

Qadri, M. A., Reid, S. and Cook, R. G. (2016), Complex conditional control by pigeons in a continuous virtual environment. *Jrnl Exper Analysis Behavior*, 105: 211-229. <https://doi.org/10.1002/jeab.190>.

Complex Conditional Control by Pigeons in a Continuous Virtual Environment

Muhammad A. J. Qadri¹, Sean Reid², & Robert G. Cook¹

Tufts University¹

University of Massachusetts, Lowell²

Running head: Conditional Control by Virtual Contexts

Corresponding author:

Dr. Robert Cook

Department of Psychology

Tufts University

490 Boston Ave

Medford, MA 02155, USA

Phone: 617 627-2456, Fax: 617 627-3181

Email: Robert.Cook@tufts.edu

Abstract

We tested two pigeons in a continuously streaming digital environment. Using animation software that constantly presented a dynamic, three-dimensional (3D) environment, the animals were tested with a conditional object identification task. The correct object at a given time depended on the virtual context currently streaming in front of the pigeon. Pigeons were required to accurately peck correct target objects in the environment for food reward, while suppressing any pecks to intermixed distractor objects which delayed the next object's presentation. Experiment 1 established that the pigeons' discrimination of two objects could be controlled by the surface material of the digital terrain. Experiment 2 established that the pigeons' discrimination of four objects could be conjunctively controlled by both the surface material and topography of the streaming environment. These experiments indicate that pigeons can simultaneously process and use at least two context cues from a streaming environment to control their identification behavior of passing objects. These results add to the promise of testing interactive digital environments with animals to advance our understanding of cognition and behavior.

Key words: pigeons, virtual reality, context control, visual tracking, object identification, occasion setting

Advances in the study of behavior have been regularly tied to developments in technology (Blough, 1977; Skinner, 1938; Thorndike, 1898). Computerized immersive reality systems for humans have produced a number of important conceptual and practical advances across a wide-ranging set of domains in recent years (Biocca & Delaney, 1995). The use of virtual environments to simulate reality promises to be one of the transformative technologies of the 21st century. The ability to vary and control the properties of immersive simulations makes this virtual approach ideal for examining any number of questions that are otherwise infeasible or impossible to explore. For example, virtual reality allows the training of dangerous military, medical, aviation, and sports skills outside of their high-risk context. It is also increasingly being used in various therapeutic settings (Bohil, Alicea, & Biocca, 2011; Hoffman, 2004).

The use of virtual environments in studying behavior, brain, and cognition in nonhuman animals has also been slowly growing over the last decade (Dombeck & Reiser, 2012). Such virtual systems offer the opportunity to create complex stimulus environments that maximize ecological validity for animals while simultaneously maintaining robust experimental control over their features.

Most of the different attempts to test virtual environments with animals have concentrated on measuring how they “navigate” through such environments. Gray, Pawlowski, and Willis (2002) tested tethered moths in an immersive reality (i.e., the animal is completely visually enclosed) with a simulated 3D landscape composed from mixtures of small white and black tiles. Neural recordings from the insects suggested they responded to and “flew” in these virtual terrains like free-flying moths. The study of virtual navigation and spatial learning in insects has since been of recurring interest across multiple insect species and input modalities using primarily neural recording techniques (Fry, Rohrseitz, Straw, & Dickinson, 2008; Peckmezian & Taylor, 2015;

Sakuma, 2002). Similar methodologies have been utilized to study motor responses and prey capture in zebrafish (Ahrens et al., 2012; Trivedi & Bollmann, 2013).

Additionally, larger and more complex vertebrate animals have also been tested in digital environments. Hölscher, Schnee, Dahmen, Setia, and Mallot (2005) tested harnessed rats “free” running on a spherical treadmill to examine their navigation through a projected virtual environment that surrounded the animal. Subsequent variations of this approach have focused on measuring rodents navigating a variety of virtual spaces (Dombeck, Harvey, Tian, Looger, & Tank, 2010; Garbers, Henke, Leibold, Wachtler, & Thurley, 2015; Harvey, Collman, Dombeck, & Tank, 2009; Katayama, Hidaka, Karashima, & Nakao, 2012; Saleem, Ayaz, Jeffery, Harris, & Carandini, 2013; Sofroniew, Cohen, Lee, & Svoboda, 2014; Thurley et al., 2014).

