

---

Spontaneous path tracing in task-irrelevant mazes:  
Spatial affordances trigger dynamic visual routines

---

Kimberly W. Wong & Brian J. Scholl

Yale University

Running Head : Spontaneous path tracing in mazes  
Addresses for : Kimberly W. Wong or Brian J. Scholl  
correspondence : Department of Psychology  
Yale University  
100 College Street #1223  
SHM CE2  
New Haven, CT 06510  
Email : kimberly.wong@yale.edu, brian.scholl@yale.edu  
Word Count : 212 (Abstract) + 6013 (Main text) = 6225  
Version : 5/17/24 — In press, *Journal of Experimental Psychology: General*

**Abstract** (212 words)

Given a maze (e.g. in a book of puzzles), you might solve it by drawing out paths with your pencil. But even without a pencil, you might naturally find yourself *mentally* tracing along various paths. This ‘mental path tracing’ may intuitively seem to depend on your (overt, conscious, voluntary) goal of wanting to get out of the maze, but might it also occur spontaneously —as a result of simply *seeing* the maze, via a kind of dynamic ‘visual routine’? Here, observers simply had to compare the visual properties of two probes presented in a maze. The maze itself was entirely task-irrelevant, but we predicted that simply *seeing* the maze’s visual structure would ‘afford’ incidental mental path tracing (à la Gibson). Across four experiments, observers were slower to compare probes that were further from each other along the paths, even when controlling for lower-level properties (such as the probes’ brute linear separation, ignoring the maze ‘walls’). These results also generalized beyond mazes, to other unfamiliar displays with task-irrelevant circular obstacles. This novel combination of two prominent themes from our field — affordances and visual routines — suggests that at least some visual routines may not require voluntary goals; instead, they may operate in an automatic (incidental, stimulus-driven) fashion, as a part of visual processing itself.

**Keywords**

Visual routines, spatial affordances, navigational affordances, curve tracing, path tracing

**Public Significance Statement** (150 words)

What goes on in your mind when you solve a maze, e.g. in a book of puzzles? Normally this might involve tracing through the paths with a pencil or finger. But not necessarily: you can also solve a maze just by looking. Here we show that this ‘mental path tracing’ also occurs when you *aren’t* trying to solve such puzzles (and with unfamiliar sorts of mazes, without entrances or exits): even when just passively viewing such displays, your mind automatically and spontaneously traces the paths between salient points. We demonstrated this by showing that the time it takes to compare two small probes that appear in a maze (which is itself entirely irrelevant to the task) is a function of how long the path between them is (as opposed to the brute linear distance between them). This reveals how even such simple stimuli may engage surprisingly sophisticated dynamic visual processing.

## Introduction: Seeing Mazes

The image in Figure 1 is probably immediately familiar: most readers, perhaps since childhood, have attempted to solve such mazes — often by tracing a path from one side to the other with a pencil or crayon. But take a moment now, and reflect on what it is like to simply *see* such a stimulus. Even without a writing instrument, you may find yourself *mentally* tracing out various paths through the maze — and it is certainly possible to solve such a maze only by looking at it. What triggers this sort of ‘mental path tracing’? The answer seems obvious: a maze of this kind is a learned cultural artifact made for a specific purpose (as a puzzle to solve), and accordingly the path tracing seems triggered by an overt (conscious, voluntary) goal — to ‘get out’ of the maze.

Here, in contrast, we ask whether such mental path tracing may also occur *spontaneously* — as a result of simply *seeing* the maze — even in the absence of any explicit goal. Or put in more theoretical terms, we ask whether maze-like stimuli automatically ‘afford’ (à la Gibson) the operation of path tracing as a dynamic ‘visual routine’, as a part of visual processing itself. In this way, our exploration of mazes may be of wider interest for the study of perception, insofar as it combines two prominent themes from our field, which have nevertheless not previously been related to each other: *visual routines* and *spatial affordances*.

### **Visual Routines**

Visual processing usually seems both incidental and instantaneous. Spend a moment looking at Figure 2, for example, and then keep reading. Now consider the question: What color were the laces? You can probably answer this question immediately, from memory — indicating that this property was extracted even before you were asked the question, just as a part of natural viewing. (We don’t have to *decide* to see colors.) But now consider a different question: Were the green tip and the blue tip part of the same shoelace, or two different shoelaces? You probably don’t know the answer yet, which indicates that this property was *not* extracted

incidentally during natural viewing. When you look back at Figure 2, of course, you can answer this question too — and you can do so merely by looking (i.e. even without using your finger to follow along a lace). But notice that even here you can't answer the question immediately: seeing which tip goes with which seems deliberate, dynamic, and temporally extended (as you mentally 'trace' from one tip to another).

The type of visual operation that underlies your ability to answer the which-tip-goes-with-which question when looking at Figure 2 has been termed a 'visual routine' (Ullman, 1984, 1996). And visual routines contrast with other forms of perception precisely in terms of the two features highlighted in that example: they are often invoked only on demand (rather than always occurring automatically, e.g. "depend[ing] on the goal of the computation"; Ullman, 1984, p. 153); and they are inherently dynamic, such that these operations often take some appreciable time to be executed (as you experience an active "tracing" operation).

Such visual routines are thought to operate over many types of perceptual queries, especially as they relate to various types of relations — e.g. when determining if one object is *inside* of another contour in the scene, or whether some particular object is to the *right or left* of another particular object. Determining such relationships may seem trivial in scenes with only a few objects, but the dynamic and task-dependent nature of visual routines becomes especially evident when scenes increase in complexity. Figure 2, for example, contains only two laces (and thus four tips), but suppose you were looking at a mass of many more laces (perhaps distributed over a wider area): here you might still be able to answer the which-tip-goes-with-which question 'on demand' (and perhaps just as easily), but it would not be as feasible even in principle to pre-compute such answers prior to the question being posed, because there would simply be too many tip pairs to consider. (Similarly, determining if object 1 is to the left of object 2 in a 2-object scene may seem trivial — but determining if object 14 is to the left of object 87 in a 100-object scene cannot be readily pre-computed.) In this way, it would often not be

possible for visual routines to operate incidentally (without a specific goal) for reasons of brute computational explosion.

