Research Article

Psychological Science
22(5) 674–681

DOI: 10.1177/0956797611404086
http://pss.sagepub.com

A Dissociation Between Judged Causality and Imagined Locations in Simple Dynamic Scenes

Florent Levillain[1,2] and Luca L. Bonatti[3]
[1]Department of Psychological and Brain Sciences, Johns Hopkins University; [2]Laboratoire "Cognitions Humaine et Artificielle," Université Paris 8; and [3]ICREA and Universitat Pompeu Fabra

Abstract

To mentally extrapolate the trajectory of a moving object that disappears from sight, different sources of information can be exploited: memory of its last visible position, its inferred movement through time, and general understanding of the causal structure of the scene. It is often assumed that these cues are integrated into unified analog mental representations. In our experiment, participants predicted the position of an object that disappeared behind an occluder and estimated the degree to which the movement was caused by another object. They made considerable errors in predicting imagined displacements. Moreover, their predictions were misaligned with their judgments of causality. They predicted the positions of the invisible moving objects better in events that they judged less causally correct than in events that they judged more causally correct. These results suggest that physical and cognitive parameters of imagined dynamic events do not merge into unitary mental models simulating actual states of the world.

Keywords

mental imagery, dynamic imagination, motion prediction, perception of causality

Received 7/9/10; Revision accepted 12/13/10

Performing actions on moving objects typically requires a high level of accuracy. The fact that people can perform tasks such as hitting or catching a ball shows that they can accurately represent the timing of a visible moving object and anticipate its future positions (Regan, 1992). Low-level visual mechanisms are available to track single and even multiple object displacements (Pylyshyn, 2001, 2003; Scholl & Pylyshyn, 1999; Scholl, Pylyshyn, & Feldman, 2001). However, in natural scenes, moving objects are not always visible. Thus, observers must often rely on mental representations of nonperceived dynamic events. In studies of mental imagery, it is generally assumed that the mind builds analog representations that are isomorphic with external events and allow the observer to "run" simulations of unseen situations (e.g., Johnson-Laird, 1983; Kosslyn, 1994); it is also assumed that these simulations even recruit the same neural tissue involved in the direct perception of visible events (Borst & Kosslyn, 2008; Klein et al., 2004; Kosslyn, Thompson, Kim, & Alpert, 1995). The ability to simulate events has been considered a hallmark of thinking; as the philosopher Colin McGinn (1989) wrote, "A thinking system, we might say, is a simulation engine—a device that mimics, copies, replicates, duplicates, imitates, parallels reality" (p. 176). This approach naturally invites the supposition that, in continuity with mechanisms for the perception of visible displacements, analog simulations provide the means to represent invisible dynamic events (Schwartz, 1999; Shepard & Cooper, 1986).

However, knowledge about dynamic events goes beyond the simple representation of invisible positions. Adults and infants can perceive high-level properties of dynamic scenes, such as causal relations (Leslie, 1982; Michotte, 1963) or agency status (Gergely, Nádasdy, Csibra, & Bíró, 1995). Thus, a plausible (although not necessary) view of the relation between perception and dynamic imagination is that information about the dynamic properties of a scene (e.g., their causal relations or the physical forces acting on objects) contribute to create a unique mental representation that simulates future states of invisible events, allowing the observer to predict them. Indeed, it has been claimed that the internalization of

Corresponding Author:
Luca L. Bonatti, ICREA and Universitat Pompeu Fabra, C. Roc Boronat, 138, Edifici Tanger, 55.110, 08018 Barcelona, Spain
E-mail: lucabonatti@mac.com

invariant properties of the environment is evolutionarily adaptive (Hubbard, 1995; Shepard, 2002).

In the experiment reported here, we tested both observers' accuracy in estimating displacement of invisible objects and the degree to which high-level properties of a dynamic scene, such as its degree of causal connectedness, are integrated into a unified mental representation. Participants estimated the timing of successive invisible displacements within dynamic scenes and judged the causal correctness of the scenes. We reasoned that if a unified mental representation grounds motion prediction by mental simulation, participants who viewed a scene that afforded reliable information about the dynamic properties of occluded objects would (a) estimate the objects' positions reasonably accurately and (b) make more accurate predictions for causally correct than for causally anomalous scenes. Alternatively, errors in estimating imagined positions, and possibly a dissociation between accuracy of on-line predictions and judgments of causal correctness, would indicate that there was no common representation of dynamic events.

