely to be judged as matched when the causal limb movement preceded the effect of body displacement in a direction consistent with the natural causal relation (lag condition).

Experiment 2b

If the temporal-asymmetry effect found in previous experiments resulted from observers' understanding of the inherent causal relation between the two motion sources involved in human body movements, the effect should be radically altered when the causal relation is changed. In Experiment 2b, we changed the cover story to specify that the dot represented a moving laser dot (as a cause) that was being followed by a person (as an effect). In this case, rather than the limb movements causing the dot motion, limb movements are inferred to be the effects of dot motions, so the direction of causality has been reversed from Experiment 2a. This type of manipulation of schematic understanding by a cover story has been used in many previous studies to distinguish the impact of causal interpretation from associative learning (e.g., Waldmann & Holyoak, 1992).

Method

Participants. Nineteen students (mean age = 20.2 years; 14 female) participated in the experiment for course credit. All participants had normal or corrected-to-normal vision. None of the observers had participated in

Experiment 2a. Data collection for this experiment stopped in the week when the expected sample size was reached.

Stimuli and procedure. The stimuli, task, and procedure in Experiment 2b were identical to those in Experiment 2a, except for a small wording change in the cover story, which now referred to "a dot-motion video with a moving laser-generated spot, which the person is following closely."

Results

In this situation, the two components of the stimuli are interpreted to represent two distinct entities, such that the moving dot (i.e., the laser dot) would now be the cause that should make the agent move in a certain way, and limb movements should be interpreted as the effect. Otherwise, stimuli were identical to those in Experiment 2a. The binding task was performed in the same manner as in Experiment 2a. Accordingly, any differences in people's patterns of judgments between Experiments 2a and 2b could be attributed only to the influence of the different causal interpretations conveyed by the cover story in each. The results of Experiment 2b are shown in Figure 4b.

A repeated measures ANOVA revealed a significant main effect of offset magnitude, $F(3, 16) = 18.97$, $p < .001$, $\eta_p^2 = 1.0$, and a significant interaction of offset magnitude and temporal direction of offset, $F(3, 16) = 16.09$, $p < .001$, $\eta_p^2 = 1.0$. When the dot motion preceded the limb movements (which was consistent with causal direction in the person-following-a-dot cover story), participants gave a high proportion of match responses, regardless of the magnitude of temporal offset. However, when dot motion followed limb movements (which was opposite the causal direction in the person-following-a-dot cover story), the magnitude of the temporal offset had significant impact on human judgments. The proportion of match responses in the lead conditions was significantly greater than the proportion of match responses in the corresponding lag conditions for offsets of 0.5 s, 1.0 s, and 1.5 s (all $ps < .001$), which indicates a stronger tolerance for temporal deviations between the two motion sources when the dot motion (cause) preceded the limb movements (effect), relative to the corresponding condition in which the effect cue preceded the cause.

We conducted a mixed-model repeated measures ANOVA to examine the difference in response patterns in Experiments 2a and 2b (in which the cover story changed but the judgment task and stimuli were the same). The dependent variable in this analysis was the temporal-asymmetry effect, calculated as the difference in the proportion of match responses between the lag condition and the corresponding lead condition (i.e., the lead

condition with the same offset magnitude). Results showed a significant interaction effect between the magnitude of temporal offsets and cover story, $F(3, 35) = 13.29$, $p < .001$, $\eta_p^2 = 1.0$. This interaction effect reflects the fact that when observers received the laser spot cover story in Experiment 2b, their judgments changed dramatically and effectively reversed their interpretation of the cause-effect relations between motion cues.