The routine of ‘curve tracing’ has most often been studied in tasks such as the which-tip-goes-with-which question in Figure 2, in which observers must determine whether two probes lie upon the same contour (or not). As its name suggests, curve tracing involves dynamic tracking along a contour in order to determine connectedness. (Sometimes curve tracing is described as an “elemental operation” that can be part of a larger visual routine, and other times as a visual routine itself. Here we adopt the latter formulation.) As a result, when participants view displays of non-overlapping curved contours, response times vary as a function of the length of the contour between the two probes, even while equating their shortest straight-line distance from each other, ignoring the curves (e.g. Jolicoeur et al., 1986). This result seems robust and generalizable, occurring even in simplified stimuli without excessive curvature (Pringle & Egeth, 1988), in various other tasks (e.g. where participants must determine whether there are any gaps along the curve between two probes; Jolicoeur et al., 1986), and even when the curve displays are only flashed briefly (Jolicoeur et al., 1991).

### **Spatial Affordances**

A separate tradition in vision research has emphasized how seeing an object may often involve spontaneously perceiving how we could interact with it: Could I pick it up? Could I throw it? Could I step over it? This point has most famously been put in terms of *affordances*: “The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill” (Gibson, 1979, p. 119). In this sense, a small rock (but not a large rock) may afford picking up or throwing; a doorknob (but not a spiderweb) may afford grasping; and a wide aperture (but not a narrower aperture) may afford passing through. Though not as familiar in vision science as are properties such as color or shape, the idea that such affordances (‘liftability’, ‘throwability’, ‘graspability’, ‘passability’) are extracted and used during perception

has been tremendously influential (for reviews see Barsingerhorn et al., 2012; Warren, 2021), even as the concept has come to be used in many different senses (Chong & Proctor, 2020).<sup>1</sup>

The notion of ‘navigational affordances’ may be especially relevant in the current contexts of mazes and curve tracing. Recent evidence suggests that when viewing a scene, the occipital place area may encode its traversable paths or routes — e.g. how to get from the current viewpoint to another salient location (Bonner & Epstein, 2017). And critically, at least three forms of recent evidence suggests that such navigational affordances are extracted spontaneously during scene perception. First, such routes are encoded even when an observer’s task has nothing to do with navigation (e.g. when they are simply detecting colors of two probes; Bonner & Epstein, 2017). Second, ERP evidence suggests that such representations (e.g. of the number of doors and thus possible paths) are formed within 200ms of viewing a scene (Harel et al., 2022). And third, recent behavioral evidence demonstrates that a particular type of navigational affordance — a path to a salient *exit* from an enclosed space — is extracted even when task-irrelevant: when observers must simply detect any change (of any type) to an image, obstacles that move to block (or unblock) the path to an exit are more readily noticed than those that do not (Belledonne & Yildirim, 2021; Belledonne et al., 2022).

### **The Current Study: Spontaneous Visual Routines Triggered by Affordances?**

The role of conscious, voluntary goals has often been emphasized in theoretical work on the nature of visual routines (Ullman, 1984; 1996). And past empirical work on curve tracing has to our knowledge almost always inherently involved an explicit goal — e.g. when participants are explicitly asked to determine whether two probes are on the same contour (in a display with multiple contours; Jolicoeur et al., 1986; cf. Pringle & Egeth, 1988).<sup>2</sup> Here, in

---

<sup>1</sup> Of course, the wider Gibsonian research program also involved many other components — including the controversial notion that such properties could be perceived *directly*, without intervening representations or symbolic computations (Fodor & Pylyshyn, 1981; Ullman, 1980). We address these below in the General Discussion.

<sup>2</sup> To our knowledge, the only partial exception to this involves Experiments 3 and 4 from Pringle and Egeth (1988). In these experiments, subjects indicated not whether two probes were on the same contour, but rather whether two probes were on the same side of a gap between two contours. On one hand, this

contrast, we ask whether this sort of dynamic ‘visual routine’ may instead operate in a more automatic and spontaneous fashion, as a part of seeing itself, even without any explicit goal (perhaps more akin to the perception of color). In other words, we ask whether certain stimuli effectively *afford* the spontaneous operation of certain visual routines. In particular, we ask in empirical terms whether the ‘navigational affordances’ of 2D maze-like stimuli (with distinct paths and barriers) are themselves sufficient to trigger mental path tracing, even when they are entirely task-irrelevant.

In Experiment 1, observers view maze-like stimuli (though without the entrances and exits of typical mazes) upon which two small probes (each a small blue square with a white cutout) appear (see Figure 3), and they must simply indicate whether the two probes are identical or not. The mazes are thus entirely task-irrelevant, but we show that they nevertheless matter: observers’ response times depend not just on the linear inter-probe distance, but also on the ‘pathwise’ distance between them, along the paths of the maze — an effect we interpret in terms of spontaneous mental path tracing. Experiments 2a and 2b each replicate this effect, while varying various aspects of the stimuli and the probe timing. Finally, Experiment 3 generalizes these effects to other forms of spatial paths and obstacles — illustrating that spontaneous mental path tracing operates as a part of visual processing itself, and is not dependent on the familiarity of cultural artifacts such as mazes.

### **Experiment 1: Task-Irrelevant Mazes**

Observers were repeatedly shown displays with two probes — each of which was a blue square with a cutout that could appear along one edge — and they simply had to indicate

---

avoids explicit reference to the “same contour”. But on the other hand, it still seems to involve a more implicit task demand to note the positions of the probes (in order to determine their spatial relationships to the gap, which was in turn defined by the curves themselves). In contrast, to foreshadow, the current experiments involve probes whose spatial locations are entirely task-irrelevant.



whether the probes were identical or not (i.e. whether the cutout was in the same position).

These probes were presented in the context of mazelike stimuli (as in Figure 3) that were entirely task-irrelevant. We asked whether response times would nevertheless vary as a function of the ‘pathwise’ distance through the mazes between the probes, even when contrasting cases have the same straight-line distance.

## **Method**

**Participants.** Fifty observers (35 males, 15 females,  $M_{\text{age}}=23.34$ ) participated using the Prolific online platform (Palan & Schitter, 2018) for monetary compensation, with this preregistered sample size determined before data collection began. (For preregistration details, see the ‘Transparency and Openness’ statement below. We had no a priori basis for expecting any particular effect size, largely because no previous experiments have tested such questions. So chose our initial preregistered sample size simply to match those from related experiments run with similar online samples from our lab.) Gender information in all experiments was collected via a multiple-choice question, with the options of “male”, “female”, or “other”. Observers were excluded (with replacement) according to two preregistered criteria. First, in a post-experimental debriefing phase, observers self-reported how well they paid attention (on a continuous scale, between 1 = “very distracted”, and 100 = “very focused”), and we excluded observers who self-reported an attention level below 75 ( $n=13$ ). Second, we also excluded observers (who were not already excluded via criterion #1) whose overall accuracy was less than 80%, or whose mean response time was at least 2 standard deviations from the mean response time of all observers ( $n=11$ ).

