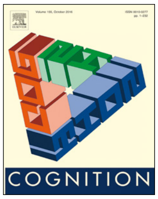
Cognition 194 (2020) 104060

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit



Original article

Causal actions enhance perception of continuous body movements

Yujia Peng[a,*], Nicholas Ichien[a], Hongjing Lu[a,b]

[a] Department of Psychology, University of California, Los Angeles, United States of America
[b] Department of Statistics, University of California, Los Angeles, United States of America

ARTICLE INFO

Keywords:
Causality
Causal action
Motion interpolation
Human action
Human interaction

ABSTRACT

Our experience of motion depends not only on spatiotemporal features of stimuli, but also on our recognition of seemingly higher-level properties, as when we see an actor's body movements as goal-directed. Here, we examined how the perception of social causation in human actions guides the perceptual interpolation of motion in the observation of body movements. Natural human-object interactions were recorded for videos in which a person prepared to catch a ball thrown by another person. We manipulated the number of image frames between key postures to yield a short clip with different frame rates, and asked participants to judge whether the catcher's action showed smooth movements or sudden changes. In the causal condition, the catcher faced toward the ball and the thrower to preserve an intention-based causal relation between the ball's movement and the catcher's action in which the former causes the catcher's intention to act. In the non-causal condition, the catcher performed the same movements to raise their hands to catch a ball, except that they faced away from the ball, creating the impression of either a psychic reaction or coincidental non-goal-directed behavior, which makes movements of the ball appear to be an implausible cause of the catcher's intention to act. Across four experiments, we found that humans were more likely to judge the catcher's body movements to be continuous in the causal condition than in the non-causal condition. The effect was maintained as long as the intention-based causal relation was present, even when only part of the chain of causal events was observed. These findings indicate that intention-based cause-effect relations in human actions guide perceptual interpolation of body movements.

1. Introduction

In our daily life, we are constantly incorporating new visual information to form a continuous impression of the dynamic world. However, the perceptual construction of smooth movements is not a trivial task, since visual inputs are actually discrete frames or disjointed clips separated by constant eye movements. Flipbooks, for example, exploit our susceptibility to apparent motion (Wertheimer, 1912), where our visual system induces the perception of dynamic scenes from the presentation of static images in rapid succession.

Apparent motion offers an illustrative case of the human visual system's tendency to interpolate the paths of perceptual objects over time, and to produce the perception of smooth motion across discrete samples of visual stimuli at different time points. It is well known that the appearance of smooth motion depends on spatiotemporal features of a visual stimulus, such as inter-frame spatial displacement and temporal sampling rate (Braddick, 1974; Burr, Ross, & Morrone, 1986). For example, previous studies have demonstrated that it is only within certain ranges of displacements and temporal sampling measured as stimulus-onset-asynchrony (SOA) between frames that a two-frame stimulus evokes a percept of smooth movement. Apparent motion is lost when the spatial and temporal parameters exceed those limits (Baker & Braddick, 1985a, 1985b; Bours, Stuur, & Lankheet, 2007; Lappin & Bell, 1976; Morgan & Ward, 1980). Further, response profiles of neurons in early visual areas (such as V1 and MT) demonstrate similar spatial and temporal limits in giving rise to the perception of apparent motion (Baker & Cynader, 1986; Churchland, Priebe, & Lisberger, 2005; Mikami, Newsome, & Wurtz, 1986; Newsome, Mikami, & Wurtz, 1986).

In addition to spatiotemporal features, the perception of smooth motion is also influenced by seemingly higher-level properties of objects and events presented in visual stimuli (Sigman & Rock, 1974; Braddick, 1980; Shiffrar & Freyd, 1990, 1993; Kim, Feldman, & Singh, 2013; Chen & Scholl, 2016). For example, in typical cases of apparent motion, an object presented in one frame tends to be perceived as moving with the shortest path to its location in the subsequent frame. However, Shiffrar and Freyd (1990, 1993) showed that this tendency disappears when the perceptual object is a human body. Specifically, when two frames of human body postures are presented with a long

---

[*] Corresponding author.
E-mail address: yjpeng@ucla.edu (Y. Peng).

https://doi.org/10.1016/j.cognition.2019.104060
Received 5 March 2019; Received in revised form 22 August 2019; Accepted 28 August 2019
Available online 10 September 2019
0010-0277/ Published by Elsevier B.V.

duration between frames, people perceive the body movements following a longer path that satisfies biomechanical constraints, rather than the shortest path which would imply a causally implausible movement through the body. Further, Chen and Scholl (2016) demonstrated that human observers are more likely to experience apparent motion when an object moves to interact with another object in a causally plausible way. In their study, observers watched a change from a complete square shape to a truncated form with a missing piece and were asked to report whether this change was sudden or gradual. The key manipulation was whether the shape of the missing piece was generated by an intrusion, where apparent motion would be causally plausible (e.g., as when an object is pushed into a lump of clay), or whether it was imposed, where apparent motion would not be causally plausible (e.g., as when a piece is cut out of a lump of clay). The results showed that observers were more likely to see apparent motion in the intrusion condition compared to the imposed condition. These findings suggest that the visual impression of dynamic stimuli is sensitive to the causal structure in the world.

Indeed, humans can spontaneously perceive cause-effect relations in some dynamic stimuli, as demonstrated by the well-known launching effect between two colliding objects (Michotte, 1946/1963). Further, such automatic perception arises not only for physical causation, but also for intentional causation in the social environment as illustrated by the Heider-Simmel animations based on moving geometric shapes (Heider & Simmel, 1944). Even as young as 9-months of age, infants perceive objects as "intentional agents" whose states can cause behavioral activities (Csibra, Gergely, Bíró, Koos, & Brockbank, 1999). Both perceptions of physical and of social causation are susceptible to changes in spatiotemporal features in the dynamic scene. For example, the perceived causation in a launching event depends on relative speeds of objects in the scene, spatial gaps between those objects, temporal gaps between objects' motions, and objects' path lengths (Scholl & Tremoulet, 2000).

Moreover, causal perception can influence subsequent perceptual judgments and inferential processes for human-involved events. For example, the cause-effect relations between limb movements and body motions in human actions provide a constraint on judgments about the naturalness of those actions as well as inferences about them (Peng, Thurman, & Lu, 2017). In addition, work on causal binding has shown that the detection of causality biases the perception of time and space (Buehner, 2012; Humphreys & Buehner, 2009, 2010). For example, Buehner and Humphreys (2009) demonstrated that when an action of pressing a key is experienced as causing the event of hearing a tone, the perceived time lapse between the two events appears shorter than when the two events are not causally related. This finding indicates that two causally related events are more likely to trigger perception of spatiotemporal contiguity.

The cause-effect relations in dynamic events can also influence memory about spatiotemporal properties of dynamic events, eliciting false memories of actions. Strickland and Keil (2011) found that causal connections between agents and objects led to false memories of action frames that were never presented. For example, adults watched videos in which an actor kicked a ball, but the videos omitted the moment in which the actor actually contacted the ball. In a later recall task, participants falsely reported seeing physical contact when the subsequent footage implied a causal relation between the actor's movements and the motion of the ball. Similarly, Bechlivanidis and Lagnado (2013, 2016) demonstrated that perceived causality can induce false memories about the temporal order of events. Having a belief that event A causes event B made participants less likely to recognize a video that had been observed seconds earlier when it violated the expected temporal order in the cause-effect relation (i.e., when the effect event B temporally preceded the cause event A). In addition, participants were more likely to report the perceived temporal order as coinciding with their causal belief, even when the order of events was presented with effects happening before their causes. These findings present compelling cases in which causality plays an influential role in consolidating memories about actions and events.

The use of intention-based causality may enable our visual system to be more adaptive to respond to dynamic events with noisy inputs in the social environment. For example, with a clear expectation in a social scene where a person walks toward another person in a crowded train station, temporarily missing visual inputs of the person due to occlusions from crowds or buildings do not weaken our impression of that person's continuous movement. Here, we propose an experimental paradigm to systematically test whether social causality in human activities influences the visual experience of observed events. Specifically, we examine whether the intention-based causal relations between agents and objects inherent in human activities influence the extent to which the visual system interpolates body motion. To answer this question, we test the hypothesis that the visual system exploits intention-based causal relations in human activities to fill in missing information between static frames, yielding the subjective experience of smooth motion in human actions.