Method

Participants

Nineteen participants completed the experiment (mean age = 24.4 years, range = 20–31 years). Ten (naive group) had learned physics only in middle or high school, 4 (intermediate group) had taken physics in college, and 5 (professional group) were physicists.

Stimuli and apparatus

We created realistic movies in which one ball (the "launcher") moved toward a second ball at the center of a scene (the "target"). One second after the beginning of each movie, the launcher appeared from one side of the scene and traveled horizontally toward the target at a constant speed of 25.8°/s, 19.3°/s, or 12.9°/s (see Videos S1–S5 in the Supplemental Material available online). The target began moving at the same speed and in the same direction at the moment the launcher stopped moving or after a delay (see the descriptions of the causality conditions later in this section).

Targets' trajectories were either occluded or visible. In the target-occluded movies, after initiating its movement, the target continued its trajectory behind an occluder, so that its actual position could only be imagined. Three vertical lines were drawn on the occluder, in one of two different configurations (Table 1). Movement direction (rightward/leftward) was counterbalanced across trials. The target-visible movies were identical to the target-occluded movies except that the position of the occluder was changed so that the target was always visible. The bars remained in the same positions as in the target-occluded movies (i.e., they were not moved with the occluder), so that participants would estimate the same movement distances in the target-visible and target-occluded conditions and accuracy in these estimates would therefore be directly comparable between the conditions. Figure 1 shows examples of the two conditions, illustrating each bar configuration.

The causality of the scenes varied. In the *contact* condition, the motion of the launcher ceased immediately after the launcher hit the target, which started to move immediately after contact. In the *delay* condition, the launcher contacted the target and stopped moving, but the target began moving after an interval of either 480 or 640 ms. In the *space* condition, the launcher stopped exactly when the target started moving, but it stopped before any contact, at a distance of either 100 or 130 pixels from the target.[1]

Finally, for both the target-visible and the target-occluded conditions, we created 20 distractor movies in which the launcher fell from above, landing in the same positions as in the experimental sequences. Because the movements of the targets and launchers were orthogonal in the distractor movies, they broke up the sequence of horizontal movements created by the experimental movies. Thus, there were 160 animations in total: 120 experimental animations, crossing five causality conditions (one contact condition, two delay conditions, two space conditions) with two target-visibility conditions

Table 1. Locations of the Vertical Bars and Hypothetical Times at Which the Targets Would Arrive at the Bars as a Function of Launcher Speed

		Hypothetical arrival time (s)		
Bar	Distance from origin	Launcher speed: 12.9°/s	Launcher speed: 19.3°/s	Launcher speed: 25.8°/s
		Configuration 1		
1	44.2°	0.44	0.28	0.24
4	60.8°	1.72	1.16	0.88
6	71.8°	2.60	1.72	1.32
		Configuration 2		
2	49.7°	0.88	0.56	0.44
3	55.3°	1.28	0.84	0.68
5	66.3°	2.16	1.40	1.12