**Apparatus.** After agreeing to participate, observers were redirected to a website where stimulus presentation and data collection were controlled via custom software written using a combination of HTML, CSS, JavaScript, PHP, and the JsPsych libraries (de Leeuw, 2023). Observers completed the experiment on either a laptop or desktop computer. (Since the experiment was rendered on observers’ own web browsers, viewing distance, screen size, and

display resolutions could vary dramatically, so we report stimulus dimensions below using pixel [px] values.)

Stimuli. All text, across the instructions and prompts, was presented in a modified version of jsPsych's default CSS style: black text on a white background drawn in the "Open Sans" font, presented at a font size that is scaled to 2.4% of the participant's viewport height. Maze stimuli were constructed via custom Python code. Mazes were constructed from a 31×31 square grid, each grid square being 26×26px, and with some grid squares colored black to form 'walls'. The entire perimeter of the grid was also black. As in the zoomed-in example from Figure 3, this yielded a mazelike stimulus but without an 'entrance' or 'exit'. Each probe consisted of a blue (hex #0000ff) square (15×15px) with a small rectangular segment (7×3px) cut out of one of its sides (top, bottom, left, or right). This cutout was not centered on the edge of the square, but was rather placed with its center 5.5px away from one corner (and thus 9.5px away from the other corner).

Procedure and Design. Each trial began with the maze-like stimulus (henceforth 'the maze') appearing at the center of the display. Two probes then appeared sequentially within the maze, the first after 800 ms, and the second after an additional delay randomly chosen between 650 and 700 ms. Both probes appeared along a (white) maze path (as in Figure 3), placed as described below. Each probe had a cutout along the same edge of its square. On half of the trials, these cutouts were in the same position along the edge (as in Figure 3a), while on other trials the cutouts were placed in different positions (as in Figure 3b). (For example, for probes with cutouts on the right side of the square, the cutout could be either near the top edge of the right side, or near the bottom edge of the right side.) On each trial, the full stimulus (maze + probes together) was randomly rotated by 0°, 90°, 180° or 270°, and was randomly flipped across its horizontal axis, its vertical axis, both axes, or neither axis. Both probes and the maze remained on screen until the observer gave a valid keypress to indicate whether the two probes were identical or not (in terms of the placement of the cutout).

The placements of the probes varied across trials in a systematic way. Throughout the experiment, the trials always came in matched pairs in which the two probes had the same linear distance from each other, placed within the same maze. However, the ‘pathwise’ distance between the probes was different for each trial in a pair, with one (having “Near” probes, as in Figure 3a) being shorter than the other (having “Far” probes, as in Figure 3b). Each observer completed 80 trials, presented in a different random order: 4 Mazes  $\times$  2 probe pairs (Far probes vs. Near probes)  $\times$  2 probe matching possibilities (identical vs. different)  $\times$  5 repetitions. (Across the four base mazes, both the straight-line distance and the pathwise distance varied.) Per the preregistered criteria, we excluded individual trials whose response times were more than 2 standard deviations from the mean response time of all observers (on average 1.58 out of 80 trials/observer).

Transparency and Openness. The raw data for all studies are included in the Supplementary Material file. The preregistered methods and analyses for each experiment can be viewed at: [https://aspredicted.org/L3Q\\_KFC](https://aspredicted.org/L3Q_KFC) (for Experiment 1), [https://aspredicted.org/DSM\\_NK9](https://aspredicted.org/DSM_NK9) (for Experiment 2a), [https://aspredicted.org/J3F\\_JX7](https://aspredicted.org/J3F_JX7) (for Experiment 2b), and [https://aspredicted.org/Y55\\_7YG](https://aspredicted.org/Y55_7YG) (for Experiment 3).

## **Results and Discussion**

As depicted in Figure 4a, response times were significantly greater on Far probe trials than on Near probe trials, as confirmed with a two-tailed paired t-test (1284.87ms vs. 1232.97ms,  $t(49)=4.06$ ,  $p<0.001$ ,  $d=0.58$ ). Since the only difference between these two trial types involved the pathwise distance between the probes as they were situated in the mazes, this suggests that observers were mentally tracing through the paths, despite their task-irrelevance.

### **Experiment 2a: Turn-Equated Paths**

The pattern of results observed in Experiment 1 cannot be explained without reference to the maze paths, since the straight-line distance between the probes was always equated across trial pairs. But there are still other ways in which the (necessarily different) maze paths could have influenced response times on Far probe vs. Near probe trials. Rather than reflecting influences of pathwise distances (as we hypothesize), for example, the difference could be due *only* to some alternative factor — such as a differential number of *turns* along those paths. We assess this here in a conceptual replication by maintaining longer vs. shorter pathwise distances while equating the number of turns (as in Figure 3, where both the Far probe and Near probe paths have 3 turns).

#### **Method**

This experiment was identical to Experiment 1, except as noted. A new set of 100 observers (50 males, 50 females,  $M_{age}=26.08$ ) were recruited, with this preregistered sample size chosen before data collection began. (This sample size was arbitrarily doubled relative to Experiment 1 on the basis of the experimental design: while the results of Experiment 1 were robust, the current experiment also controlled for additional factors that were not measured or manipulated in Experiment 1, as described below.) Observers were excluded (with replacement) according to the same two preregistered criteria: self-reported attention level ( $n=26$ ), and outlying overall performance ( $n=13$ ). Stimuli were identical to those used in Experiment 1, except that mazes spanned a  $31 \times 15$  grid, and matching Far and Near trials were generated with the new restriction that the inter-probe paths had the same number of turns. (Mazes and probes were still randomly reflected as in Experiment 1, but they were not randomly rotated due to the new rectangular dimensions of the grid.) Per the preregistered criteria, we excluded individual trials whose response times were more than 2 standard deviations from the mean response time of that observer (on average 3.56 out of 80 trials/observer).

## **Results and Discussion**

As depicted in Figure 4b, response times were again significantly greater on Far probe trials than on Near probe trials (1282.50ms vs. 1259.07ms,  $t(99)=2.55$ ,  $p=0.012$ ,  $d=0.25$ ). These results effectively replicate those from Experiment 1, while confirming that the observed difference must reflect spontaneous path tracing through the task-irrelevant mazes. (We note that the effect size in this experiment was smaller than that from Experiment 1, suggesting that other factors such as the number of turns may also have influenced the earlier results — but of course the key point is that such factors cannot explain, or even contribute to, the current results, since they were carefully equated.)

