Physically Implied Surfaces

Patrick C. Little & Chaz Firestone

Johns Hopkins University

Version: 9/19/20 — in press, *Psychological Science*

Address for Correspondence:
Patrick Little (pat.little@nyu.edu) or Chaz Firestone (chaz@jhu.edu)
Department of Psychological and Brain Sciences
Johns Hopkins University
Baltimore, Maryland 21218

Abstract (150 words)

In addition to seeing objects that are directly in view, we also represent objects that are merely *implied* (e.g., by occlusion, motion, and other cues). What can imply the presence of an object? Here, three experiments (preregistered N = 360 adults) explore the role of *physical interaction* in creating impressions of objects that are not actually present. After seeing an actor collide with an "invisible wall" or step onto an "invisible box", subjects gave facilitated responses to actual, visible surfaces that appeared where the implied wall or box had been—a "Stroop"-like pattern of facilitation and interference that suggested automatic inferences about the relevant implied surfaces. Follow-up experiments ruled out confounding geometric cues and anticipatory responses. We suggest that physical interactions can trigger representations of the participating surfaces, such that we automatically infer the presence of objects implied only by their physical consequences.

Keywords: illusory contours, intuitive physics, causal perception, magic, mime

Statement of Relevance (144 words)

How do we know which objects are around us? It might seem as simple as having light from those objects hit our eyes. Yet we also respond to “illusory objects” that are merely implied, such that no light from them reaches us at all. Consider the vivid experiences that mimes induce when they seem to interact with walls, ropes or boxes: We often get a surprisingly clear sense of the size, shape, and location of those objects, even though they don’t actually exist. The present work takes this experience “into the lab”, by exploring how the mind rapidly and automatically represents the surfaces that people and objects seem to interact with. How we represent objects in our environment can be driven not only by the objects themselves, but also by what happens *to* them.

Our experience of the world goes beyond the light reaching our eyes. For example, an object's perceived color is determined not only by the wavelengths of the light it reflects, but also by the inferred conditions of its illumination; an object's perceived size depends not just on its angular extent, but also its apparent distance; and of course nearly all visual illusions reflect a discrepancy between retinal stimulation and our subsequent perceptual experience—as when two equal lines appear different in length, or a static image appears to move.

However, there may be no better illustration of this principle than when we have impressions of objects that aren't even "visible" in the first place, because they cast no light onto our eyes. For example, when an object is partially hidden by an occluding surface, we may infer its continuity behind the surface even though no light from that part of the object reaches us. And in the phenomenon of illusory contours, we experience surfaces that don't exist at all but rather are only implied by other cues, such as coincidental clipping of multiple figures, or unified motion against a background (Figure 1).

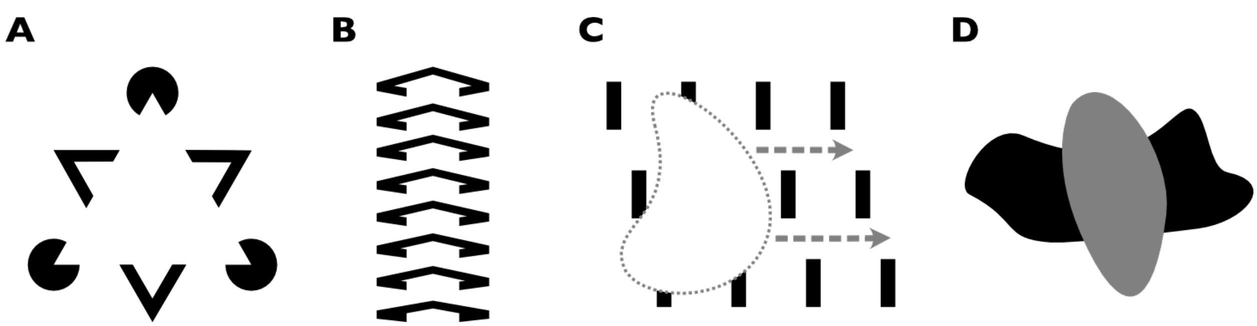

Figure 1. Geometric and kinetic cues can give rise to impressions of objects and surfaces that are not actually present. Note the illusory upright triangle in panel A and the "tower" in panel B. Panel C represents a dynamic display, where the illusory shape is revealed over time as visible texture elements are accreted and deleted. Panel D demonstrates amodal completion (of the black object, behind the gray occluder). (Adapted from Kanizsa, 1976; Tse, 1998; Palmer et al., 2006; Singh & Fulvio, 2007). Note that panels A–C are instances of "modal" completions, whereas panel D depicts an "amodal" completion.

Such phenomena are crucial to our experience of a coherent and unified environment. Indeed, object representation is rarely straightforward in the real world, which presents us with a variety of impoverished viewing conditions. As such, these processes may operate at essentially every moment we see the world, and have rightly played a prominent role in theorizing about the nature, function, and development of object processing (e.g., Kellman & Shipley, 1991; Kellman & Spelke, 1984; Nakayama, He, & Shimojo, 1995; Rock & Anson, 1979).

