Melting ice with your mind: Representational momentum for physical states

Alon Hafri[a,1,*], Tal Boger[a,b,1], Chaz Firestone[a,*]

[a] *Johns Hopkins University, Baltimore, MD 21218*
[b] *Yale University, New Haven, CT 06520*

Abstract

When a log burns, it transforms from a block of wood into a pile of ash. Such state-changes are among the most dramatic ways objects change, going beyond mere changes of position or orientation. How does the mind represent changes of state? A foundational result in visual cognition is that memory extrapolates the positions of moving objects—a distortion called "representational momentum." Here, five experiments (N=400 adults) exploited this phenomenon to investigate mental representations in "state-space." Participants who viewed objects undergoing state-changes—e.g., ice melting, logs burning, or grapes shriveling—remembered them as more changed (e.g., more melted, burned, or shriveled) than they actually were. This pattern extended to several types of state-changes, went beyond their low-level properties, and even adhered to their natural trajectories in state-space. Thus, mental representations of a dynamic world actively incorporate change, in surprisingly broad ways: Whether in position or state, memory extrapolates how objects change.

Keywords: State-Changes, Intuitive Physics, Event Cognition, Memory Distortion, Visual Memory

---

[*] To whom correspondence should be addressed.

Email addresses: `alon@jhu.edu` (Alon Hafri), `tal.boger@yale.edu` (Tal Boger), `chaz@jhu.edu` (Chaz Firestone)

[1] Equal contribution.

In press, Psychological Science *August 27, 2021*

Introduction

The world is dynamic, not static: Objects *change*, challenging the mind to represent both their stability (as persisting individuals over time) and dynamicity (as entities whose appearance may shift from one moment to the next). In solving this challenge, the mind not only encodes an object's present appearance, but also predicts its future. For example, when playing catch, we combine our knowledge of the ball's current location with our prediction of where it will go next (Fink et al., 2009; Hecht and Bertamini, 2000). A foundational result in visual cognition demonstrates that this "forward momentum" is so ingrained in object representation that it distorts memory for changing objects: People misremember objects as displaced "forward in time" along their trajectories, a phenomenon known as representational momentum (Freyd, 1983; Freyd and Finke, 1984; Hubbard, 2005).

However, objects move not only in physical space, but also in "state-space": ice melts, logs burn, grapes shrivel, and so on. Such transformations represent a fundamentally distinct category of change (Aristotle, 1984, Physics, Book III), differing dramatically from changes in location or orientation. For example, when a ball moves, most of its features remain constant; the relevant change is simply its relation to its external environment. By contrast, state-changes are characterized by a complete transformation of an object's internal and external properties: When a log burns or an ice cube melts, its shape, texture, color, and many other essential qualities often change drastically, such that the object's final state may barely resemble its initial state. Furthermore, such changes are not uniform transformations of a single image property (e.g., color, size): state-changes look different depending on the type, such as melting, burning, or shriveling.

How does the mind represent changes of physical state? It has long been known that such changes organize mental representations in a variety of domains, including semantic memory, language, and cognitive development (Croft, 2015; Gropen et al., 1991; Hindy et al., 2015; Jackendoff, 1990; Lakusta and Landau, 2005; Levin, 1993; Muentener and Carey, 2010; Sakarias and Flecken, 2019; Solomon et al., 2015; Talmy, 2000; Vendler, 1957); indeed, it has recently been suggested that these and other changes serve a foundational role in event representations more generally (Altmann and Ekves, 2019). For example, eight-month-old infants show sophisticated knowledge of state-changes and the kinds of agents likely to cause them (Muentener and Carey, 2010). State-changes also shape linguistic representations, including the syntactic structures that verbs can take and the meanings such structures convey. For example, in English, many state-change verbs—e.g., "melt" or "deform"—participate in causative alternation structures (such that one can transform a sentence like *I* *melted the ice* into *The ice melted*, while still describing the same event), but other types of verbs do not (e.g., verbs of communication, as in *I told the story* vs. the ungrammatical *The story told*; Jackendoff, 1990; Levin, 1993).

The present experiments: Melting ice in memory

Whereas it is increasingly understood how state-changes are represented in higher-level cognition, it remains unclear whether they reach down into more foundational processes of visual cognition and memory. On one hand, previous work has speculated that they might (Finke et al., 1986; Freyd, 1987; Hubbard, 2017b, 2015a,b); for example, Finke et al. (1986) suggested that the mind might extrapolate *any* transformation forward in time. On the other hand, it is possible that the variation and complexities of physical state-changes might lead them to recruit different cognitive processes from other dynamic changes. Here, we explore these possibilities empirically by asking whether mental representations of state-changes share a behavioral profile with other dynamic changes.

To address this question, we tested whether state-changes exhibit representational momentum, such that memory extrapolates the future appearance of objects undergoing changes of state (Figure 1). We created physically realistic animations of familiar objects undergoing state-changes—ice melting, grapes shriveling, logs burning, and so on—and played them to participants before stopping the animations at a given frame. We predicted that participants would represent such changes dynamically, and thus that the last frame they remembered seeing would be "forward in time" relative to the one they actually saw. In other words, we predicted that the mind might proactively melt, shrivel, and burn the objects it encounters, incorporating such extrapolation into memory itself.

Experiment 1 explored representational momentum for state-changes in the way just described. Experiment 2 asked whether such representations are flexible, by contrasting forward-playing animations with backward-playing ones. Experiment 3 asked whether the mind represents state-changes dynamically even without dynamic input, by using static images. Finally, Experiments 4a and 4b replicated the previous results with a forced-choice response method. Demos of these experiments can be viewed at `https://perceptionresearch.org/dynamicstates`.