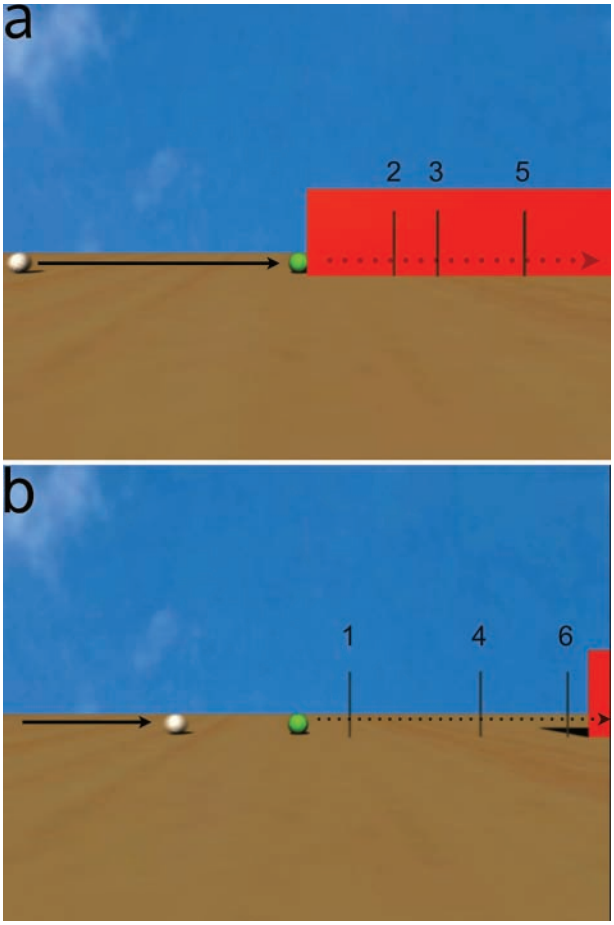

Fig. 1. Overall structure of the experimental movies. In the target-occluded condition, the movement of the target (the green ball) could only be imagined; the example shown here (a) represents the launcher (the white ball) in its initial position. In the target-visible condition, the movement of the target was always visible; the example shown here (b) represents the stopping position of the launcher in the space condition with the larger spatial interval. These examples illustrate the two different configurations of the vertical bars.

(visible, occluded), two bar configurations, three speeds, and two direction of movement, plus 40 distractor trials. At the end of each movie, a graded scale appeared so that participants could make causal judgments.

The experiment ran on a PowerMac G4, programmed with PsyScope X (for more information on this open-source software, see http://psy.cns.sissa.it). The animations were projected on a 200-cm × 135-cm screen by an Epson EMP 8100 projector placed behind the participants. The scenes filled the entire visual space, covering a visual angle of 70°. Response times (i.e., estimates of when the target reached the vertical bars) were collected with a button box that was placed together with a numerical keypad on a table in front of the participants.

Procedure

Participants sat in a dimmed room, 2.5 m from the screen. Each session began with an example of the contact condition as a practice trial. Participants were instructed to visually track the launcher and press the button on the button box when they felt that the target reached the position corresponding to each bar. They were told to press the button only three times (once for each bar), and to do so both if the target remained visible and if it disappeared behind the occluder. They were also informed that the balls moved at identical and constant velocities, so that the speed of the launcher was entirely predictive of the speed of the target.

Participants were also instructed to indicate the perceived strength of the causal relation between the launcher and the target by entering a number on the keypad when the scale appeared on the screen (0 = *not at all causal*, 9 = *completely causal*). No explicit connection was drawn between predicting the target's position and making the causality judgment. No feedback was given during the experiment.

Participants initiated each trial by pressing a button. The movies were presented in two blocks of 80, arranged in a pseudorandom order, with the constraint that the same causality condition could not occur more than three times in a row. The target-occluded movies were presented in the first block, and the target-visible movies in the second block. The experiment lasted approximately 1 hr.

Results

A trial was excluded from analysis (0.23% of the data) if the participant did not indicate the target's position exactly three times, if the participant pressed the button before the disappearance of the target, or if a response time was more than 2.5 standard deviations from the mean response time for that trial's condition. The two variables of interest were mean timing error (MTE) and causality judgments. MTE was calculated as the difference between the response time indicating when the target was estimated to reach a bar and the actual time when the target crossed that bar. A positive value indicates that the response was entered after the target crossed the bar, and a negative value indicates that the response was given before the actual arrival time. The frame in which an occluded target reached each bar was determined off-line as the first frame in which the target made contact with the bar in the 3–D model generating the movie. All analyses we report are two-way repeated measures ANOVAs, and all post hoc tests are alpha-adjusted with Bonferroni correction.

As movement direction had no main effect and did not interact with any other independent variable, we collapsed across this factor. We also collapsed across the two temporal and spatial intervals of the delay and space conditions, respectively, as restricted analyses showed that they had no influence on MTE. Figure 2 shows the average MTE at each bar position, for each combination of causality and target-visibility conditions.

