Retinotopic adaptation reveals distinct categories of causal perception

Jonathan F. Kominsky[a],[*], Brian J. Scholl[b]

[a] Harvard University, United States of America
[b] Yale University, United States of America

ABSTRACT

We can perceive not only low-level features of events such as color and motion, but also seemingly higher-level properties such as causality. A prototypical example of causal perception is the 'launching effect': one object (A) moves toward a stationary second object (B) until they are adjacent, at which point A stops and B starts moving in the same direction. Beyond these motions themselves — and regardless of any higher-level beliefs — this display induces a vivid visual impression of causality, wherein A is seen to cause B's motion. Do such percepts reflect a unitary category of visual processing, or might there be multiple distinct forms of causal perception? While launching is often simply equated with causal perception, researchers have sometimes described other phenomena such as 'triggering' (in which B moves faster than A) and 'entraining' (in which A continues to move alongside B). We used psychophysical methods to determine whether these labels really carve visual processing at its joints, and how putatively different forms of causal perception relate to each other. Previous research demonstrated retinotopically specific adaptation to causality: exposure to causal launching makes subsequent ambiguous events in that same location more likely to be seen as non-causal 'passing'. Here, after replicating this effect, we show that exposure to triggering also yields retinotopically specific adaptation for subsequent ambiguous launching displays, but that exposure to entraining does not. Collectively, these results reveal that visual processing distinguishes some (but not all) types of causal interactions.

1. Introduction

There is perhaps no concept more central to our understanding of the physical world than that of causality. As such, the study of cause and effect has been a central project in many disciplines, from physics (for a review, see Brukner, 2014) to philosophy (for a review, see Paul & Hall, 2013) to psychology (for a review, see Sloman & Lagnado, 2015). In human cognition, the notion of causality may be so central that it structures not only how we understand the world, but also how we see the world in the first place. Accordingly, it has been suggested that the currency of visual processing consists not only of seemingly low-level features (such as color and shape) but also various types of physical causality (for reviews, see Scholl & Tremoulet, 2000; White, 2017; cf. Rips, 2011).

The vast majority of work on causal perception has actually involved just one particular sort of causal event: the *launching* effect (Michotte, 1946/1963) — which occurs when one object (A) moves toward a second stationary object (B) until they are adjacent, at which point A stops and B begins moving in the same direction (and at the same speed, or moderately slower). Beyond these brute kinematics, this display induces a vivid impression of causality, wherein A is seen to cause B's motion. Indeed, in Michotte's introduction to such demonstrations, he points out that as long as certain constraints are met (such as B not moving too far, and not moving faster than A), such events are perceived as if every aspect of B's motion is fully determined by A's impact (something Michotte referred to as the *ampliation* of A's movement). This type of causal launching is depicted schematically in Fig. 1a. (Of course, it is always difficult to depict such intrinsically dynamic events in static figures. For this reason, dynamic animations of such figures — and of the key stimuli employed in the current studies — are available online at http://perception.yale.edu/causal-adaptation/.)

The perception of causality that arises when viewing a launching display has many features in common with other forms of visual processing, while contrasting in several key ways with higher-level causal judgment. On one hand, beyond arising early in infancy (e.g. Leslie & Keeble, 1987; Mascalzoni, Regolin, Vallortigara, & Simion, 2013; Newman, Choi, Wynn, & Scholl, 2008) and seeming to operate cross-culturally (Morris & Peng, 1994), such percepts seem exquisitely sensitive to otherwise-subtle spatiotemporal features of the relevant visual stimulation. For example, even moderate temporal or spatial gaps between A and B obliterate the perception of causality (Michotte, 1946/1963), and this percept is also influenced by subtle visual properties of

a. Launching

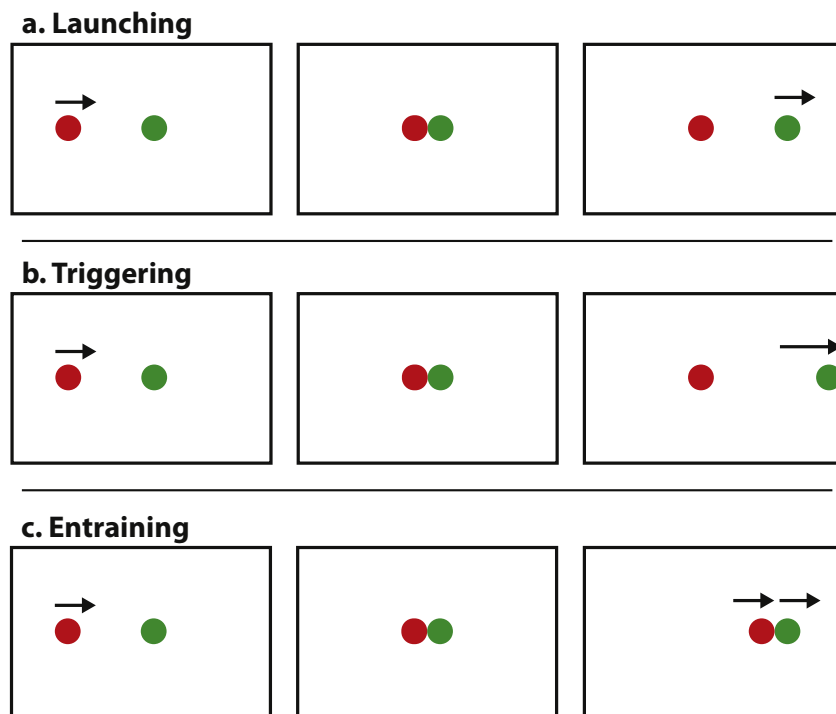## b. Triggering

c. Entraining

Fig. 1. (a) A typical launching event. The red object (A) moves until it is adjacent with the green object (B), at which point A stops and B immediately starts moving in the same direction. This yields a visceral impression that A *causes* B's movement. (Arrows indicate motion, and were not present in the actual displays.) (b) A typical triggering event, in which B moves much faster than A. (c) A typical entraining event, in which A continues to move alongside B. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

other nearby visual stimuli (Bae & Flombaum, 2011; Choi & Scholl, 2004; Scholl & Nakayama, 2002) — including temporal properties that depend on synchrony on the order of at most a couple hundred milliseconds (Choi & Scholl, 2006b). And Michotte, in his book, describes many dozens of demonstrations exploring the (often surprisingly precise) influence of other particular sorts of visual details. Moreover, another indication that the impression of causality when viewing launching events is due to visual processing is just the fact that such percepts can in turn alter the perception of other more basic visual properties, including the perception of spatial distance (Buehner & Humphreys, 2010; Scholl & Nakayama, 2004) and perhaps temporal order (Bechlivanidis & Lagnado, 2016). And indeed, the detection of visual causality can be so foundational that it even influences visual awareness itself — insofar as launching displays break through continuous flash suppression into conscious experience more readily than do non-causal control displays (Moors, Wagemans, & de-Wit, L., 2017).

On the other hand, despite this sensitive tuning to subtle visual details, causal perception seems markedly unaffected by higher-level knowledge and judgment. Indeed, in a way you can experience this for yourself when viewing the online demonstrations: even as you know with certainty that there are no actual causal relationships among these collections of pixels, this knowledge does nothing to blunt the power of the resulting causal percepts. And this impotence of higher-level cognition is also apparent in the context of particular experiments, as when you see causal launching while simultaneously judging that B's motion is in fact caused by some other factor entirely distinct from A (Schlottmann & Shanks, 1992) — a pattern that is generally indicative of visual illusions (see Firestone & Scholl, 2016; van Buren & Scholl, 2018).

1.1. Beyond launching?

Although the launching effect has dominated research in this domain, it has been recognized — ever since the days of Michotte — that there may also be other distinct forms of causal perception. Indeed, even Michotte (1946/1963) himself described other types of putatively causal events such as *triggering* (in which B moves much faster than A; see Fig. 1b), and *entraining* (in which A does not stop but rather moves

alongside B, appearing to push it along; see Fig. 1c). And more recent investigations have focused on the perception of potentially related events such as *pulling* (e.g. White & Milne, 1997), *bursting* (White & Milne, 1999), and *shattering* (Hubbard & Ruppel, 2013). To our knowledge, however, none of these other putatively different forms of causal perception have ever been explored in more than a handful of papers (unlike launching, which has been explored in many dozens of empirical studies), and little is known about them beyond the fact that observers will readily use such terms to freely describe and rate such animations.

This state of affairs leaves us with a surprisingly foundational open question about this topic — and one that lies at the heart of the present experiments: do all of these sorts of demonstrations reflect a single, unitary underlying process of 'causal perception' resulting in a single sort of representation (and a single sort of percept) whose differences are merely superficial variations on a theme? Or do they reflect independent visual routines, perhaps resulting in distinct types of causal representations, and importantly different classes of percepts? For example, is the sort of causality we that perceive in launching events the same sort of causality that we perceive in, say, entraining or triggering events? In short, is causal perception at root *one thing*, or are there actually multiple causal perception*s*?

Existing research does not seem to provide any clear answer, and indeed has seldom asked such questions. On one hand, the very fact that researchers have found it worth memorializing these various examples with different labels suggests some form of categorical distinction. But on the other hand, some researchers have striven to explain disparate examples (e.g. pulling and launching) in terms of a single overarching explanation (e.g. involving the perceived transmission of physical forces; White, 2010, 2011, 2014, 2015). Nevertheless, despite this lack of clarity, the question itself (causal perception, or causal perception*s*?) seems fundamental to our understanding of what causal perception is in the first place — similar perhaps to questions about whether other notions such as attention reflect unitary constructs or instead reflect multiple distinct forms of selection (e.g. Chun, Golomb, & Turk-Browne, 2011).

1.2. A new tool for studying causal perception: retinotopically specific visual adaptation

One reason why it has proven difficult to answer (or even ask) questions about whether there are distinct types of causal perception may simply be that most existing dependent measures are unable to capture such distinctions (for discussion, see Rips, 2011). Most existing studies simply use descriptions or ratings, but these seem clearly unsuited to the task of carving the mind at its joints in this respect. (Consider again the case of attention: just because we — as both scientists and as laypeople — use the same word "attention" in so many different contexts certainly does not entail that all the referents of this word reflect a single unitary cognitive resource or process.) Recently, however, a new method based on visual adaptation has appeared that may allow for such an investigation.

In broad terms, visual adaptation is a phenomenon wherein exposure to a particular stimulus induces a temporary change in the perception of (and sensitivity to) subsequent related stimuli — often resulting in a form of aftereffect that continues even after the initial (adapting) stimulus is no longer present (for a review, see Webster, 2016). Perhaps the most famous example of this is the 'waterfall illusion', in which exposure to movement in one direction (such as the downward movement of a waterfall) subsequently enhances sensitivity to movement in the opposite direction — even giving rise to an illusory motion aftereffect in a subsequent static display (Addams, 1834). This pattern of adaptation is observed with almost any sort of visual stimulus, from low-level factors such as color and orientation (e.g. Kohn, 2007) to higher-level categories such as faces (e.g. Webster, Kaping, Mizokami, & Duhamel, 2004) and even number (e.g. Fornaciai, Cicchini, & Burr, 2016). And one indication that many such forms of adaptation must reflect visual processing per se is simply that many of these types of adaptation (including the waterfall illusion) operate retinotopically, such that the enhanced sensitivity and the resulting aftereffects occur (only, or more strongly) in that region of the visual field that was initially adapted. (This strikes us as a largely unambiguous and uncontroversial way to identify visual processing, since we know of no type of higher-level judgment that yields any sort of retinotopically specific effect; cf. Scholl & Gao, 2013.)

One of the most astounding recent examples of retinotopically specific adaptation of which we are aware involves the perception of causality in launching events (Rolfs, Dambacher, & Cavanagh, 2013). This study employed an ambiguous sort of stimulus similar to launching, but in which A and B overlap by some amount (perhaps fully) before A stops and B begins moving (see Fig. 2a). Whereas a 0%-overlap animation is reliably perceived as unambiguous causal launching, a 100%-overlap animation is often perceived as entirely non-causal 'passing' — in which one object simply moves past another stationary object, with no interaction at all (Choi & Scholl, 2004; Scholl & Nakayama, 2002, 2004). Rolfs et al. (2013) adapted observers to a long sequence of launching events, and reported that subsequent

animations in which A and B overlapped were more likely to be seen as non-causal passing rather than causal launching — but only when the two animations were presented in the same retinal location. And this pattern of results was specific to launching, per se, since it did not occur when observers were adapted to non-causal 'slip' events in which A moved fully past B before B started moving. (These results are also consistent with neuroscientific studies: fMRI explorations of launching stimuli — when contrasted with other key stimuli that control for lower-level visual properties — implicate area V5/MT [e.g. Blakemore et al., 2001], and this area is known to be retinotopically organized [e.g. Kolster, Peeters, & Orban, 2010].)

1.3. The current study

In the current experiments, we exploit the phenomenon of retinotopically specific adaptation in order to ask whether there are multiple distinct categories of causal perception — focusing in particular on launching, triggering, and entraining. In doing so, we aim to follow the underlying logic of many adaptation studies. In the words of a recent review: "[S]pecific adaptations remain a powerful tool for dissecting vision by exposing the mechanisms that are adapting. That is, 'if it adapts, it's there.'" (Webster, 2016, p. 547). This logic applies to the existence of particular categories in the first place: if causal perception per se can be adapted, then it must "be there" in visual processing. (Some recent work has attempted to reinterpret such demonstrations in terms that do not involve causality — e.g. Arnold, Petrie, Gallagher, & Yarrow, 2015 — and these reinterpretations have stemmed in large part from doubts about whether this adaptation is truly retinotopically specific. But to foreshadow, we find strong and unambiguous support for such retinotopic specificity in the present project — and we return to the proper interpretations of these effects in the General discussion.)

Critically, this logic of adaptation also applies in terms of the *transfer* of adaptation (and the lack thereof) across category boundaries. If adaptation persists despite some change to the stimulus, then we may conclude that this change did not involve a property that is intrinsic to the underlying category. But if adaptation does not persist, then we may conclude that this change crossed over a categorical 'joint' in visual processing. Or in other words: "A hallmark of these changes is that they are selective, reducing sensitivity for stimuli similar to the adaptor but not for sufficiently different patterns. Characterization of these selective changes reveals the coding strategies in the visual system" (Webster, 2016, p. 548).