Whereas the smaller animals were easily tested in fully immersive environments, research with larger mammals focused on simulating virtual environments that are presented frontally on a computer monitor and navigated using a joystick. For example, Washburn and Astur (2003) trained four rhesus monkeys to find specific objects in simple 3D virtual mazes. Similar to the rodent research, this digital approach has been utilized to examine place perception in primates (Sato, Sakata, Tanaka, & Taira, 2004, 2006, 2010). A summary of all of these different research approaches across species is that animals regularly behave in such environments in a manner consistent with a human 3D-like experience.

Birds are also excellent candidates for testing in virtual or immersive systems. They are highly visual, relying on this modality to actively navigate and interact with most aspects of their world (Cook & Murphy, 2012; Cook, Qadri, & Keller, 2015). Further, these object-based interactions occur at a spatial scale similar enough to humans that software and hardware are readily available.

Over the years, our laboratory has been strongly interested in how pigeons process various types of motion and “behavior” from objects and “agents” as created from digital technology (Asen & Cook, 2012; Cavoto & Cook, 2006; Cook, Beale, & Koban, 2011; Koban & Cook, 2009; Qadri, Asen, & Cook, 2014; Qadri, Sayde, & Cook, 2014). One outgrowth of these scientific questions was a considerable interest in exploring how pigeons interact with digital objects in a more complete virtual environment. Whereas others have created various kinds of 2D virtual tasks (Miyata & Fujita, 2010; Wasserman, Nagasaka, Castro, & Brzykcy, 2013), we aimed to create a task that might utilize the foraging-like behaviors of pigeons in the natural world by examining their experiences in a 3D-like simulated environment.

Thus, the goal of the current project was to begin investigating how pigeons experience a continuously streaming digital environment. Our behavioral approach differed from previous ones by not concentrating on navigation or spatial cognition as the principal means of measuring the animal’s experience. In contrast, we explored the pigeons’ “understanding” of the virtual environment by requiring different behaviors in each context. Specifically, the pigeons had to use the continuously streaming digital environment presented on the monitor to conditionally and correctly identify whether to select various objects that appeared at different positions within this environment. Unlike the navigation approach, this technique has the advantage of letting one determine and measure which features of the contextual environment control the animal’s behavior at any moment.

Each session consisted of a small number of trials that each contained an extended block of sequentially presented 3D objects appearing within different computerized, semi-naturalistic “terrains.” The birds’ task was to identify and peck at the object designated as correct for that terrain. Pecks to this correct target object resulted in reinforcement (a *hit* in signal detection

terms). Pecks to any distractor objects in that context (a *false alarm*) resulted in an additional 8 s delay to the appearance of the next object in the block. *Correct rejections* (not selecting distractor objects) and *misses* (not selecting target objects) to the objects had no programmed consequences. We used a trial-based structure of extended runs within an environment to give the birds an attentional rest from continuously searching over the long duration of a session. Furthermore, short of an object disappearing when it was correctly pecked, this was an open-loop system in that the animals' movements within the operant chamber did not affect the appearance of the streaming environment (compared to some of the closed-loop approaches referenced above, where the animals' actions appropriately and proportionally changed the environment).

By using this conditional object identification search task, we examined which and how many properties of a virtual environment influenced the birds' behavior across a series of object presentations within a trial. We report the results of two experiments. In Experiment 1, we trained pigeons to identify and conditionally select two objects based on the surface material of the streaming terrain. In Experiment 2, we increased the demands of the discrimination by adding a dynamic topography to the streaming terrains and requiring the birds to conditionally select and identify four different objects based on a conjunctive relation among the surface material and topography.