We analyzed the difference in individual MTE means between the target-visible and target-occluded conditions by means of an ANOVA with target occlusion (visible, occluded),

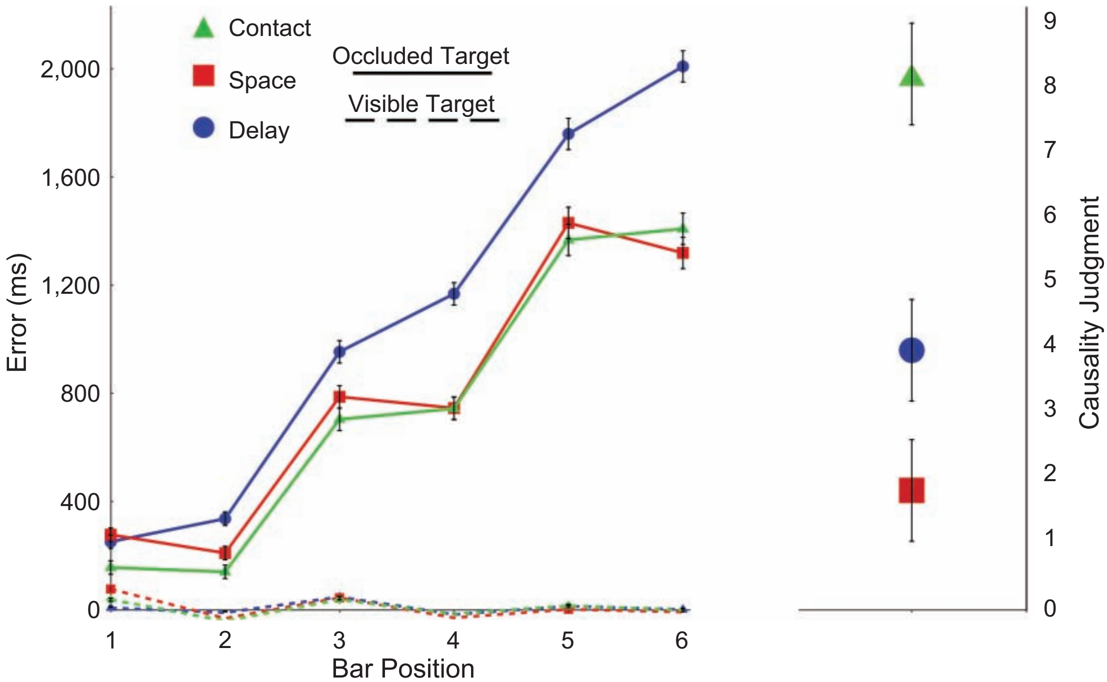

Fig. 2. Mean timing errors (left panel) and causal-strength judgments (right panel). Timing error is plotted as a function of bar position, separately for the contact, space, and delay conditions, both when the target was visible and when it was occluded. Causality judgments were made on a scale from 0 (*not at all causal*) to 9 (*completely causal*). Error bars represent 95% within-participants confidence intervals (Loftus & Masson, 1994) in the respective conditions.

speed (25.8°/s, 19.3°/s, 12.9°), bar (1, 2, 3, 4, 5, 6), and scene causality (contact, delay, space) as independent variables. Occlusion had a strong effect, $F(1, 18) = 100$, $p < .0001$, $p_{rep} = .99$, $\eta_p^2 = .77$. When the target was visible, participants correctly determined the exact moment of arrival at each position. This result not only shows that they could accurately predict contact points when direct visual feedback was available, but also shows that the task of tracking successive positions with the spatiotemporal parameters used was totally feasible. However, in the target-occluded condition, when participants had to imagine the target's position, they made important systematic errors.

Post hoc tests revealed that in the target-visible condition, participants were as precise in monitoring contact with the last bar as in monitoring contact with the first bar; error did not increase with traveled distance. However, the pattern of results in the target-occluded condition was markedly different. First, participants overestimated the time to contact at each tested position, which indicates that overestimation error began right when visual feedback for the target position was terminated. Second, although overestimation increased during occlusion, the increase was not continuous, as a simulation model would predict. Indeed, in the target-occluded condition (but not in the target-visible condition), participants could not differentiate between the close positions tested in the different bar configurations ($MTE_{Bar2} - MTE_{Bar1} = 0.85$ ms, $p = 1$; $MTE_{Bar4} - MTE_{Bar3} = 72.11$ ms, $p = .4$; $MTE_{Bar6} - MTE_{Bar5} = 61$ ms, $p = 0.8$). That is, error increased between responses, but not between equally spaced positions. This result suggests that participants had at best a rather limited ability to predict the position of an invisible target.