Here, we aim to "dissect" causal perception by exploring whether retinotopically specific adaptation transfers across putatively different categories of causal events. In particular, we ask whether the perception of launching, triggering, and entraining events will each induce a retinotopically specific adaptation effect in an ambiguous launching display. We focus in particular on triggering and entraining because they are perhaps the best possible candidates for truly distinct categories of
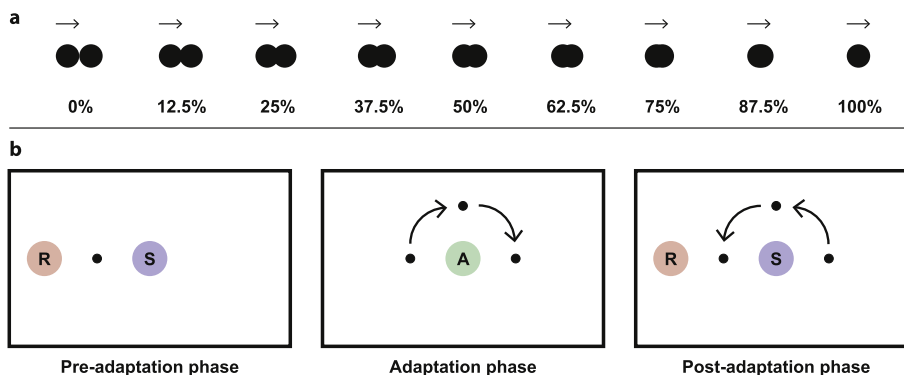

Fig. 2. Schematic depiction of the adaptation methods employed in these experiments. (a) The nine degrees of overlap used in these displays. At 0% overlap, events are typically perceived as unambiguous causal launching. As the degree of overlap increases, these events are more likely to be seen as non-causal passing. (b) The locations of presentation in each phase of the experiments. The two locations where the test animations were presented matched the location of the adaptation sequence ('A', center panel) either retinotopically ('R', left and right panels) or spatiotopically ('S', left and right panels). See text for details.

causal perception. In the first place, they are extremely similar to launching in their lower-level visual properties. (In contrast, while we could in principle ask about transfer to or from other putative causal categories such as 'bursting', those displays differ to a large degree along many different lower-level dimensions — e.g. the size and shape of their spatial envelope, and the number of distinct contours involved — any one of which could independently influence the transfer of adaptation.) But at the same time, triggering and entraining have been terminologically (and in part theoretically) distinguished as separate categories of causal perception ever since the earliest work on this topic. Michotte (1946/1963) himself suggested that observers perceive triggering ("déclenchement") when B moves roughly three times faster than A (see Fig. 1b, and also Natsoulas, 1961) — and this event is putatively distinguished from launching in part because it is no longer the case that B's entire motion seems to be fully determined by A's impact. And of course entraining (as in Fig. 1c) involves A moving at entirely distinct times during the event (i.e. post-impact), compared to launching. Yet, entraining is (like launching) also an event in which B's motion seems to be fully determined by contact with A — not only at the moment of the initial collision but also in their continued later contact.

So do these distinctions carve vision at natural joints, or not? We clearly cannot answer this question simply by analyzing the events' structures a priori — since, as we just reviewed, triggering and entraining are similar to launching in some ways, but different in other ways. Neither can we answer this question simply by noting the use of verbal labels (in reports and ratings) alone, since it is not clear whether such labels reflect perception or high-level categorization (even ignoring the possibility that Michotte's subjects were unfairly pressured into describing the animations in different ways; see Boyle, 1972; Joynson, 1971). But we can potentially answer this question by exploring whether retinotopically specific adaptation transfers among these categories. In Experiment 1, we first replicate an experiment from Rolfs et al. (2013) demonstrating that adaptation to launching is both robust and truly retinotopically specific — here using a larger sample of naïve (rather than psychophysically trained) observers. In Experiment 2, we then explore whether exposure to triggering and entraining will each also yield retinotopically specific adaptation for subsequent ambiguous launching displays. To foreshadow: we find that one of these putatively different event types *does* transfer in this way, while the other does not. And in Experiment 3, we show that this difference cannot be explained by properties of these events other than the nature of the causal interaction.

2. Experiment 1: adapting to launching

To our knowledge, the study by Rolfs et al. (2013) remains the only extant demonstration of retinotopically specific adaptation for perceived causality, and so it seemed wise to begin by replicating their study.[1] Rather than using a small number of psychophysically trained observers, however, we aimed to test the power and robustness of such effects by exploring adaptation in a greater number of naïve observers. As in Rolfs et al. (2013), our observers adapted to causal launching, and then subsequently viewed ambiguous passing displays at either the same or different retinal location (see Fig. 2b). Adaptation predicts more perceived non-causal passing following adaptation, and retinotopically specific adaptation predicts that this increase in perceived passing should be stronger for test events in the same retinal location.

---

[1] A related study demonstrated this same phenomenon in children (some with autism spectrum disorder), but did not assess retinotopic specificity (Karaminis et al., 2015). And one other study (explored in more detail in the General Discussion) successfully observed the general adaptation effect (albeit not in a direct replication), but not with any retinotopic specificity (Arnold et al., 2015).

2.1. Method

2.1.1. Participants

Twelve naïve observers (5 female, 7 male) from the Yale University community participated in two 1-hour sessions each in exchange for a modest monetary payment. This sample size was chosen before the experiment began, to be exactly triple the number of observers tested in Experiment 3 of Rolfs et al. (2013) — with the tripling due to the fact that we tested naïve observers rather than trained psychophysical observers, and so expected considerably noisier data. To reach this number, we tested 30 total observers (17 female, 13 male), with 14 excluded (after the first session) for failing the first pre-adaptation block test (as described in detail below), and 4 for failing to complete the *sec*ond session (3 for not responding, 1 for hardware failure). We subsequently analyzed the data from the 12 observers who completed both sessions (following Rolfs et al., 2013), and also the 16 who had completed at least a single session (with the extra 4 observers beyond the preplanned sample size of 12 simply being those who had completed their first session before we had obtained 12 who successfully completed both sessions).

2.1.2. Apparatus

Observers were seated in a large barber's chair in a well-lit room and were instructed to sit in a position where they could keep their head as still as possible for the duration of the (approximately 1-hour) session, with both the monitor (a 60 Hz Dell LCD) and the eye-tracker (an SMI RED500 system, both on a moving boom) then centered approximately 60 cm from the observer — with all spatial measurements below computed based on this viewing distance. At this distance, the active region of the monitor subtended 40° × 23°, and all stimuli were presented using custom software written in Python using the PsychoPy libraries (Peirce et al., 2019), along with the SMI iViewX API (SMI, 2014). Fixations were recorded at a rate of 60 Hz in order to ensure good syncing with the display.

2.1.3. Stimuli

Observers repeatedly viewed 167 ms animations, each involving two 1.5° black discs (which we will refer to as A and B) presented on a gray background. A moved toward B at 45°/sec, covering 3.75° in 83 ms. A then stopped, after which B immediately moved in the same direction, at the same speed, covering the same distance. Across trials, we varied how much A and B overlapped before A stopped and B began moving. There were 9 possible degrees of overlap, evenly distributed between 0% (no overlap; A stops as soon as it becomes adjacent to B) and 100% (full overlap; A and B are completely co-located before A stops and B begins moving), with the full range of overlaps depicted in Fig. 2a. This variation in overlap was always achieved only by varying the starting location of B, such that the distances and speeds of the movements of A and B were constant across all trials, as was the point at which the leading edge of A stopped relative to the fixation dot (as described below). So, in events with greater overlap, B was initially presented slightly closer to A — such that A and B overlapped even though A always moved the same distance before it stopped — and B correspondingly stopped moving at the end of the trial relatively closer to A.

2.1.4. Procedure

Each experimental session began with a manually verified nine-point eye-tracker calibration (using a single black dot moving from point to point against a gray background), followed by three primary phases: pre-adaptation test, adaptation, and post-adaptation test.

2.1.4.1. Pre-adaptation test phase. Observers initiated each trial by looking at a black 0.5° circular fixation dot, with instructions to fixate throughout the trial. This fixation dot was located 5° to the left of the center of the display for half the observers (as depicted in Fig. 2b), and
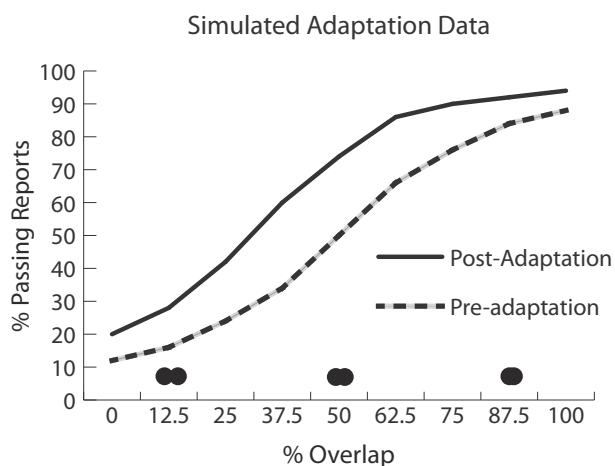
Fig. 3. Hypothetical data that would indicate adaptation to perceived causal launching. The horizontal axis depicts the possible degrees of overlap (see Fig. 2b), the vertical axis depicts the percentage of perceived passing (vs. launching) reports in (both pre- and post-adaptation) test events. The dashed line represents a hypothetical function relating these factors before adaptation, and the solid line represents the same function following exposure to an adaptation sequence of launching events — in which observers are subsequently biased to see more non-causal passing.

5° to the right of the center for the other half, and it remained in this location throughout all of the trials in this phase. Once the eye-tracker detected a fixation within 2.5° of the center of the fixation dot, the trial began. Each trial involved the presentation of a single animation (as described above), with the 2D axis along which A and B moved set at a random angle. The center of each event (i.e. the point at which the leading edge of A stopped) was located 5° from the fixation point to either the left or right (left panel, Fig. 2b). Immediately after the motion ended, both discs disappeared, and observers pressed one of two keys to indicate whether they had perceived a 'launch' (in which one object collided with another, causing the second object to move) or a 'pass' (in which one moving object appeared to pass over or through a second, stationary object). The next trial then began as soon as the observer was looking at the fixation dot (with a minimum inter-trial delay of 500 ms). Observers completed 180 trials in a randomized order — 90 trials at each fixation-relative location, with 10 for each of the 9 different degrees of overlap. These were preceded by 18 practice trials (one of each combination of fixation-relative location and degree of overlap), the results of which were not included in the analyses.

2.1.4.2. Adaptation phase. Observers were informed that during the next phase, they would simply view a series of collision events without making any responses. At the start of this phase, the fixation dot moved from its location in the pre-adaptation phase to a new location 5° above the center of the screen (which we will here call the 'intermediate fixation location'). When a fixation was detected at this new location, the fixation dot then moved again to a location 5° from the center of the screen on the opposite side of its location in the pre-adaptation phase. Once fixation was detected at this new fixation position, 320 animations (each in a different random orientation) were presented in quick succession, separated by a 100 ms period in which the discs disappeared. Each animation was identical to a 0%-overlap event as presented during the pre-adaptation phase, except that it was twice as long (334 ms), since immediately after B ended its movement, the animation played in reverse, such that B would move back toward A until they were adjacent, at which point A would move away in the same direction. These animations were centered in the display (which was itself 5° to the left or right of the fixation dot; center panel, Fig. 2b). This entire sequence of adaptation animations lasted approximately 2.5 min, though it would pause between events if ever observers were

not fixating within 2.5° of the fixation dot (resuming as soon as the fixation was again detected).

2.1.4.3. Post-adaptation test phase. Immediately after the adaptation sequence finished, the fixation dot again moved to the intermediate fixation location, and then (after fixation was detected there) back to its pre-adaptation location. As soon as fixation was detected at the pre-adaptation location, observers began completing test trials that were identical to those from the pre-adaptation test phase, with one important difference: after each response, the fixation dot returned to the intermediate fixation location and then (following fixation detection) to the adaptation phase location, at which point observers passively viewed a "top-up" sequence of 16 events that were identical to those from the adaptation phase. (This top-up sequence then ended just as did the adaptation phase itself, after which the next trial of this post-adaptation test phase began.) Just as in the pre-adaptation test phase, observers completed 180 trials in a randomized order — 90 trials at each fixation-relative location, with 10 for each of the 9 different degrees of overlap. As a result, 90 of these trials matched the spatiotopic location of the adaptation stream (i.e., the center of the screen), and the other 90 matched the retinotopic location of the adaptation stream (i.e., the same side of the fixation dot where the adaptation stream had been presented; right panel, Fig. 2b).

2.2. Results

2.2.1. Analyzing adaptation

We measured adaptation (both spatiotopically and retinotopically, as described below) in terms of the influence of the adapting stimuli on the function relating launching/passing reports to the degree of overlap. In general, we expect this function to resemble that depicted by the dashed line in Fig. 3. At 0% overlap, there is no ambiguity, and so we expect relatively few passing reports. At 100% overlap, in contrast, the cues to launching are weakest — especially since they are presented quickly and in the periphery — and so we expect the most passing reports. We can then characterize causal perception in general via the line that connects these two endpoints as a function of the degree of overlap — and we can subsequently measure the effects of adaptation by exploring how the adapting sequence changes this function. In particular, we expect that this will yield an *aftereffect*, wherein the function is skewed in the opposite direction from the adapting stimuli — so that adapting to launching yields more perceived passing (just as adapting to red yields more perceived green). In the context of Fig. 3, such an aftereffect would be realized by the dashed pre-adaptation curve changing into something like the solid post-adaptation curve — in which there is more perceived passing for each individual degree of overlap (especially for the intermediate degrees of overlap) — such that the space between those two lines reflects the presence and magnitude of the adaptation effect.

2.2.2. Participant exclusions

Each individual animation used in this experiment could be perceived as either 'launching' or 'passing'. Indeed, readers who view the online movies will find that each individual animation is clearly one or the other. (If you have difficulty perceiving passing, simply move one of the animations into the periphery, and then the passing percept will become very apparent.) At typical speeds, these are straightforward dichotomous percepts, and people have no trouble seeing or distinguishing them (e.g. Choi & Scholl, 2006b; Scholl & Nakayama, 2002, 2004). However, in order to replicate the stimuli used by Rolfs et al. (2013), these animations had to be presented at a very high speed relative to most previous work on this topic: whereas the 'launching' and 'passing' animations used in previous studies often lasted on the order of 1500 ms (e.g. Choi & Scholl, 2006b), each individual animation in this experiment took only 167 ms from start to finish. This came at a high cost for our untrained observers: for some observers, the events

were simply too fast to reliably distinguish launching from passing at all. Indeed, in early piloting, we found that some observers completely failed to distinguish between launching and passing at these speeds and simply responded either uniformly (providing the identical response for every individual animation) or randomly (providing roughly 50% passing reports for every degree of overlap).