### **Experiment 2b: Simultaneous Probe Presentation**

In Experiments 1 and 2a, the two probes on each trial appeared sequentially. Could this have somehow encouraged mental path tracing (as attention was captured by the second probe, having already been focused on the first probe)? We explore this here in another conceptual replication, in which both probes appeared simultaneously.

## **Method**

This experiment was identical to Experiment 2a, except as noted. A new set of 100 observers (61 males, 39 females,  $M_{age}=28.44$ ) were recruited, with this preregistered sample size chosen before data collection began. Observers were excluded (with replacement) according to the same two preregistered criteria ( $n=14$  and  $n=17$ , respectively). Stimuli were identical to those used in Experiment 2a, except that both probes appeared simultaneously. Per the preregistered criteria, we again excluded individual trials whose response times were more than 2 standard deviations from the mean response time of that observer (on average 3.33 out of 80 trials/observer).

## **Results and Discussion**

As depicted in Figure 4c, response times were again significantly greater on Far probe trials than on Near probe trials (1304.04ms vs. 1282.68ms,  $t(99)=3.14$ ,  $p=0.002$ ,  $d=0.31$ ). These results effectively replicate those from Experiment 1, while demonstrating that the spontaneous mental path tracing was not somehow triggered by (and does not require) sequential probe presentation.

### **Experiment 3: Beyond Mazes**

The mental path tracing observed in the previous three experiments was spontaneous insofar as the mazes themselves were completely task-irrelevant. Nevertheless, the maze-like nature of the stimuli could have been essential in two ways. First, the fact that the visible ‘walls’ of the maze were all vertical and horizontal could have somehow encouraged people to shift their attention between the probes along horizontal and vertical paths. Second, despite their lack of entrances and exits, those stimuli were still clearly recognizable as mazes. As such, it remains possible that the mental path tracing reflects a sort of ingrained habit, due to the cultural familiarity of the stimulus, despite its irrelevance to the task.

To explore these possibilities, we replicated our study using a completely different — and completely unfamiliar — sort of visual scene. Observers completed the same task, but now the probes were presented amidst task-irrelevant displays of discs with various sizes and colors, as depicted in Figure 5. As before, the straight-line distance between probes was equated across pairs of trials, which nevertheless had Near or Far probe paths which avoided the discs. (And this new type of display also allowed us to control for several other factors, such as the number of disc-boundaries that the straight-line path would intersect, and the total amount of overlap between the discs and the straight-line paths.)

## **Method**

This experiment was identical to Experiment 2a except as noted.

**Participants.** A new set of 100 observers (60 males, 40 females,  $M_{age}=27.07$ ) participated, with this preregistered sample size determined before data collection began. Observers were excluded (with replacement) according to the same two preregistered criteria ( $n=5$  and  $n=9$ , respectively).

**Stimuli.** Displays each contained 4 discs, each with a 1px outline: a green (#438c59) disc with a dark green (#1f4913) outline (radius 71px), a magenta (#8c4367) disc with a dark magenta (#441048) outline (radius 64.5px), a purple (#5f438c) disc with a dark purple (#141b49) outline (radius 55px), and a yellow (#8c8543) disc with a dark yellow (#492814) outline (radius 51.5px). These discs were placed in a central 600×600px region as described below, with a minimum spacing (between their nearest edges) of 13px. Probes always appeared diagonally from each other, along roughly opposite diagonals, with the first probe appearing (again after 800 ms) near the top, and the second probe appearing (always after an additional 400 ms) near the bottom (so that e.g. if the first probe was positioned towards the upper right corner of the central region, the second probe was positioned towards the lower left corner).

**Procedure and Design.** As in the previous experiments, the placements of the probes relative to the discs varied across trials in a systematic way — again using matched pairs of trials in which the two probes had the same linear distance from each other. Here, in addition, the two probes in a matched pair of trials appeared in identical positions on the display, with only the placements of the discs being varied — such that curving paths between the same two probe positions (i.e. to avoid touching the discs) could be either Near (as in Figure 5a) or Far (as in Figure 5b). (In fact, these different disc placements were always created simply by swapping the positions of two of the discs. In Figure 5, for example, you can see that the positions of the purple and magenta discs have been swapped, while the positions of the green and yellow discs are identical. And these particular placements were chosen so that the intersections between

linear inter-probe paths and discs were always equated across the two trials in a pair, in terms of both the number of intersecting discs, and the total extent of the intersection.)

Each observer completed 32 trials, presented in a different random order: 2 probe pairs (Far probes vs. Near probes)  $\times$  2 possible reflections (horizontal reflection vs. no reflection)  $\times$  2 probe matching possibilities (identical vs. different)  $\times$  4 repetitions. Within this design, each observer was randomly assigned to one of 4 baseline displays, each with a different arrangement of discs. The first 3 trials were practice trials, the results of which were not recorded. Per the preregistered criteria, we again excluded individual trials whose response times were more than 2 standard deviations from the mean response time of that observer (on average 1.31 out of 32 trials/observer).

### **Results and Discussion**

As depicted in Figure 5c, response times were again significantly greater on Far probe trials than on Near probe trials (1016.21ms vs. 996.74ms,  $t(99)=2.18$ ,  $p=0.032$ ,  $d=0.22$ ). These results effectively replicate all of the earlier experiments, showing that such effects generalize to very different displays which (a) do not have salient rectilinear paths, and (b) do not constitute familiar cultural artifacts (such as mazes).

## **General Discussion**

The central empirical result of the four experiments reported here is straightforward: people seemed to mentally trace paths between salient landmarks in maze-like stimuli, even when there was no need to do so — since the mazes themselves were entirely task-irrelevant. In particular, observers were faster to compare two probes (presented either sequentially or simultaneously) when the *pathwise* distance (i.e. through the maze) between them was shorter. These results were first obtained (in Experiments 1, 2a, and 2b) using stimuli that were evocative of common cultural artifacts (similar to mazes in puzzle books, as in Figure 3), but



they also generalized to entirely unfamiliar stimuli (with circular obstacles between probes, as in Figure 5). And critically, these results had to reflect the paths between the objects (avoiding crossing over the walls of the maze, or the circular obstacles), since other factors were held constant — e.g. the straight-line distances between the probes (in all experiments), the number of turns required along the maze paths (in Experiments 2a and 2b), the exact degree to which straight-line paths intersected with circular obstacles (in Experiment 3), etc. (These other factors may still play independent roles in response times, but they cannot explain the key path-length effects.)