Speed had no effect on error, nor did speed and target visibility have an interactive effect on error. However, the interaction of speed and causality condition was significant, $F(4, 18) = 5.1$, $p \leq .0004$. Also, the significant three-way interaction of speed, causality condition, and target visibility, $F(4, 18) = 4.5$, $p \leq .0012$, showed that speed had an effect only when causality was violated in imagination.

To further clarify how the experimental manipulations affected participants' imagination of dynamic scenes, we ran restricted analyses on the target-occluded condition, collapsing across the two bar configurations, which did not differentially affect MTE. An ANOVA with response order (Response 1, or R1; Response 2, or R2; Response 3, or R3), speed, and causality condition as independent variables showed a strong effect of response order, $F(2, 18) = 99.13$, $p \leq .0001$, $p_{rep} = .99$, $\eta_p^2 = .85$. The overestimation increased at each successive response ($MTE_{R2} - MTE_{R1} = 621$ ms, $p \leq .0001$; $MTE_{R3} - MTE_{R2} = 698$ ms, $p \leq .0001$). The causal nature of the movies was also relevant, $F(2, 18) = 91.9$, $p \leq .0001$, $p_{rep} = .99$, $\eta_p^2 = .99$. Post hoc analyses showed that participants were less accurate in the delay condition than in contact condition ($MTE_{delay} - MTE_{contact} = 325$ ms, $p \leq .0001$), and less accurate in the delay condition than in the space condition ($MTE_{delay} - MTE_{space} = 284$ ms, $p \leq .0001$). However, they were equally (in)accurate in the contact and space conditions ($MTE_{space} - MTE_{contact} = 31$ ms, $p = .3$).

Speed had no main effect, but it interacted with causality condition, $F(4, 18) = 9.4$, $p \leq .0001$. Post hoc analyses showed that the interaction was due to the fact that MTEs in the space and contact conditions were not distinguishable when the target and launcher moved at speeds of 12.9°/s and 19.3°/s, but at speeds of 25.8°/s, scenes with spatial violations generated higher errors than causally correct scenes, $p \leq .0001$. In contrast, scenes with temporal violations generated a much higher error than scenes with spatial violations or causally correct scenes at all three speeds (all contrasts, $p \leq .00001$) Thus, in all tested conditions, a violation of the temporal relation of causality resulted in an error that was greater than the error observed when there was a spatial violation, or no violation at all.

The causal nature of the scenes also influenced participants' explicit judgments of causality (see Fig. 2), as revealed by an ANOVA with causal judgment as the dependent variable and causal condition as the independent variable, $F(2, 18) = 69.5$, $p \leq .0001$, $p_{rep} = .99$, $\eta_p^2 = .79$. As expected, the contact condition was judged more causally correct ($M = 8.25$, $SD = 0.71$) than the delay condition ($M = 3.99$, $SD = 2.86$) and the space condition ($M = 1.83$, $SD = 2.35$). Causality judgments differed significantly between all pairs of conditions (contact vs. space: $p \leq .0001$; contact vs. delay: $p \leq .0001$; delay vs. space: $p \leq .002$), but the difference was greatest between the contact and space conditions: Whereas causality was considered strongest in the contact condition, it was considered null in the space condition. It is notable that participants rated events involving temporal violations as more causal than those involving spatial violations. Thus, paradoxically, in the implicit task of predicting the position of an invisible target, participants performed better in the space condition than in the delay condition, even though they explicitly judged causality to be weaker in the space condition than in the delay condition. That is, prediction abilities and assessment of causality did not align.

Participants had to base their predictions about the position of the invisible target on the velocity of the visible launcher. Thus, it could be argued that although participants simulated the movement of the target, the representational momentum of the launcher, which has been argued to affect responses to Michotte-like scenes (Hubbard, Blessum, & Ruppel, 2001), influenced these simulations. Specifically, because the target in the space condition began to move exactly when the launcher stopped, the target could have "inherited" the representational momentum of the launcher. In contrast, in the delay condition, the launcher stopped before the target started moving, and thus there was much less opportunity for the target's speed to be simulated via representational momentum.[2] We were able to test this alternative explanation by exploiting the distractor trials. Because the target moved orthogonally to the launcher in these trials, there could be no transfer of representational momentum. Hence, according to this alternative explanation, causality condition should have had no effect on MTE in the distractor trials. In other words, the data should show an interaction between trajectory (orthogonal in distractor trials, identical in experimental trials) and causality condition.