It is not possible for us to directly measure aftereffects due to visual adaptation in observers who respond either randomly or uniformly (where both patterns are completely independent of the degree of overlap). Put in terms of Fig. 3: we cannot directly evaluate how the line between the two extreme points changes when those two extremes do not exist in the first place. Therefore, in order to directly measure the effects of adaptation on the launching/overlap function, we were only able to examine observers whose responses *did* vary systematically by overlap in the first place — and so we designed a set of criteria to exclude observers for whom this was not the case. In particular, at the end of the pre-adaptation test phase, the program automatically calculated the proportion of 'passing' responses at each degree of overlap, and observers only continued to the adaptation phase (and then to the post-adaptation test phase) if their responses in the pre-adaptation test phase met the following three conditions: (a) the rate of 'passing' responses for the two smallest degrees of overlap (0% and 12.5%) was < 50%; (b) the rate of 'passing' responses for the two greatest degrees of overlap (87.5% and 100%) was > 50%; and (c) the difference between these two averages was at least 20%.

As noted above, these criteria led us to exclude a very large proportion of our observers (14/30) — and indeed a number of these observers indicated during their subsequent debriefing that the events were simply too fast for them to see what had happened. (As noted below, we subsequently took several steps to limit this high exclusion rate in Experiment 2, which cut this rate by more than half.) These exclusions were presumably not needed in the original experiments by Rolfs et al. (2013) since they used trained psychophysical observers (including some of the authors), whereas we used naïve undergraduate observers.

2.2.3. Trial exclusions

Following Rolfs et al. (2013), we also excluded individual trials from the analyses on the basis of observers' ability to maintain fixation. This seemed critical: since our primary question involves the possibility of retinotopically specific aftereffects, the data could only be meaningful in this respect if the relevant stimuli were in fact presented in the same retinal location. Eye velocity was calculated for each fixation within a trial (when observers were asked to fixate), starting from the fixation that triggered the start of the trial, and measured in terms of the 2D distance between one fixation and the next. The median velocity for all the test trials in each session was then calculated, and trials were excluded from analysis if they contained either missing fixations (indicating a blink or eye-tracker failure) or movement between two fixations that exceeded the median velocity for test trials in that session by five standard deviations (indicating a saccade). This led to the exclusion of 15% of individual trials across all observers (ranging from a minimum of 4% to a maximum of 31%).

2.2.4. Adaptation

We conducted our primary analyses on the first sessions of the 16 observers who completed at least a single one-hour session. (As reported in Appendix A, we found qualitatively similar results when we analyzed both sessions of the 12 observers who completed two sessions. We focused on the single-session data here for ease of comparison to Experiment 2, in which all observers completed only a single session.)

The results of this experiment are depicted in Fig. 4. In this figure, each horizontal axis reflects the nine different degrees of overlap, and the vertical axis reflects the percentage of trials perceived as passing. As with the idealized data from Fig. 3, the resulting lines all exhibited a steep positive slope, corresponding to the impact of the degree of

overlap on causal perception. (Keep in mind that due to our exclusion criteria, the slope of this line had to be positive to some degree.) Fig. 4a depicts the results for the half of the test events that were presented at the same retinal location as the adaptation stream. Fig. 4b depicts the results for the other half of the test events, which were presented at the same spatiotopic location (but a *different* retinotopic location) as the adaptation stream. In both cases, the dashed line depicts responses from the pre-adaptation phase, while the solid line represents results in the post-adaptation phase. The resulting red shading in Fig. 4a thus corresponds to the magnitude of the aftereffects at the same retinal location, while the blue shading in Fig. 4b corresponds to the magnitude of the non-specific aftereffects (which could involve aftereffects that apply across the entire visual field). And as should be clear from these graphs (by the fact that the dashed lines were consistently below the solid lines), both sorts of adaptation existed.

Isolating the retinotopically specific adaptation thus requires that we compare the magnitude of the red shading in Fig. 4a to the magnitude of the blue shading in Fig. 4b — and indeed one can appreciate at a glance that the former was greater. To capture this more directly, we calculated the difference between the solid line and dashed line at each degree of overlap, and these difference scores are depicted for the retinotopically identical location by the red line in Fig. 5, and for the spatiotopically identical location by the blue line in Fig. 5 (such that the blue and red lines represent the magnitudes of the blue and red shading from Fig. 4). In this graph the difference between the red line and the blue line thus indicates the magnitude of the *retinotopically specific* adaptation effect, as indicated by the green shaded area. And indeed, the fact that this green shading was consistently present in Fig. 5 indicates immediately that there was substantial retinotopically specific adaptation at every degree of overlap.

These impressions were verified with the following statistical tests. We computed the adaptation effect for each degree of overlap (for trials at the same retinotopic location as the adapting stream, and for trials at the same spatiotopic location as the adapting stream), and we then created a difference score for each degree of overlap of retinotopic adaptation – spatiotopic (nonspecific) adaptation — resulting in a single score for retinotopically specific adaptation at each degree of overlap (corresponding to the green shaded area in Fig. 5). We then binned the degrees of overlap into three regions: 'low overlap' (0%–25%), 'intermediate overlap' (37.5%–62.5%), and 'high overlap' (75%–100%). We expected that the adaptation effect would be clearest for animations with intermediate overlap, since these animations were most ambiguous (in terms of trial-by-trial variability) to begin with. (In contrast, at the lowest and highest overlaps, we expected that observers should be at ceiling [nearly all passes] or floor [nearly all launches], and so these events should be relatively unaffected by adaptation.)

The presence of retinotopically-specific adaptation was thus assessed for each degree of overlap via a single-sample *t*-test comparing the magnitude of the difference score against 0. As predicted, there was a significant degree of retinotopically-specific adaptation for animations with intermediate overlap ($M = 13.3\%$, $SD = 24.5$, $t(15) = 2.17$, $p = .047$, $d = 0.54$).[2] In addition, and unexpectedly, there was also a significant degree of retinotopically-specific adaptation effect for animations with both low overlap ($M = 15.5\%$, $SD = 18.0$, $t(15) = 3.45$, $p = .004$, $d = 0.86$) and high overlap ($M = 10.1\%$, $SD = 18.8$, $t(15) = 2.15$, $p = .048$, $d = 0.54$).

---

[2] To be clear, when we refer to a mean retinotopically specific adaptation effect of 13.3%, what we mean is that 13.3% more of the overall number of reports at this degree of overlap and at the retinally matched location were passing reports following adaptation, and this is the increase that can be attributed to retinotopically specific adaptation effects while accounting for nonspecific adaptation effects, calculated as the difference of differences as described above.
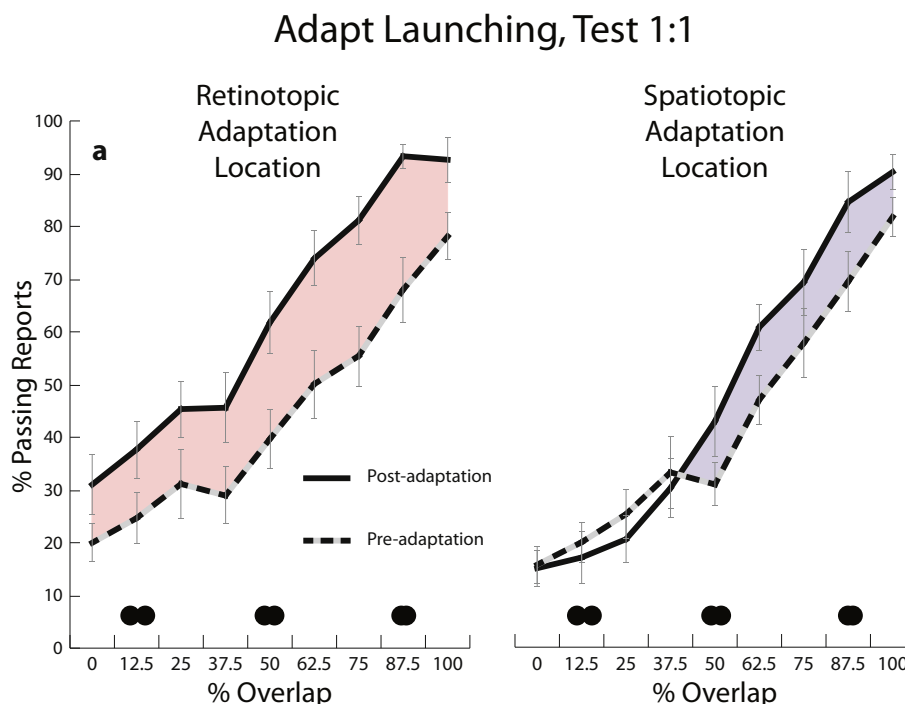
Adapt Launching, Test 1:1

Fig. 4. Percentages of perceived non-causal passing both before and after adaptation at the (a) retinotopic and (b) spatiotopic presentation locations in Experiment 1. The shaded regions indicate adaptation effects. Error bars represent ± 1 SEM.

2.3. Discussion

We successfully replicated the retinotopically-specific adaptation effect for causal launching, as first reported by Rolfs et al. (2013), here with 16 untrained observers completing a single one-hour session.[3] This sort of robust effect then allows us to test other kinds of putatively causal events in Experiment 2.

3. Experiment 2: adapting to launching, triggering, and entraining

If the labels of 'launching', 'triggering', and 'entraining' accurately carve the mind at its joints, then we may expect that adapting to one of these events may not influence subsequent perception of the others (just as, say, adapting to the facial features corresponding to one race may not influence the subsequent perception of faces corresponding to another race; Jaquet, Rhodes, & Hayward, 2008). However, if these labels simply identify superficial differences in a broader category of perceived causality (just as adapting to motion a few degrees clockwise from vertical will still influence the subsequent perception of motion a few degrees counterclockwise from vertical; e.g., Levinson & Sekuler, 1976) — that is, if there is really only one underlying form of causal perception — then adapting to any of these sorts of putatively causal events should influence the subsequent perception of launching. To test this, we adapted observers to launching, triggering, or entraining, and subsequently tested them on ambiguous launching. (We also included a condition in which observers were adapted to triggering and subsequently tested on ambiguous 1:3 events. But an analogous test was not readily possible with these displays for entraining, as there is no clear

'passing' equivalent with which to create an ambiguous entraining event.)

3.1. Method

3.1.1. Participants

Eighty observers from the Harvard University Psychology Department Study Pool (20 in each of 4 conditions; 57 female, 21 male, 2 declined to identify) participated in a single one-hour session for course credit. (Recall that in Experiment 1 we tested 12 observers who completed two sessions, but ended up with 16 observers who completed one session. Because the pattern of results was qualitatively similar in these two samples — as detailed in Appendix A — we decided to use a single session in this experiment, with a slightly increased sample size.) To reach this preregistered sample of 80 observers, we recruited 94 naïve observers (68 female, 24 male, 2 declined to identify). Of these, 13 (10 female, 3 male) were excluded (via the same now-preregistered criteria used in Experiment 1) for failing to distinguish between launching and passing events in the pre-adaptation phase, and one additional observer elected to withdraw from the experiment after the pre-adaptation phase due to a headache.

3.1.2. Apparatus, stimuli, and procedure

This experiment was preregistered (with the details available at https://osf.io/g3fus), and was identical to Experiment 1 except for three changes. First, we used a different apparatus. We switched to a head-stabilized SR Research Eyelink 1000 eye-tracker, and used a more typical arrangement of a chair and chinrest to minimize head movement. Stimuli were presented on a Dell 22″ flat-panel LCD operating at a 1920 × 1080 resolution and a 60 Hz refresh rate, and the viewing distance was 89 cm (with the extents reported below based on this new viewing distance). (As the PsychoPy script used degrees of visual angle as its primary unit of measure when drawing stimuli, the description of the stimuli from Experiment 1 still accurately captures the relative dimensions, locations, and velocities of all features that were carried over from that experiment.)
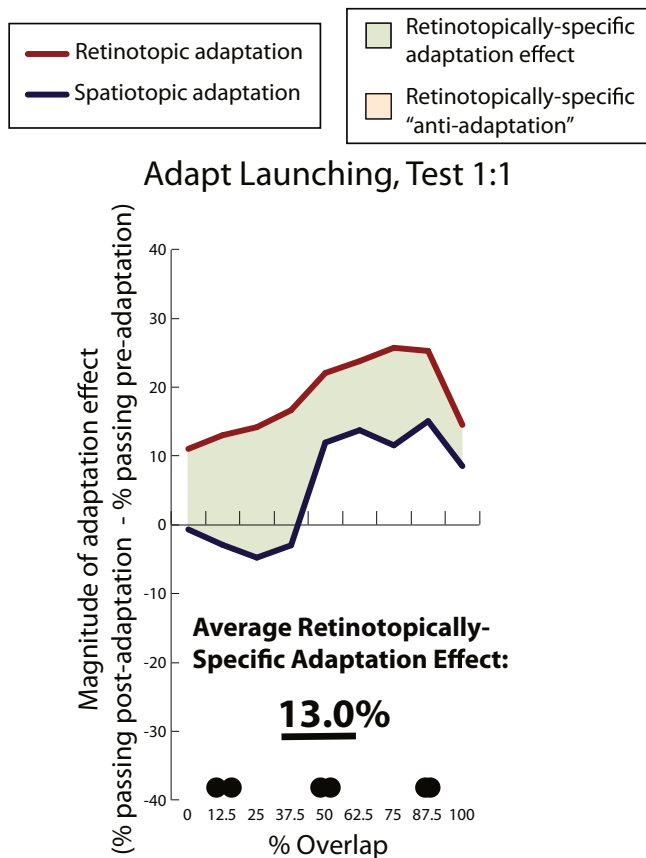
---

[3] Corresponding to the idealized effect depicted in Fig. 3, the analyses reported in the preceding paragraphs strike us as the most direct and straightforward way to analyze these sorts of aftereffects. However, this sort of analysis diverges from the one used by Rolfs et al. (2013), who employed a much less direct Bayesian curve-fitting approach. Thus, to emphasize the qualitative equivalence of our results, we also re-analyzed our data using the same approach as Rolfs et al. (2013), as reported in Appendix B.