Experiment 1: Representational Momentum in State-Space

Does memory extrapolate the changing states of objects? Experiment 1 showed participants animations of different objects undergoing changes of state (e.g., ice

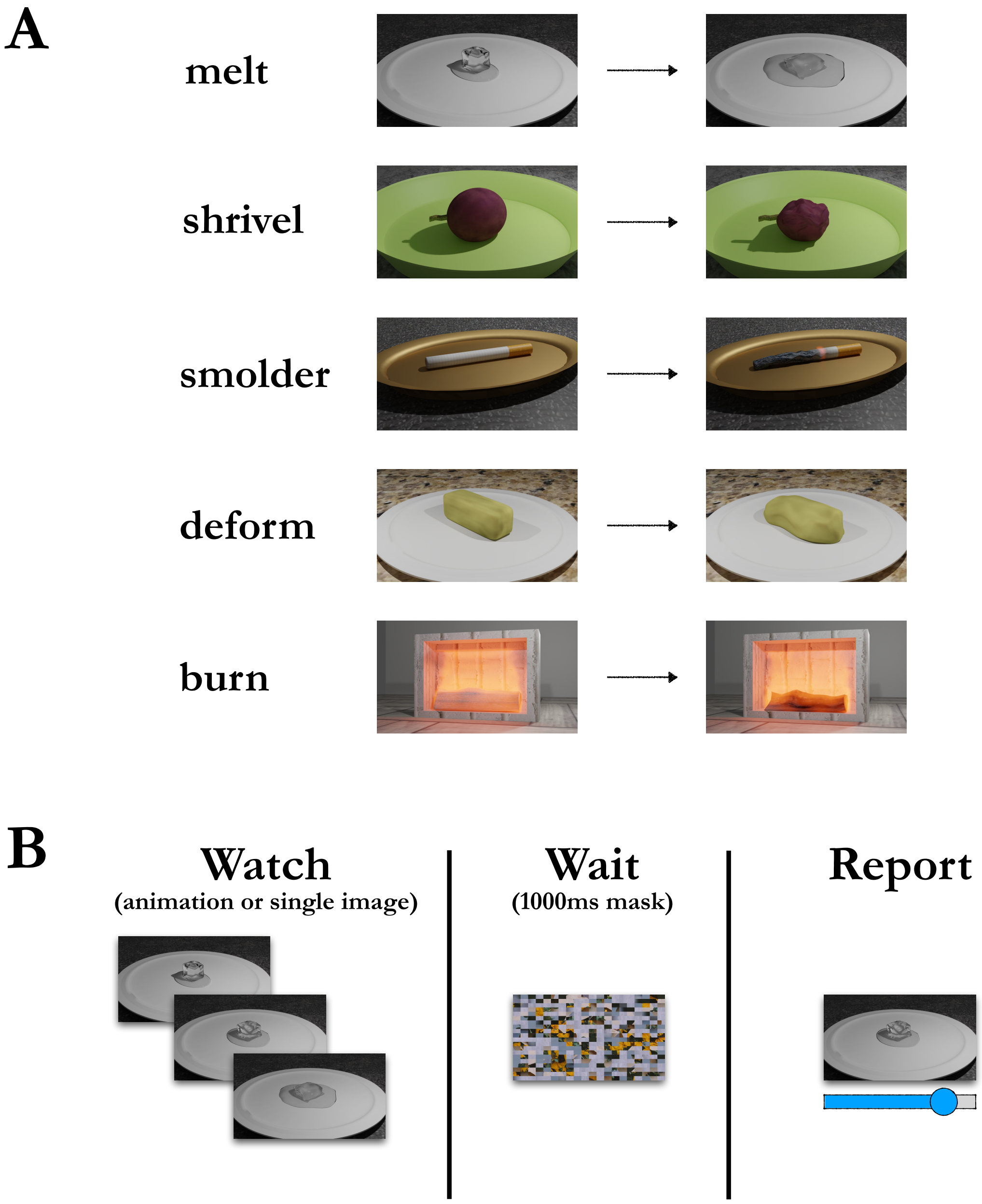

Figure 1: Design of Experiments 1–3. **(A)** We explored several state-changes, each involving very different image properties. **(B)** Participants saw animations (Experiments 1 and 2) or single static images (Experiments 3) of state-changes; the animation or static image was masked, and participants identified the final frame they had seen using a slider that advanced through all frames of the animation.

melting, log burning) and asked them to identify the last frame they saw before the animation was stopped.

Methods

Open Science Practices

All data, code, analyses, stimuli, and pre-registrations (for this experiment and all others reported here) are available at `https://perceptionresearch.org/dynamicstates`. This webpage also includes demos of each experiment, so that readers can experience these tasks as participants did. The sample sizes and analysis plans (as well as other details) for all experiments were pre-registered.

Participants

50 adult participants were recruited from the online platform Prolific. (For a discussion of the reliability of this and other online subject pools, see Peer et al., 2017). This was chosen to be as large or larger a sample size in comparison to previous visual cognition studies of this sort (typically $n < 40$; e.g., De Freitas et al., 2016; Freyd and Finke, 1984; Johnston and Jones, 2006; Thornton, 2014). Sample sizes were pre-registered for this and all other experiments. All studies were approved by the [university name withheld] Institutional Review Board.

Stimuli and Procedure

To depict physical state-changes while retaining full control of timing and other visual factors, we simulated and rendered state-changes under realistic physics using Blender v2.82 (`https://www.blender.org`). We created five different state-change stimuli, each involving very different objects and physical changes: melting, shriveling, smoldering, deforming, and burning (Figure 1A). Each animation lasted 240 frames and was presented at 30 fps (eight seconds total). Note that, while some of the state-changes we explore here correspond to transitions between physical states of matter (e.g., a melting ice cube, which transforms from a solid to a liquid), other changes involve chemical reactions (e.g., combustion) or other physical processes such as osmosis (e.g., shriveling). For present purposes, we consider all such processes to fall under the umbrella term “state-changes,” though future work could further explore distinctions between these types of change.