We designed four experiments to examine the role of social causality in guiding perceptual interpolation of motion in human actions. We recorded videos of human-object interactions in a natural environment (a thrower directing a ball to a catcher). For short clips in which the catcher prepared to receive the ball, the number of image frames between key postures was manipulated to introduce short and long interframe durations, defined as SOA. The duration of short SOAs was 33.3 ms/frame; that of long SOAs was 100 ms/frame. In other words, the short SOA condition presents a critical period with a high frame rate, and the long SOA condition presents a critical period with a low frame rate. In the causal condition, the facing direction of the catcher was kept intact to preserve an intention-based causal relation between the ball's movement and the catcher's action in which the former causes the catcher's intention to act. In the non-causal condition, the catcher performed the same movements to raise their hands to catch a ball, except that they faced away from the ball, creating the impression of either a psychic reaction or coincidental non-goal-directed behavior, which makes movements of the ball appear to be an implausible cause of the catcher's intention to act. The video stimuli used in the experiments can be viewed at http://cvl.psych.ucla.edu/causal-illusion-motion.html. Participants were asked to judge whether the catcher's action showed smooth body movements or sudden changes.

In Experiment 1, we presented human interactions with two agents throwing and catching a ball. We hypothesized that social causation in human actions influences the interpolation of discrete pieces of motion information, so that observers would be more likely to perceive smooth movements when observing causal than non-causal actions. In Experiment 2, we further tested if the effect would generalize to a situation in which only a partial chain of causal events was observed, as the intermediate cause (i.e., the ball movements) was not presented in the stimuli. We hypothesized that the casual interpolation effect would be elicited by social causation between two agents, even upon the presentation of an incomplete causal chain, though the effect size might be attenuated. In Experiment 3, we examined whether the causal interpolation effect is present for human-object interaction when viewing only movements of the ball and the catcher. In Experiment 4, upright videos were compared to upside-down videos to test whether the causal interpolation effect for human actions depends on visual familiarity, such as body orientation. Body inversion is known to disrupt visual processes for detecting and recognizing human actions (Pavlova & Sokolov, 2000). If the causal interpolation effect depended on visual familiarity to upright body orientation, rather than being driven solely by cues signaling goal-directed social causality (such as temporal contingency among the events and intention-based causal relation), we would expect the effect to be attenuated when action processing is disrupted by inversion of the body.

2. Experiment 1

Experiment 1 was designed to assess how a causal action between agents influences interpolation in the perception of smooth human actions. Causal actions were generated with two agents facing each other while throwing and catching a ball. Non-causal actions were generated with the same agent facing away from the moving object and the other agent. We hypothesized that in the causal action condition, discretized human actions would be more likely to be perceived as smooth motion sequences.

2.1. Methods

2.1.1. Participants

Forty-nine University of California, Los Angeles (UCLA) undergraduate students (mean age = 20.31; 39 female) participated in the experiment for course credit. All experimental procedures were approved by the Committee for Protection of Human Subjects at UCLA. All participants had normal or corrected-to-normal vision.

2.1.2. Stimuli

Action videos were filmed with a camera in the gym with a temporal resolution of 30 frames/s. Two pairs of actors (one male pair and one female pair) were enrolled and each pair performed three throwing-catching actions (i.e. bounce pass, chest pass, and underhand throw), with each actor being the thrower once and catcher once. Seven video clips were selected as experimental stimuli.

For each video, a short critical period was selected during which the catcher's arms showed the largest rising momentum during preparation to catch the ball. Each video lasted for 567 ms. The duration was 333 ms before the critical period, and 33 ms after the critical period. The critical period began when the catcher's arms started to rise, and it ended right before the actor's hands touched the ball. The duration of the critical period was 200 ms. In the long-SOA condition, only the first and the last frame of the catcher's body movements were presented, all the middle frames were omitted. The presentation duration of the first and the last frames were lengthened to each cover half of the critical period at 100 ms per frame. In the short-SOA condition, all six frames showing body movements of the catcher were displayed, each was presented for 33 ms. Note that the duration of the critical period was the same (200 ms) for both long-SOA and short-SOA displays. The movements of the ball were also the same and were kept intact in both long-SOA and short-SOA displays (Fig. 1).

As shown in Fig. 2, the causal condition showed the catcher facing toward the ball and the thrower as the ball movement causes the catcher to move his or her body in preparation. To generate non-causal actions, image frames were processed using Matlab and Adobe Photoshop to horizontally reverse the facing direction of the catcher. The catcher was flipped horizontally to face away from the ball and the thrower in the entire video, while keeping the background and the ball movement intact.

2.1.3. Procedure

Participants were seated 35 cm in front of a monitor with a 1024 × 768 resolution and 60 Hz refresh rate. All the stimuli were generated by MATLAB Psychtoolbox (Brainard, 1997). Participants were instructed, “You will view two actors playing sports (such as passing a basketball). The person who throws the ball is the thrower. The person who catches the ball is the catcher. The task is to judge whether the catcher actor shows a smooth action or a non-smooth sudden posture change. For a smooth action, the actor smoothly moves from one posture to another. For a non-smooth action, the actor suddenly moves from one posture to another.”

On each trial, a white fixation cross was presented at the center of the screen on a black background. Participants were asked to focus on the fixation cross throughout the experiment and to use their peripheral vision to see the video without making saccades. The center of the video was presented 13.7 degrees to the left or to the right of the fixation point with a height of 18 degrees. Showing the video in peripheral vision reduced the possibility that observers would track movements of the catcher without paying attention to other parts of the display. A random half of the trials presented the video on the left of the fixation and the other random half on the right. The catcher actor was always presented on the side relatively farther away from the fixation point. For example, if the video was presented on the right side, the ball flew from left to right and the catcher was located on the right side of the ball. After the video display, participants were asked to press one of two buttons to judge whether the video demonstrated actions with smooth body movements or sudden posture changes.

Participants were first presented with two blocks of practice trials to familiarize them with the task. In each of the practice blocks, participants saw “correct” on the screen plus a beep after each correct response, and they saw “incorrect” without a beep after each incorrect response. Each practice block consisted of eight trials, with half of the trials showing causal actions and half showing non-causal actions. A separate video was used as the stimulus for the practice block; this video was not presented in the test. In the first block of practice, videos were slowed down to show the entire video with the frame rate of 66.6 ms/frame and to display the critical period for 666 ms. This manipulation was intended to allow participants to become familiar with the experimental setting and to understand the difference between smooth motion and sudden posture changes in body movements. In the second block of practice trials, videos were presented at a frame rate of 33.3 ms/frames, and the duration of the critical period was 200 ms, as in the test session.

The test session followed the practice blocks. Test trials were identical to those in the second practice block with two exceptions: participants received no feedback on test trials, and test trials employed six new videos that were not used in practice blocks. A total of five test blocks were administered, each with 24 trials (causal/non-causal × long-/short SOA × 6 actions). In each block, the presentation order of videos was randomly shuffled. Proportions of responses in judging actions as smooth motion were recorded for each condition.

2.2. Results

A 2 (SOA) by 2 (causality) repeated-measures ANOVA was conducted with the dependent variable being the proportion of “smooth” judgments, averaged across the five test blocks. Results (Fig. 3) showed a significant main effect of causality, $F(1,48) = 20.869$, $p < .001$, suggesting that the causal condition yielded a significantly higher proportion of judging actions as smooth actions compared to the non-causal condition. The main effect of SOA was highly significant, $F(1,48) = 227.289$, $p < .001$. As expected, the smooth motion signal was much weaker in the long-SOA display, since the stimulus included only two static postures with the largest spatial displacements between frames in the critical period. The interaction effect was not significant, $F(1,51) = 0.155$, $p = .696$. The proportion of judging the long-SOA video as a smooth action increased from 15.2% in the non-causal condition to 18.8% in the causal condition, yielding a Cohen's effect size value of $d = 0.32$. The proportion of judging the short-SOA video as a smooth action increased from 79.5% in the non-causal condition to 85.6% in the causal condition, yielding a Cohen's effect size value of $d = 0.65$.