Fig. 5. The magnitudes of the adaptation effects at each degree of overlap (i.e., the percentage of passing reports post-adaptation – the percentage of passing reports pre-adaptation) at each location in Experiment 1. The green shading indicates the magnitude of retinotopically specific adaptation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Second, in an effort to focus observers' attention on the contrast between launching and passing in such fast-moving animations (and so to hopefully reduce the number of excluded observers), we showed several half-speed examples of both (0% overlap) launching and (100% overlap) passing animations at the beginning of the experiment (3 from each category, each lasting 334 ms — whereas the experimental animations in Experiment 1 each took 167 ms — all played left-to-right, and all presented 5° above the fixation dot). Observers were told that these were demonstrations of launching and passing, and saw three of each event in rapid succession. As noted above, this change cut the exclusions due to pre-adaptation test performance by more than half, relative to Experiment 1.

Third, observers were randomly assigned to one of four conditions. In three of these conditions, observers were tested on the same ambiguous launching events used in Experiment 1, but saw one of three different adaptation streams: a launching adaptation stream, a triggering adaptation stream, or an entraining adaptation stream. In the triggering adaptation event (depicted in Fig. 1b), disc A moved 3.75° in 200 ms, at which point A stopped and B moved for 3.75° in 67 ms (thus moving three times faster than A). The entraining adaptation stream was identical to the launching adaptation stream, except that A continued moving with B after making contact, maintaining the same speed and direction and covering the same distance (as depicted in Fig. 1c). An additional group was also adapted to triggering, but then tested on ambiguous 1:3 speed ratio events (with the same degrees of overlap as the ambiguous 1:1 events, but with the speeds of the triggering adaptation stream).

Notably, entraining events are not symmetrical in time in the way that launching events are. As such, it is not possible to play them forward and backward continuously, as the causal agent moves with the causal patient to the end of its trajectory, rather than remaining at the point of contact. Therefore, knowing that the adaptation stream for the entraining condition would necessarily be unidirectional, we made *all* of the adaptation streams unidirectional (such that the first frame of the next animation appeared as soon as the last frame of the previous animation had been cleared from the display). To compensate for this, the initial adaptation stream was extended to 400 events, but overall observers still saw fewer collisions in this study compared to Experiment 1's 320 adaptation events with reversals (which resulted in 640 distinct collisions). Top-up adaptations were left at 16 events, thus reducing the number of collisions in the top-up adaptation streams by half compared to previous experiment. Since this experiment included a replication of Experiment 1 in its design, we were confident that it would be apparent whether the reduced exposure to the adapting stimuli was still sufficient to produce the relevant aftereffects (which in fact it was).

3.2. Results

We excluded 9% of individual trials across all observers (with a range of between 1% and 17% across individuals, and between 7% and 10% across conditions).

The resulting magnitude of adaptation at each test location is depicted for each condition in Fig. 6. Inspection of this figure reveals a clear 3-part pattern. First, as depicted in Fig. 6a, the launching/launching condition (i.e. the condition in which observers adapted to launching and were then tested on ambiguous launching) effectively replicated both Experiment 1 and Rolfs et al. (2013): again, there was reliable retinotopically specific adaptation (as indicated by the green shading) — not just for the intermediate degrees of overlap (as expected), but for nearly all degrees of overlap (just as was the case for Experiment 1). Second, as depicted in Fig. 6b and c, a similar degree of highly robust retinotopically specific adaptation was also observed for both of the conditions in which observers adapted to triggering. Third, as depicted in Fig. 6d, there was no hint whatsoever of any retinotopically specific adaptation for the entraining/launching condition. (Figures illustrating the raw rates of passing reports at each degree of overlap — corresponding to Fig. 4 for Experiment 1 — can be found in Appendix C.)

These impressions were verified by the following statistical analyses. We first analyzed the data exactly as in Experiment 1 — looking only at the intermediate degrees of overlap where we had initially expected an effect. Then, because we unexpectedly observed reliable effects for the entire range of overlaps in Experiment 1, we also looked at all of the overlaps together. (Both of these analyses, including all of the details below, were fully preregistered together at https://osf.io/g3fus.)

Retinotopically specific adaptation was first assessed separately for each condition by comparing the magnitude of adaptation at the retinotopic location to the magnitude of adaptation at the spatiotopic location — effectively statistically comparing the red vs. blue lines for each panel in Fig. 6. This revealed reliable retinotopically specific adaptation for the launching/launching condition (intermediate overlaps: $t(19) = 5.30$, $p < .001$, $d = 1.18$; all overlaps: $t(19) = 6.17$, $p < .001$, $d = 1.38$), the triggering/triggering condition (intermediate overlaps: $t(19) = 3.75$, $p = .011$, $d = 0.84$; all overlaps: $t(19) = 4.26$, $p = .003$, $d = 0.95$), and the triggering/launching condition (intermediate overlaps: $t(19) = 3.53$, $p = .017$, $d = 0.79$; all overlaps: $t(19) = 3.86$, $p = .008$, $d = 0.86$) — but not for the entraining/launching condition (intermediate overlaps: $t(19) = 0.13$, $p > .9$, $d = 0.03$; all overlaps: $t(19) = 1.05$, $p > .9$, $d = 0.24$). (These p-values are Bonferroni-corrected for eight comparisons.)

We then looked for differences across conditions by conducting a one-way ANOVA, along with pairwise comparisons (using Tukey HSD
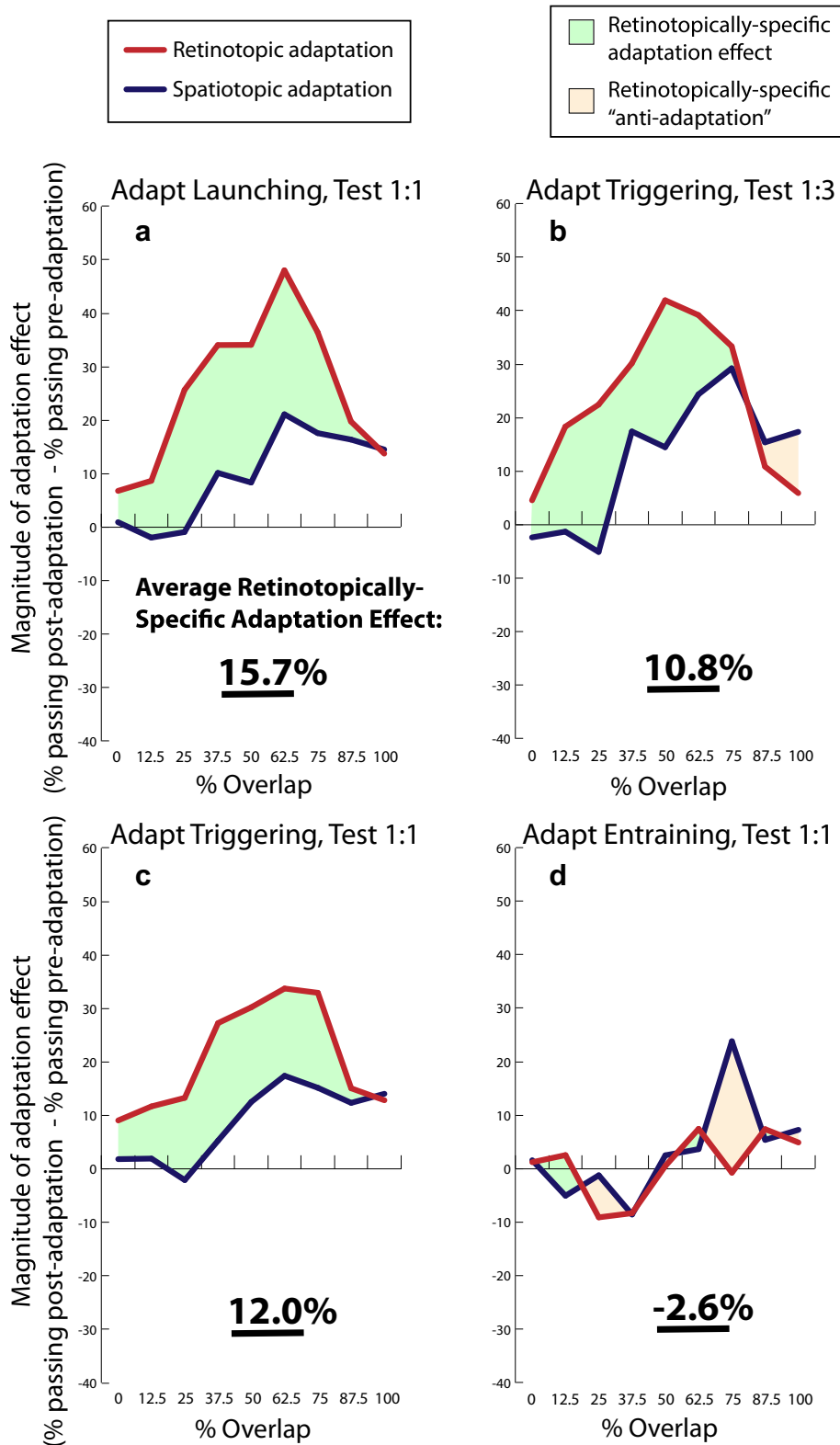
Fig. 6. The magnitudes of the adaptation effects at each degree of overlap (i.e., the percentage of passing reports post-adaptation – the percentage of passing reports pre-adaptation) for each condition and location in Experiment 2, for (a) the launching/launching condition (i.e. in which observers were adapted to launching, and then later tested on an ambiguous launching event), (b) the triggering/triggering condition, (c) the triggering/launching condition, and (d) the entraining/launching condition. The green shading indicates the magnitude of retinotopically specific adaptation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

corrections). This yielded a reliable main effect for both intermediate overlaps ($F(3, 76) = 4.46$, $p = .006$, $\eta_p^2 = 0.15$) and all overlaps ($F(3, 76) = 8.92$, $p < .001$, $\eta_p^2 = 0.26$). The pairwise comparisons then yielded reliable or marginal differences only for the three comparisons involving the entraining/launching condition, for both intermediate overlaps (vs. launching/launching: $t(38) = 3.55$, $p = .004$, $d = 1.12$;

vs. triggering/triggering: $t(38) = 2.50$, $p = .07$, $d = 0.79$; vs. triggering/launching: $t(38) = 2.45$, $p = .06$, $d = 0.78$) and all overlaps (vs. launching/launching: $t(38) = 5.15$, $p < .001$, $d = 1.63$; vs. triggering/triggering: $t(38) = 3.78$, $p = .004$, $d = 1.20$; vs. triggering/launching: $t(38) = 3.67$, $p = .001$, $d = 1.16$). No other pairwise comparisons were reliable, for either intermediate overlaps (all $ts < 1.04$, all $ps > .74$, all

ds < 0.33) or all overlaps (all ts < 1.36, all ps > .57, all ds < 0.43).[4]

3.3. Discussion

The results of this experiment seemed especially clear and straightforward: Launching adapted launching (as expected, and replicating both Experiment 1 and Rolfs et al., 2013), but triggering also adapted launching, while entraining did not. And these effects were extremely unsubtle: the effect observed when triggering adapted launching was in excess of 10% (and almost 20% for the intermediate overlaps) — whereas the null effect observed when entraining adapted launching was not only unreliable, but was numerically in the opposite direction (and was <1% for the intermediate overlaps). These effects thus demonstrate that not all causal events are created equal — but neither are they all distinct.

4. Experiment 3: adaptation to launching without stopping

We have suggested that perceived entraining does not adapt perceived launching because they represent different categories of causal perception. But there is another possibility (as suggested by insightful anonymous reviewers): Entraining was also the only adaptation event in which A did not stop at the moment of (first) contact between A and B. So might it be that this difference in the presence of "stopping" (relative to launching or triggering) was the key factor that prevented the transfer of adaptation from entraining to ambiguous launching? To find out, we simply had observers adapt to a causal event that is like launching, but in which A never comes to a stop: A moves until it is adjacent with B, at which point B begins moving at A's former speed, while A continues at a fraction of its former speed. This preserves the essential features of launching (i.e. that A unambiguously causes B to move in a manner akin to an elastic collision), but it does not involve A coming to a stop. Thus, if the entraining/launching dissociation reported in Experiment 2 reflects "stopping adaptation", then the retinotopically-specific adaptation effect should disappear in the present experiment. But if this dissociation reflects a lack of transfer across categorically different types of causal events, then these launching-without-stopping events should generate the same type of retinotopically-specific adaptation effect we have observed for launching adaptation events in Experiments 1 and 2.

4.1. Method

4.1.1. Participants

Twenty observers from the Rutgers University – Newark Psychology Department study pool (14 female, 6 male) participated in a single one-hour session for course credit. To reach this preregistered sample of 20 observers, we recruited 32 naïve observers (22 female, 8 male, 2 declined to identify). Of these, 11 were excluded for failing to distinguish between launching and passing events in the pre-adaptation phase, using the same criteria as in Experiments 1 and 2, and 1 withdrew from the study due to exhaustion shortly after starting the post-adaptation phase.

---

[4] As noted in the Introduction, the existence of retinotopically specific adaptation seems like an especially reliable guide to the existence of visual processing, since we know of no higher-level decision process that operates in a retinotopic reference frame. As such, we focus throughout this paper only on the retinotopically specific adaptation effects (and the lack thereof). However, it is also interesting to note that the entraining/launching condition produced no reliable adaptation whatsoever — at either the retinotopic location or the spatiotopic location. In contrast, the other three conditions each reliably produced adaptation at both locations (though of course always with more adaptation at the retinotopic location). We do not include these analyses here for the reasons just noted, but they can be easily computed from the raw data provided online at https://osf.io/u5n8r/.

4.1.2. Apparatus, stimuli, and procedure

This experiment was preregistered (with the details available at https://osf.io/g3fus), and was identical to Experiment 1 except for three changes. First, we used a different apparatus — an SR Research Eyelink 1000 eye-tracker operating in arm-mount mode, tracking monocularly, with no head support. Observers wore a sticker on their forehead to allow the eye-tracker to locate them. Stimuli were presented on a 23.5″ flat-panel LCD operating at a 1920 × 1080 resolution and a 60 Hz refresh rate, and the viewing distance was 89 cm (with the extents reported below based on this new viewing distance). The eye-tracker was mounted below the screen, which was mounted on a movable arm. (As the PsychoPy script used degrees of visual angle as its primary unit of measure when drawing stimuli, the description of the stimuli from Experiment 1 still accurately captures the relative dimensions, locations, and velocities of all features that were carried over from that experiment.)