All stimuli were 704×396 pixels in the participant’s Web browser. Due to the nature of online studies, we cannot know the exact viewing distance, screen size, luminance (etc.) of these stimuli as they appeared to participants. However, any distortions introduced by a given participant’s viewing distance or monitor settings would have been equated across all stimuli and conditions for that participant.

On each trial of the study (Figure 1B), participants viewed an animation of one of the state-changes, which was stopped before completion and then masked for 1000ms with a box-scrambled mask (20×20 blocks, randomly selected from 7 possible masks of natural scenes). Following this, participants' task was simply to identify the last frame of the animation that they saw before it was stopped. Participants controlled a slider that stepped through the animation frame-by-frame, such that the participant could move the slider to select the target frame. (The starting position of the slider was randomized on every trial.) The left end of the slider represented the beginning of the animation, and the right end was the end of the animation. When satisfied that the image on the screen matched the final frame they had seen earlier in the trial, participants clicked a button to move on to the next trial.

To ensure that the task was clear, participants first completed an "easy" trial during the instruction phase in which they had to reproduce the exact frame at which an animation was stopped. The target frame to reproduce remained on screen throughout this practice trial (so that the correct answer was clear); participants could not proceed with the study until they performed this trial as instructed.

There were three blocks of experimental trials, each containing the five state-changes in a random order (15 trials total). Each animation was stopped either 25%, 50%, or 75% before completion (randomized order, once for each state-change), and was then masked immediately after it was stopped such that the animation did not proceed further. The full animations (and Blender code to render them) are available on OSF (`https://osf.io/gz9a3`); demos of this experiment and the others reported in this paper can be viewed at `https://perceptionresearch.org/dynamicstates`.

For each trial, we calculated the "frame error": the signed difference between the frame chosen by participants and the actual target frame. For example, if the last frame that appeared was frame 180, a response of 188 would be a frame error of +8. We predicted that participants would misremember the last frame they saw as farther forward in time than it actually was, and report it as such, resulting in a positive frame error.

Results

In accordance with our pre-registered analysis plan, we excluded participants if they did not contribute a complete dataset, or if their mean slider responses (averaged across state-changes) were not lower for earlier target frames and higher for later target frames. (We reasoned that participants not giving lower frame responses for earlier target frames were likely not performing the required task.) There were 43 participants after these exclusions.

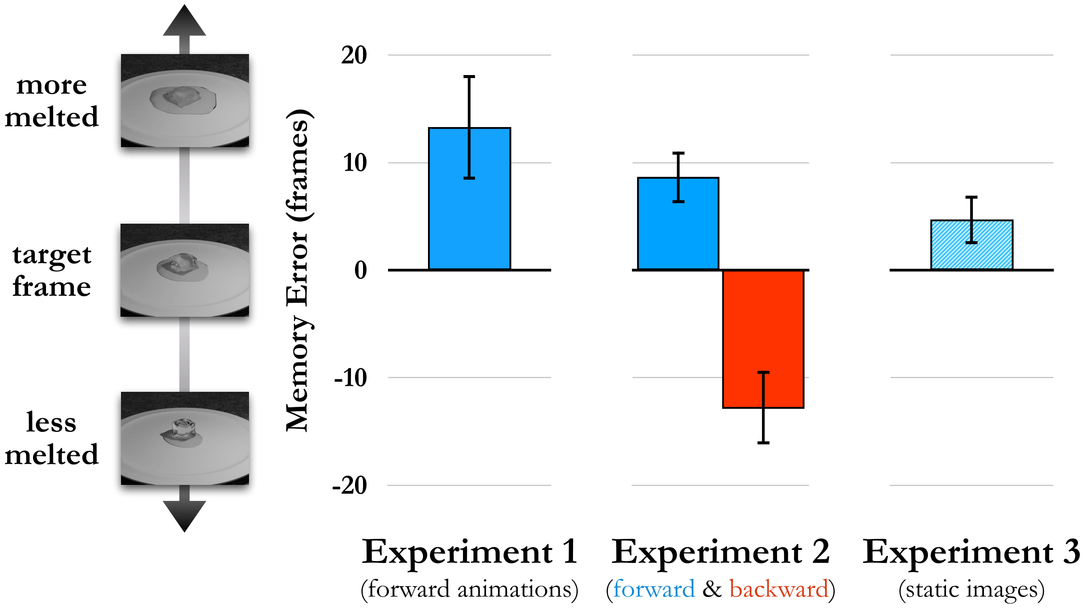

Figure 2: Results for Experiments 1–3. In Experiment 1, participants saw forward animations and reported the final frame as "later" (e.g., more melted) than it actually appeared. In Experiment 2, participants saw forward and backward animations and showed representational momentum in the direction of the animations, in both cases (i.e., forward animations were remembered as more melted, while backward animations were remembered as more "unmelted"). In Experiment 3, participants saw only a single static frame, and showed a bias in the "forward" or physically natural direction of change. (Error bars are ±95% confidence intervals; all bars differ from 0 with $p < .001$.)

As predicted, we observed a significant positive frame error, such that participants reported a frame "forward" in time relative to the true final frame (M=13.25 frames out of 240, or 442ms out of 8s of the animation; $t(42) = 5.34$, $p < .001$; $d = 0.81$; 95% CI = [8.24, 18.26]; Figure 2). In other words, participants reported the ice as more melted than it really appeared. This pattern occurred for every type of state-change shown (melting: $M = 21.99$, $t(42) = 6.08$, $p < .001$, $d = 0.93$, 95% CI = [14.69, 29.29]; shriveling: $M = 7.63$, $t(42) = 2.19$, $p = 0.034$, $d = 0.33$, 95% CI = [0.59, 14.66]; smoldering: $M = 8.96$, $t(42) = 4.41$, $p < .001$, $d = 0.67$, 95% CI = [4.86, 13.07]; deforming: $M = 7.67$, $t(42) = 2.00$, $p = .052$, $d = 0.31$, 95% CI = [-0.06, 15.40]; burning: $M = 20.00$, $t(42) = 5.18$, $p < .001$, $d = 0.79$, 95% CI = [12.21, 27.80]).