3. Experiment 2

In Experiment 1, we found evidence that causal interactions between actors facilitated the perception of smooth movements. In Experiment 2, we investigated whether the effect could be generalized from human-object interactions to human-human interactivity when the ball movements that link the two people are not presented.

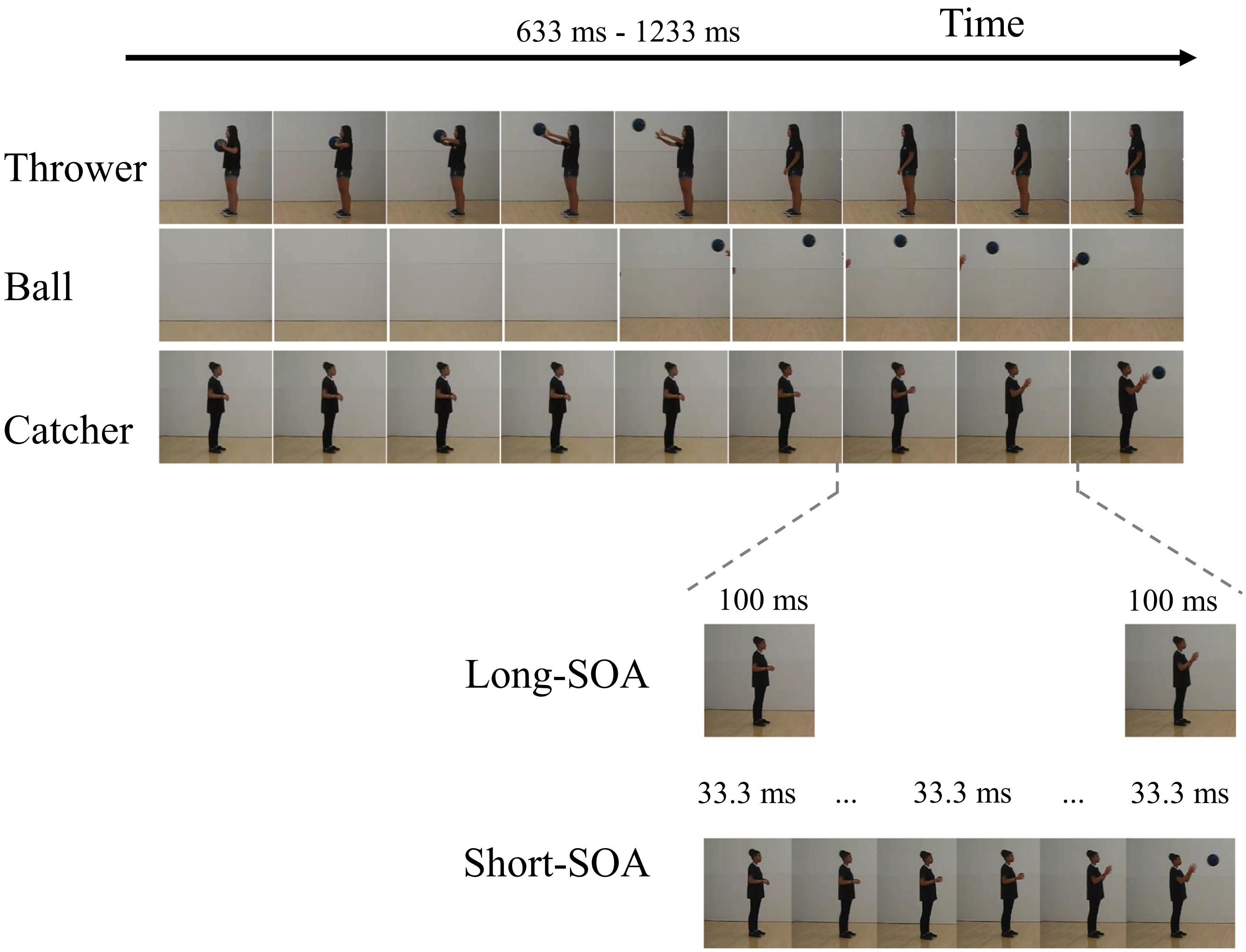


Fig. 1. Illustrations of the critical period in the long-SOA display with two frames (100 ms/frame) with a sudden posture change, and in the short-SOA display consisting of six frames (33.3 ms/frame). Both display conditions were generated from the same natural movements of the catcher and with the same duration, except different numbers of frames were included between the two key postures in the critical period.

3.1. Methods

3.1.1. Participants

Forty-eight new UCLA students (mean age = 20.48; 33 female) participated in the experiment for course credit. All participants had normal or corrected-to-normal vision.

3.1.2. Stimuli and procedure

The experiment employed the same basic videos as in Experiment 1, showing two actors pass balls. The stimuli included the body movements of the thrower and the catcher (Fig. 4). A white occluder was presented at the center of the video to cover the movements of the ball. In the instructions, participants were asked to respond to the movements of the catcher while paying attention to the entire video. The causal manipulation in Experiment 2 was the same as Experiment 1: the facing direction of the catcher was horizontally reversed to generate the non-causal condition. The procedure for Experiment 2 was the same as that for Experiment 1.

3.2. Results and discussions

As shown in Fig. 5, the proportion of smooth responses again revealed a significant main effect of causality ($F(1,47) = 5.471$, $p = .024$). This result suggests that the causal relation between the two actors' actions impacted the visual experience of the catcher's body movements, as perceiving the catcher's movements elicited perception of more smooth and coherent motion. The proportion of judging the long-SOA video as a smooth action increased from 15.6% in the non-causal condition to 18.8% in the causal condition, yielding a Cohen's effect size value of $d = 0.40$. The proportion of judging the short-SOA video as a smooth action increased from 84.4% in the non-causal condition to 85.5% in the causal condition, yielding a Cohen's effect size value of $d = 0.13$. No interaction effect was found, $F(1,47) = 2.072$, $p = .157$. These results replicated the causal interpolation effects observed in Experiment 1, despite the fact that only part of the chain of causal events was presented in the stimuli.

To compare the causal interpolation effects observed in Experiments 1 and 2, an independent t-test compared the causal interpolation effects (i.e., the proportion of judgments of the catcher's action as smooth motion in causal conditions minus the proportion of judgments of the catcher's action as smooth motion in non-causal conditions) averaging across short-SOA and long-SOA conditions. Experiment 1 showed a slightly stronger interpolation effect than Experiment 2, with a marginally significant difference ($t(95) = 1.907$, $p = .060$). This difference was likely due to the stronger causal mechanisms available in Experiment 1 in which the full causal chain was shown with throwing

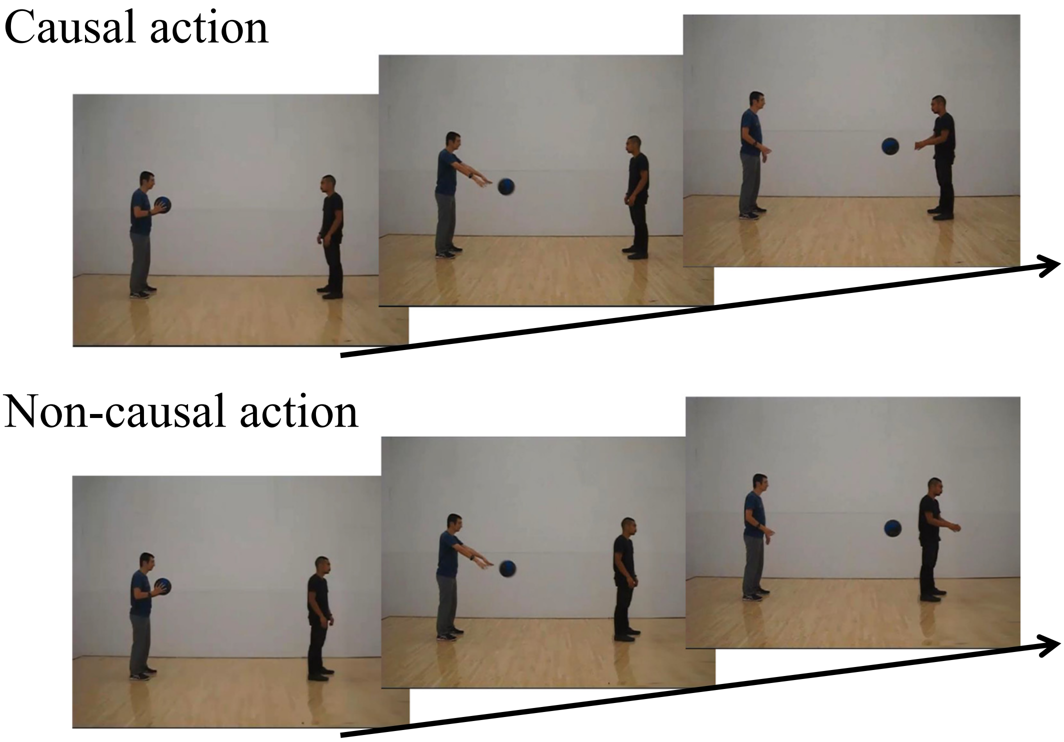


Fig. 2. An illustration showing sample frames in the causal condition with the catcher facing toward the thrower and the non-causal condition with the catcher facing away from the thrower in Experiment 1.