Second, the adaptation event was similar to that used in the launching-adaptation condition of Experiment 2, except that object A did not stop at the moment of contact. Instead, it continued moving in the same direction, but at 25% of its original speed (and thus 25% the speed of B). Thus, like entraining, A did not stop at the moment of contact — but unlike entraining, there was a clear separation between A and B following their contact.

Third, in a (largely unsuccessful) effort to reduce attrition, we elected to slow down all of the events (both adaptation and test). In test events, each object moved at 28°/s, covering 3.75° in 133 ms. In adaptation events, A moved at this speed prior to collision and B at this speed after collision, while after collision A moved at 7°/s. To ensure that the study could be completed in a single one-hour session, we correspondingly reduced the number of test trials in the pre-adaptation and post-adaptation block, from 180 (10 at each combination of overlap and location) to 144 (8 at each combination of overlap and location).

4.2. Results

We excluded 12% of individual trials across all observers (with a range of between <1% to 32% across individuals). The resulting magnitude of adaptation at each test location is depicted in Fig. 7. As should be clear from this figure, the retinotopically-specific adaptation effect was again robust. As in Experiment 2, we preregistered analyses of the adaptation effect collapsing across all overlaps, and collapsing across the middle three overlaps. Both analyses showed a significant retinotopically-specific adaptation effect (all overlaps: $t(19) = 2.69$, $p = .015$, $d = 0.42$; middle overlaps: $t(19) = 2.88$, $p = .009$, $d = 0.40$).

4.3. Discussion

The key result of this control experiment was simply that retinotopically-specific adaptation of causal perception in the launching displays (in which A launches B) is still robust even when A continues moving (slower than B) after the impact. This rules out the possibilities (a) that such adaptation effects for launching and triggering displays (in Experiments 1 and 2) depended on A stopping, and (b) that the lack of transfer in adaptation from entraining to launching (in Experiment 2) could be due to the mere lack of stopping in the entraining event.

5. General discussion

Does the perception of causality reflect a single, unitary form of visual processing, or might there be independent categories of causal perception — perhaps matching the sorts of labels that have been previously generated based on explicit reports and ratings, such as launching, triggering, and entraining? The present study attempted to answer this question by exploiting the very phenomenon that has most convincingly demonstrated that causal launching indeed reflects a form of visual processing, per se (rather than higher-level judgment):
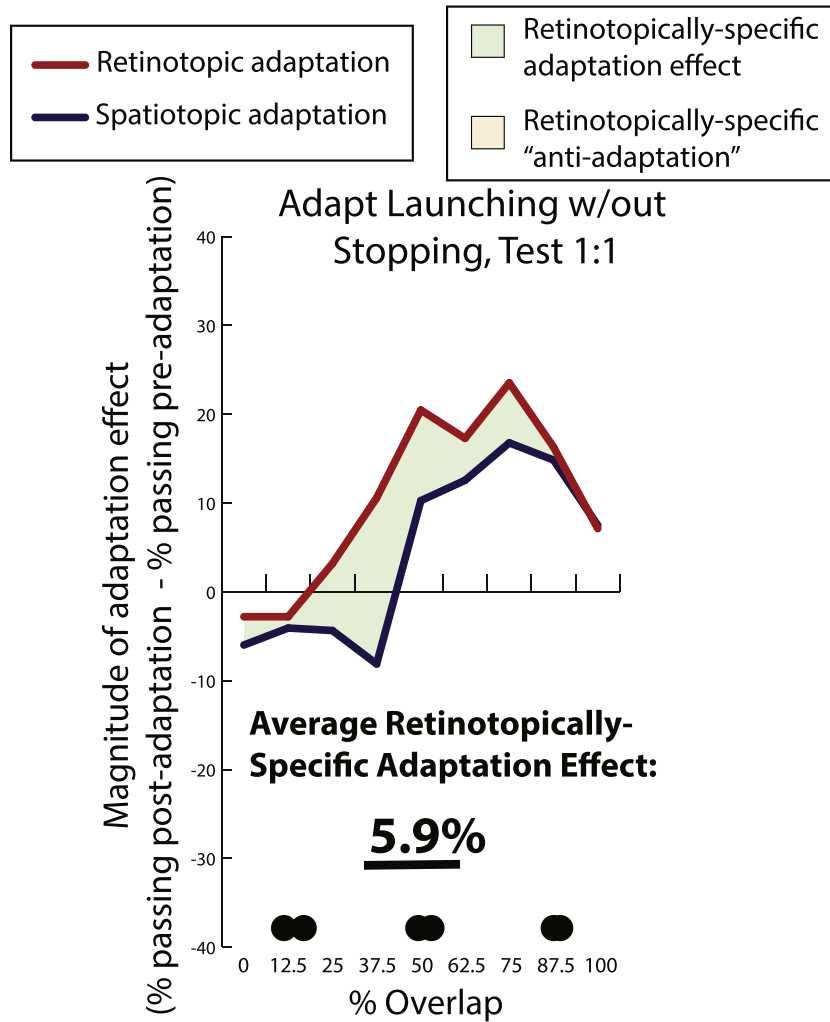
Fig. 7. The magnitudes of the adaptation effects at each degree of overlap (i.e., the percentage of passing reports post-adaptation – the percentage of passing reports pre-adaptation) at each location in Experiment 3. The green shading indicates the magnitude of retinotopically specific adaptation. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

retinotopically-specific adaptation (Rolfs et al., 2013). Our data revealed a clear answer to this question — albeit one that was both nuanced and surprising: entraining and launching seemed to be distinct categories of causal perception (since the former did not adapt the latter), whereas triggering and launching seemed to involve the same category of causal perception (since the former did adapt the latter). These conclusions follow exactly from the previously quoted logic of adaptation transfer: "A hallmark of these changes is that they are selective, reducing sensitivity for stimuli similar to the adaptor but not for sufficiently different patterns. Characterization of these selective changes reveals the coding strategies in the visual system" (Webster, 2016, p. 548). In these terms, the current study indicates that the 'coding strategy' for causality in visual processing is neither monolithic (in which case every event type would have adapted launching) nor piecemeal (in which only launching would have adapted launching).

This particular taxonomy — wherein triggering pairs with launching, but entraining does not — may seem rather surprising from at least one perspective. Michotte (1946/1963) proposed that the essential feature of launching was that the motion of object B was *fully determined* by the impact of A — a feature that is shared with entraining, but not with triggering. (In triggering, B's motion is only *initiated* by A's impact — since that impact cannot account for B's subsequent greater velocity.) On this basis, we may speculate that Michotte himself might have predicted that adapting to entraining would transfer

to launching, but that adapting to triggering would not.[5] But of course we found the opposite pattern.

5.1. Launching vs. triggering vs. entraining

At root, this project is asking: is causal perception *one thing*, or are there actually multiple causal perception*s*? And when phrased in this simple dichotomous manner, our data suggest a clear answer: causal perception is *not* monolithic, and there are in fact at least two distinct categories of causal perception. Slightly less simplistically, we have also asked whether the different intuitive 'flavors' of causal perception may all in fact reflect independent visual routines. And here the answer is

---

[5] In his typically nuanced way, Michotte also suggested that launching and entraining might be derived from different Gestalt principles (continuation and common fate, respectively) which produce the impression that B's movement is determined by A. And on that basis, he might have made the alternative prediction that there would be no adaptation transfer between *any* of these three events. While we may speculate that the present results may thus have surprised him from either perspective, our reading of his work suggests that he thought ampliation was the relevant foundational, categorical property — and thus that perceived launching and perceived entraining "can both be brought together under one and the same basic concept, that of ampliation of the movement" (Michotte, 1946/1963, p. 217).

more nuanced: Some might, but some might not. Following the logic of adaptation, launching and triggering seem to share a "coding strategy" in visual processing, but launching and entraining do not.

We have suggested that these kinds of questions are foundational in part because they have so rarely been addressed in past work, beyond the use of relatively vague terminological intuitions. In fact, we know of only two other empirical papers that even attempted to address such questions, and in this section we briefly consider how our results compare to those previous projects.

At the dawn of research on causal perception, and even (and especially) in Michotte's work, launching was often discussed in the same breath with other putative forms of causal perception such as triggering and entraining. But as noted above, the vast majority of recent work has only ever studied the launching effect, per se. One of the rare exceptions that directly addressed Michotte's entraining effect also happened to contrast it with launching, and its results are consistent with the current project, insofar as entraining seemed to work differently. This project involved infant cognition, and one of the central pieces of evidence that infants do in fact perceive launching as causal in the first year of life is that they are sensitive to the *reversal* of causal roles: if habituated to A launching B, they will reliably dishabituate to B launching A — whereas similar reversals of non-causal events (such as the same events but with temporal gaps) do not yield dishabituation (Leslie & Keeble, 1987). Another project (Bélanger & Desrochers, 2001) subsequently replicated these findings with launching (a rare and laudable example of direct replication in the infant literature!), but also failed to find a similar pattern of dishabituation with entraining, tested in the very same manner. They concluded that these two categories of causal perception are importantly different, and we might speculate that this is because only launching forces infants to encode the causal roles of 'agent' and 'patient'. The imposition of a categorical distinction here is exactly in line with the present results, and it raises the interesting question about whether this agent/patient distinction may also lie at the root of how these phenomena patterned with respect to visual adaptation.

In contrast, the only other recent paper that to our knowledge has even attempted to address the perception-or-perception*s* question reached a very different conclusion — and this apparent conflict seems especially salient to us because one of us wrote this other paper (Kominsky et al., 2017). This study contrasted launching and triggering in the context of both visual search and infant dishabituation. In the visual search task, adults saw three simultaneous events. One was a launching event with a 1:1 speed ratio, the other was a 3:3 launching event, and then there was one 'asymmetric' display — either a 1:3 triggering event, or a 3:1 launching event. In this context, observers were reliably faster to find the asymmetric event when it was triggering. In the infant dishabituation task, 7–9-month-old infants who were habituated to a 1:1 launching event dishabituated to a 1:3 triggering event, but not to a 3:1 launching event. In both of these tasks, the key result was thus that triggering seemed meaningfully distinct from launching (and in both tasks these effects disappeared when a delay was introduced in the middle of the events).

This study thus concludes that launching and triggering are categorically distinct — whereas the current project suggests that they are part of the same underlying "coding strategy" in visual processing. What might explain this difference? In the first place, we note that there is no direct contradiction in these studies, since they used such different tasks. The logic of (the transfer of) visual adaptation pertains not to what differences are visually discriminable in the first place, but rather to what differences the visual system treats as categorically distinct. Just because adapting to motion in a certain direction may transfer to the perception of motion at a mildly different direction (e.g. Levinson & Sekuler, 1976) does not imply that observers cannot notice that difference. And just because adapting to a face of one race transfers to a different face of that same race (e.g. Jaquet et al., 2008) does not imply that observers cannot tell those two faces apart. Both visual search and infant dishabituation may simply be sensitive to differences that go

beyond the sorts of categorical 'joints' in visual processing that are revealed by adaptation (see Webster, 2016). And it seems like this must be true in some sense: visual search advantages are found not only for categorical distinctions (e.g. finding a curved contour amidst straight lines) but also for meaningful distinctions that fall within that category (e.g. finding a very curved contour amidst mildly curved contours). And infants will readily dishabituate not only to a categorical change (e.g. seeing a person after viewing artifacts), but also to a change between two members of the same overarching visual category (e.g. seeing a grating with a higher spatial frequency after viewing gratings with lower spatial frequencies). So it may just be that search and dishabituation are sensitive to a wider variety of meaningful psychological kinds than is visual adaptation. In particular, this earlier work cannot be readily explained by low-level visual differences, due to the experimental design: cases that promote success (such as 1:3 triggering) had the same degree of difference as did cases that did not promote success (such as 3:1 launching) in both paradigms.

With only three studies (including the current project) that have examined the relationships between such events, there is a clear need for more work on such questions. And given the current results, perhaps the most obvious avenues for future work concern the nature of the (putatively categorically distinct) entraining events. In particular, whereas the current studies asked about the degree to which exposure to launching vs. triggering vs. entraining yields retinotopically specific adaptation for subsequent ambiguous launching displays, future studies could ask about other cells in this initial 'what-adapts-what' matrix — e.g. whether exposure to entraining adapts subsequent ambiguous triggering displays, or whether exposure to launching adapts some sort of (not yet developed) ambiguous entraining displays.

5.2. But is it really causal perception?

We opened this paper by noting that causality is probably the single most central concept to our understanding of the world. Yet curiously, almost every attempt to suggest that this property is somehow deeply ingrained into the mind has met with ardent resistance, and in particular with attempts to explain such possibilities away in terms of other non-causal alternatives. And this has similarly been true of the work that inspired the current project.

Our experiments adopted the methods and logic of Rolfs et al. (2013), but one of the only other studies to have employed similar methods has also questioned whether this work really implicates causality at all. Inspired by the fact that the ambiguous launching event involves degrees of partial overlap, Arnold et al. (2015) attempt to explain such results in terms of perceived 'squishiness' rather than causality. We enthusiastically agree that this sort of material property may be a part of relatively low-level visual perception (e.g. Fleming, 2014), and we think that squishiness is an innovative and inspired property to explore in this respect. However, we are deeply skeptical that this explanation can possibly work for the kinds of stimuli used here (or in Rolfs et al.). A key feature of the launching/passing contrast, after all, is one of object tracking: in causal launching, one object is seen to stop, while another starts moving — whereas in noncausal passing, one object is seen to pass over or through a second object that remains entirely stationary. But squishiness cannot accommodate this latter possibility (which was quite explicitly delineated in these terms in our instructions): this property can yield a 'hard' launch or a 'squishy' launch, but it cannot cause one object to pass right through another (perhaps even by definition).

In practice, however, the bulk of the arguments from Arnold's group — both in the 2015 paper and in a recent preprint (Gallagher & Arnold, 2019) — involve the putative retinotopic specificity in Rolfs et al.'s studies, which they question on both theoretical and empirical grounds. Theoretically, they suggest that the observation that causality can be appreciated in such a wide variety of contexts (including over vast distances, as in the connection between lightning and thunder) "does

not sit comfortably with the suggestion that mechanisms that detect such relationships are located at low-levels of the visual hierarchy, and have retinotopically-mapped receptive fields" (Arnold et al., 2015, p. 7). But this seems to us to entirely conflate causal perception and causal reasoning: of course we can *understand* the connection between lightning and thunder, but that doesn't mean that there aren't far stricter constraints (much more in keeping with the nature of lower-level visual processing) on what it takes in induce a vivid *percept* of causality. (There may still be active debates about the existence of causal perception, but surely nobody thinks that there is *only* causal perception, without any other form of higher-level causal thought.) And in fact previous work has clearly shown both (a) how causal perception can be entirely absent for (and indeed can directly conflict with!) such a higher-level judgment (Schlottmann & Shanks, 1992); and (b) that causal perception *is* extremely tightly constrained along several psychophysical dimensions (e.g. Choi & Scholl, 2006b).