***Physically* implied surfaces?**

What kinds of factors can trigger representations of objects that aren't there? On one hand, implied objects and the rules governing their creation can be quite rich and sophisticated, incorporating information about perspective, volume, and rigidity (as in, e.g., Figure 1B; Tse, 1998, 1999; also Meyer & Dougherty, 1990). On the other hand, the kinds of input that are known to give rise to such impressions tend to involve only basic geometric and kinetic factors. For example, coincidental clipping (Figure 1A) can be characterized in terms of geometric properties, and even known dynamic cues (Figure 1C) are still explained in terms of straightforward patterns of motion. Could more sophisticated cues play a role?

Here, we explore this possibility by examining the role of *physical interactions*—bumping, bouncing, colliding, supporting, and other such events—in the representation of objects and surfaces. One clue that events of this sort might play this role comes from a potentially surprising source: stage performers—especially mimes—who can induce vivid impressions of implied objects (such as a wall, rope, or box) simply by seeming to *physically interact* with them. Even though no light from such objects reaches us (since they do not exist in the first place), a sufficiently convincing physical interaction with an imaginary object—such as appearing to lean on, run into, or climb over a wall—may lead us to infer (and almost "see") that one is there. Such impressions seem to lie somewhere between full-blown visual processing and mere higher-level reasoning—a kind of automatic imagination in which we *can't help* but represent the implied participating objects (Nanay, 2010). Indeed, instances of this phenomenon have been so compelling that they have ignited popular media interest (e.g., the viral "invisible box challenge"; https://osf.io/w6fty/). More generally, attention and cognition are especially tuned to physical interactions in the world. For example, infants are sensitive to causality in displays of collisions (Leslie & Keeble, 1987), and adults readily infer properties of objects involved in physical interactions (Hamrick et al., 2016; Todd & Warren, 1982; Ullman et al., 2018; Wu et al., 2015), represent the future states of moving and colliding objects (Gerstenberg et al., 2017; see also Guan & Firestone, 2020; Hubbard & Bharucha, 1988; Peng et al., 2020), and even form impressions of causal relations mediated by unseen, force-transmitting connecting elements (as in classic "pulling" stimuli; Michotte, 1963; White & Milne, 1997; see also Scholl & Nakayama, 2004).

These observations suggest that objects and surfaces might be inferred from physical interactions themselves, in ways that go beyond previously known cues. For example, when an actor appears to lean against an invisible wall, the resulting impression of the wall does not derive from the actor being clipped or occluded (cf. Figure 1A), nor from the motion properties that generate other illusory surfaces (e.g., Andersen & Braunstein, 1983; Kellman & Cohen, 1984). Instead, we infer the presence of the wall as an *explanation* for the otherwise-mysterious kinematics of the actor's behavior—as if the mind is asking, "how else could he be leaning like that?" (Rock & Anson, 1979).

The present experiments: "Invisible objects" implied by physical interaction

Here, we aimed to capture this experience in a laboratory setting. We set aside whether such objects are properly "seen" (see General Discussion), and instead explore the *automaticity* of such representations—their tendency to arise spontaneously, without instruction, and in ways that interfere with other responses. We showed subjects videos of actors colliding with "invisible walls" or stepping onto "invisible boxes", in ways that created vivid impressions of the surfaces they appeared to interact with. After viewing these events, a visible line appeared that either matched or didn't match the orientation of the surface implied by the interaction; subjects' task was simply to report the line's orientation, which was completely unpredictable from the preceding event. We reasoned that, if the mind automatically infers the surfaces implied by such physical interactions, then visible lines that match (or don't match) these implied surfaces would show a Stroop-like pattern of facilitation (or interference) with subsequent responses (MacLeod, 1991; Stroop, 1935), even with no statistical connection between the events (such that the nature of the physical interactions are completely task-irrelevant; for recent use of such designs in visual cognition studies, see Konkle & Oliva, 2012; Long & Konkle, 2017). Collectively, these experiments explore how physical interactions trigger inferences about the implied participating objects, and in ways that intrude upon subsequent responses.

Experiment 1: Running Into An Invisible Wall

Can physical interaction create impressions of "invisible surfaces"? Experiment 1 tested this possibility using a facilitated reporting paradigm, involving videos of a real human actor physically interacting with implied objects.

Method

Open Science Practices

All data, code, analysis, stimuli, and pre-registrations (for this experiment and all others reported here) are available at https://osf.io/sq9td/. The sample sizes and analysis plans (as well as other details) for all experiments were pre-registered.

Participants

A convenience sample of 120 adult participants was recruited from Amazon Mechanical Turk. (For a discussion of this subject pool's reliability, see Crump et al., 2013.) This was chosen as a generous sample size in comparison to previous visual cognition studies of this sort (typically n<40; e.g., Long & Konkle, 2017, Palmer et al., 2006, White & Milne, 1997), and was pre-registered for this and all other experiments.