In theoretical terms, we have described these results via a combination of two foundational themes from vision research: *visual routines* (à la Ullman) and *affordances* (à la Gibson). In this context, the current results suggest that certain types of spatial affordances automatically trigger the operation of certain visual routines, in a spontaneous stimulus-driven manner. (This is consistent with an earlier suggestion from Pringle and Egeth, 1988, that curve tracing “might well be automatic” [p. 726]; cf. Footnote 2.) This conclusion extends our understanding of both visual routines and affordances in certain critical ways — in each case suggesting that certain key features of these frameworks can be divorced from others.

### **Expanding Visual Routines (Dynamic Visual Operations, but Without Voluntary Goals?)**

Perhaps the two key properties of visual routines, as reviewed above, are that they are (a) inherently dynamic and thus temporally extended, while also (b) operating only “on demand” (as when determining which tip goes with which in Figure 2; Ullman, 1984, 1996) — as opposed to visual processes that occur seemingly instantaneously without any specific goal (as when determining what color the laces were in Figure 2). These features are both salient properties of traditional visual routines, even beyond path-tracing. When looking at a complex mass of twisting contours, for example, you can tell whether a given dot is inside a closed contour or not (as in Figure 2c of Ullman, 1984), and you can tell just by looking, but (a) it takes an appreciable

amount of time to determine this (via the operation of ‘coloring’), and (b) this only occurs when you have the explicit intention to determine this inside/outside relation — which may not occur at all while you casually view such a stimulus.

Here we suggest that these two properties can be divorced from each other. In particular, the current experiments with mazes have plenty of synergy with the first property, but not with the second: these results suggest that the visual routine of path-tracing may *not* require explicit, voluntary goals; instead, it may operate (still in an inherently dynamic way!) in a *spontaneous* manner, as a part of visual processing — with the ‘demand’ being triggered by the spatial affordances of the stimuli themselves. (Of course, in this context, spontaneity does not imply irresistibility: observers might still be able to compare two probes without tracing along the paths if they intentionally tried to do so; what the current results show, however, is that such path tracing occurs by default, without any explicit intention.)

Putting this same point in more operationalized terms, note that previous experiments on the operation of curve tracing could not speak to this possibility, since they always involved an explicit, voluntary goal which *required* the curves to be traced (e.g. in order to determine whether two probes were on the same curve; Jolicoeur et al., 1986; Jolicoeur et al., 1991; McCormick & Jolicoeur, 1992; cf. Pringle & Egeth, 1988 and see Footnote 2). In contrast, in the present experiments, not only did observers *not* have the explicit goal of path tracing (since the mazes were task-irrelevant), but the operation of this visual routine may have actually harmed performance: observers could have simply ‘traced’ in a straight line from one probe to the other — ignoring the irrelevant walls or discs — yet instead they effectively took unnecessary mental detours around these obstacles.

### **Expanding Affordances (Functional Stimulus-Driven Percepts, but Without Direct Perception?)**

The notion of affordances is famously plastic (Chong & Proctor, 2020), but perhaps their two key properties in most interpretations (as inspired by the work of Gibson; e.g. Gibson, 1979)

are that they are (a) stimulus-driven percepts that go beyond low-level features, while also (b) being perceived ‘directly’. (In some conceptions, they are also intrinsically linked to potential *actions*.) The maze stimuli in the present experiments have both low-level features (various contours of different orientations) and seemingly higher-level functional features (such as the ‘traversability’ between two probes in the maze). For Gibson it is the latter sort of property that is primary, since it is what functionally matters to an organism: “psychologists assume that objects are composed of their qualities. But I now suggest that what we perceive when we look at objects are their affordances, not their qualities. We can discriminate the dimensions of difference if required to do so in an experiment, but what the object affords us is what we normally pay attention to” (1979, p. 126). At the same time, however, the perception of such affordances is thought to be *direct* — a function of brute “resonance” with the environment, in a type of “information pickup”, without intermediate inferential steps. This notion of ‘directness’ is notoriously controversial and difficult to unpack — as it has been argued that there is simply no feasible story for how this “pickup” could occur without mediating representations and intermediate computational steps (Fodor & Pylyshyn, 1981; Ullman, 1980). But at least one animating aspect of ‘direct’ perception is that it is thought to be *immediate*: “there can be direct or immediate awareness of objects and events when the perceptual system resonates so as to pick up information” (Gibson, 1967, p. 168).

Here we suggest that these two properties can be divorced from each other. In particular, the current experiments with mazes have plenty of synergy with the first property, but not with the second. These results suggest that simply viewing the maze-like stimuli themselves (without any particular goal) already involves the extraction not only of various lower-level features (e.g. the colors of the contours), but also of higher-level functional properties, in the form of navigational affordances. This is just the same sense in which, for Gibson, viewing a stone involves the extraction not only of its various lower-level features (e.g. its color), but also of affordances such as its ‘throwability’ and ‘graspability’. However, the

perception of such properties in the current context is clearly *not* direct in the sense of being immediate: like all visual routines, it is a dynamic temporally extended operation, realized via path tracing.<sup>3</sup>

Putting this same point in more operationalized terms, note that previous experiments involving affordances could not speak to this possibility, since they nearly always involved affordances that seemed immediately perceptible — e.g. the properties of being “fall-off-able” or “bump-into-able” (Gibson, 1979, p. 128), the climbability of stairs (Warren, 1984), or the passability of an aperture (Warren & Whang, 1987). In contrast, in the current experiments, the “navigability” of being able to trace from one probe to the other (without being blocked by an obstacle) cannot be immediately recovered without first executing the temporally extended routine of path tracing.

### **Constraints on Generality**

The effects reported here generalized across multiple types of obstructions in otherwise-navigable paths between two probes — including both walls of mazes (in familiar cultural artifacts), and arrays of discs (in unfamiliar scenes). However, there remain several open questions about the degree to which these results might generalize in other ways. Of course future work could explore such effects in different subject populations. Based on the possibility that visual routines and affordances both reflect core aspects of scene perception, however, we would predict relatively few qualitative and systematic differences across people — and at least in the current data, we did not observe any salient systematic differences across age or gender. It might also be of interest for future work to explore how and whether such effects may depend on observers paying close and careful attention to the displays. (Of course, we can’t be sure

---

<sup>3</sup> Another contrast between the current experiments and classical Gibsonian proposals involves the degree to which affordances must involve *oneself*. For Gibson, affordances are central because of how they allow *us* to act, and thus affordances may differ for different people: a certain stairway with a large riser height may be perceived as *climbable* (or a certain hefty stone may be perceived as *graspable*) for an adult, but not for a young child, etc. Here, in contrast, the navigational affordance was only between two ‘third-person’ probes, and did not directly involve oneself (though of course one may readily imagine oneself in such a maze, moving from one point to the other).

about this without testing it directly, but our hunch is that this study is somewhere in the middle when it comes to the degree of attentiveness that is necessary. On one hand, subtle response-time measures often suffer from inattention; but on the other hand, the current results were robust enough to be observed in online samples — who are often paying less focused attention than is possible to ensure during in-lab experiments.) Finally, it may also be especially important for future work to explore how these effects may or may not generalize to other types of affordances (beyond spatial affordances such as navigability) and visual routines (beyond path tracing).