Moreover, the results were not driven by a mere tendency to respond toward the slider's center: Although frame error was highest for animations that stopped earlier ($M = 22.69$, $t(42) = 5.60$, $p < .001$; $d = 0.85$; 95% CI = [14.52, 30.86]), we still found positive frame errors for animations stopped halfway through ($M = 13.00$, $t(42) = 4.34$, $p < .001$; $d = 0.66$; 95% CI = [6.95, 19.05]), and we even observed a positive trend for animations that stopped at frames corresponding to the "later" end of the slider ($M = 4.07$, $t(42) = 1.92$, $p = 0.061$; $d = 0.29$; 95% CI = [-0.20, 8.33]), where a tendency to respond towards the center of the slider should have favored the *opposite* effect (stacking the deck against our prediction). These results suggest that the mind extrapolates state-changes beyond what is actually observed: representational momentum for state-changes.

Experiment 2: Flexibility of Extrapolation

Some changes of state are "irreversible": An ice cube can melt into a puddle, but a puddle can't "unmelt" into an ice cube. (The best it can do, perhaps, is freeze in place.) Does the mind flexibly extrapolate state-changes along directions we have rarely (if ever) encountered (i.e., not only melting, but also "unmelting")? Experiment 2 tested this by including trials where state-change animations played in reverse; in such backward animations, participants saw a puddle "unmelt" into an ice cube.

Methods

50 new participants were recruited for Experiment 2, which was identical to Experiment 1 except for the addition of three blocks of experimental trials in which the animations played in reverse (with order of forward/backward sets counterbalanced across participants). We also counterbalanced slider direction (left-earlier/right-later

or vice-versa) across participants to control for possible directional biases in using the slider.

Results

In accordance with our pre-registered analysis plan, we excluded participants if they did not contribute a complete dataset, or if their mean slider responses (averaged across state-changes) were not lower for earlier target frames and higher for later target frames. This left 48 participants.

We again observed positive frame errors for forward animations ($M = 8.61$ frames, $t(47) = 4.95$, $p < .001$; $d = 0.71$; 95% CI = [5.11, 12.12]; Figure 2), consistent with the results of Experiment 1. Intriguingly, backward animations also showed frame errors along the direction of the animation; these were negative frame errors, as they were in a direction opposite to the physically natural direction depicted in forward animations ($M = -12.81$ frames, $t(47) = -7.34$, $p < .001$; $d = -1.07$; 95% CI = $[-9.32, -16.30]$). In other words, when shown an animation of ice "unmelting" (a backward animation), participants remembered the ice as more "unmelted" than it really was. For backward-playing animations, this provides evidence for representational momentum in the same way that positive frame errors in forward-playing animations provide evidence for representational momentum. (Although a trend for greater frame errors in the direction of animation was observed for backward compared to forward animations ($t(47) = 1.76$, $p = .085$; $d = 0.25$; 95% CI = $[-0.60, 9.00]$), results from Experiments 3, 4a, and 4b reported below strongly suggest that the default direction of change in the mind is "forward," i.e., the physically natural direction of change.)

Beyond demonstrating flexible representation of state-changes, these results also suggest that our earlier findings weren't driven by mere familiarity with a given pattern of physical change. If the memory biases observed in Experiment 1 were simply driven by prior experience seeing ice cubes melt and logs burn (etc.), one would not have expected the same effects to arise for unmelting and unburning. So, the fact that similar effects *do* arise for unmelting and unburning suggests that the effects go beyond simply recreating events one has seen before, and instead involves actively representing and extrapolating state-changes *as they occur*.

Experiment 3: Static Images

We have suggested that the present effects arise because the mind represents state-changes per se. But our previous results might be explained by a lower-level mechanism. In particular, our dynamic animations necessarily included not only high-level information about changing states, but also lower-level visual changes that are inevitably correlated with those state-changes (e.g., optic flow or motion-energy). In that case, the effects might not have been driven by participants running forward the state-changes themselves (i.e., mentally melting the ice), but rather by ordinary representational momentum for the *motion* present in the animations—e.g., the expansion of the puddle formed by the melting ice.

Experiment 3 addressed this possibility by asking whether a single static frame can elicit representational momentum in state-space, as has been previously shown for location memory (e.g., Bertamini, 1993; Finke et al., 1986; Freyd, 1987). This design not only ruled out effects of low-level motion, but also allowed us to investigate whether the mind privileges one direction over the other (e.g., representing the physically natural "forward" direction—melting, rather than unmelting—by default).

Methods

Participants

100 new participants were recruited. This sample size was larger than in the previous two experiments because we expected the representational momentum effects to be more subtle for static than dynamic stimuli.

Stimuli and Procedure

The stimuli, task, and conditions of Experiment 3 matched Experiment 1, except that participants viewed a single static frame (for 1000ms) instead of dynamic animations. Moreover, we included three 50%-frame trials per state-change rather than just one, and only analyzed those trials (and pre-registered this analysis), as we expected that a tendency to respond towards the slider's center—which would result in biased results at non-50% frames—might obscure the more subtle representational momentum effects we anticipated for static images. The 25%-frame and 75%-frame trials (one of each per state-change) were included in the experiment (but were not analyzed) to decrease the possibility that participants would realize the frames of interest were always at exactly 50%. Thus, there were five blocks of stimuli, each containing the five state-changes in a random order (25 trials total). For each state-change, the order of the target frame image (25%, 75%, or the three 50% frames) was randomized. As in Experiment 2, we also counterbalanced slider direction (left-earlier/right-later or vice-versa) across participants.