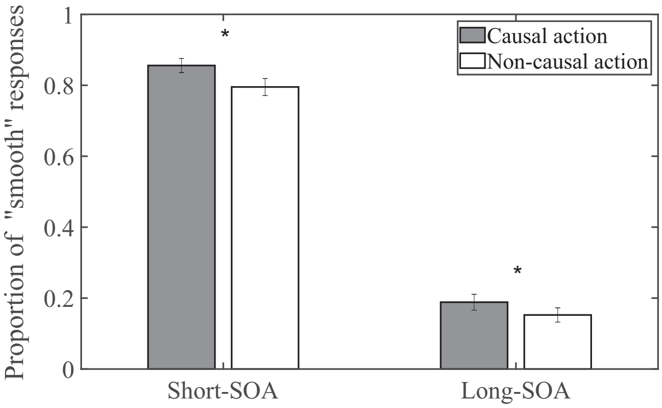


Fig. 3. Results of Experiment 1, showing a greater proportion of responses judging the catcher's action as smooth motion in the causal condition than in the non-causal condition. Asterisks indicate significant pairwise comparisons between causal and non-causal actions under short- or long-SOA conditions ($p < .05$).

action, ball movement and catching action, whereas the middle events in the causal chain (i.e., the flying ball) were not visible in Experiment 2.

In addition to reducing social causation, the non-nausal (flipped catcher) condition in Experiments 1 and 2 entailed a slightly greater distance between the actors. To rule out the possibility that these different distances can account for the interpolation effect, we conducted a control experiment with 53 new participants. We measured the distance between the two actors in the frame in which the catcher showed the most extended arm posture in the causal condition for each video. In the non-causal action condition, after horizontally flipping the catcher as in Experiment 2, the catcher was moved closer to the thrower to match the distance in the non-causal condition. To match the distance between the actors to that in the causal condition, the central part of the video (i.e., the ball flying in the air) was cut from the original video to move the thrower closer to the flipped catcher. Due to the closer distance, the width of the whiteboard occluder was reduced in the control experiment relative to in Experiment 2, although the occluder width in the control experiment was kept the same in the causal and non-causal condition. To avoid the ball reappearing after flying through the narrowed occluder, a shortened test clip was used to reduce the critical period from 200 ms to 133 ms, thereby showing less spatial displacement between the key postures. In addition, we selected videos starting earlier in time to maintain the same duration for a trial. Hence, the selected videos included more frames with little body motion of actors, and also less dramatic movements between the two key postures. These manipulations tended to yield less available dynamic information.

This control experiment revealed a causal interpolation effect, with a significant main effect of causality in the first block ($F(1,52) = 4.081$, $p = .049$), suggesting that the unequal proximity of the arms of the actions in the two conditions cannot account for the interpolation effect. However, we also found that the causal interpolation effect was attenuated with more repetitions of blocks, and after merging all five blocks the main effect did not reach significance ($F(1,52) = 1.767$, $p = .190$). The weakened effect in the control experiment is likely due to the fact that less dynamic motion information was provided in the stimuli, resulting in a weakened sense of social causation for the interaction between agents.

4. Experiment 3

In Experiment 2, we replicated the results in Experiment 1 that causal actions between agents induced stronger tendencies of perceiving smooth human body movements, even when the ball movements were covered, which likely weakens the availability of causal chain mechanisms. However, the effect size was reduced in Experiment 2 probably due to an occluded causal link between two agents. In Experiment 3, we aimed to test if the same effect could be replicated in situations with an agent interacting with only a moving object. Non-

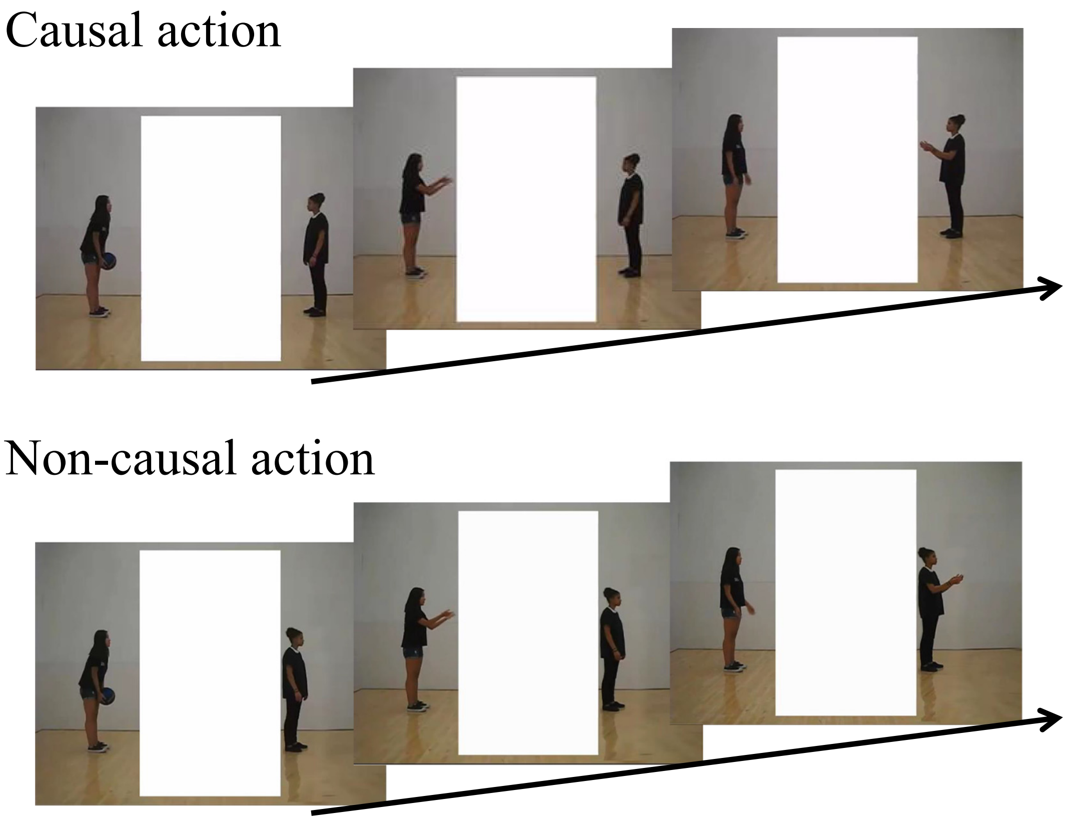


Fig. 4. An illustration showing sample frames in Experiment 2, in which the ball movements were occluded.

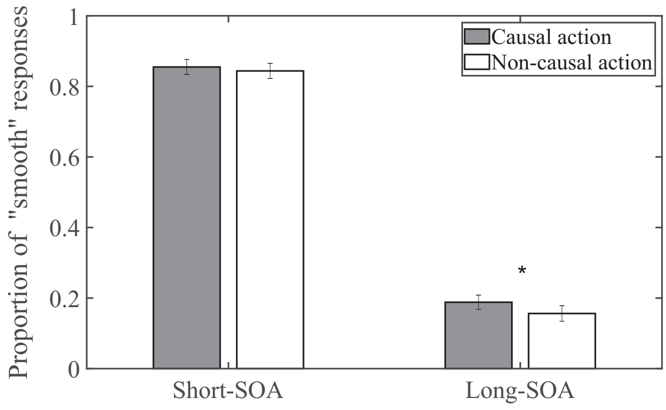


Fig. 5. Results of Experiment 2, showing a greater proportion of responses judging the catcher's action as smooth motion in the causal condition than in the non-causal condition. The asterisk indicates a significant pairwise comparison between causal and non-causal actions under the long-SOA condition ($p < .05$).

causal actions were generated with the same agent facing away from the moving object. We hypothesized that in the causal action condition, discretized human actions would be more likely to be perceived as smooth motion sequences.