Empirically, Arnold and colleagues have questioned the retinotopic specificity of this effect because they have failed to replicate that aspect of this work multiple times (albeit not in entirely direct replications; Arnold et al., 2015; Gallagher & Arnold, 2019). They carefully note that "we cannot dismiss these possibilities [of retinotopic specificity]" on the basis this sort of null effect, and they go on to note that: "We would therefore encourage other researchers to reexamine these issues independently" (Arnold et al., 2015, p. 8). That is in essence exactly what we have done here — and indeed this paper represents what is to our knowledge the first direct replication of the methods employed by Rolfs et al. (2013). In this context, our results provide a ringing endorsement of the conclusions of Rolfs and colleagues. In particular, we observed extremely robust, reliable retinotopically-specific adaptation in this study — in five independent groups of observers (one in Experiment 1, three in Experiment 2, one in Experiment 3). And we have based all of our arguments and conclusions only on this retinotopically-specific adaptation (as opposed to more generic forms of adaptation, which can more readily be mirrored by effects of higher-level judgment).

5.3. Conclusions: adaptation as a window on "what's there" in causal perception

In many ways, the adaptation effect introduced by Rolfs et al. (2013) is what we think the study of perceived causality has been waiting for, for at least several decades. The strength of research on causal perception — and the reason why it has generated such sustained interest for so long (yielding >150 papers; see http://perception.yale.edu/Causality-Papers/) — has always been its associated phenomenology. Launching events just *look* so strikingly distinct (no pun intended), and so categorically different from other types of spatiotemporal patterns! But of course phenomenology is also notoriously difficult to measure, and so in some ways the study of causal perception has always seemed hobbled by its reliance on subjective visual experience — yielding so many studies that simply "measure" it via free reports or overt ratings, which of course can be very easily

contaminated by higher-level judgments about what we *should* consider as causal. (Indeed, even those studies that have employed psychophysical methods to demonstrate how causal perception is so sensitive to subtle spatiotemporal features of the relevant visual stimulation have still typically relied on simple subjective reports, with other more objective measures being few and far between; e.g. Choi & Scholl, 2006a; Scholl & Nakayama, 2004.)

This state of affairs has led to persistent uncertainty from some quarters about whether causal "perception" really reflects visual processing at all (e.g. Rips, 2011), and it correspondingly makes answering the sorts of questions addressed here difficult. But in this context, the discovery that causal perception per se can induce visual adaptation (Rolfs et al., 2013) seems to cleanly slice through all such controversy. For while many researchers may continue to doubt the distinction between perception and cognition in other ways (for commentary see Firestone & Scholl, 2016), the retinotopically specific nature of this adaptation seems clearly and unambiguously associated with visual processing.

Accordingly, we think this tool has exceptional promise for driving research in this fascinating area, and we hope that the present results will help contribute to this project. Insofar as studies of such adaptation transfer can reveal "what's there" in vision (Webster, 2016), we suggest on the basis of the present work that only some putative types of causal perception may be independently "there" in visual processing.

Author note

Credit authorship contribution statement

Jonathan F. Kominsky: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Brian J. Scholl:** Conceptualization, Methodology, Writing - review & editing, Visualization, Supervision, Funding acquisition, Resources.

Appendix A. Supplementary analyses of experiment 1

In the main text we report analyses only of the first-session data from the 16 observers who completed at least one session, in order to match the subsequent experiment (which tested observers in only a single session). Here we report the same analyses for both sessions of the 12 observers who completed two experimental sessions. Overall, 13.9% of individual trials were excluded from analyses on the basis of eye-tracking data. The figures corresponding to Figs. 4 and 5 from the main text are included as Figs. A1 and A2. A cursory glance once again gives the clear impression of substantial retinotopically specific adaptation at nearly every degree of overlap (as depicted by the green shading in Fig. A2). These results were confirmed by the same analyses conducted in the main text for the first-session-only data. The presence of retinotopically specific adaptation was again assessed for each degree of overlap via a single-sample t-test comparing the magnitude of the difference score against 0. There was a significant retinotopically specific adaptation effect for events with intermediate overlap ($M = 19.0\%$, $SD = 19.2$, $t(11) = 3.42$, $p = .006$, $d = 0.99$), and for events with low overlap ($M = 18.7\%$, $SD = 10.4$, $t(11) = 6.24$, $p < .001$, $d = 1.80$), but not for events with high overlap ($M = 5.3\%$, $SD = 11.4$, $t(11) = 1.60$, $p = .14$, $d = 0.46$). These results are similar to those reported in the main text, but with substantially lower variance for these estimates of the magnitude of adaptation, due to each estimate including approximately twice as many trials.

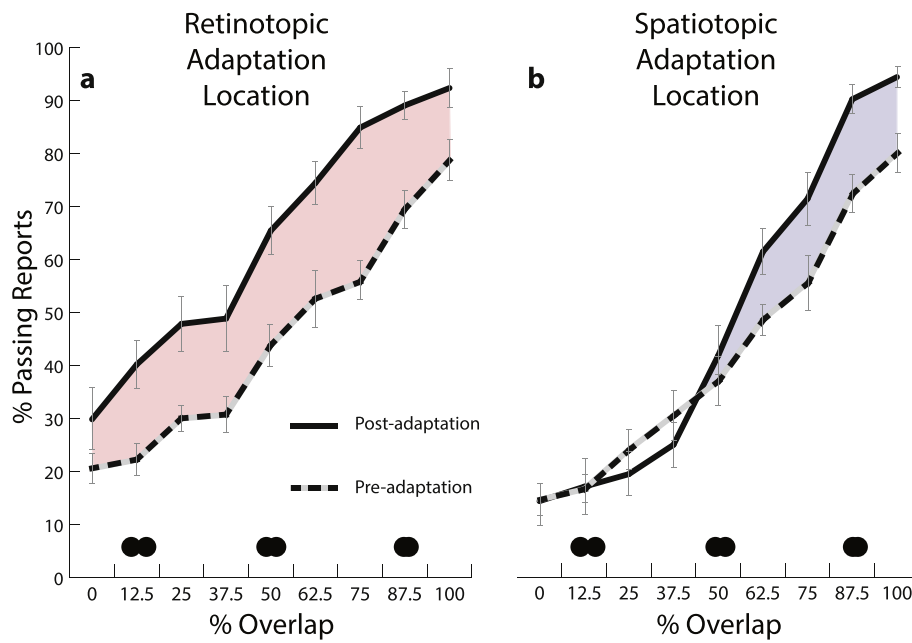## Adapt Launching, Test 1:1, both sessions N=12

Fig. A1. Percentages of perceived non-causal passing both before and after adaptation at the (a) retinotopic and (b) spatiotopic presentation locations in Experiment 1, from the sample of 12 observers who completed 2 sessions each. The shaded regions indicate adaptation effects. Error bars represent ±1 SEM.
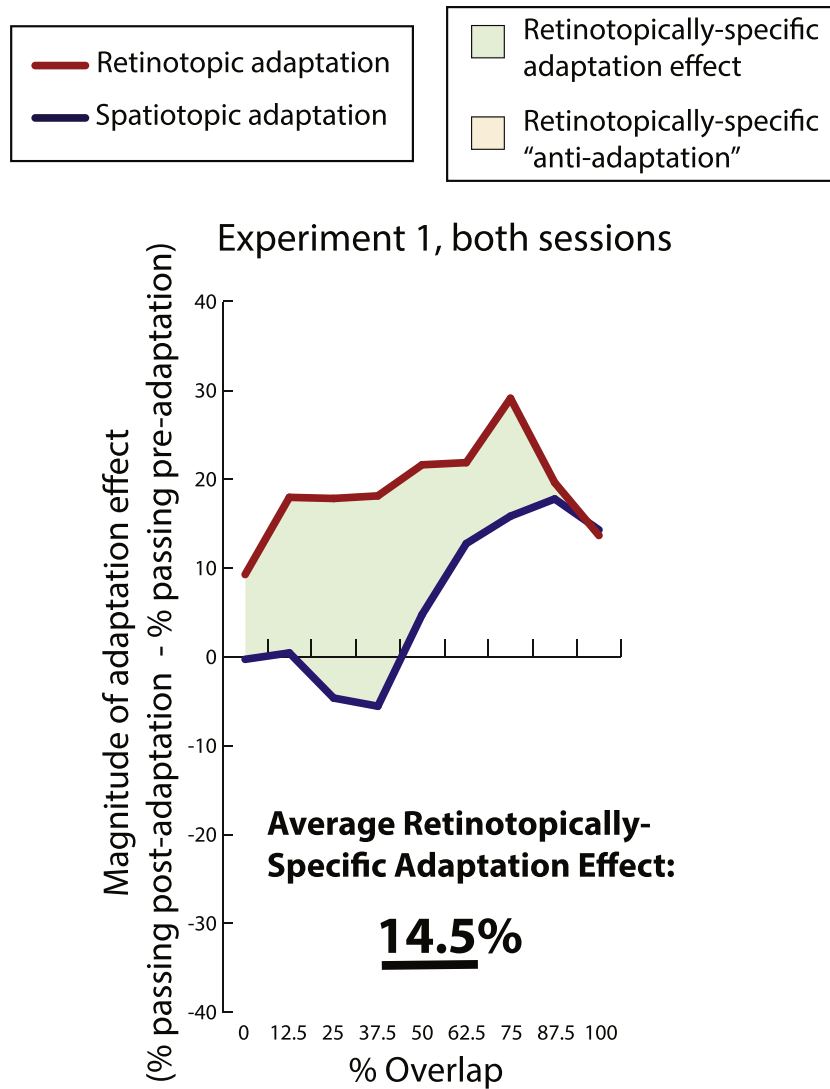
Fig. A2. The magnitudes of the adaptation effects at each degree of overlap (i.e., the percentage of passing reports post-adaptation – the percentage of passing reports pre-adaptation) at each location in Experiment 1, from the sample of 12 observers who completed 2 sessions each. The green shading indicates the magnitude of retinotopically specific adaptation.

Appendix B. Bayesian curve-fitting analysis of experiment 1

Rolfs et al. (2013) drew on their substantial expertise in psychometric function fitting to analyze their results. While their analysis allowed for the calculation of a point of subjective equality (PSE) — i.e. the degree of overlap at which an event is equally likely to be seen as a launch or a pass — we feel that this is an overly indirect approach to answering the key underlying question of whether there is retinotopically specific visual adaptation in these displays. As such, the main text simply measured and analyzed this sort of adaptation directly, as depicted in Figs. 4 and 5. As outlined in the main text, simple *t*-tests can then clearly identify the relevant effects (which are also immediately apparent in the graphs themselves), and they are not prone to some of the pitfalls of the more complex Bayesian curve-fitting analysis, as described below. The curve-fitting analysis further requires the experimenter to make assumptions about the data, and we found that the assumptions that are required to use their analysis are not appropriate for our data, and in fact are likely inappropriate for most data produced by naïve observers in this context. However, in the interests of demonstrating a complete replication of their results in our Experiment 1, we also report our results here using their approach (using the same matlab analysis script used in the original study, as provided by Martin Rolfs), and we then discuss our reasons for not preferring this approach. The details of the original analysis, and the template for how it is reported here, can be found in the supplemental experimental procedures of Rolfs et al. (2013).

Analysis 1: twelve observers that completed two sessions each

We elected to start with the dataset that was most similar to the one used in Rolfs et al. (2013), from the 12 observers who completed two sessions, and so saw the same number of trials as did the observers in Rolfs et al. (2013). We computed PSEs (in terms of percentage overlap) by fitting cumulative Gaussian functions with four parameters (mean, standard deviation, lower and upper asymptotes) to each observer's responses in each block/location combination using maximum-likelihood estimation with no prior assumptions about the mean or standard deviation of the overall distribution. The priors for the mean and standard deviations of the asymptotes were set to 0 and 0.05, respectively, reflecting an assumption that at the most extreme degrees of overlap tested, observers should report the expected percept (launching at 0% overlap and passing at 100%

overlap) almost uniformly. Following curve fitting to each observer's responses, we computed 95% CIs and s.e.m. for these PSEs using standard bootstrapping practices, resampling N PSEs from N observers on each repetition, then averaging across the 4 resampled PSEs to create the mean of the bootstrap sample. We computed s.e.m. over 10,000 independent repetitions of this sampling process and then computed 95% CI from that (1.96 * s.e.m.).

To determine the change in PSE at a given location (ΔPSE), we computed 95% CIs from 10,000 independent bootstrap samples of an observer's PSEs at that presentation location and at each sample computed the difference between the post-adaptation and pre-adaptation PSE at each location. To determine whether this ΔPSE was significantly different from 0, we determined what fraction of the bootstrapped distribution lay beyond 0, assuming a normal distribution. To determine whether the effect was retinotopically specific, we applied the same method to the difference of differences, i.e., $\Delta PSE_{ret-specfic} = \Delta PSE_{ret} - \Delta PSE_{spa}$. We computed the Bayes Factor (BF) of each ΔPSE, describing the odds that a given ΔPSE was different from a null hypothesis of a uniform distribution around 0 with a range of $[-1,1]$. BFs < 1 favor the null hypothesis (i.e. no adaptation aftereffect), and BFs > 1 favor the hypothesis that ΔPSE is different from 0 (i.e. an adaptation aftereffect). While we had a somewhat larger sample size, we still elected to use the correction for small sample sizes used by Rolfs et al. (2013), with the standard deviations of the likelihood estimations calculated as s.e.m. multiplied by a factor of $1 + 20/N^2$, where N is the number of observers in the analysis. We continued to use this correction primarily because it is unclear what the criteria for a "small sample size" should be in this context.