Stimuli and Procedure

To create the "colliding" and "stepping" stimuli, we filmed an actor running into a (real) wall and stepping onto a (real) box, and then digitally removed these objects from the videos to produce the impression that the actor was interacting with an invisible surface. To these modified videos, we added visible candidate "surfaces" whose locations and orientations were either *congruent* or *incongruent* with the actor's behavior. In particular, 300ms after the interaction (i.e., after the actor bounced off the wall or stepped off the box), a black line (6px in thickness) appeared that could either be vertical (181px long) or horizontal (96px long), and thus either congruent with the surface implied by the interaction (i.e., a horizontal line after stepping on the box, or a vertical line after running into the wall) or incongruent (i.e., a vertical line after stepping on the box, or a horizontal line after running into the wall). Due to the nature of online experiments, we cannot specify here the exact size, viewing distance, or brightness (etc.) of the images as they appeared to subjects, because we could not know each subject's particular viewing conditions or display parameters. However, any distortions introduced by a given participant's viewing distance or monitor settings would have been equated across all stimuli and conditions.

Participants' task was simply to report the orientation of the line that appeared, regardless of what interaction came before (Figure 2A). Note that this design differs from classical biological motion experiments (e.g., Kozlowski & Cutting, 1977; Cutting & Kozlowski, 1977; Troje & Chang, 2013) in that the focus was on properties of the interacted-with surface, rather than the actor. (For discussion of biological motion experiments that *do* involve manipulated objects, see Experiment 2; Runeson & Frykholm, 1981; Stoffregen & Flynn, 1994.)

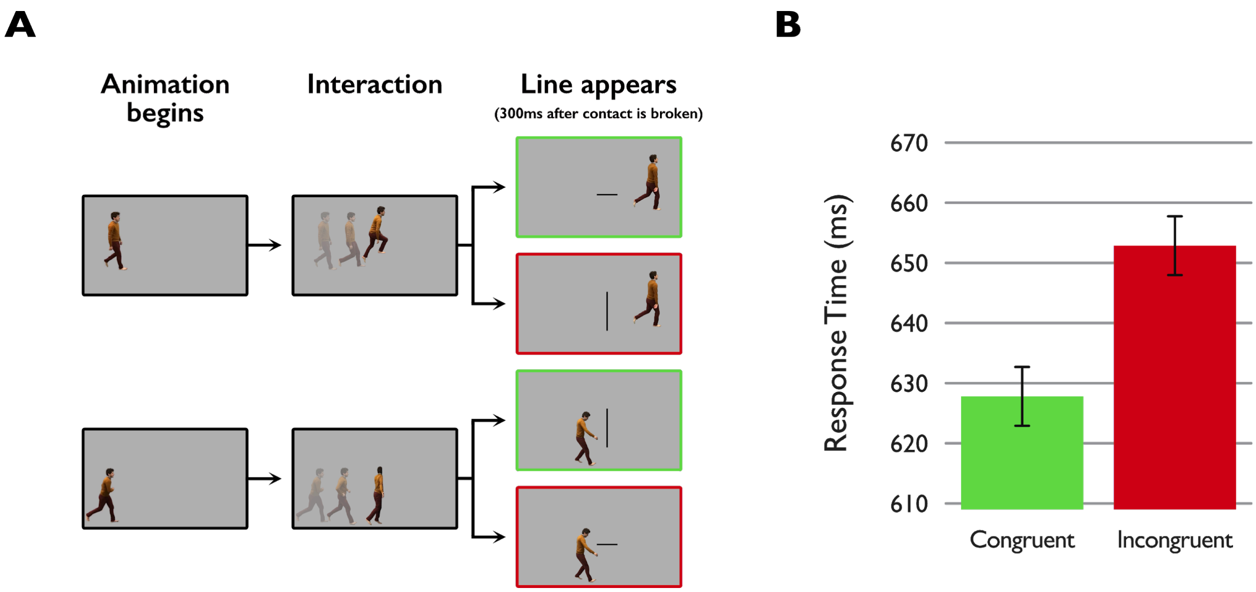

Figure 2. Design and results of Experiment 1. (A) On each trial, subjects saw an actor collide with an “invisible wall” or step onto an “invisible box”. After the collision, a line appeared; subjects simply had to report the line’s orientation. (B) Subjects were faster to report the line’s orientation when it matched the orientation of the surface implied by the actor’s behavior. Error bars are ±1 standard error of the mean difference between conditions.

Both the original and edited videos are available in our full methods archive; shorter demos of these displays as they appeared in the experiment can be viewed at https://perceptionresearch.org/mime.

Importantly, the preceding physical events were entirely non-predictive of the line’s orientation: on half of the trials, the line was congruent with the physical interaction, and on half it was incongruent, such that the actor’s behavior was a completely unreliable cue to the orientation of the line. There were thus four primary trial types, corresponding to the two types of interactions (colliding with a wall, or stepping onto a box), and the two line orientations (vertical, or horizontal): Wall-Vertical and Box-Horizontal (the *Congruent* trial types), and Wall-Horizontal and Box-Vertical (the *Incongruent* trial types). There were 40 trials total: 10 each of the four trial types. (Half of the trials of each trial type were "mirrored" such that half of the time the actor approached the invisible wall or box from the left of the display, and half of the time from the right of the display; however, we collapse over these variants from here onward.) To respond, subjects simply pressed the key that was assigned to that orientation ('1' or '2', randomly assigned for each subject), and were given one second to do so once the line appeared; if they did not respond in this time, a "Too Slow" feedback message was shown, and the trial was not included in the response time analysis.