To test this alternative explanation, we first checked that distractors engaged participants in processing the causal nature of the scenes. An ANOVA with causality judgment as the dependent variable and causality condition as the independent variable showed that they did, $F(2, 18) = 38.7$, $p < .001$. As we found for the experimental scenes, participants judged the events in the space and delay conditions as less causal than the events in the contact condition ($M_{space} = 2.18$, $SD = 2.8$; $M_{delay} = 3.51$, $SD = 2.4$; $M_{contact} = 7.33$, $SD = 1.2$). We then ran an ANOVA with MTE in the target-occluded condition as the dependent variable and trajectory (orthogonal, identical), causality condition, and response order as independent variables. Again, causality condition had a very strong effect, $F(2, 18) = 57.5$, $p \leq .0001$, and errors were higher in the delay condition than in both the space and the causality conditions (ps $\leq .000001$), but errors in the space and causality conditions did not differ ($p = .2$). Yet there was no effect of trajectory, $F(1, 18) = 0.05$, $p = .8$, and trajectory and causality condition did not have an interactive effect, $F(2, 18) = 0.35$, $p = .7$. That is, errors did not differ between the distractor and experimental movies. This result excludes an explanation in terms of transfer of representational momentum.

Finally, because many of our participants were highly skilled physicists, we could also control for expertise effects. An ANOVA with level of expertise as the independent variable (naive, intermediate, professional) revealed that expertise had no effect on prediction accuracy, $F(2, 18) = 0.35$, $p = .71$, or on causality judgments, $F(2, 18) = 1.35$, $p = .29$.

Discussion

How does the cognitive system compensate when information about the trajectory of a moving object is incomplete? The literature on mental imagery and on the prediction of motion frequently appeals to mental analog representations as a potential substrate for spatial computations and understanding of dynamic events. Here, we have provided evidence that this approach may not offer an adequate account of how observers represent dynamic stimuli.

We devised a motion-prediction task that directly probed participants' on-line ability to predict the future position of a moving object, rather than testing memory for past positions, as is done in other paradigms (e.g., Hubbard, 1995; Hubbard et al., 2001). We found that participants were highly accurate at determining the exact moment of arrival of a moving target only when the target was continuously visible. When it was occluded and they had to rely on imagination, their estimates were highly inaccurate, even though the scenes were very simple and brief; the amount of error was as high as 70% of the scenes' durations (see Table 1 and Fig. 2). These results are in accordance with and expand on previous results obtained with different paradigms (e.g., Gilden, Blake, & Hurst, 1995). Neither participants' knowledge that the movement of the visible launcher was completely predictive of the invisible movement nor their thorough understanding of the underlying physical laws could reduce or prevent the error. Furthermore, the error did not appear to increase continuously as traveled distance increased: Not only were participants unable to simulate a simple uniform movement, but they could not even distinguish close intervals when these intervals were not estimated by immediate sequential responses; this was true regardless of the speed of the objects and the causal correctness of the scenes. Taken together, this quantum increase in error is difficult to reconcile with a simulation theory of imagined movement.

Finally, by coupling the movement-prediction task with explicit judgments of causal strength, we showed that intuitive judgments of causality do not accord with predictions of imagined object movements. Participants' predictions of the displacements of occluded objects were better for scenes that participants judged as highly causally incorrect (space condition) than for scenes they judged as more causally correct (delay condition). Moreover, we showed that this dissociation cannot be explained by an influence of representational momentum on simulations of the targets' movement. This dissociation casts strong doubt on the possibility that a common substrate integrates simulations of physical variables and high-level knowledge about the simulated scenes into analog representations that replicate and parallel reality.

How lower- and higher-level properties of visual representations are integrated is a widely discussed topic. Some authors argue that visual perception is cognitively impenetrable: Low-level visual mechanisms tie the mind to the world in a manner that is unaffected by what an organism thinks or knows about the world (Pylyshyn, 1986, 1999). Others argue that the high-level interpretation of a scene influences low-level perceptual properties. Of particular relevance to our topic is Buehner and Humphreys (2010) argument that "the higher-level concept of causality has a profound influence on humans' perception of both space and time" (p. 44), influencing low-level properties, such as the perception of the size of an object. Our results add to this debate, suggesting instead that causality does not influence even imagined object movements. Indeed, even Buehner and Humphreys's experiments contain traces of the same dissociation we found. They suggested that space-time perception is warped along the causality dimension because, in a memory task, participants underestimated the actual size of an object involved in causal scenes. However, scenes judged as highly noncausal induced object-size recollections that were closer to those induced by fully causal scenes than to those induced by scenes that were similar to the scenes of our delay condition and were similarly judged as violating causality only mildly. That is, a causal violation judged to be more severe exerted less "spatial warping" than a less severe violation.