Results

In accordance with our pre-registered analysis plan, we excluded participants if they did not contribute a complete dataset, or if their mean slider responses (averaged across state-changes) were not lower for earlier target frames and higher for later target frames. This left 94 participants.

We again observed a positive frame error: Even when shown only a single static image from the "middle" of the state-changes, participants misremembered them in their physically natural direction ($M = 4.64$ frames, $t(93) = 4.06$, $p < .001$; $d = 0.42$; 95% CI = [2.37, 6.92]; Figure 2).[2] Thus, (a) representational momentum arises in state-space even without any lower-level dynamic cues to indicate a direction of change, and (b) the extrapolated direction in state-space is *forward* by default, suggesting that this process incorporates physically natural constraints on such changes.

Experiments 4a and 4b: Forced Choice

The previous experiments suggested that the mind extrapolates state-changes forward, even without dynamic input suggesting such changes. However, by using a slider as the response modality, these experiments may have allowed participants to "play" the animation forward, such that the "momentum" we observed may have had nothing to do with a memory distortion in state-space but rather with the actual responses they gave. (Indeed, on this alternative account, the effect could literally be due to the physical momentum of their hands moving a mouse!) In a final set of experiments, we replicated Experiment 3 using a forced-choice paradigm to rule out even this alternative.

Methods

The design of Experiments 4a and 4b was similar to Experiment 3, in that participants observed a single static frame on each trial. However, unlike Experiment 3, after the mask appeared, participants were shown two possible frames (rather than a slider), and they were instructed to choose the frame that matched the target frame that they observed earlier in the trial (Figure 3A). In fact, neither frame was correct (though participants were not informed of this): one was always earlier than the true target frame, and the other was always later than the true target frame (by the same magnitude in each direction). We asked whether extrapolation would still be observed here, despite the difference in probing method.

---

[2] Though we only report our pre-registered analysis of the 50%-frame trials here, we present analyses for all frames (for this and all experiments) in the Supplemental Material. All such analyses are consistent with the effects we report in the manuscript—i.e., representational momentum for state-changes—both with and without including data from all trials.

Participants

Two groups of 100 participants each were recruited from Prolific for both Experiment 4a and Experiment 4b (i.e., 200 participants total). We chose sample sizes of 100 in both experiments to match the 100 used in Experiment 3, given that both experiments contained static stimuli instead of dynamic stimuli.

Stimuli and Procedure

In contrast with the slider-based response method of Experiments 1–3, the method of probing memory here was a two-alternative forced-choice task. The two options were either earlier or later than the target frame (by 30 frames in each direction, determined via pilot testing).

To ensure that the change in response method (from slider to forced-choice) was the only difference between Experiment 4a and Experiment 3, Experiment 4a kept the same design as Experiment 3, including the sampling and analysis of frames. Participants viewed four 50%-frame trials per state-change rather than just three. As in Experiment 3, we only analyzed those 50%-frame trials (and pre-registered this analysis); the 25%-frame and 75%-frame trials (one of each per state-change) were included in the experiment to decrease the possibility that participants would realize the frames of interest were always at exactly 50%. There were two "epochs" in the study, each containing three blocks (with each block containing the five state-changes in random order). Each epoch contained two 50% trials for each state-change; the third trial for each state-change in the epoch, either 25% or 75%, was randomly assigned (e.g., for melting, the 25% trial may have appeared in the first epoch and the 75% trial in the second; for smoldering, the 75% trial may have appeared in the first epoch and the 25% trial in the second). Thus, there were 30 trials in total. Position of the later frame, left or right, was counterbalanced for each state-change and epoch (i.e., each state-change and epoch had half its trials with the later image on the left).

Experiment 4b differed more substantially: In addition to using the forced-choice procedure described above, it also eliminated any selective sampling in both the experimental design and the analyses, in order to ensure that the effects were not particular to potential idiosyncrasies of the single 50% frame analyzed. In this experiment, we sampled from a uniform distribution of frames for every participant and state-change: On any given trial, participants saw a frame chosen randomly from the full range of possible frames for a state-change. Then, as in Experiment 4a, they had to choose between two frames that were offset by 30 frames in each direction from the true target frame. As in Experiment 4a, there were 6 blocks, each containing one static frame of each state-change in a random order. The target frames shown for each state-change were sampled from a uniform distribution between frames 31–209; crucially, this ensured that the +30 and −30 offsets for probe frames would stay within the bounds of the 240 total frames for each state-change, and thus that it would be possible to choose either the earlier or later option even at the extremes. Frames were sampled such that the mean frame shown for each state-change was 120 (or 50% through the state-change). This average of 120 was accomplished by choosing 3 frames randomly for each state-change, and then setting the remaining 3 frames to be 240 minus the initial 3 frames chosen. For example, if the frames chosen for ice melting were 36, 97, and 170, then frames 204, 143, and 70 were also included, which together average to 120. The order of these frames was randomized within-block.

We expected that, when forced to choose between an earlier and later frame, participants would choose the later frame more often than the earlier frame.

Results

In accordance with our pre-registered analysis plan, we excluded trials with a response time that was considered too fast (< 400ms). (This was conservative, as it only excludes trials in which it was unlikely that the participant could have fully registered the images and planned their response.) We also excluded participants if they did not contribute a complete dataset, or if more than 10% of their trials were excluded for being too fast, reasoning that participants with too many fast responses were likely not performing the required task. This left 99 participants in each of Experiments 4a and 4b. Considering these remaining participants, 0.70% of trials were excluded for being too fast in Experiment 4a, and 0.17% of trials were excluded in Experiment 4b.