We reasoned that, if the mind automatically infers the "invisible surfaces" that must be present to explain a physical interaction, then responses to the visible line would be primed or facilitated by first seeing an interaction that implied such a surface—even when such interactions were completely non-predictive of the visible line's orientation.

Results

In accordance with our pre-registered analysis plan, we excluded any subjects who failed to provide a complete dataset, or who responded accurately (and within the time limit) on fewer than 80% of trials. Of the 83 subjects remaining, accuracy was 91.4%, and mean response time was 640ms. From these subjects, 0.15% of trials were excluded for subjects' responses being too fast (responding in less than 200ms). Thus, the task was fairly easy and straightforward, as expected.

Crucially, subjects responded faster when the real, visible line had the same orientation as the invisible surface that had just been implied by the actor's stepping or colliding (628ms vs. 653ms; $t(82)=5.10$, $p<.001$; d=0.56; 95% CI for the difference between conditions = [15.28, 34.85]; Figure 2B). Additionally, they were no less accurate in the Congruent condition than the Incongruent condition—92.3% vs. 90.5% (which, if anything, was more accurate). This suggests that the surface implied by the physical interaction was actively represented by the mind, such that it could alter later responding. In other words, seeing an actor collide with a vertical wall produced a sufficiently vivid impression of a vertical surface that subjects were primed to respond to a *real* vertical surface that subsequently appeared. (Or, alternatively, subjects were *slower* to respond to a surface whose orientation conflicted with the implied interaction, in ways analogous to the phenomenon of Stroop interference, which has recently been applied in visual cognition more generally to explore similar questions of automaticity; Konkle & Oliva, 2012; Long & Konkle, 2017.) We took this result as evidence that physical interactions automatically or spontaneously trigger representations of the surfaces that they seem to imply.

Experiment 2: Idealized Stimuli With "Postdictive" Processing

We are interpreting these facilitated responses as reflecting the *physics* of the interaction between the actor and the implied surface. However, such physical interaction was perhaps confounded with the *shape* of the actor's body. For example, in Wall trials, the actor necessarily assumed a vertical posture upon colliding with the wall; perhaps, then, subjects responded more quickly to subsequent vertical lines only because the actor literally appeared more vertical on those trials (in that his body deformed on contact with the vertical wall), rather than because of the implied surface that the actor collided with.

Experiment 2 addressed this issue using idealized stimuli that implied surface orientation only "postdictively" (Eagleman & Sejnowski, 2000). Subjects saw a rigid disc fall toward—and then "bounce" off of—an invisible surface whose orientation was implied only by the disc's exit trajectory, which could either be straight up (implying a horizontal surface), or angled (implying an oblique surface). Crucially, the disc contacted the surface only once, and it was its behavior *after leaving* the surface that retroactively specified the orientation of the surface it must have interacted with. If the same pattern holds here as in Experiment 1, this would isolate the present phenomenon to the physics of the interaction per se, rather than any confounding geometric cues.

Method

Participants

120 adult participants were recruited from Amazon Mechanical Turk. All details of this sample size (as well as the analysis plan and exclusion criteria mentioned below) were pre-registered.

Stimuli and Procedure

This experiment was identical to Experiment 1 except that the stimuli were now animated displays of a gray disc which "dropped" down with realistic acceleration under gravity before suddenly appearing to "bounce" off of an invisible surface. On half of the trials, the disc bounced straight back up, implying a horizontal surface; on the other half of trials, the disc bounced off at an angle, implying an oblique surface. Importantly, the disc contacted the implied surface only once, and in only a single location (instead of a spatiotemporally extended interaction). This ensured that properties of the unseen surface could be inferred only from the motion of a different object, and not on the basis of any visual information from the surface or its boundary. Indeed, this aspect of the design removed a confound that was present not only in our Experiment 1 but perhaps also in previous work exploring representations of objects that are manipulated by biological motion actors (Runeson & Frykholm, 1981; see also Stoffregen & Flynn, 1994, which cleverly includes an object that isn't visible at all but which nevertheless still provides multiple samples of its boundary). As in Experiment 1, a visible line (10px in thickness and 250px in length) appeared 300ms after the bounce (though here the line was a faint shade of gray only slightly lighter [#B5B5B5; half of trials] or darker [#ABABAB; half of trials] than the neutral gray background [#B0B0B0]); the line was either horizontal or rotated 15° clockwise. There were 80 trials total: 20 each of the four trial types, of which half had the lighter gray line appear and half the darker gray line. Participants' task was simply to report with a keypress the orientation of the line that appeared, regardless of what interaction came before, within a time limit of two seconds (Figure 3A).