4.1. *Methods*

4.1.1. *Participants*

Fifty new UCLA undergraduate students (mean age = 21.1; 40 female) participated in the experiment for course credit. All participants had normal or corrected-to-normal vision.

4.1.2. *Stimuli and procedure*

The experiment employed the same basic videos as in Experiment 1 except that the thrower was not presented in the videos. As shown in Fig. 6, the causal condition showed the catcher facing toward the ball as the ball movement causes the catcher to move his or her body in preparation. In non-causal actions, the facing direction of the catcher was horizontally reversed. The catcher was flipped horizontally to face away from the ball in the entire video, while keeping the background and the ball movement intact. The procedure for Experiment 3 was the same as that for Experiment 1 and 2.

4.2. *Results and discussions*

We conducted a 2 (SOA: short- vs. long-SOA) by 2 (causality: causal action vs. non-causal action) repeated-measures ANOVA on the proportion of responses judging the catcher's action as smooth motion. As shown in Fig. 7, results revealed a significant main effect of causal action, $F(1,49) = 12.419$, $p = .001$. These results indicate that subjects were more likely to judge actions as smooth motion in the causal condition compared to the non-causal condition. The main effect of the SOA was significant, $F(1,49) = 170.448$, $p < .001$. The two-way interaction effect between causality and SOA was not significant, $F(1,49) = 1.316$, $p = .257$. The proportion judging the long-SOA video as a smooth action increased from 19.0% in the non-causal condition to 23.7% in the causal condition, yielding a Cohen's effect size value of $d = 0.49$. The proportion of judging the short-SOA video as a smooth action increased from 77.7% in the non-causal condition to 80.5% in the causal condition, yielding a Cohen's effect size value of $d = 0.29$.

To compare the causal interpolation effects in Experiment 1 and the present experiment, we conducted an independent t-test on the causal interpolation effect averaging across short- and long-SOA conditions. The comparison did not show a significant difference between the causal interpolation effect observed in the two experiments (t

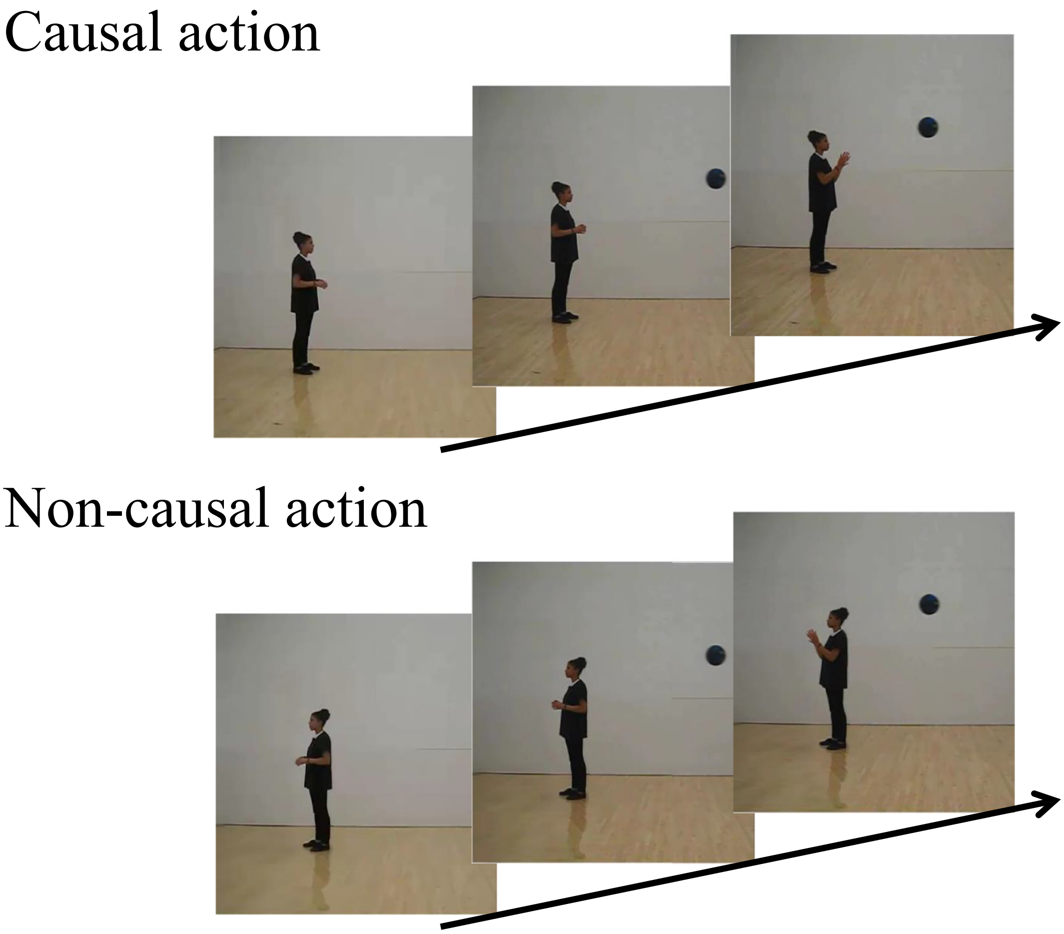


Fig. 6. An illustration showing sample frames in Experiment 3: a causal action with the catcher facing toward the ball, and a non-causal action with the catcher facing away from the ball.

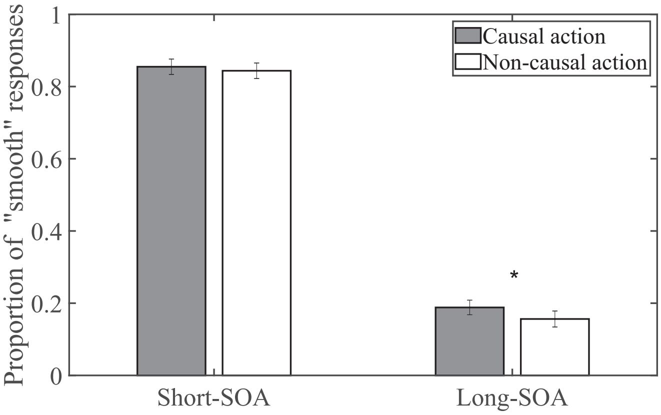


Fig. 7. Results of Experiment 3, showing a greater proportion of responses judging the catcher's action as smooth motion in the causal condition than in the non-causal condition. Asterisks indicate significant pairwise comparisons between causal and non-causal actions under short- or long-SOA conditions ($p < .05$).

(97) = 0.707, p = .481). This result shows that the direct causal relation between the ball movement (the cause) and the catcher's action (the effect) plays an important role in generating the interpolation effect. This finding suggests that when the causal object at the end of the causal chain is present, the addition of other events in the early part of the causal chain may not result in a larger interpolation effect.

5. Experiment 4

Experiment 4 aimed to investigate whether the influence of causal actions on motion interpolation depends on other visual cues. Body orientation is a well-known cue that can disrupt action processing (Pavlova & Sokolov, 2000; Thurman & Lu, 2013, 2014), as observers show worse recognition performance when actions are presented upside-down. If the interpolation effect revealed in the previous experiments was primarily linked to intention-based causal relations, then we expect that the causal interpolation effect would still be obtained in the upside-down condition, as both upright and upside-down displays preserve the temporal contingency among the events and the intention-based causal relation between humans and objects. However, if the effect depends on critical visual cues for action processing (such as upright body orientation) rather than intention-based causal relations, we would expect that the difference between causal and non-causal conditions in the upside-down videos would be attenuated or even eliminated.

5.1. Methods

5.1.1. Participants

Fifty-three new UCLA undergraduate students (mean age = 20.8; 43 female) participated in the experiment for course credit. All participants had normal or corrected-to-normal vision.