### **Conclusions**

It is perhaps unsurprising that the current sorts of experiments have not been previously conducted, since key parts of the two foundational motivations for this project seem incompatible on their surfaces, in two closely related ways: (a) Visual routines are inherently dynamic and temporally extended, while affordances are thought to be directly and immediately extracted. And (b) Many affordances are extracted as a part of spontaneous perception itself, while visual routines are thought to be triggered only by specific goals. The current experiments suggest that divorcing these theoretical properties from each other may yield empirical benefits: some dynamic visual routines may be spontaneously triggered by navigational affordances — as in the visual perception of mazes.

## References

- Barsingerhorn, A. D., Zaal, F. T., Smith, J., & Pepping, G. J. (2012). On possibilities for action: The past, present and future of affordance research. *Avant*, 3(2), 54–69.
- Belledonne, M., & Yildirim, I. (2021). Automatic computation of navigational affordances explains selective processing of geometry in scene perception: Behavioral and computational evidence. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 43, No. 43).
- Belledonne, M., Bao, Y., & Yildirim, I. (2022). Navigational affordances are automatically computed during scene perception: Evidence from behavioral change blindness and a computational model of active attention. *Journal of Vision*, 22(14), 4128.
- Bonner, M. F., & Epstein, R. A. (2017). Coding of navigational affordances in the human visual system. *Proceedings of the National Academy of Sciences*, 114(18), 4793–4798.
- Chong, I., & Proctor, R. W. (2020). On the evolution of a radical concept: Affordances according to Gibson and their subsequent use and development. *Perspectives on Psychological Science*, 15(1), 117–132.
- De Leeuw, J. R., Gilbert, R. A., & Luchterhandt, B. (2023). jsPsych: Enabling an open-source collaborative ecosystem of behavioral experiments. *Journal of Open Source Software*, 8(85), 5351.
- Fodor, J. A. & Pylyshyn, Z. W. (1981). How direct is visual perception? Some reflections on Gibson's 'ecological approach'. *Cognition*, 9(2), 139–196.
- Gibson, J. J. (1967). New reasons for realism. *Synthese*, 17, 162–172.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton, Mifflin and Company.

- Harel, A., Nador, J. D., Bonner, M. F., & Epstein, R. A. (2022). Early electrophysiological markers of navigational affordances in scenes. *Journal of Cognitive Neuroscience*, 34(3), 397–410.
- Jolicoeur, P., Ullman, S., & Mackay, M. (1986). Curve tracing: A possible basic operation in the perception of spatial relations. *Memory & Cognition*, 14(2), 129–140.
- Jolicoeur, P., Ullman, S., & Mackay, M. (1991). Visual curve tracing properties. *Journal of Experimental Psychology: Human Perception and Performance*, 17(4), 997–1022.
- McCormick, P. A., & Jolicoeur, P. (1992). Capturing visual attention and the curve tracing operation. *Journal of Experimental Psychology: Human Perception and Performance*, 18(1), 72–89.
- Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17(2), 22–27.
- Pringle, R., & Egeth, H. E. (1988). Mental curve tracing with elementary stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 14(4), 716–728.
- Ullman, S. (1980). Against direct perception. *Behavioral and Brain Sciences*, 3(3), 373–415.
- Ullman, S. (1984). Visual routines. *Cognition*, 18(1-3), 97–159.
- Ullman, S. (1996). Visual cognition and visual routines. Chapter 9 of *High Level Vision*. MIT Press.
- Warren, W. H. (1984). Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance*, 10(5), 683–703.
- Warren, W. H. (2021). Information is where you find it: Perception as an ecologically well-posed problem. *i-Perception*, 12(2), Article 20416695211000366.
- Warren, W. H., & Whang, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3), 371–383.

### **Acknowledgements**

This project was funded by ONR MURI #N00014-16-1-2007 awarded to BJS. For helpful conversation and/or comments on earlier drafts, we thank Howard Egeth, Mario Belledonne, Ilker Yildirim, Joan Danielle Ongchoco, Julia Ongchoco, Marlene Berke, and the members of the Yale Perception and Cognition Laboratory. Portions of this work were presented at the 2021 *Object, Perception, Attention, and Memory (OPAM)* meeting, and the 2022 and 2023 *Vision Sciences Society (VSS)* meetings.

### **Author CRediT Contributions**

**Kimberly W. Wong:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Validation, Data curation, Visualization, Writing - Original Draft, Writing - Review & Editing

**Brian J. Scholl:** Conceptualization, Methodology, Visualization, Writing - Original Draft, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

### **Competing Interests Statement**

The authors declare no competing interests.

### **Data Availability Statement**

The raw data for all studies are included in the Supplementary Material file.



## Figure Captions

Figure 1. An example maze. See text for details.

Figure 2. Two shoelaces. See text for details.

Figure 3. A zoomed-in example of (a) a Pathwise Near Probes trial and (b) its corresponding Pathwise Far Probes trial. In each case, the probes have an equal straight-line distance from each other, but (a) has a shorter pathwise distance (7 steps) than does (b) (10 steps). In (a) the two probes are identical (with their white cutout “notches” both at the bottom), but in (b) the two probes are different (with their notches in different positions). The examples provided here both have 3 turns in the path when tracing from one probe to the other. A fully zoomed-out version of this maze is provided in the Supplementary data file.

Figure 4. Average response times in (a) Experiment 1, (b) Experiment 2a, and (c) Experiment 2b. Error bars reflect 95% confidence intervals. Asterisks indicate significant differences between conditions (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ).

Figure 5. Sample stimuli and results from Experiment 3: (a) a Pathwise Near Probes trial and (b) its corresponding Pathwise Far Probes trial. In each case, the probes have an equal straight-line distance from each other, but (a) has a shorter pathwise distance than does (b). (c) Average response times. Error bars reflect 95% confidence intervals. The asterisk indicates a significant differences between conditions (\* $p < .05$ ).