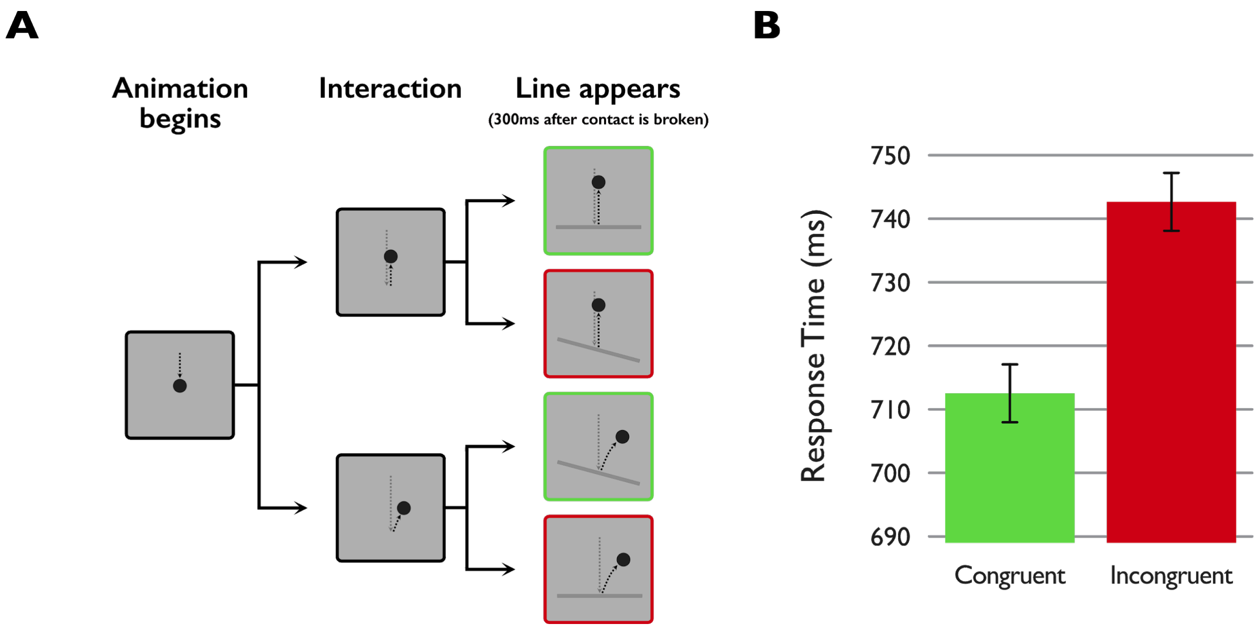

Figure 3. Design and results of Experiment 2. (A) On each trial, subjects saw a disc bounce off an "invisible surface" (which it contacted only once) and then exit along a particular trajectory. After this, a line appeared, and subjects simply had to report the line's orientation, regardless of the preceding events. (B) Subjects were faster to report the line's orientation when it matched the orientation of the surface implied by the disc's exit trajectory. Error bars are ± 1 standard error of the mean difference between conditions. Note that, in the experiments themselves, the lines were a fainter shade of gray than appears here.

<u>Results</u>

In accordance with our pre-registered analysis plan, we excluded any subjects who failed to provide a complete dataset, or who responded accurately on fewer than 80% of trials. Of the 106 subjects remaining, accuracy was 93.9% and mean response time was 728ms. From these subjects, 0.35% of trials were excluded for being too fast (responding within 200ms).

As in Experiment 1, subjects were faster to report the orientation of the line when it had the same orientation as the surface implied by the disc's bounce (714ms vs. 743ms; $t(105)=6.41$, $p<.001$;

d=0.63; 95% CI for the difference between conditions = [20.32, 38.51]). Unlike in Experiment 1, this pattern can only be explained by the *physics* of the interaction—in particular, by the exit trajectory of the disc.

Experiment 3: No Delay

The phenomenon we seek to capture would suggest an interaction between the surface representation inferred by subjects and their subsequent judgments of the visible surface that then appeared. However, the previous experiments could have an alternative explanation. In particular, it was possible that, on at least some trials, subjects prepared a response (based on the preceding physical interaction) during the 300ms before the visible line even appeared, which could manifest in faster responses for congruent trials in ways that wouldn't require subjects to have perceived the visible line at all (e.g., if they "made up their mind" before the line was even shown). Though even this interpretation would still suggest a previously unknown influence of physical interactions, Experiment 3 asked whether such events can still intrude on judgments of other stimuli even without any delay period in which to prepare a response in advance of the line's appearance.

Method

Participants

120 adult participants were recruited from Prolific (Peer et al., 2017). All details of this sample size (as well as the analysis plan and exclusion criteria mentioned below) were pre-registered.

Stimuli and Procedure

As in Experiment 2, the stimuli were falling discs that bounced off of invisible surfaces—and there were again "congruent" and "incongruent" trials (Figure 4), for a total of 80 trials.

However, rather than appearing after a 300ms delay, the visible line appeared at the very same moment the disc made contact with the implied surface. In previous experiments including a delay between the interaction with the invisible surface and the appearance of the visible line, subjects might have prepared a response (based on that trial's physical interaction) before the line even appeared. But here, with the line appearing at the moment of contact, that was not possible; instead, subjects could not prepare a response until the line actually appeared—which is the central way in which this experiment differed from Experiments 1–2.

Even beyond the previous experiments, this design ensured that the disc's exit trajectory was truly task-irrelevant—not only in the weak sense that there was no statistical association between the surface implied by the bouncing behavior and the actual surface that appeared, but also in the stronger sense that the subject could in principle see which visible surface actually appeared before even knowing which surface was implied by the bounces. Thus, this experiment was an especially strong test of the automaticity of implied surfaces from physical interactions: If the bouncing behavior still interfered with subjects' responses, this would suggest that they *can't help* but compute the orientation of the surface implied by physical interactions, and that such inferences directly intrude upon judgments about what subjects see.