These results, as well as ours, suggest that there is no simple way to understand the influence of causality on other cognitive processes, nor any simple way to categorize causality in humans' cognitive landscape. Impressions of causality lie at the "intersection between perception and cognition" (Scholl & Nakayama, 2002, p. 493). Clearly, some low-level properties that induce impressions of the presence or absence of causality, such as contact between objects or temporal asynchronies between the movements of two objects, have a profound effect on other lower-level cognitive processes, such as memory for an object's size or reconstruction of a dynamic sequence. But it remains a wide-open question whether it is causality per se, as opposed to such low-level properties, that influences memory, perception, or imagination.

How can one account for participants' poor prediction of imagined movements, as well as for the dissociation between assessment of causality and prediction of movement? Work by Pylyshyn and his collaborators suggests one radical possibility: The visual indexes assigned to objects during tracking encode no object properties at all (Pylyshyn, 2007), or at most encode objects' last visible locations (Keane & Pylyshyn, 2006), and if location information is encoded, it cannot be updated by computations extrapolating trajectories. Indeed, in a multiple-object tracking paradigm, Keane and Pylyshyn (2006) showed that participants track moving objects that disappear better when the objects reappear at the location of their disappearance than when they reappear at the location where they should be, according to an extrapolation updating their locations. Our data are compatible with Pylyshyn's hypothesis. This framework suggests that the position of an object will be encoded correctly when continuous visual feedback updates location information, but not when such feedback is lacking. Therefore, our participants' inaccurate yet systematic errors in estimating the location of an invisible object may indicate that the imagined estimations come from postperceptual processes, possibly involving neither spatial representations nor analog simulation of trajectories. Indeed, although the visual system tends to anticipate the position of moving objects (Nijhawan, 2008), imagined position is characterized by severe time overestimation, which suggests that a mechanism unrelated to visual tracking per se may be involved.

Some studies suggest that subjective durations of events involving moving objects tend to be longer than subjective durations of events involving stationary stimuli (Brown, 1995; Kanai, Paffen, Hogendoorn, & Verstraten, 2006). We suggest that time dilation, although not sufficient to explain all our results, may be partially responsible for the direction of the timing errors we observed. If participants tried to reproduce the time necessary for a target to cover a certain distance and experienced time dilation induced by movement, they would overestimate how long it would take for the target to cover the distance. In other words, rather than extrapolating an invisible object's position by means of an analog mental simulation of real physical forces, participants may have used an internal clock to estimate the position of the object. Such an estimate would be so coarse and detached from sources of knowledge about dynamic events that it would not integrate an event's basic causal structure. Further experiments are needed to explore this possibility. Note that because time estimates recruit brain networks other than those involved in spatial representation (e.g., Meck, 2005), our explanation suggests that the mechanisms involved in the representation of physical scenes may be functionally and neurally distinct from the mechanisms involved in the representation of imagined dynamic movement.

In conclusion, our results point toward the existence of different independent systems that may be involved in imagining dynamic movement. One of these systems may translate the passing of time into rough estimates of object positions, and another may compute causal relations in the world. These results show that dynamic imagination, if it exists at all, is severely limited. They call into question the idea that the essence of a thinking system is the ability to generate richly detailed analog representations.

Acknowledgments

We thank J. Mehler and the members of the Language, Cognition and Development Lab in Trieste for discussions and their help in setting up our experimental paradigm; Z. Pylyshyn, M. Buehner, and M. Nathan for their insightful comments; and L. Filippin and A. Isaja for their technical support.

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 supported by the Spanish Ministry of Science and Innovation (Ref. PSI2009-08232) and by McDonnell Foundation Grant 220020096.