Experiment 1

The objective of the first experiment was to create and refine an operant setting in which we could test a streaming virtual environment with pigeons. In the experiment, two pigeons saw two different flat "terrains." The left panel of Figure 1 shows a screen capture of the "grass" terrain and the right panel shows the "snow" terrain. In both terrains, the pigeons received a 3D perspective view of the flat landscape, with the observer's perspective continuously moving forward through the terrain. By the end of training, two 3D-rendered objects (cubes and

pyramids) appeared within these two terrains. These objects would be encountered as the pigeons' perspective approached the objects in the terrain. They first appeared as small, "distant" objects in one of three positions (left, center, and right of the midline) on the "horizon" between the "land" and "sky." As the perspective approached the objects, they grew in apparent size and changed in position on the screen, and they disappeared off the screen to the bottom as the perspective passed by them. Each trial consisted of an extended block of these objects appearing periodically within the terrain. The pigeons' task with each encounter with an object was to identify whether or not it was the rewarded target object for that context and, if so, to select it by pecking it.

The training of this object identification task progressed in several steps and was conducted more out of caution than of necessity. The pigeons were first taught to peck the objects as they appeared on one terrain. Once object pecking was established, the perspective within the terrain began to move, while the object was centrally positioned. This was then expanded to add the two other object paths to the left and right, as well as differences in the rate of transit over the terrain. At this point, the second object and terrain were included in training. Each of these various stages of initial training proceeded quickly.

True object discrimination started with the introduction of either of the two objects appearing on either of the two terrains in separate trials. At this time, the objects differed in both their shape and color (redundant cue discrimination), and only one object type would be seen multiple times within each trial. Once the pigeons were accurately selecting the correct object within each terrain between trials, the color differences between the objects were removed. This now required the pigeons to discriminate the objects based on their shape properties (shape discrimination). Finally, in the third and last phase of the experiment, both objects were

presented within a trial. Thus, the pigeons had to differentially respond to each object within the continuous presentation of a single terrain.

Method

Animals

Two male pigeons (*Columba livia*) were tested. One (#1D) was an 8 yr. old White Carneaux pigeon and the other (#2L) was a 13 yr. old Silver King. Both had experience with a large-item visual memory task similar to one previously reported (Cook, Levison, Gillett, & Blaisdell, 2005). They were maintained at 80% - 85% of their free-feeding weights and individually caged in a colony room (12-hour L/D cycle) with free access to water and grit.

Apparatus

Animals were tested in an illuminated flat-black computer-controlled testing chamber equipped with an LCD color monitor (NEC Accusync LCD51VM). The monitor was located just behind a 33 × 20 cm viewing window equipped with a touchscreen (EZscreen EZ-150-Wave-USB) in the front panel of the chamber. Grain was delivered via a centrally placed food hopper beneath this window. A 28-V houselight was centrally located in the chamber ceiling and was always illuminated. All experimental events were controlled by Microsoft Visual Basic.

Virtual Environment and Objects

The continuous environment displayed on the monitor during each extended trial was comprised of a flat terrain rendered with one of two surface materials (see Figure 1; also see supplemental material). One surface material was primarily green with specks of brown and designed to visually resemble “grass.” The second was primarily white with specks of brown and gray and designed to resemble “snow.” These two “grounds” terminated about midway up the screen (~ 13 cm) at a “horizon” with a blue, partially cloudy “sky.” One of four cloud patterns was selected for each trial to be visible in the sky. The simulated light source illuminating the

environment came from above and behind the rendering camera. Because shadows would rarely be visible in this configuration and to promote smooth and rapid stimulus generation and playback, shadows were omitted.

The camera rendered a perspective projection of this scene. As the camera transited forward, the ground terrain scrolled by with appropriate and expanding perspective deformation of the ground and any objects passing within the camera's field of view. This created an optic flow that from the viewer's perspective looked as if they were moving over the terrain (the camera was placed at 1.5 virtual distance units high). Because of their relative distance, the sky and clouds remained stationary. The translating perspective moved forward at one of two speeds. The faster speed permitted a point on the ground to move from the horizon to off of the bottom of the screen in about 12 s. The slower speed took approximately 23 s.