We once again observed evidence that memory for objects changing state is extrapolated forward in time. Even when shown only a single static image from the state-change events in Experiment 4a, participants misremembered them in their physically natural direction, more often selecting the later probe frame than the earlier probe frame ($M = 58.86\%$ of trials on which the later probe frame was selected, $t(98) = 7.44$, $p < .001$; $d = 0.75$; 95% CI = [56.50%, 61.23%]; Figure 3B). Furthermore, this was not just a result of seeing the "middle" of the state-changes; in Experiment 4b, where participants saw frames that were chosen uniformly across the entire range of state-change frames, they again selected the later probe frame more often than the earlier probe frame ($M = 57.37\%$ of trials on which the later probe frame was selected, $t(98) = 7.41$, $p < .001$; $d = 0.74$; 95% CI = [55.39%, 59.34%]; Figure 3B). Whereas the results of Experiments 1–3 may have been explained by the natural biases of the slider, this possibility cannot explain the results in the

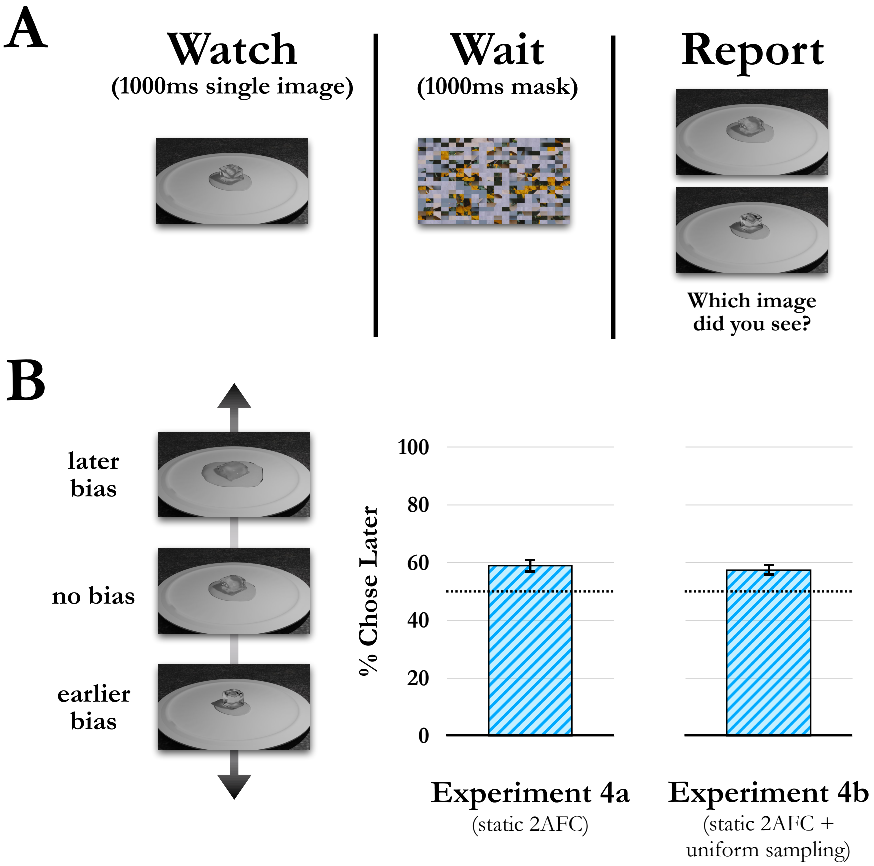

Figure 3: Design and results of Experiments 4a and 4b. **(A)** Participants saw single static images of state-changes; the target image was masked, and participants were tasked with selecting the target image from between two options. Participants were not informed that neither frame was correct: one was earlier than the target frame and one was later. **(B)** In both studies, participants consistently selected "later" frames, showing a bias in the forward or physically natural direction of change. In Experiment 4a, results are from the 50%-frame (the "middle" of the state-changes, as in Experiment 3), while in Experiment 4b, results reflect uniform sampling across the entire range of frames. (Error bars are ±95% confidence intervals; both bars differ from 50% with $p < .001$.)

current experiment, where participants were forced to choose between two discrete options. Thus, even with a different response method, participants demonstrated representational momentum for state-changes.

General Discussion

The present experiments suggest that state-change representations share a behavioral profile with more traditionally studied dynamic event representations, in that memory distorts such changes forward in time. The dynamic nature of object representation is thus surprisingly general: Our minds represent not only where an object is likely to have moved, but also how an object is likely to have *transformed*.

Dynamic Distortions

Importantly, the memory distortions observed here go beyond merely *predicting* the future states of changing objects. It is not so surprising that one can predict how a melting ice cube will look at some later time, just as one can predict the future appearance of all sorts of objects and events. What is distinctive about the present results, however, is that participants actively mistook a later stage of these state-changes for what they actually observed. Thus, even if the representational momentum effects reported here were driven by predictions of some sort (Hubbard, 2019), they go beyond simply *making* those predictions and instead intrude upon more foundational processes of memory itself. In other words, these effects are a case of inferences *causing* memory distortions for state-changes, in a manner similar to memory distortions for physical locations (Freyd, 1987; Hubbard, 2006).

Moreover, the existence of representational momentum for state-changes was not a foregone conclusion. On one hand, representational momentum is clearly established for location (e.g., Freyd, 1983; Freyd and Finke, 1984; for a review, see Hubbard, 2005), and it has been extended to other properties such as pitch (Freyd et al., 1990; Johnston and Jones, 2006), action (Chatterjee et al., 1996; Hudson et al., 2016; Verfaillie and Daems, 2002), and even social position (Kakkar et al., 2019). On the other hand, it has *not* been conclusively demonstrated for other continuous properties, such as luminance (Brehaut and Tipper, 1996), hue (Callahan-Flintoft et al., 2020), and emotional expression (Thornton, 2014). (Indeed, the lack of forward momentum in these cases is another reason that the present effects go beyond mere "prediction," since it is quite easy to predict the future luminance value of an object that is smoothly increasing in brightness.) Thus, not only is representational momentum for state-changes a genuinely new discovery about how such changes are represented, but its existence supports theories holding that the nature of dynamic representation is quite general (Finke et al., 1986; Freyd, 1987; Hubbard, 2015a,b, 2017a,b).