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

Notes

1. The supplemental videos illustrate the following combinations of variables—Video S1: contact condition, target occluded, 19.3°/s launcher speed, Bar Configuration 1, rightward movement; Video S2: space condition, target occluded, 12.9°/s launcher speed, Bar Configuration 1, rightward movement, 100-pixel spatial interval; Video S3: space condition, target visible, 25.8°/s launcher speed, Bar Configuration 2, leftward movement, 130-pixel spatial interval; Video S4: delay condition, target occluded, 12.9°/s launcher speed, Bar Configuration 1, rightward movement, 480-ms temporal interval; Video S5: delay condition, target visible, 25.8°/s launcher speed, Bar Configuration 2, leftward movement, 640-ms temporal interval.
2. We thank Marc Buehner for raising this possibility.

References

Borst, G., & Kosslyn, S.M. (2008). Visual mental imagery and visual perception: Structural equivalence revealed by scanning processes. *Memory & Cognition*, *36*, 849–862.

Brown, S.W. (1995). Time, change, and motion: The effects of stimulus movement on temporal perception. *Perception & Psychophysics*, *57*, 105–116.

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

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*,165–193.

Gilden, D., Blake, R., & Hurst, G. (1995). Neural adaptation of imaginary visual motion. *Cognitive Psychology*, *28*, 1–16.

Hubbard, T.L. (1995). Environmental invariants in the representation of motion: Implied dynamics and representational momentum, gravity, friction, and centripetal force. *Psychonomic Bulletin & Review*, *2*, 322–338.

Hubbard, T.L., Blessum, J.A., & Ruppel, S.E. (2001). Representational momentum and Michotte's 'launching effect' paradigm (1946/1963). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 294–301.

Johnson-Laird, P.N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.

Kanai, R., Paffen, C.L., Hogendoorn, H., & Verstraten, F.A. (2006). Time dilation in dynamic visual display. *Journal of Vision*, *6*(12), Article 8. Retrieved from http://www.journalofvision.org/content/6/12/8.full

Keane, B.P., & Pylyshyn, Z.W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*, 346–368.

Klein, I., Dubois, J., Mangin, J.-F., Kherif, F., Flandin, G., Poline, J.-B., et al. (2004). Retinotopic organization of visual mental images as revealed by functional magnetic resonance imaging. *Cognitive Brain Research*, *22*, 26–31.

Kosslyn, S.M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press.

Kosslyn, S.M., Thompson, W.L., Kim, I.J., & Alpert, N.M. (1995). Topographical representations of mental images in primary visual cortex. *Nature*, *378*, 496–498.

Leslie, A.M. (1982). The perception of causality in infants. *Perception*, *11*, 173–186.

Loftus, G.R., & Masson, M.E.J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*, 476–490.

McGinn, C. (1989). *Mental content*. Oxford, England: Basil Blackwell.

Meck, W.H. (2005). Neuropsychology of timing and time perception. *Brain and Cognition*, *58*, 1–8.

Michotte, A. (1963). *The perception of causality*. Oxford, England: Basic Books.

Nijhawan, R. (2008). Visual prediction: Psychophysics and neurophysiology of compensation for time delays. *Behavioral & Brain Sciences*, *31*, 179–198.

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case of impenetrability of visual perception. *Behavioral & Brain Sciences*, *22*, 341–365.

Pylyshyn, Z. (2003). *Seeing and visualizing: It's not what you think*. Cambridge, MA: MIT Press.

Pylyshyn, Z.W. (1986). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Pylyshyn, Z.W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, *80*, 127–158.

Pylyshyn, Z.W. (2007). *Things and places: How the mind connects with the world*. Cambridge, MA: MIT Press.

Regan, D. (1992). Visual judgements and misjudgements in cricket and the art of flight. *Perception*, *21*, 91–115.

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

Scholl, B.J., & Pylyshyn, Z.W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*, 259–290.

Scholl, B.J., Pylyshyn, Z.W., & Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*, *80*, 159–177.

Schwartz, D.L. (1999). Physical imagery: Kinematic versus dynamic models. *Cognitive Psychology*, *38*, 433–464.

Shepard, R.N. (2002). Perceptual-cognitive universals as reflections of the world. *Behavioral & Brain Sciences*, *24*, 581–601.

Shepard, R.N., & Cooper, L.A. (1986). *Mental images and their transformations*. Cambridge, MA: MIT Press.