Results

This analysis largely replicated the results of Rolfs et al. (2013), though with some added noise. There was a significant adaptation effect at the retinotopic location ($\Delta PSE_{ret} = -0.377$, $SE = 0.098$, $p < .001$, BF = 42.21). For the spatiotopic location, interpretation hinges on the Bayes Factor, as a frequentist analysis did indicate a small but significant adaptation effect, but the Bayes Factor indicates that the pattern of results was twice as likely to be generated by the null hypothesis ($\Delta PSE_{spa} = -0.112$, $SD = 0.049$, $p = .023$, BF = 0.51). A repeated-measures ANOVA indicated a main effect of phase ($F(1, 11) = 11.05$, $p = .007$, $\eta_p^2 = 0.50$), and location ($F(1, 11) = 7.93$, $p = .017$, $\eta_p^2 = 0.42$), and, critically, a significant interaction ($F(1, 11) = 13.09$, $p = .004$, $\eta_p^2 = 0.54$). The ΔPSEs are summarized in Fig. A3, modeled on (and produced by the same code as) Rolfs et al. (2013)'s Figs. 2C and 3C.
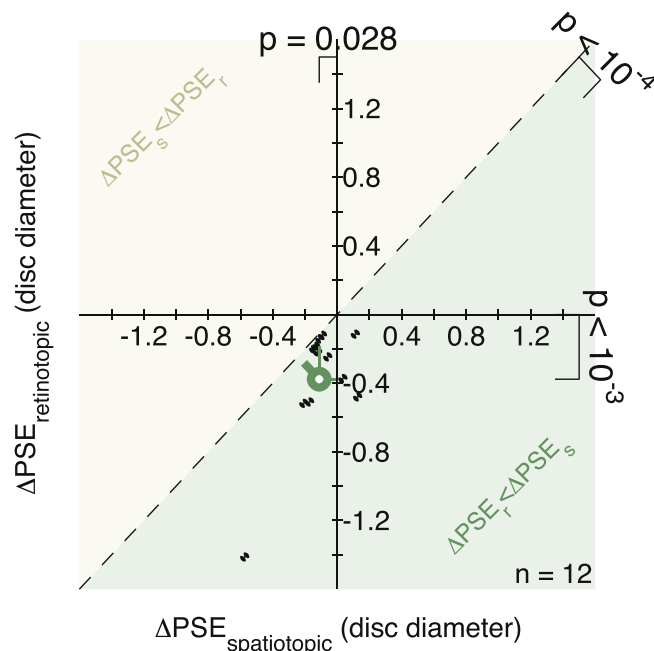

Fig. A3. ΔPSEs for each of the 12 two-session observers in Experiment 1, calculated using a Gaussian(0, 0.05) prior. Points in the green-shaded area are observers who showed greater retinotopic than spatiotopic adaptation (i.e., $\Delta PSE_{ret} > \Delta PSE_{spa}$).

Concerns

Our primary motivation for preferring the types of analyses reported in the main text is that they seem considerably more direct, simple, and transparent, in terms of the calculation of the key effects. But we are also hesitant to apply the current analysis strategy to our data for other reasons:

Nonsensical outliers. The first issue should be apparent near the very bottom of Fig. A3, where a single observer has a ΔPSE_{ret} of less than -100%. This seeming paradox is not an error in the analysis or the data, but is rather an exaggeration of an unusual pattern of responding: In the post-adaptation block, this observer reported passing at least 50% of the time for every animation at every overlap presented at the retinotopic location — even those animations with 0% overlap. This resulted in a calculated PSE for the retinotopic location in the post-adaptation phase of -85% overlap. In other words, this analysis predicts that this observer would be equally likely to perceive an animation with -85% overlap as a launch or a pass. However, exactly what this animation could possibly be is unclear. (If they had responded this way in the pre-adaptation phase, they would have been excluded before they reached the adaptation phase, but this pattern is not necessarily an indicator of improper responding.) If taken at face value, it seems that the adaptation stream affected this observer's percepts so profoundly that even typically unambiguous launching events were perceived as passing most of the time. Furthermore, this pattern is present in both of this observer's sessions. In other words, they provided a pattern of responses that satisfied our exclusion criteria in two independent pre-adaptation test phases, and in both post-adaptation test phases they showed this extreme adaptation effect. Notably, this only occurred for the retinally matched location, since their responses in the post-adaptation phase to

events presented at the spatiotopically matched location showed some adaptation (as seen by the fact that this point falls in the bottom-left quadrant), but produced a plausible (if low) PSE in the post-adaptation phase (8%).

One could argue that this outlier should simply be excluded from the analyses based on its nonsensicality, but it is only possible to do so based on entirely post-hoc criteria — and while this exclusion would remove an outlier strongly in favor of the predicted effect (an extreme retinotopically specific adaptation effect), it is difficult to justify exclusions of this sort more broadly, as the same logic could be used to remove outliers in the other direction. Notably, without the exclusion criteria noted in the main text (ensuring that observers could distinguish launches from passes in their first pre-adaptation phase), this problem would be much more widespread, as any observer who failed those criteria would produce a similarly unintelligible PSE.

Assumptions and asymptotes. We are also concerned with the fact that our observers' responses more broadly fail to meet the assumptions that are intrinsic to the analyses conducted by Rolfs et al. (2013). In particular, these analyses fit the responses to a Gaussian curve that asymptotes at rates of "pass" reports of 0% and 100% with a standard deviation of 5%. Such assumptions may be appropriate for data from four highly trained observers. But as revealed by a momentary glance at the endpoints of the lines in Fig. 4, these assumptions were grossly inappropriate in our data. Our naïve observers (unsurprisingly) provided much noisier responses, and even at the most extreme overlaps of 0% and 100% they did not respond with uniform launching or passing reports. As a result, fitting a curve to these data with the initial assumptions leads to those asymptotes falling outside the degrees of overlap used in this experiment.

Modified assumptions. However, in the interests of more accurately analyzing our results using this general analysis strategy, we also conducted an additional analysis using a cumulative Beta function with estimates for the asymptotes of 1 and 3 for mean and standard deviation, respectively (i.e. an almost completely uninformed prior, allowing the curves to more accurately conform to the actual data). The results are summarized in Fig. A4. The results are largely identical, although a few PSEs have moved around. There is a significant adaptation effect at the retinotopic location ($\Delta PSE_{ret} = -0.371$, $SE = 0.102$, $p < .001$, BF = 24.62), but now no effect at the spatiotopic location ($\Delta PSE_{spa} = -0.109$, $SE = 0.061$, $p = .074$, BF = 0.30). A repeated-measures ANOVA found significant effects of phase ($F(1, 11) = 9.19$, $p = .011$, $\eta_p^2 = 0.46$), and location ($F(1, 11) = 5.10$, $p = .045$, $\eta_p^2 = 0.32$), and a significant interaction ($F(1, 11) = 12.03$, $p = .005$, $\eta_p^2 = 0.52$).
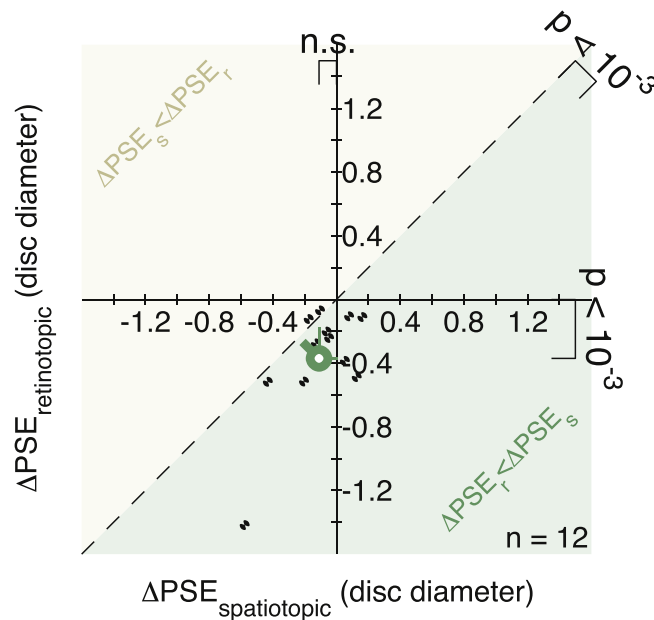

Fig. A4. ΔPSEs for each of the 12 two-session observers in Experiment 1, calculated using a Beta(1, 3) prior for the asymptotes of the cumulative likelihood function rather than a Gaussian(0, 0.05) prior. While largely similar to Fig. A3, there are some notable differences, such as two observers who showed more spatiotopic than retinotopic adaptation.

Analysis 2: sixteen observers, one session each

We also conducted this analysis examining the first sessions of all 16 observers (as reported in the main text), using the same Beta(1,3) prior, to verify that the effect was present in this more limited sample. The results are summarized in Fig. A5. In this analysis, there was a strong adaptation effect at the retinotopic location ($\Delta PSE_{ret} = -0.375$, $SE = 0.097$, $p < .001$, BF = 81.71), and no adaptation effect at the spatiotopic location ($\Delta PSE_{spa} = -0.000$, $SE = 0.098$, $p > .9$, BF = 0.13). A repeated-measures ANOVA found a main effect of phase ($F(1, 15) = 6.41$, $p = .023$, $\eta_p^2 = 0.30$), and a significant interaction between phase and location ($F(1, 15) = 7.51$, $p = .015$, $\eta_p^2 = 0.33$). However, as Fig. A5 clearly shows, in this larger sample with noisier individual estimates, there were also more extreme outliers.
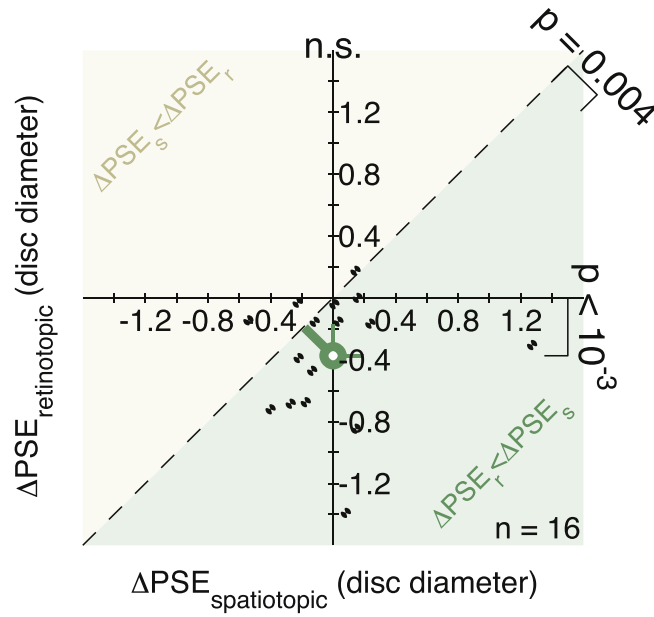
Fig. A5. ΔPSEs for each of the 16 single-session observers in Experiment 1, using a Beta(1, 3) prior for the asymptotes of the cumulative likelihood function. The retinotopic specificity of the effect is less uniform, but still significant. Notably, two observers show no retinotopic adaptation at all (y-coordinate ≥ 0).

Summary

The key results of this study — viz. the retinotopically specific adaptation to causal launching — were just as clear and robust with the original sorts of analyses used by Rolfs et al. (2013) as they were with the analyses reported in the main text. But we suggest that our preferred direct analyses make the nature of these effects far more transparent, and they do not suffer from some of the difficulties identified here.

Appendix C. Detailed results from experiments 2 and 3

The figures below (Figs. A6–A10) depict the rate of passing reports at each degree of overlap for each condition of Experiment 2, as well as Experiment 3. These figures thus correspond to the data presented Fig. 4 (in the main text) for Experiment 1.
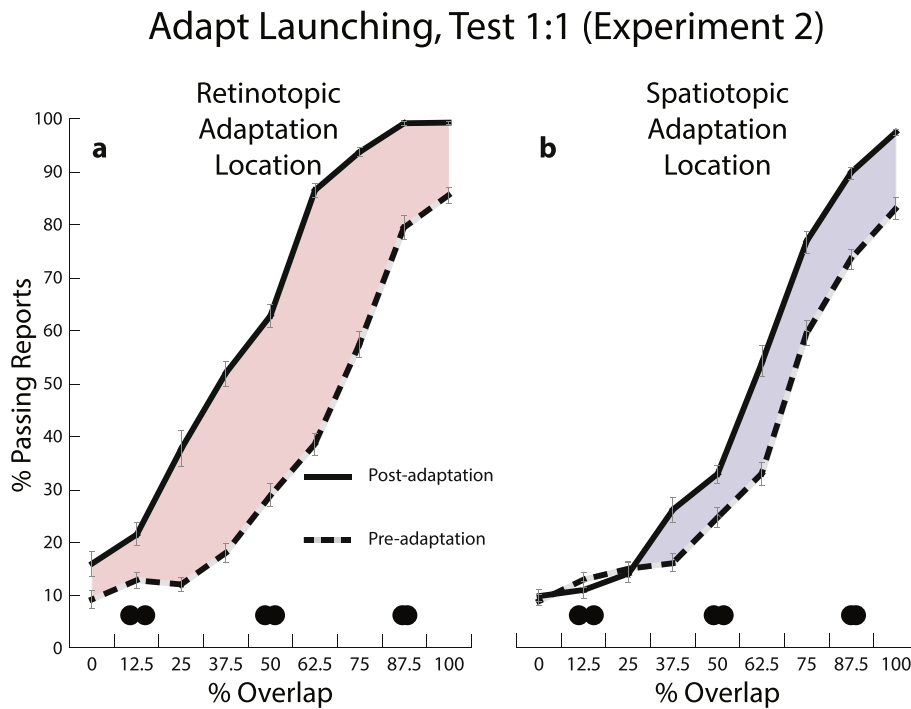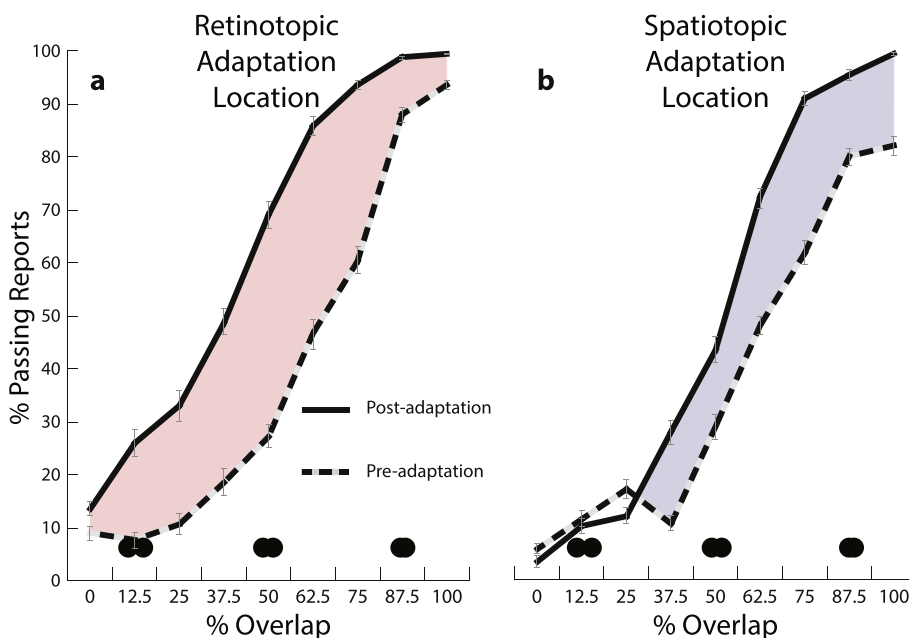