Results

In accordance with our pre-registered analysis plan, we excluded any subjects who failed to provide a complete dataset, or who responded accurately on fewer than 80% of trials. Of the 114 subjects remaining, accuracy was 95.4% and mean response time was 721ms. From these subjects, 0.04% of trials were excluded for being too fast (i.e., responding within 200ms).

As before, participants reported the orientation of the line more quickly when it had the same orientation as the surface implied by the disc's bounce (717ms vs. 727ms; $t(113)=2.66$, $p=.009$;

d=0.25; 95% CI for the difference between conditions = [2.40, 16.37]). This suggests that the bouncing behavior influenced subjects' responses about which stimuli were shown, even when there was no reason (either in principle or in practice) for it to do so. In other words, whereas Experiments 1–2 may have involved prospective anticipation of which surfaces might appear, the present result suggests that physical interactions also influence subjects' responses about surfaces *after* they appear.

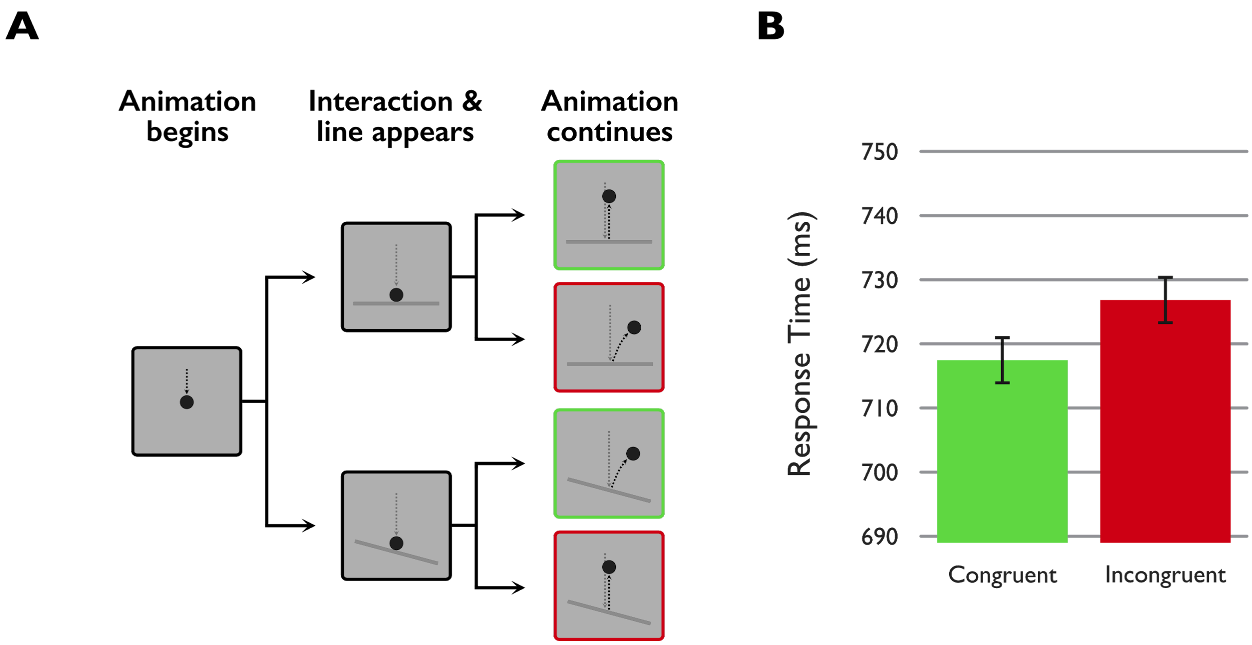

Figure 4. Design and results of Experiment 3. (A) On each trial, subjects saw a disc bounce off an "invisible surface" (which it contacted only once) and then exit along a particular trajectory. At the moment the disc "made contact" with the invisible surface, a line appeared, and subjects simply had to report the line's orientation, regardless of the preceding events. (B) Subjects were faster to report the line's orientation when it matched the orientation of the surface implied by the disc's exit trajectory. Error bars are ± 1 standard error of the mean difference between conditions.

General Discussion

What can make us represent an object that isn't there? Here we have suggested that physical interaction can automatically trigger inferences about "invisible" objects or surfaces.[1] When an actor collided with or stepped onto an invisible object, this physical event produced a vivid impression of the implied participating surface, facilitating responses to actual, visible surfaces matching those implied by the actor (Experiment 1). This phenomenon could not be explained by geometric confounds or spatiotemporally extended interactions, and even occurred postdictively (Experiments 2–3); it also generalized between idealized displays and more naturalistic stimuli with real actors. Of course, like so much visual cognition research, the present task involves only computer displays rather than the real world itself; and these experiments only explored a small region of the possible design space (e.g., in terms of the temporal delay between interaction and line, the behavioral measures used, etc.). Future work could expand on both of these dimensions to further increase the generalizability of these findings, or explore lower-level forms of processing (e.g., not only altered responses or judgments but also enhanced contrast sensitivity; Teufel, Dakin, & Fletcher, 2018; though we note that it is controversial whether such methods could apply to the phenomena we explore here; see, e.g., Salvano-Pardieu et al., 2010).