Intuitive Reasoning about Physical States

The present results add to a growing literature on intuitive physical reasoning. Recent work reveals that the mind represents future arrangements of physical scenes, as if pressing "play" on a simulation of that scene (e.g., block-towers; Battaglia et al., 2013; Fischer et al., 2016; Firestone and Scholl, 2016; Kubricht et al., 2017; Ullman et al., 2017; also Guan and Firestone, 2020). Our findings go beyond these sorts of results in at least two ways. First, they suggest that such intuitive physical reasoning can operate not only over the arrangement and movement of objects, but also over their physical composition. And second, they suggest that the cognitive mechanisms underlying such intuitions not only support higher-level *inferences* about how physical scenes will unfold, but also actively distort memory for them.

Future work could explore whether state-changes in physical reasoning are represented in ways that are less reflective of the continuous nature of real-world changes and more similar to how state-changes are represented in other domains (such as language), where a core distinction is made between gradual, process-based changes of the kind explored here (e.g., *the balloon expanded*) and "instantaneous" transitions (e.g., *the balloon exploded*; Croft, 2015; Vendler, 1957). For example, even if one watches a slowed-down video of an exploding balloon to see the pieces scatter, the mind may still treat this state-change as categorical and instantaneous. Perhaps there are even "attractors" in state-space, much like those that have been established in physical space (e.g., cardinal biases; Huttenlocher et al., 1991; Newcombe and Huttenlocher, 2000; Palmer, 1980; Tversky, 1981). Indeed, prior work suggests that certain locations in state-space are particularly salient in the mind (Croft, 2015; Lakusta and Landau, 2005; Sakarias and Flecken, 2019). In that case, one might imagine that nearly-melted ice gets treated as fully melted by the mind, or that an ice cube that has only barely begun melting may get mentally reverted to an unmelted ice cube. Some exploratory analyses from Experiment 4 are in line with this intriguing possibility (see Supplemental Material), which may be investigated in future work.

General Implications and Open Questions

The implications of these results may go beyond new findings about state-changes or memory distortions, by interacting with more general theories of event perception and memory (e.g., Event Segmentation Theory [Zacks et al., 2007] or the Theory of Event Coding [Hommel et al., 2001]; see also Kim et al., 1995; for a review, see Zacks, 2020). One uniting factor of such proposals is that the mind represents or detects cognitively salient aspects of the event at hand. Our results add to this literature by suggesting that surprisingly complex state-changes—including fundamental changes to material or matter—are not only incorporated into higher-level reasoning about events that we have experienced (or otherwise represented) but also play an active role in on-line event representation.

An open question concerns the generality of such state-change representations in the mind. We found that the directionality of state-changes is quite flexible, even for directions rarely encountered (e.g., "unmelting" ice); but a related question is whether state-change representations are constrained by the kinds of objects that usually undergo such changes. For example, grapes shrivel differently from other fruits, and ice doesn't normally shrivel at all; yet state-changes like shriveling are in principle quite general, applicable to many types of objects (just like affine changes such as rotation and translation; Schmidt et al., 2016; Ward et al., 2018). Perhaps representational momentum for state-changes would also generalize to rarely observed object/state-change associations—e.g., shriveling butter, or smoldering ice.

Concluding Remarks

The discovery of representational momentum for state-changes complements related work in domains such as cognitive development, semantic memory, and linguistics (Altmann and Ekves, 2019; Hindy et al., 2015; Jackendoff, 1990; Lakusta and Landau, 2005; Levin, 1993; Muentener and Carey, 2010), extending this research for the first time into the domain of visual cognition and memory. By demonstrating that the mind dynamically represents the physical changes of objects—and even incorporates their probable future states into memory—we show that state-changes not only organize how we think and speak about the world but also constrain how we remember it in the first place.

Acknowledgments

For helpful discussion and/or comments on previous drafts, the authors thank John Trueswell, and members of the JHU Perception & Mind Laboratory. This work was supported by NSF BCS-2021053 awarded to C.F. and NSF SBE Postdoctoral Research Fellowship SMA-2021053 awarded to A.H.

References

Altmann, G. T. M. and Ekves, Z. (2019). Events as intersecting object histories: A new theory of event representation. *Psychological Review*, 126(6):817–840.

Aristotle (1984). In Barnes, J., editor, *The Complete Works of Aristotle, Volumes I and II*. Princeton University Press, Princeton.

Battaglia, P. W., Hamrick, J. B., and Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, 110(45):18327–18332.

Bertamini, M. (1993). Memory for position and dynamic representations. *Memory & Cognition*, 21(4):449–457.

Brehaut, J. C. and Tipper, S. P. (1996). Representational momentum and memory for luminance. *Journal of Experimental Psychology: Human Perception and Performance*, 22(2):480–501.

Callahan-Flintoft, C., Holcombe, A. O., and Wyble, B. (2020). A delay in sampling information from temporally autocorrelated visual stimuli. *Nature Communications*, 11(1):1–11.

Chatterjee, S. H., Freyd, J. J., and Shiffrar, M. (1996). Configural Processing in the Perception of Apparent Biological Motion. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4):916–929.

Croft, W. (2015). Force Dynamics and Directed Change in Event Lexicalization and Argument Realization. In de Almeida, R. G. and Manouilidou, C., editors, *Cognitive Science Perspectives on Verb Representation and Processing*, pages 103–129. Springer.

De Freitas, J., Myers, N. E., and Nobre, A. C. (2016). Tracking the changing feature of a moving object. *Journal of Vision*, 16(3):22.

Fink, P. W., Foo, P. S., and Warren, W. H. (2009). Catching fly balls in virtual reality: A critical test of the outfielder problem. *Journal of Vision*, 9(13):14–14.