Furthermore, this temporal-asymmetry effect was maintained for a larger range of offset magnitudes (from 0.5 s to 1.5 s) for Experiment 2b than for Experiment 2a. The strength and robustness of the effect in Experiment 2b are probably due to participants' qualitative interpretation of a "following" action. Participants may perceive a causal relation between the two motion cues as long as the movement of one entity follows the same trajectory as that of another entity, without necessarily being constrained to a specific value of temporal delay between the two movements. Whereas the causal relation between limb movements of a person and displacement of their body (Experiments 1 and 2a) is theoretically more closely coupled in time, the action of an agent that is following a separate object (Experiment 2b) can be much more temporally variable, as long as the motion of the agent consistently lags behind that of the object.

In summary, Experiments 2a and 2b used identical visual stimuli but induced different causal beliefs about the relation between limb movements and a distinct moving object (the dot). The opposite pattern of judgments obtained in Experiment 2a compared with Experiment 2b suggests that when considering movements attributed to a single walker, by default observers link bodily movements of articulated limbs (causes) to motions of body locations through the environment (effects). But when considering the movements of one agent with respect to a separate distal object, observers can flexibly assign the effect role to bodily movements of the agent, interpreting them as being caused by intent to follow a moving target. The contrast between the patterns of results observed in Experiment 2a and Experiment 2b provides strong evidence that the temporal-asymmetry effect is based on the observer's attribution of the causal relation between two movement cues and does not reflect a mere association or low-level physical properties of the displays.

To assess the possibility that the temporal-asymmetry effect might be due to statistical learning of temporal regularities (i.e., that limb movements should precede body displacement in time, without necessarily assuming limb movements cause the latter motion), we performed a further post hoc control experiment that introduced a noncausal association relation between the two motion cues. In the cover story, participants were told:

Imagine that two actors aim to synchronize their body movements and they walk on two identical

terrains, each with steps. You are given two sources of information: (1) a posture-change video from a motion tracking system, which records Actor 1's posture change over time, and (2) a dot-motion video from a GPS system, which tracks the location of Actor 2.

Participants were asked to judge whether the posture-change video of Actor 1 and the dot-motion video of Actor 2 matched one another, such that the two actors moved in synchrony. All other aspects of the experiment were identical to those of Experiment 2a. Twenty-two UCLA students participated in this study. In contrast to the results of Experiment 2a, none of the offset conditions revealed a difference in the proportions of "match" responses between the lag and lead conditions, including the critical 1.0-s temporal-offset condition (lead: $M = .37$, $SD = .23$; lag: $M = .37$, $SD = .20$), $t(21) < 0.001$, $p = 1.0$. These findings indicate that mere temporal association was not sufficient to elicit a temporal-asymmetry effect in the absence of a direct causal link between the motion cues.

General Discussion

The basic finding in the present study is that participants were more likely to bind human actions in a causally consistent way. Although observers commonly observe the occurrence of the two types of motion in synchrony, they can tolerate deviation from this normal synchronicity. Our results show that the degree of tolerance is constrained by the directionality of the causal relation between motions.

In fact, such a temporal-asymmetry effect has been observed (albeit not previously noted by researchers) even in the classic ball-collision paradigm introduced by Michotte (1946/1963), which elicits immediate and irresistible causal impressions that one moving ball causes a second ball to move or launch. It is well known that the causal impression is reduced after the introduction of a spatial or temporal gap. These two manipulations can be interpreted as temporal offsets with different directions in time. In the spatial-gap condition, the effect ball (the one that is launched) moves before the causal ball (the launcher) reaches the contact location (i.e., effect precedes the cause), analogous to the lead condition in the present article. In the temporal-gap condition, motion of the effect ball is delayed after the causal ball arrives at the contact location (i.e., cause precedes the effect with an abnormal temporal gap), analogous to our lag condition. We examined human data from a recent study using this paradigm (Sanborn, Mansinghka, & Griffiths 2013). Given the speeds of the moving balls used in their study, a 1-cm spatial gap corresponded to a positive 16-ms temporal offset (i.e., the effect ball moved for 16 ms

before the causal ball arrived in the contact location), which yielded a causal rating of about .4 in the lead condition. In comparison, a 16-ms temporal gap yielded a causal rating of .75 in the lag condition. Thus, in conditions that equated the magnitude of temporal offset, people gave much higher causal ratings to the lag condition, presumably because it preserved the expected causal order of events.