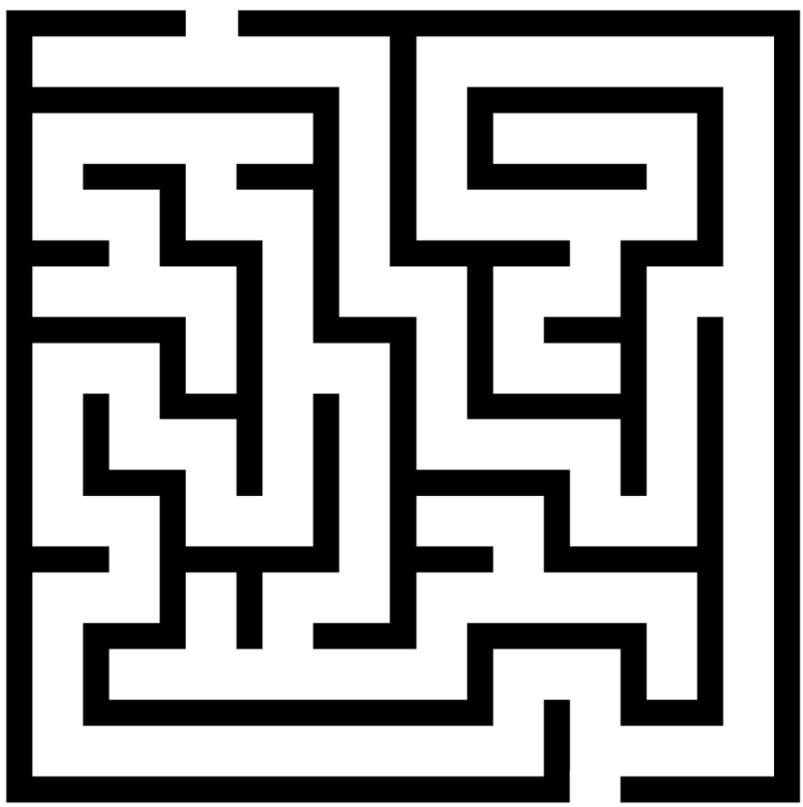


Figure 1

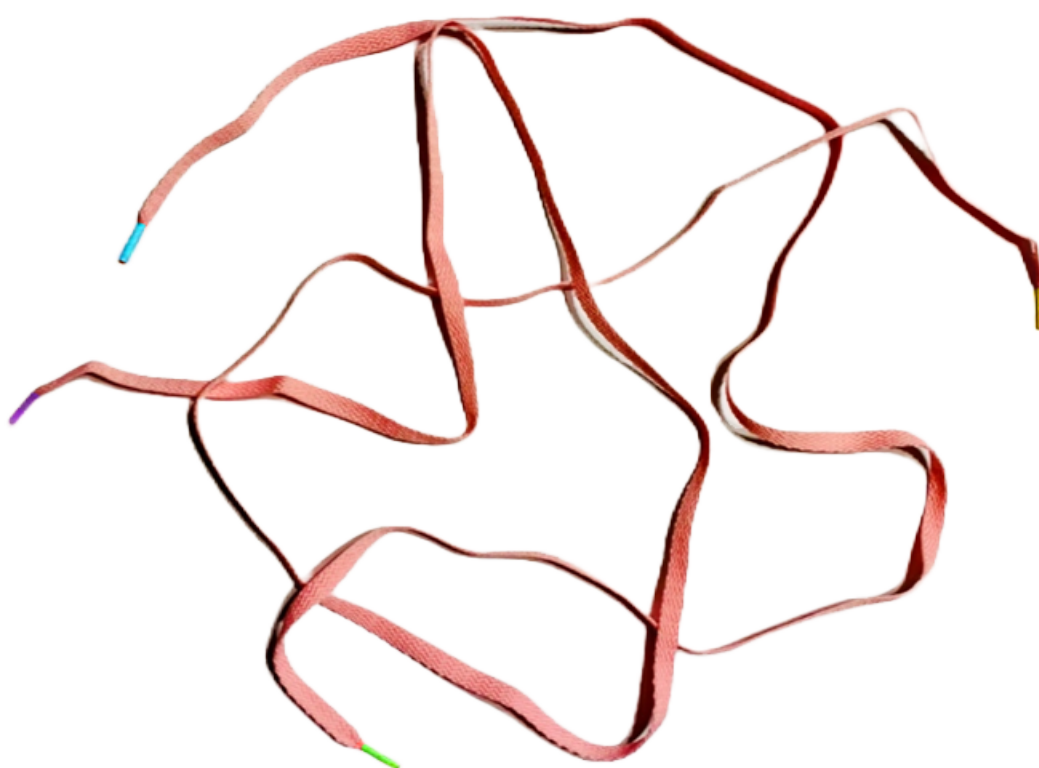
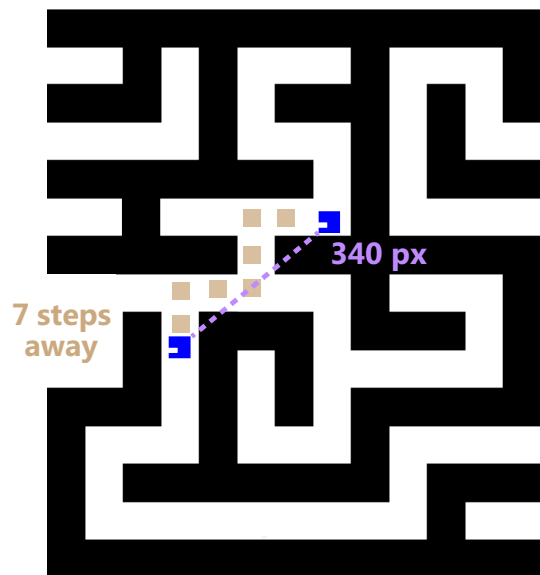