Finke, R. A., Freyd, J. J., and Shyi, G. C. (1986). Implied velocity and acceleration induce transformations of visual memory. *Journal of Experimental Psychology: General*, 115(2):175–188.

Firestone, C. and Scholl, B. (2016). Seeing stability: Intuitive physics automatically guides selective attention. *Journal of Vision*, 16(12):Article 689.

Fischer, J., Mikhael, J. G., Tenenbaum, J. B., and Kanwisher, N. (2016). Functional neuroanatomy of intuitive physical inference. *Proceedings of the National Academy of Sciences*, 113(34):E5072–E5081.

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception & Psychophysics*, 33(6):575–581.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review*, 94(4):427–438.

Freyd, J. J., DeKay, M. L., and Kelly, M. H. (1990). Representational momentum in memory for pitch. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(6):1107–1117.

Freyd, J. J. and Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(1):126–132.

Gropen, J., Pinker, S., Hollander, M., and Goldberg, R. (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition*, 41(1-3):153–195.

Guan, C. and Firestone, C. (2020). Seeing what's possible: Disconnected visual parts are confused for their potential wholes. *Journal of Experimental Psychology: General*, 149(3):590–598.

Hecht, H. and Bertamini, M. (2000). Understanding projectile acceleration. *Journal of Experimental Psychology: Human Perception and Performance*, 26(2):730–746.

Hindy, N. C., Solomon, S. H., Altmann, G. T., and Thompson-Schill, S. L. (2015). A cortical network for the encoding of object change. *Cerebral Cortex*, 25(4):884–894.

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24(5):849–878.

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review*, 12(5):822–851.

Hubbard, T. L. (2006). Bridging the gap: Possible roles and contributions of representational momentum. *Psicológica*, pages 1–34.

Hubbard, T. L. (2015a). Forms of momentum across time: Behavioral and psychological. *The Journal of Mind and Behavior*, pages 47–82.

Hubbard, T. L. (2015b). The varieties of momentum-like experience. *Psychological Bulletin*, 141(6):1081–1119.

Hubbard, T. L. (2017a). Momentum in music: Musical succession as physical motion. *Psychomusicology: Music, Mind, and Brain*, 27(1):14–30.

Hubbard, T. L. (2017b). Toward a general theory of momentum-like effects. *Behavioural Processes*, 141:50–66.

Hubbard, T. L. (2019). Momentum-like effects and the dynamics of perception, cognition, and action. *Attention, Perception, & Psychophysics*, 81:2155–2170.

Hudson, M., Nicholson, T., Ellis, R., and Bach, P. (2016). I see what you say: Prior knowledge of other's goals automatically biases the perception of their actions. *Cognition*, 146:245–250.

Huttenlocher, J., Hedges, L. V., and Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological review*, 98(3):352–376.

Jackendoff, R. (1990). *Semantic structures.* MIT Press, Cambridge, MA.

Johnston, H. M. and Jones, M. R. (2006). Higher order pattern structure influences auditory representational momentum. *Journal of Experimental Psychology: Human Perception and Performance*, 32(1):2–17.

Kakkar, H., Sivanathan, N., and Pettit, N. C. (2019). The impact of dynamic status changes within competitive rank-ordered hierarchies. *Proceedings of the National Academy of Sciences*, 116(46):23011–23020.

Kim, N.-G., Effken, J. A., and Shaw, R. E. (1995). Perceiving persistence under change and over structure. *Ecological Psychology*, 7(3):217–256.

Kubricht, J. R., Holyoak, K. J., and Lu, H. (2017). Intuitive physics: Current research and controversies. *Trends in Cognitive Sciences*, 21(10):749–759.

Lakusta, L. and Landau, B. (2005). Starting at the end: The importance of goals in spatial language. *Cognition*, 96(1):1–33.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation.* University of Chicago Press, Chicago.

Muentener, P. and Carey, S. (2010). Infants' causal representations of state change events. *Cognitive Psychology*, 61(2):63–86.

Newcombe, N. and Huttenlocher, J. (2000). *Making space: The development of spatial representation and reasoning.* MIT Press, Cambridge, MA.

Palmer, S. E. (1980). What makes triangles point: Local and global effects in configurations of ambiguous triangles. *Cognitive psychology*, 12(3):285–305.

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

Sakarias, M. and Flecken, M. (2019). Keeping the result in sight and mind: General cognitive principles and language-specific influences in the perception and memory of resultative events. *Cognitive Science*, 43(1):e12708.

Schmidt, F., Spröte, P., and Fleming, R. W. (2016). Perception of shape and space across rigid transformations. *Vision Research*, 126:318–329.

Solomon, S. H., Hindy, N. C., Altmann, G. T. M., and Thompson-Schill, S. L. (2015). Competition between mutually exclusive object states in event comprehension. *Journal of Cognitive Neuroscience*, 27(12):2324–2338.

Talmy, L. (2000). *Toward a Cognitive Semantics.* MIT Press, Cambridge, MA.

Thornton, I. M. (2014). Representational momentum and the human face: An empirical note. *XJENZA On line*, pages 101–110.

Tversky, B. (1981). Distortions in memory for maps. *Cognitive psychology*, 13(3):407–433.

Ullman, T. D., Spelke, E., Battaglia, P., and Tenenbaum, J. B. (2017). Mind games: Game engines as an architecture for intuitive physics. *Trends in Cognitive Sciences*, 21(9):649–665.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66:143–160.

Verfaillie, K. and Daems, A. (2002). Representing and anticipating human actions in vision. *Visual Cognition*, 9(1-2):217–232.

Ward, E. J., Isik, L., and Chun, M. M. (2018). General transformations of object representations in human visual cortex. *The Journal of Neuroscience*, 38(40):8526–8537.

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology*, 71:165–191.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., and Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological bulletin*, 133(2):273–293.