A temporal-asymmetry effect has now been observed for a range of causal events, including object interactions, action-initiated changes in object status (e.g., a button press triggering a flashed disk; Desantis & Haggard, 2016; Rohde, Greiner, & Ernst, 2014), and body movements (as observed in the present study). We expect that similar effects would be observed for other causal events involving nonbiological stimuli, such as rotating wheels (as cause) and cars moving forward (as effect), as long as people have an understanding of the causal relation involved in the physical system. From this perspective, human body movements do not have a special status; rather, they serve as one example of causal events that yield the temporal-asymmetry effect. However, human actions may elicit a temporal-asymmetry effect with different timing properties relative to other physical systems that involve inanimate objects. For example, in the ball-collision and action-initiated-flash situations, the magnitude of temporal offset that yields an asymmetry is very short (< 200 ms), which is consistent with causal mechanisms that operate quickly and perhaps spontaneously (Scholl & Tremoulet, 2000). In contrast, the asymmetry effect found for body movements in the present study is substantially longer (offset of about 1.0 s). This larger temporal window indicates much greater tolerance for temporal misalignment between limb movements and body displacements, which may reflect the increased complexity of the causal mechanisms involved in perceiving body movements.

One possible alternative explanation of the temporal-asymmetry effect observed in the present study is that people's judgments are guided by temporal regularities learned in the environment and perhaps transferred by analogy to the similar situations described in the cover stories used in Experiment 2. For example, people are likely to have clear expectations of temporal order for a relation such as "chasing" or "following." In general, a causal relation implies a certain temporal order (cause before effect), but perhaps causality is just one special case (albeit an important one) of effects more directly attributable to temporal knowledge *per se*.

Although this alternative explanation might account for the findings of Experiment 2, which involved cover stories that probably conveyed temporal as well as causal knowledge, it does not offer a compelling explanation of the results of Experiment 1. Experiment 1 did

not manipulate a cover story; participants simply observed point-light displays of walkers and judged their naturalness as a function of temporal offsets. In daily life, people often observe the cooccurrence of two types of motion at the same time (e.g., planting a foot, extending the legs, and moving the body forward cooccur). If learned temporal regularity, rather than causality, was the key factor, we would expect to have observed perceived naturalness peaks at zero offset, with a symmetric decline as the temporal offset was increased. Instead, we found the asymmetrical pattern predicted by a causal interpretation. Hence, the causal relation between limb movements and body displacements provides a parsimonious explanation for the findings regarding both action perception and inference. However, we note that using solitary actions of a single actor makes it difficult to isolate causation from temporal regularity (i.e., to generate conditions analogous to Michotte's temporal-gap experiment). Future studies may be able to further distinguish effects of causality versus temporal order *per se* by investigating patterns of motion cues involving interactions between two actors.

Actions can be interpreted as a willful expression of body movements in the environment, caused by intentional patterns of limb movement. Sensitivity to causal dynamics in body movements may play a general role in tracking perceptual animacy, which supports the ability to visually distinguish living from nonliving entities (Thurman & Lu, 2013, 2014). In addition, as relational binding in general enhances representational power (Lu, Chen, & Holyoak, 2012), the perceived causal relation between the two types of movement cues makes it possible for people to understand why the body moves the way it does. Causal understanding of actions enables explanation of the past as well as prediction of the future (Cheng, 1997), which provides a fundamental constraint on perception and inference from human bodily movements.

Action Editor

Marc J. Buehner served as action editor for this article.