Figure 2

**a. Sample “Near” Probes**



**b. Sample “Far” Probes**

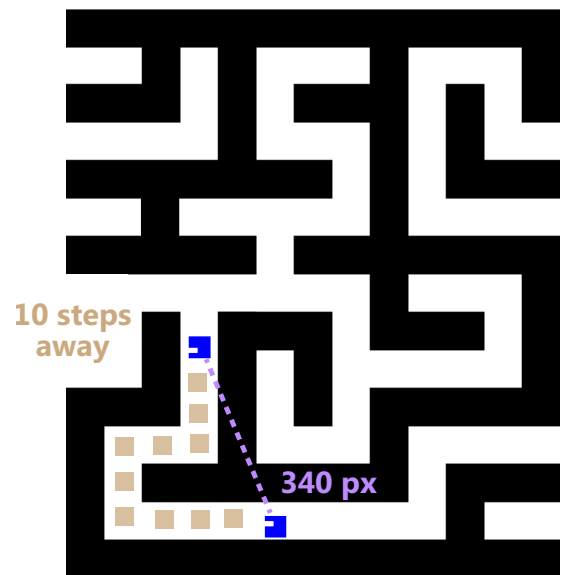
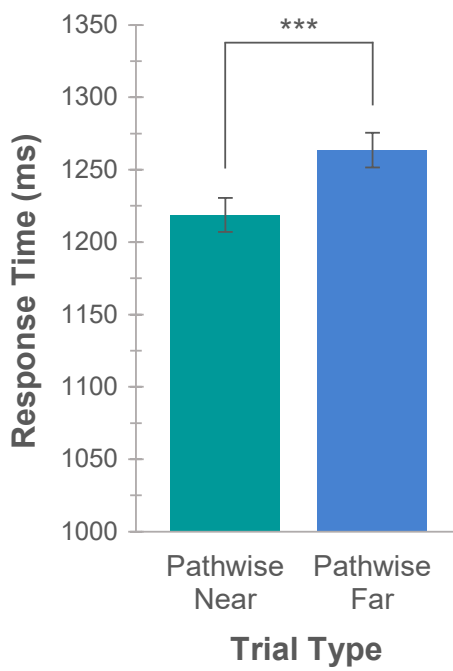
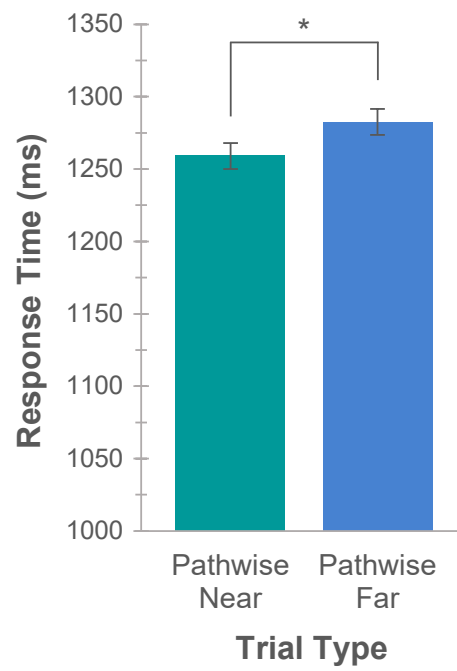


Figure 3

**a. Expt 1**



**b. Expt 2a**



**c. Expt 2b**

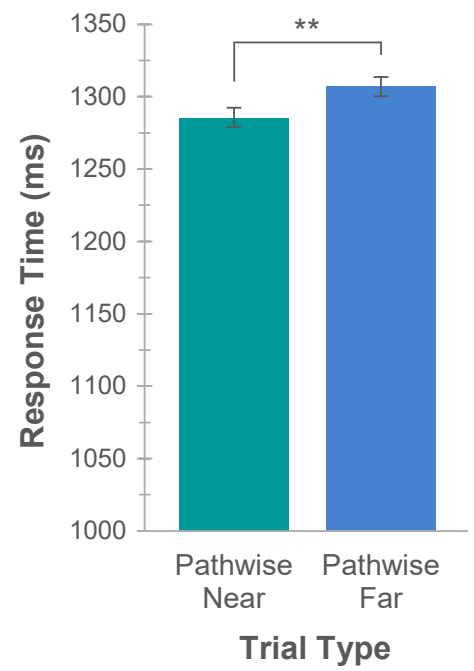
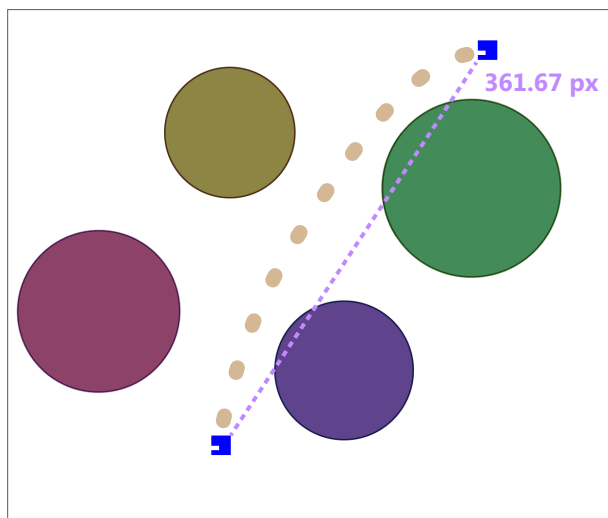
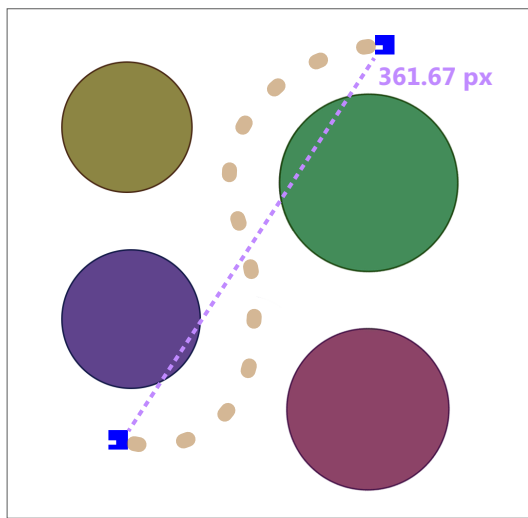


Figure 4

**a. “Near” Probes**



**b. “Far” Probes**



**c. Expt 3**

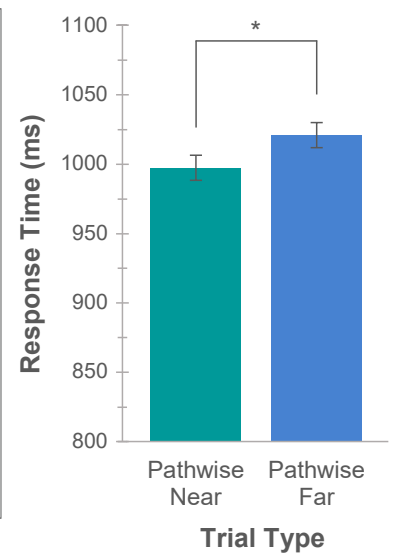


Figure 5