5.1.2. Stimuli and procedure

Experiment 4 used the same stimuli from the causal and non-causal

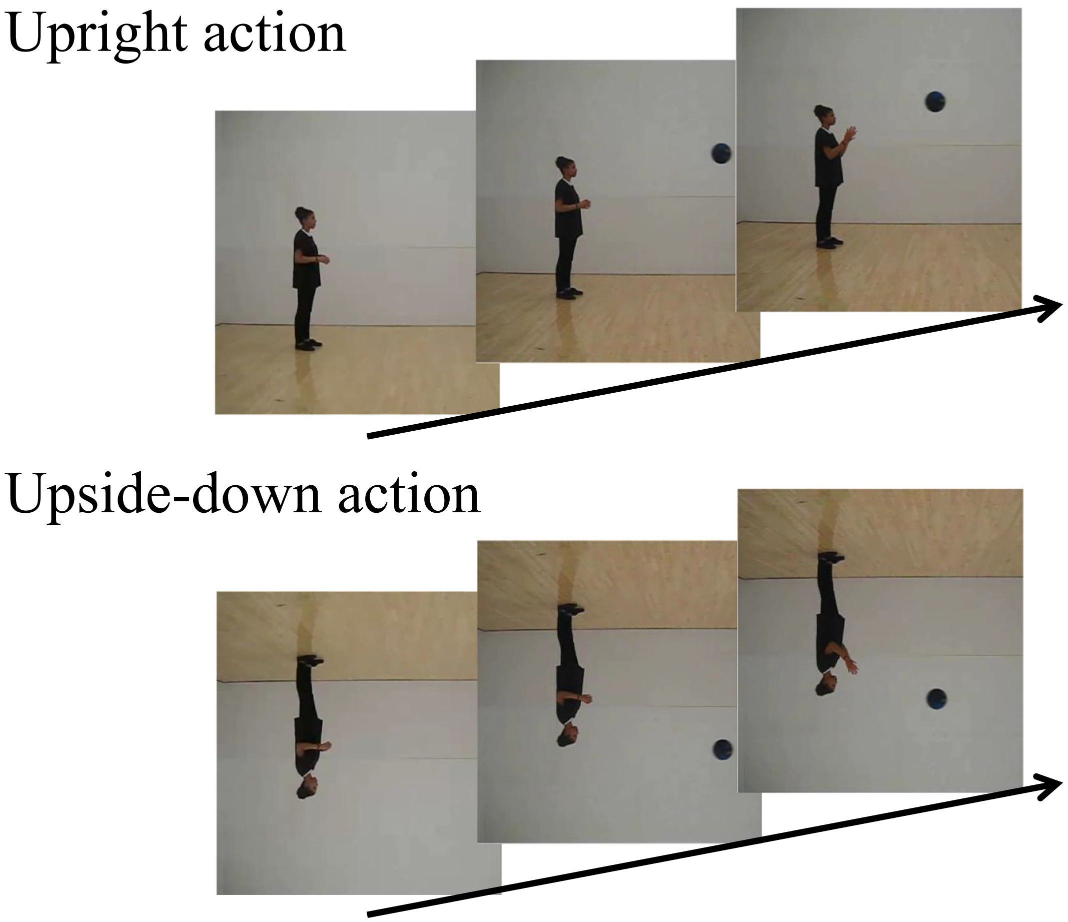


Fig. 8. An illustration showing sample frames of an upright and an upside-down action in the causal condition of Experiment 4. The non-causal condition used the facing-away catcher as previous experiments.

condition in Experiment 3 for the upright condition for half of the trails. The other half of trials used upside-down videos (Fig. 8). A total of 48 trials were administered (causal/non-causal × upright/upside-down × long-/short SOA × 6 actions). The task and procedure of Experiment 4 were otherwise the same as in Experiment 3.

5.2. Results and discussion

We conducted a 2 (SOA) by 2 (causality) by 2 (orientation: upright vs. upside-down) repeated-measures ANOVA on the proportion of responses in which the catcher's action was judged to be smooth motion. As shown in Fig. 9, the main effect of causality was significant ($F(1,52) = 5.909$, $p = .019$), revealing a causal interpolation effect for both upright and upside-down videos. The main effect of orientation was significant ($F(1,52) = 11.72$, $p = .001$), showing that the upright condition received a significantly higher proportion of responses for smooth motion than did the upside-down condition. This finding indicates a bias such that participants were more likely to report perceiving smooth actions in the upright than upside-down videos. However, none of two-way or three-way interactions were significant. Together, the results suggest that both body orientation (upright vs. upside-down) and social causation affect visual experience of seeing smooth movements. The absence of a two-way interaction effect between causality and body inversion confirms that the interpolation effect is primarily driven by causal relations between human actions and object movements. These causal relations were signaled by temporal contingency and perceived intention, rather than visual familiarity of the upright body orientation that we typically observe in daily life. Hence, the results of Experiment 4 indicate that the interpolation effect is triggered by the perception of social causality in a general fashion. Indeed, this effect is so robust that it can also be observed using upside-down action stimuli, which are highly unfamiliar to most participants.

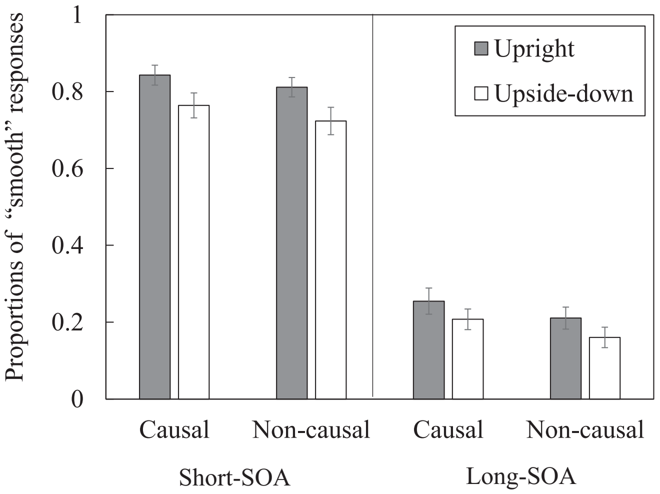


Fig. 9. Results of Experiment 4, showing a greater proportion of responses judging the catcher's action as smooth motion in the causal condition than in the non-causal condition, and also more responses in the upright condition than in the upside-down condition.

6. General discussion

The present paper reported converging evidence that causal relations between an agent and a physical object, or between agents, increased the likelihood that people would perceive smooth actions even when the stimuli showed a sudden change between two frames. This result suggests that intention-based causality acts as a temporal "glue" to fill in observers' visual experience by interpolating discrete image frames to produce the perception of smooth, continuous motion. The reported patterns of results are broadly consistent with the hypotheses of invoking higher-level visual processing within Braddick's (1980) two-process theory of apparent motion and the top-down influence in human action perception (Lu, Tjan, & Liu, 2006). Here, prior knowledge of causal relations involved in human actions is incorporated in higher-level visual processing, so that the recognition of events as causally connected facilitates the production of smooth motion from discrete visual inputs. The influence of intention-based causality may be stronger in situations where uncertainty about the visual input is high, as we found that Block 1 showed the relatively large effect size across all the experiments. We conjecture that this effect would be prominent when dynamic stimuli are presented in peripheral vision or embedded in noise. The effect may be weakened after repetitive exposures to the stimuli, as causal adaptation effect may occur to attenuate perceived causality (Rolfs, Dambacher, & Cavanagh, 2013) and perceptual learning may enhance performance for visual tasks by selecting only task-relevant information (Huang, Lu, Tjan, Zhou, & Liu, 2007).

The main findings in the present paper are consistent with previous evidence that a causal understanding of observed human actions helps to fill in important visual information left out from a sequence of events and to form a continuous perception (Strickland & Keil, 2011). The current result can also be viewed as the extension of previous findings showing the impact of physical causality on the perception of shapes. The representation of an object's implicit causal history has been shown to induce a transformational apparent motion (Tse, Cavanagh, & Nakayama, 1998) of simple objects (Chen & Scholl, 2016), akin to the "causal filling in" effect reported by Strickland and Keil (2011). The inference of implicit causal history of objects not only changes motion perception but also essentially has an impact on the visual shape representation (Spröte, Schmidt, & Fleming, 2016).