Author Contributions

H. Lu and Y. Peng developed the study concept. All the authors contributed to the study design. Testing and data collection were performed by Y. Peng. Y. Peng performed the data analysis and interpretation under the supervision of H. Lu. Y. Peng drafted the manuscript, and S. Thurman and H. Lu provided critical revisions. All the authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Hakwan Lau, Joseph Burling, Keith J. Holyoak, and Scott P. Johnson for helpful comments and Diana Santana, Kejia Wang, and Roshni Desai for their assistance in data collection.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was funded by National Science Foundation Grant BCS-1353391 (to H. Lu) and by a China Scholarship Council scholarship (to Y. Peng).

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797617697739>

Open Practices



All data have been made publicly available via the Open Science Framework and can be accessed at <https://osf.io/qh896/>. The complete Open Practices Disclosure for this article can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797617697739>. This article has received the badge for Open Data. More information about the Open Practices badges can be found at <https://www.psychologicalscience.org/publications/badges>.

References

- Abraham, E., Levan-Goldschmidt, E., & Stevenson, M. B. (1976). Action imitation: The early phase of infancy. *Child Development, 47*, 1032–1044.
- Bechlivanidis, C., & Lagnado, D. A. (2013). Does the “why” tell us the “when”? *Psychological Science, 24*, 1563–1572.
- Bechlivanidis, C., & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order. *Cognition, 146*, 58–66.
- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405.
- Desantis, A., & Haggard, P. (2016). Action-outcome learning and prediction shape the window of simultaneity of audiovisual outcomes. *Cognition, 153*, 33–42.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*, 243–259.
- Hume, D. (1888). *Hume's treatise of human nature* (L. A. Selby-Bigge, Ed.). Oxford, England: Clarendon Press. (Original work published 1739)
- Lu, H., Chen, D., & Holyoak, K. J. (2012). Bayesian analogy with relational transformations. *Psychological Review, 119*, 617–648.
- Masselink, J., & Lappe, M. (2015). Translation and articulation in biological motion perception. *Journal of Vision, 15*(11), Article 10. doi:10.1167/15.11.10
- Michotte, A. E. (1963). *The perception of causality* (T. R. Miles, Trans.). London, England: Methuen & Co. (Original work published 1946)
- Price, H. (1992). Agency and causal asymmetry. *Mind, 101*, 501–520.
- Rohde, M., Greiner, L., & Ernst, M. O. (2014). Asymmetries in visuomotor recalibration of time perception: Does causal binding distort the window of integration? *Acta Psychologica, 147*, 127–135.
- Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological Review, 120*, 411–437.
- Scholl, B. J., & Nakayama, K. (2002). Causal capture: Contextual effects on the perception of collision events. *Psychological Science, 13*, 493–498.
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*, 299–309.
- Shu, T., Thurman, S., Chen, D., Zhu, S.-C., & Lu, H. (2016). Critical features of joint actions that signal human interaction. In A. Papafragou, D. Grodner, D. Mirman, & J. Trueswell (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 574–579). Retrieved from <https://mindmodeling.org/cogsci2016/cogsci16/proceedings.pdf>
- Thurman, S. M., & Lu, H. (2013). Physical and biological constraints govern perceived animacy of scrambled human forms. *Psychological Science, 24*, 1133–1141.
- Thurman, S. M., & Lu, H. (2014). Perception of social interactions for spatially scrambled biological motion. *PLoS ONE, 9*(11), Article e112539. doi:10.1371/journal.pone.0112539
- Thurman, S. M., & Lu, H. (2016). Revisiting the importance of common body motion in human action perception. *Attention, Perception, & Psychophysics, 78*, 30–36.
- van Boxtel, J. J., & Lu, H. (2013). A biological motion toolbox for reading, displaying, and manipulating motion capture data in research settings. *Journal of Vision, 13*(12), Article 7. doi:10.1167/13.12.7
- Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*, 222–236.
- White, P. A. (1999). Toward a causal realist account of causal understanding. *American Journal of Psychology, 112*, 605–642.
- White, P. A. (2006). The causal asymmetry. *Psychological Review, 113*, 132–147.