[1] Note that we do *not* claim to have demonstrated enhanced "perception" of the line that appeared (nor a top-down effect of cognition on perception; Firestone & Scholl, 2016b); all we report is a kind of response interference. Indeed, the two authors of this paper disagree about whether the word "visual" captures even the phenomenology of the present demonstrations (e.g., the display in Experiment 1, which also appears at https://perceptionresearch.org/mime). PCL thinks not, because the surface representations evoked by these physical interactions lack seemingly essential properties of visual perception (such as color and spatial extent). CF agrees that such representations are missing several qualities of full-blown visual experience, but nevertheless thinks they have something relevant in common with normal vision; after all, one is not tempted in any way to say that one *hears*, *smells*, or *tastes* the invisible box that the actor steps on—but one is at least somewhat tempted to say that one *sees* it. However, each author thinks the other makes a good case; and the central question of this paper ("Does the mind automatically infer the surfaces implied by physical interactions?") is independent of whether (a) the induced representation is properly visual, and certainly (b) the nature of any subsequent response interference.

Importantly, these results (involving facilitated or impaired judgments of a visible line) do not establish that perception of the line itself was altered (e.g., that a horizontal line "looks different" or is somehow "harder to see" when preceded by an actor running into an invisible vertical surface), nor even that the induced surface representation is properly "visual". Instead, our claim is simply that inferences about the surfaces implied by such interactions proceed spontaneously and even automatically, such that they interfere with otherwise-straightforward perceptual judgments. In other words, we take our results to show that the mind infers the surfaces implied by physical interactions even without any requirement to do so—and indeed even when doing so actively impairs performance.

The "reach" of physics

This work adds to a growing literature exploring mental representations of *physical events*. Whereas classical work on physical reasoning focused on slower and more deliberate judgments about physical systems (e.g., McCloskey et al., 1980), more recent work explores aspects of physical processing that may be faster and more intuitive, including the attentional processes involved in representing the future states of colliding objects (Gerstenberg et al., 2017; see also Chen & Scholl, 2016; Guan & Firestone, 2020; Kominsky et al., 2017), falling towers (Battaglia et al., 2013; Firestone & Scholl, 2016a), or swinging pendula (Smith et al., 2013). The present results suggest that such physical representations not only affect judgments of causality, stability, or time (Buehner & Humphreys, 2009), but also trigger representations of objects and surfaces themselves, even when they do not physically exist in the first place.

From the stage to the lab: Automaticity and performance

This work also shows how insights from entertainers and stage performers can inform psychological research—an aspiration for many years that is only recently being realized (e.g., Barnhart et al., 2018; Ekroll, Sayim, & Wagemans, 2017; Yao, Wood, & Simons, 2019). Here, we were inspired by these striking experiences, but we studied them by measuring subjects' *performance* on an indirect task in ways that allowed us to evaluate the automaticity of such processing. Even previous work investigating impressions of interacted-with objects has typically studied such impressions by asking subjects *about* the very experiences under study; but this raises the possibility that such impressions arise only in task-specific ways, or even that subjects may not have had such impressions if they hadn't been asked to report them. By contrast, subjects' task here—responding to a line—was *different* from the phenomenon of interest (the surface implied by a physical interaction). Success on this task didn't require any attention to the actor's particular behavior or especially the disc's exit trajectory, yet such events still influenced performance on the line-identification task, suggesting that subjects spontaneously represented such surfaces.

Together, this work suggests that how we represent objects in our environment—and even whether we represent them at all—can be driven not only by properties of the objects themselves, but also by what happens *to* them.

Author Contributions

PCL and CF jointly designed the experiments. PCL programmed and ran the experiments, and analyzed the data, with input from CF. PCL and CF jointly wrote the manuscript.

Acknowledgments

For helpful discussion and/or comments on previous drafts, we thank Mick Bonner, Yi-Chia Chen, Jorge Morales, Joan Ongchoco, Ian Phillips, and other members of the JHU Perception & Mind Laboratory. PCL was supported by a STAR award from the JHU Office of Undergraduate Research. CF was supported by the JHU Science of Learning Institute and National Science Foundation BCS #2021053.

References

Andersen, G. J., & Braunstein, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, *34*, 356–362.

Barnhart, A. S., Ehlert, M. J., Goldinger, S. D., & Mackey, A. D. (2018). Cross-modal attentional entrainment: Insights from magicians. *Attention, Perception & Psychophysics*, *80*, 1240–1249.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*, 18327–18332.

Buehner, M. J., & Humphreys, G. R. (2009). Causal binding of actions to their effects. *Psychological Science, 20*, 1221-1228.

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

Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral Research. *PLoS One*, *8*, e57410.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, *9*, 353–356.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science*, *287*, 2036-2038.

Ekroll, V., Sayim, B., & Wagemans, J. (2017). The other side of magic: The psychology of perceiving hidden things. *Perspectives on Psychological Science, 12*, 91–106.

Firestone, C., & Scholl, B. (2016a). Seeing stability: Intuitive physics automatically guides selective attention. *Journal of Vision*, *16*, 689.