Could the causal interpolation effect be driven by factors other than social causality? The effect was unlikely to be caused by a failure of understanding the task, as we provided two practice blocks to participants at the beginning of each experiment. Across four experiments, on average, participants showed high accuracy in both practice block 1 with slower frame rate ($M = 0.91$, $SD = 0.13$) and practice block 2 with a normal frame rate ($M = 0.83$, $SD = 0.18$), demonstrating a good understanding of the task. Moreover, the causal interpolation effect was unlikely to be caused by differences between low-level visual information of causal and non-causal events, such as proximity between actors, proximity between the ball and the catcher, and symmetry of movements. For proximity between actors, we ran a control experiment to match the distance between actors in the causal and non-causal conditions. After controlling for the proximity of actors, the same causal interpolation effect was replicated. The results of Experiment 2 help to rule out the potential confounding variable of proximity between movements of the ball and the actor, as the causal interpolation effect was replicated even when ball movements were not visible. In addition, previous studies have shown that contextual movement is likely to induce apparent motion in the same direction of context but inhibit apparent motion in the opposite direction of context (Dawson, 1987; Green, 1983). In our study, the ball moved in the opposite horizontal direction from the catcher's arm movements in the causal condition, but in the same horizontal direction in the non-causal condition (due to the horizontal flipping). Based on this past research, we would expect that the non-causal condition would be more likely to elicit the perception of apparent motion, as the ball and the catcher's hand move in the same horizontal direction. But for the causal condition, in which the ball and the catcher's arm moved in opposite directions, and with greater proximity, the close and opposing ball movements would be expected to reduce the apparent motion of the catcher's arm.. This prediction is contrary to what we observed in the study. Hence, the proximity factor is unlikely to explain our main finding of stronger perception of smooth motion in the causal than in the non-causal condition. Lastly, symmetry between the catcher and the thrower is unlikely to have contributed to the effect because movements from the catcher and the thrower are not symmetric at any particular time point, as there is a delay between the arm movements of the two actors. However, our experimental conditions could not completely rule out a possible impact of an inward bias toward the center of the actor's visual field on the interpolation effect, as in most communicative interactions people are facing toward each other and toward the center of our visual field. To fully address this possibility, in future research one would need to examine stimuli without clear facing directions, such as object movements.

The impact of causality on continuous movements is potentially related to *temporal binding* and *spatial binding*. *Temporal binding* is a well-documented phenomenon where the time between two events appears shorter as a function of some relation between those two events (Buehner & Humphreys, 2009; Engbert & Wohlschläger, 2007; Engbert, Wohlschläger, Thomas, & Haggard, 2007; Humphreys & Buehner, 2009, 2010; Moore & Haggard, 2008; Wohlschläger, Haggard, Gesierich, & Prinz, 2003). Haggard, Clark, and Kalogeras (2002) were the first to demonstrate this phenomenon, and they interpreted this phenomenon by appealing to a coupling between the visual system and the motor system where the temporal binding of actions and effects heightens their association in order to facilitate action-outcome learning (Haggard, Aschersleben, Gehrke, & Prinz, 2002). Later, Buehner and Humphreys demonstrated that the crucial relation between the two events is causal: When one event is represented as causing another event, the time between the two events appears shorter than when the two events are not causally related. Similarly, *spatial binding* is where two objects appear closer in space when they are causally linked than when they are not (Buehner & Humphreys, 2010). Buehner and Humphreys (2009, 2010) explain both of these phenomena by invoking their theory of Bayesian ambiguity reduction. Appealing to Bayes Theorem, Buehner and Humphreys reason that two causally related events are more likely to instantiate spatiotemporal contiguity. They argue that the perceptual system uses prior knowledge of causal relations to help resolve ambiguities faced with taking noisy perceptual input to produce the subjective experience of visual motion. As a result, event kinds in causal relationships are more likely to appear bound in time and space.

Cause-effect relations instantiated by human body movements and its connection to social perception may not only help to connect discrete events in the perceptual process, but it may also facilitate the process of making inferences and predictions about actions. A causal framework may help the visual system to infer the past. For example, human observers get a vivid feeling of seeing the immediate past of objects or human postures presented in static frames (Kourtzi, 2004). This phenomenon suggests that the visual system can infer and reconstruct the causal history of objects and human actions. In addition, social causality in human actions may also help the visual system to predict the future. Su and Lu (2017) used skeletal biological motion displays and found a flash-lag effect, such that when a briefly-flashed dot was presented physically in perfect alignment with a continuously-moving limb, the flashed dot was perceived to lag behind the position of the moving joint. This finding suggests that the representation of human actions is anticipatory. It has also been found that infants as young as five months are able to gaze toward the future direction implied by the static posture of a runner (Shirai & Imura, 2014, 2016), suggesting the early emergence in infancy of an ability to predict dynamic human actions from still pictures. A "causal filling in" mechanism could have benefitted from evolutionary selection pressure by aiding the continuous perception of animal motions despite occlusion by trees or other obstacles.

Recognition of causal connections in human social interactions also helps the process of visual reconstruction. A study by Papeo, Stein, and Soto-Faraco (2017) found that two bodies facing each other were recognized more easily than the two bodies facing away from each other. This finding suggests that two-body dyads serve as a functional unit with a strong causal connection in social situations, and this structured configuration translates perception of scenes with multiple bodies into representations of social interactions. Previous research has shown that the presence of one agent that is demonstrating communicative actions increased the likelihood of detecting a second agent's movements embedded in noise, or the “second-agent effect” (Manera, Del Giudice, Bara, Verfaillie, & Becchio, 2011; Neri, Luu, & Levi, 2006). The improvement can be explained in a framework of causal actions, where the perception of others' action is constructed not only from the visual input, but also from the intrinsic predictive activities. The presence of one agent in a causal interaction in the social context impacts the prior expectation of seeing the second agent. When the expectation derived from prediction is strong enough, it elicits the illusory perception of a second agent even without the valid bottom-up visual input, described as “Bayesian ghost” by Manera et al. (2011).

The current study provides evidence of the important role played by intention-based causality in the perception of smooth motion. Social causation instantiated by human actions, and their interactions with objects and other agents, have a strong influence on motion perception for body movements. Perception of social causation involved in actions facilitates visual interpolation of discrete dynamic events to provide a continuous perception of human-involved activities, where the influence of causality in higher-level visual processing interacts with low-level visual processing in action perception. Hence, causality can serve as a basis for the visual system's anticipation of the future, as well as its retrospection of the recent past, for both physical events and social events.

Acknowledgments

This research was funded by National Science Foundation grant BSC-1655300 to HL and China Scholarship Council (CSC) Scholarship to YP. We thank Brian Scholl for helpful comments; Eun Ji Song, Tabitha Safari, and Jiming Sheng for helping with filming actions; and Eun Ji Song, Andrew Kwik, and Korosh Bahrami for their assistance in data collection.

References

Baker, C. L., & Cynader, M. S. (1986). Spatial receptive-field properties of direction-selective neurons in cat striate cortex. *Journal of Neurophysiology, 55*(6), 1136–1152. https://doi.org/10.1152/jn.1986.55.6.1136.

Baker, C. L., & Braddick, O. J. (1985a). Temporal properties of the short-range process in apparent motion. *Perception, 14*(2), 181–192. https://doi.org/10.1068/p140181.

Baker, C. L., & Braddick, O. J. (1985b). Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Research, 25*(6), 803–812. https://doi.org/10.1016/0042-6989(85)90188-9.

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.

Bours, R. J. E., Stuur, S., & Lankheet, M. J. M. (2007). Tuning for temporal interval in human apparent motion detection. *Journal of Vision, 7*(1), 2. https://doi.org/10.1167/7.1.2.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research, 14*(7), 519–527. https://doi.org/10.1016/0042-6989(74)90041-8.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. Philosophical Transactions of the Royal Society of London. *B, Biological Sciences, 290*(1038), 137–151. https://doi.org/10.1098/rstb.1980.0087.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436. https://doi.org/10.1163/156856897X00357.

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.