Fig. A6. Percentages of perceived non-causal passing both before and after adaptation at the (a) retinotopic and (b) spatiotopic presentation locations in Experiment 2's Adapt Launching, Test 1:1 condition. The shaded regions indicate adaptation effects. Error bars represent ±1 SEM.
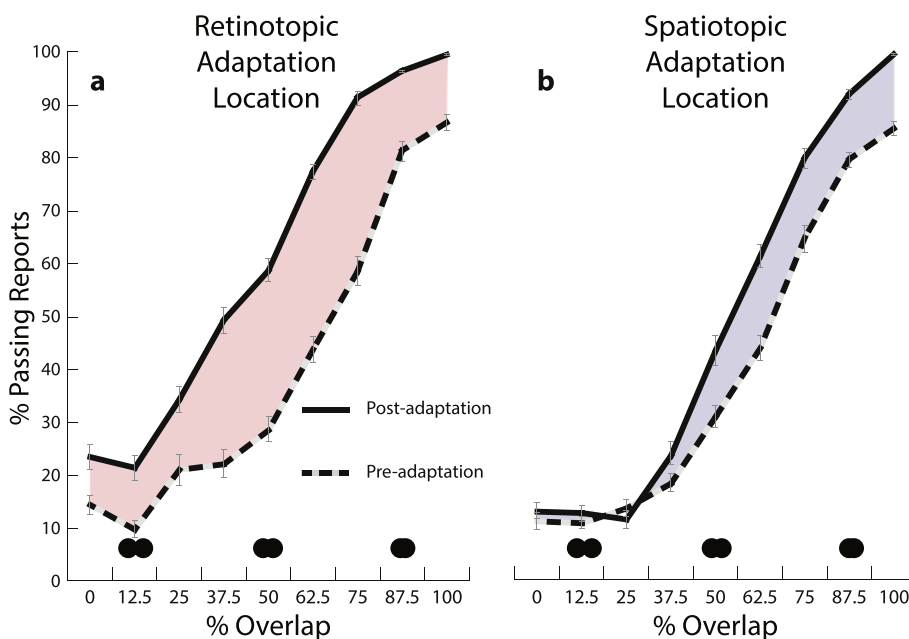
Adapt Triggering, Test 1:3

Fig. A7. Percentages of perceived non-causal passing both before and after adaptation at the (a) retinotopic and (b) spatiotopic presentation locations in Experiment 2's Adapt Triggering, Test 1:3 condition. The shaded regions indicate adaptation effects. Error bars represent ±1 SEM.

Adapt Triggering, Test 1:1

Fig. A8. Percentages of perceived non-causal passing both before and after adaptation at the (a) retinotopic and (b) spatiotopic presentation locations in Experiment 2's Adapt Triggering, Test 1:1 condition. The shaded regions indicate adaptation effects. Error bars represent ±1 SEM.

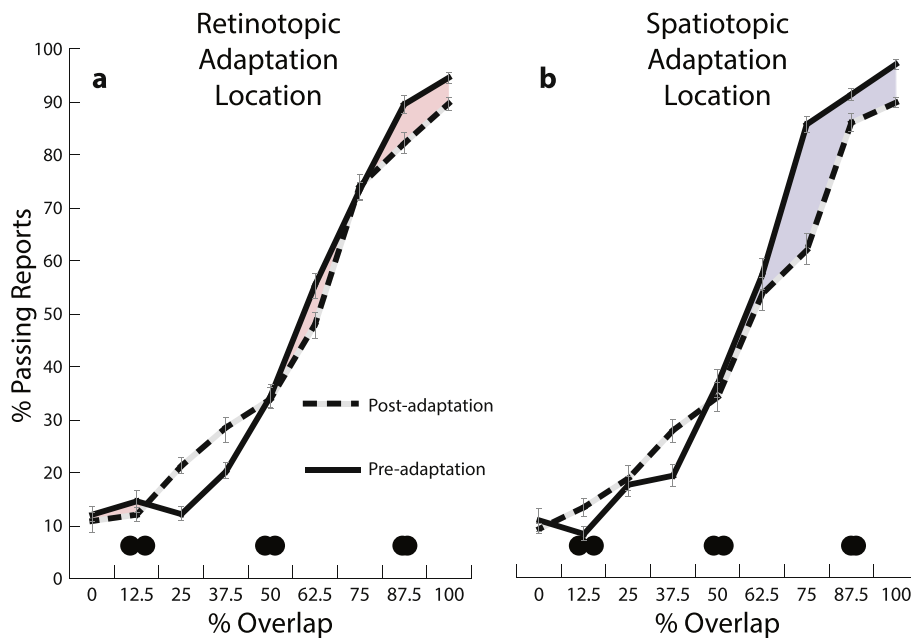## Adapt Entraining, Test 1:1

Fig. A9. Percentages of perceived non-causal passing both before and after adaptation at the (a) retinotopic and (b) spatiotopic presentation locations in Experiment 2's Adapt Entraining, Test 1:1 condition. The shaded regions indicate adaptation effects. Error bars represent ±1 SEM.
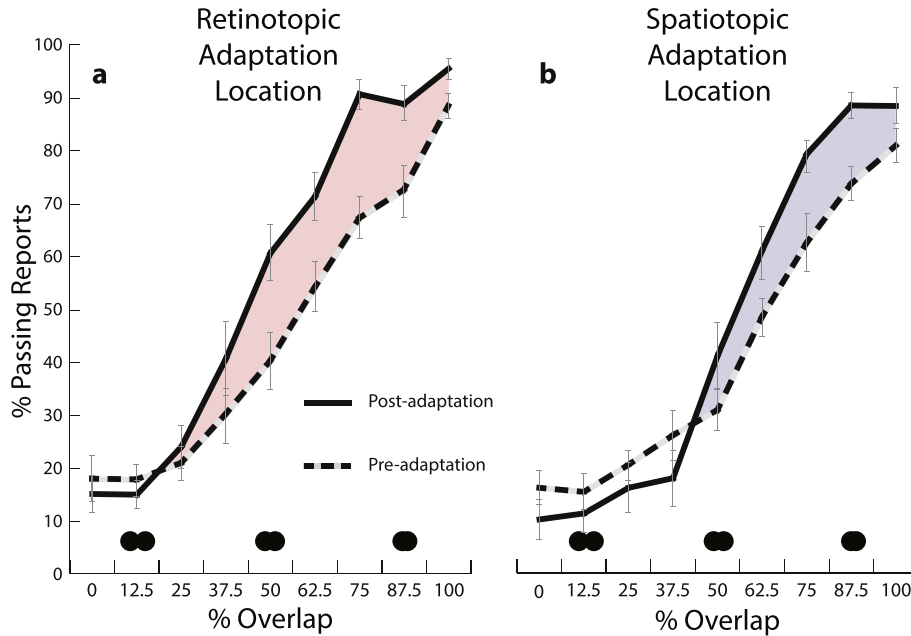
Adapt Launching w/out stopping, Test 1:1

Fig. A10. Percentages of perceived non-causal passing both before and after adaptation at the (a) retinotopic and (b) spatiotopic presentation locations in Experiment 3. The shaded regions indicate adaptation effects. Error bars represent ±1 SEM.

References

SMI IView X API (2014). *Computer software*. Teltow, Germany: SMI.

Addams, R. (1834). An account of a peculiar optical phenomenon seen after having looked at a moving body. *London and Edinburgh Philosophical Magazine and Journal of Science, 5*, 373–374.

Arnold, D., Petrie, K., Gallagher, R., & Yarrow, K. (2015). An object-centered aftereffect of a latent material property: A squishiness visual aftereffect, not causality adaptation. *Journal of Vision, 15*(4), 1–9.

Bae, G. Y., & Flombaum, J. I. (2011). Amodal causal capture in the tunnel effect. *Perception, 40*, 74–90.

Bechlivanidis, C., & Lagnado, D. A. (2016). Time reordered: Causal perception guides the interpretation of temporal order. *Cognition, 146*, 58–66.

Bélanger, N. D., & Desrochers, S. (2001). Can 6-month-old infants process causality in different types of causal events? *British Journal of Developmental Psychology, 19,* 11–21.

Blakemore, S.-J., Fonlupt, P., Pachot-Clouard, M., Darmon, C., Boyer, P., Meltzoff, A. N., ... Decety, J. (2001). How the brain perceives causality: An event-related fMRI study. *NeuroReport, 12,* 3741–3746.

Boyle, D. G. (1972). Michotte's ideas. *Bulletin of the British Psychological Society, 25,* 89–91.

Brukner, Č. (2014). Quantum causality. *Nature Physics, 10,* 259–263.

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

van Buren, B., & Scholl, B. J. (2018). Visual illusions as a tool for dissociating seeing from thinking: A reply to Braddick (2018). *Perception, 47,* 999–1001.

Choi, H., & Scholl, B. J. (2004). Effects of grouping and attention on the perception of causality. *Perception & Psychophysics, 66,* 926–942.

Choi, H., & Scholl, B. J. (2006a). Measuring causal perception: Connections to representational momentum? *Acta Psychologica, 123,* 91–111.

Choi, H., & Scholl, B. J. (2006b). Perceiving causality after the fact: Postdiction in the temporal dynamics of causal perception. *Perception, 35,* 385–399.

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology, 62,* 73–101.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Behavioral and Brain Sciences, e229,* 1–77.

Fleming, R. (2014). Visual perception of materials and their properties. *Vision Research, 94,* 62–75.

Fornaciai, M., Cicchini, G. M., & Burr, D. C. (2016). Adaptation to number operates on perceived rather than physical numerosity. *Cognition, 151,* 63–67.

Gallagher, R., & Arnold, D. (2019). Comparing the aftereffects of motion and causality across visual co-ordinates. *BioRxiv.* https://doi.org/10.1101/499764 preprint.

Hubbard, T. L., & Ruppel, S. E. (2013). Ratings of causality and force in launching and shattering. *Visual Cognition, 21,* 987–1009.

Jaquet, E., Rhodes, G., & Hayward, W. G. (2008). Race-contingent aftereffects suggest distinct perceptual norms for different race faces. *Visual Cognition, 16,* 734–753.

Joynson, R. B. (1971). Michotte's experimental methods. *British Journal of Psychology, 62,* 293–302.

Karaminis, T., Turi, M., Neil, L., Badcock, N., Burr, D., & Pellicano, E. (2015). Atypicalities in perceptual adaptation in autism do not extend to perceptual causality. *PLoS One, 10,* Article e0120439.

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97,* 3155–3164.

Kolster, H., Peeters, R., & Orban, G. A. (2010). The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *Journal of Neuroscience, 30,* 9801–9820.

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.

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

Levinson, E., & Sekuler, R. (1976). Adaptation alters perceived direction of motion. *Vision Research, 16,* 779–781.

Mascalzoni, E., Regolin, L., Vallortigara, G., & Simion, F. (2013). The cradle of causal reasoning: Newborns' preference for physical causality. *Developmental Science, 16,* 327–335.

Michotte, A. (1946/1963). *The perception of causality.* New York, NY: Basic Books (Original work published 1946).

Moors, P., Wagemans, J., & de-Wit, L. (2017). Causal events enter awareness faster than non-causal events. *PeerJ, 5,* Article e2932. https://doi.org/10.7717/peerj.2932.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology, 67,* 949–971.

Natsoulas, T. (1961). Principles of momentum and kinetic energy in the perception of causality. *The American Journal of Psychology, 74,* 394–402.

Newman, G. E., Choi, H., Wynn, K., & Scholl, B. J. (2008). The origins of causal perception: Evidence from postdictive processing in infancy. *Cognitive Psychology, 57,* 262–291.

Paul, L. A., & Hall, E. J. (2013). *Causation: A user's guide.* Oxford: Oxford University Press.

Peirce, J. W., Gray, J., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51,* 195–203.

Rips, L. J. (2011). Causation from perception. *Perspectives on Psychological Science, 6,* 77–97.

Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology, 23,* 250–254.

Schlottmann, A., & Shanks, D. R. (1992). Evidence for a distinction between judged and perceived causality. *The Quarterly Journal of Experimental Psychology: A, Human Experimental Psychology, 44,* 321–342.

Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment. In M. D. Rutherford, & V. A. Kuhlmeier (Eds.). *Social perception: Detection and interpretation of animacy, agency, and intention* (pp. 197–230). Cambridge, MA: MIT Press.

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

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

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4,* 299–309.

Sloman, S. A., & Lagnado, D. (2015). Causality in thought. *Annual Review of Psychology, 66,* 223–247.

Webster, M. A. (2016). Visual adaptation. *Annual Review of Vision Science, 1,* 547–567.

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428,* 557–561.

White, P. A. (2010). The property transmission hypothesis: A possible explanation for visual impressions of pulling and other kinds of phenomenal causality. *Perception, 39,* 1240–1253.

White, P. A. (2011). Visual impressions of force exerted by one object on another when the objects do not come into contact. *Visual Cognition, 19,* 340–366.

White, P. A. (2014). Perceived causality and perceived force: Same or different? *Visual Cognition, 22,* 672–703.

White, P. A. (2015). Visual impressions of generative transmission. *Visual Cognition, 23,* 1168–1204.

White, P. A. (2017). Visual impressions of causality. In M. Waldmann (Ed.). *Oxford handbook of causal reasoning* (pp. 245–264). Oxford University Press.

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.

White, P. A., & Milne, A. (1999). Impressions of enforced disintegration and bursting in the visual perception of collision events. *Journal of Experimental Psychology: General, 128,* 499–516.