Firestone, C., & Scholl, B. J. (2016b). Cognition does not affect perception: Evaluating the evidence for ‘top-down’ effects. *Behavioral & Brain Sciences, 39*, e229.

Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2017). Eye-tracking causality. *Psychological Science*, *28*, 1731-1744.

Guan, C., & Firestone, C. (2020). Seeing what’s possible: Disconnected visual ‘parts’ are confused for their potential ‘wholes’. *Journal of Experimental Psychology: General*, *143*, 590–598.

Hamrick, J. B., Battaglia, P. W., Griffiths, T. L., & Tenenbaum, J. B. (2016). Inferring mass in complex scenes by mental simulation. *Cognition*, *157*, 61-76.

Hubbard, T. L., & Bharucha, J. J. (1988). Judged displacement in apparent vertical and horizontal motion. *Perception & Psychophysics, 44*, 211–221.

Kanizsa, G. (1976). Subjective contours. *Scientific American*, *234*, 48-53.

Kellman, P. J., & Cohen, M. H. (1984). Kinetic subjective contours. *Perception & Psychophysics*, *35*, 237–244.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, *15*, 483–524.

Kominsky, J. F., Strickland, B., Wertz, A. E., Elsner, C., Wynn, K., & Keil, F. C. (2017). Categories and constraints in causal perception. *Psychological Science*, *28*, 1649–1662.

Konkle, T., & Oliva, A. (2012). A familiar-size Stroop effect: real-world size is an automatic property of object representation. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 561–569.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, *21*, 575–580.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality?. *Cognition, 25*, 265–288.

Long, B., & Konkle, T. (2017). A familiar-size Stroop effect in the absence of basic-level recognition. *Cognition*, *168*, 234–242.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin, 109*, 163–203.

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: Naive beliefs about the motion of objects. *Science*, *210,* 1139-1141.

Meyer, G. E., & Dougherty, T. J. (1990). Ambiguous fluidity and rigidity and diamonds that ooze!. *Perception*, *19*, 491-496.

Michotte, A. (1963). *The perception of causality*. New York: Basic Books.

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. In S. M. Kosslyn & D. N. Osherson (Eds.), *Visual cognition: An invitation to cognitive science* (pp. 1–70). Cambridge, MA: MIT Press.

Nanay, B. (2010). Perception and imagination: amodal perception as mental imagery. *Philosophical Studies, 150*, 239–254.

Palmer, E. M., Kellman, P. J., & Shipley, T. F. (2006). A theory of dynamic occluded and illusory object perception. *Journal of Experimental Psychology: General*, 135, 513–541.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153–163.

Peng, Y., Ichien, N., & Lu, H. (2020). Causal actions enhance perception of continuous body movements. *Cognition*, *194*, 104060.

Rock, I., & Anson, R. (1979). Illusory contours as the solution to a problem. *Perception*, *8*, 665–681.

Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 733–740.

Salvano-Pardieu, V., Wink, B., Taliercio, A., Fontaine, R., Manktelow, K. I., & Ehrenstein, W. H. (2010). Edge-induced illusory contours and visual detection: Subthreshold summation or spatial cueing? *Visual Cognition*, *18*, 591–616.

Scholl, B. J., & Nakayama, K. (2004). Illusory causal crescents: Misperceived spatial relations due to perceived causality. *Perception, 33*, 455–469.

Singh, M., & Fulvio, J. M. (2007). Bayesian contour extrapolation: Geometric determinants of good continuation. *Vision Research, 47*, 783-798.

Smith, K. A., Battaglia, P., & Vul, E. (2013). Consistent physics underlying ballistic motion prediction. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 35, No. 35).

Stoffregen, T. A., & Flynn, S. B. (1994). Visual perception of support-surface deformability from human body kinematics. *Ecological Psychology*, *6*, 33–64.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.

Teufel, C., Dakin, S. C., & Fletcher, P. C. (2018). Prior object-knowledge sharpens properties of early visual feature-detectors. *Scientific Reports, 8*, 10853.

Todd, J. T., & Warren Jr, W. H. (1982). Visual perception of relative mass in dynamic events. *Perception, 11*, 325–335.

Troje, N. F., & Chang, D. H. F. (2013). Shape-independent processes in biological motion perception. In K. L. Johnson & M. Shiffrar (Eds.), *People watching: Social, perceptual, and neurophysiological studies of body perception* (pp. 82–100). New York, NY: Oxford University Press.

Tse, P. U. (1998). Illusory volumes from conformation. *Perception*, *27*, 977–992.

Tse, P. U. (1999). Volume completion. *Cognitive Psychology*, *39*, 37–68.

Ullman, T. D., Stuhlmüller, A., Goodman, N. D., & Tenenbaum, J. B. (2018). Learning physical parameters from dynamic scenes. *Cognitive Psychology*, *104*, 57–82.

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

Wu, J., Yildirim, I., Lim, J. J., Freeman, B., & Tenenbaum, J. (2015). Galileo: Perceiving physical object properties by integrating a physics engine with deep learning. In *Advances in Neural Information Processing Systems* (pp. 127–135).

Yao, R., Wood, K., & Simons, D. J. (2019). As if by magic: An abrupt change in motion direction induces change blindness. *Psychological Science*, *30*, 436–443.