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological Science, 20*(10), 1221–1228. https://doi.org/10.1111/j.1467-9280.2009.02435.x.

Buehner, M. J., & Humphreys, G. R. (2010). Causal contraction: Spatial binding in the perception of collision events. *Psychological Science, 21*(1), 44–48. https://doi.org/10.1177/0956797609354735.

Burr, D. C., Ross, J., & Morrone, M. C. (1986). Smooth and sampled motion. *Vision Research, 26*(4), 643–652. https://doi.org/10.1016/0042-6989(86)90012-X.

Chen, Y.-C., & Scholl, B. J. (2016). The perception of history: Seeing causal history in static shapes induces illusory motion perception. *Psychological Science, 27*(6), 923–930. https://doi.org/10.1177/0956797616628525.

Churchland, M. M., Priebe, N. J., & Lisberger, S. G. (2005). Comparison of the spatial limits on direction selectivity in visual areas MT and V1. *Journal of Neurophysiology, 93*(3), 1235–1245. https://doi.org/10.1152/jn.00767.2004.

Csibra, G., Gergely, G., Bíró, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of ‘pure reason’ in infancy. *Cognition, 72*(3), 237–267. https://doi.org/10.1016/S0010-0277(99)00039-6.

Dawson, M. R. (1987). Moving contexts do affect the perceived direction of apparent motion in motion competition displays. *Vision Research, 27*(5), 799–809. https://doi.org/10.1016/0042-6989(87)90077-0.

Engbert, K., & Wohlschläger, A. (2007). Intentions and expectations in temporal binding. *Consciousness and Cognition, 16*(2), 255–264. https://doi.org/10.1016/j.concog.2006.09.010.

Engbert, K., Wohlschläger, A., Thomas, R., & Haggard, P. (2007). Agency, subjective time, and other minds. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1261–1268. https://doi.org/10.1037/0096-1523.33.6.1261.

Green, M. (1983). Inhibition and facilitation of apparent motion by real motion. *Vision Research, 23*(9), 861–865. https://doi.org/10.1016/0042-6989(83)90053-6.

Haggard, P., Aschersleben, G., Gehrke, J., & Prinz, W. (2002). Action, binding and awareness. *Common mechanisms in perception and action,* 266–285.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5*(4), 382–385. https://doi.org/10.1038/nn827.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology, 57*(2), 243–259. https://doi.org/10.2307/1416950.

Huang, X., Lu, H., Tjan, B. S., Zhou, Y., & Liu, Z. (2007). Motion perceptual learning: When only task-relevant information is learned. *Journal of Vision, 7*(14), 1–10. https://doi.org/10.1167/7.10.14.

Humphreys, G. R., & Buehner, M. J. (2009). Magnitude estimation reveals temporal binding at super-second intervals. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1542–1549. https://doi.org/10.1037/a0014492.

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.

Kim, S. H., Feldman, J., & Singh, M. (2013). Perceived causality can alter the perceived trajectory of apparent motion. *Psychological Science, 24*(4), 575–582. https://doi.org/10.1177/0956797612458529.

Kourtzi, Z. (2004). But still, it moves. *Trends in Cognitive Sciences, 8*(2), 47–49. https://doi.org/10.1016/j.tics.2003.12.001.

Lappin, J. S., & Bell, H. H. (1976). The detection of coherence in moving random-dot patterns. *Vision Research, 16*(2), 161–168. https://doi.org/10.1016/0042-6989(76)90093-6.

Lu, H., Tjan, B. S., & Liu, Z. (2006). Shape recognition alters sensitivity in stereoscopic depth discrimination. *Journal of Vision, 6*(1), 75–86.

Manera, V., Del Giudice, M., Bara, B. G., Verfaillie, K., & Becchio, C. (2011). The second-agent effect: Communicative gestures increase the likelihood of perceiving a second agent. *PLoS One, 6*(7), e22650. https://doi.org/10.1371/journal.pone.0022650.

Michotte, A. (1946/1963). The perception of causality (trans. T. R. Miles & E. Miles). New York: Basic Books.

Mikami, A., Newsome, W. T., & Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *Journal of Neurophysiology, 55*(6), 1328–1339. https://doi.org/10.1152/jn.1986.55.6.1328.

Moore, J., & Haggard, P. (2008). Awareness of action: Inference and prediction. *Consciousness and Cognition, 17*(1), 136–144. https://doi.org/10.1016/j.concog.2006.12.004.

Morgan, M. J., & Ward, R. (1980). Conditions for motion flow in dynamic visual noise. *Vision Research, 20*(5), 431–435. https://doi.org/10.1016/0042-6989(80)90033-4.

Neri, P., Luu, J. Y., & Levi, D. M. (2006). Meaningful interactions can enhance visual discrimination of human agents. *Nature Neuroscience, 9*(9), 1186–1192. https://doi.org/10.1038/nn1759.

Newsome, W. T., Mikami, A., & Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. *III. Psychophysics and physiology of apparent motion. Journal of Neurophysiology, 55*(6), 1340–1351. https://doi.org/10.1152/jn.1986.55.6.1340.

Papeo, L., Stein, T., & Soto-Faraco, S. (2017). The two-body inversion effect. *Psychological Science, 28*(3), 369–379. https://doi.org/10.1177/0956797616685769.

Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics, 62*(5), 889–899. https://doi.org/10.3758/BF03212075.

Peng, Y., Thurman, S., & Lu, H. (2017). Causal action: A fundamental constraint on perception and inference about body movements. *Psychological Science, 28*(6), 798–807. https://doi.org/10.1177/0956797617697739.

Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology, 23*(3), 250–254. https://doi.org/10.1016/j.cub.2012.12.017.

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.

Shiffrar, M., & Freyd, J. J. (1990). Apparent motion of the human body. *Psychological Science, 1*(4), 257–264. https://doi.org/10.1111/j.1467-9280.1990.tb00210.x.

Shiffrar, M., & Freyd, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science, 4*(6), 379–384. https://doi.org/10.1111/j.1467-9280.1993.tb00585.x.

Shirai, N., & Imura, T. (2014). Implied motion perception from a still image in infancy. *Experimental Brain Research, 232*(10), 3079–3087. https://doi.org/10.1007/s00221-014-3996-8.

Shirai, N., & Imura, T. (2016). Emergence of the ability to perceive dynamic events from still pictures in human infants. *Scientific Reports, 6*(1), https://doi.org/10.1038/srep37206.

Sigman, E., & Rock, I. (1974). Stroboscopic movement based on perceptual intelligence. *Perception, 3*(1), 9–28. https://doi.org/10.1068/p030009.

Spröte, P., Schmidt, F., & Fleming, R. W. (2016). Visual perception of shape altered by inferred causal history. *Scientific Reports, 6*(1), https://doi.org/10.1038/srep36245.

Strickland, B., & Keil, F. (2011). Event completion: Event based inferences distort memory in a matter of seconds. *Cognition, 121*(3), 409–415. https://doi.org/10.1016/j.cognition.2011.04.007.

Su, J., & Lu, H. (2017). Flash-lag effects in biological motion interact with body orientation and action familiarity. *Vision Research, 140*, 13–24. https://doi.org/10.1016/j.visres.2017.06.010.

Thurman, S. M., & Lu, H. (2013). Physical and biological constraints govern perceived animacy of scrambled human forms. *Psychological Science, 24*(7), 1133–1141. https://doi.org/10.1177/0956797612467212.

Thurman, S. M., & Lu, H. (2014). Perception of social interactions for spatially scrambled biological motion. *PLoS One, 9*(11), e112539. https://doi.org/10.1371/journal.pone.0112539.

Tse, P., Cavanagh, P., & Nakayama, K. (1998). *The role of parsing in high-level motion processing.* High-level motion processing: Computational, neurobiological, and psychophysical perspectives249–266.

Wertheimer, M. (1912). Experimentelle Studien uber das Sehen von Bewegung. *Zeitschrift Fur Psychologie, 61*. Retrieved from https://ci.nii.ac.jp/naid/10024048816/.

Wohlschläger, A., Haggard, P., Gesierich, B., & Prinz, W. (2003). The perceived onset time of self- and other-generated actions. *Psychological Science, 14*(6), 586–591. https://doi.org/10.1046/j.0956-7976.2003.psci_1469.x.