Two objects were used for the object identification task. These consisted of a 3D rendered cube and pyramid (with a square base the same size as the cube). The objects had a fixed orientation throughout their presentation, with a primary surface parallel with the horizon facing toward the pigeons (see Figure 1). All objects were sized to fit within a cube 1 virtual unit on an edge, but their apparent size varied depending on their location in the terrain. Objects were approximately 0.4 cm in height when they first appeared on the horizon, and grew in size as they moved down and off the screen. When placed centrally, the front face of the cube object (which was 1 virtual unit on an edge) would grow to 5.7 cm before going off the screen. The base colors of the objects were green (RGB: 0, 255, 0) or red (RGB: 255, 0, 0). The actual color of their surfaces at any one point in a presentation depended on its location in the environment, as the illumination and their changing size created differential lighting on the objects over their duration on the display.

Objects moved in three paths (left, center, and right) depending on where they initially appeared on the horizon. A centrally placed object would appear in the midline of the environment and move in a path directly down and off-screen at the bottom of the display, as if going underneath the camera. An object placed to the left or right would appear approximately 1.3 cm off-center and proceed on a straight path with an approximately 10.5° angle from the horizontal before passing to the left or right of the camera. Objects using these two paths appeared to go off the screen immediately left or right of the camera's perspective. Objects in the left or right paths were visible for shorter durations and grew to smaller sizes relative to objects in the central one. Whereas a central object would grow to 5.7 cm, an object in the left or right path would only grow to 3.0 cm. Some part of a central object would be visible for approximately 16.3 s at the faster speed, while some part of a lateral object would be visible for approximately 14.0 s (partial visibility made these times longer than the point of the terrain value provided above). The continuous environments and the objects presented within a trial were created in real time using the Torque gaming engine (interfaced via Blockland to Visual Basic) rendering at a resolution of 800×600 pixels.

Procedure

All trials within a session were initiated by a single peck to a centrally located gray 2 cm square on a black screen. This started the virtual environment for that trial. The first object of a trial appeared on the horizon 8 s later. Trials consisted of six successive object presentations within each terrain. For the purposes of measuring object choice behavior, an object was considered selected when the pigeons pecked it five times (i.e., FR 5). Selecting a correct object (i.e., cubes on the flat grass terrain and pyramids on the flat snow terrain) removed the object from the display and permitted 2.9 s access to mixed grain from the central hopper located below the display. Hopper activations did not delay the timing of other events. Pecks to an object were

touchscreen activations within a limited region around the object. This region was initially about twice the size of the object but was gradually reduced over successive sessions as the birds learned to more precisely peck directly at the objects. Selecting an incorrect object (i.e., pyramids on the flat grass terrain and cubes on the flat snow terrain) did not remove the object during its presentation. Such pecks did cause an additional 8 s delay before the next object was presented within a trial. Following each object presentation, there was an inter-object period that lasted at least 5 s where only the moving terrain was presented. Pecks to the display during these object-free periods resulted in an additional 3 s delay before the next object presentation. Thus, individual trials could last for some time depending on a bird's pecking behavior.

Initial training. The pigeons were exposed to this procedure over several incremental steps. They were first trained to peck a static red square on a black background for food reward and then a static perspective of the 3D red cube on the grass terrain. This quickly advanced to the camera moving through the grass terrain, transiting at a slower speed than finally tested, with cubes appearing only in the central path. The cube could then appear in any of its three possible paths, and the camera progressed at the two speeds tested in the experiment. Finally, the pigeons were trained to peck at the green pyramid presented on the snow terrain. Over the several sessions it took to implement these different steps, the pigeons quickly and unremarkably learned to track the objects at each point.

Redundant cue discrimination. At this juncture, explicit discrimination training on the object identification task started. The pigeons experienced trials that contained either red cubes or green pyramids appearing on either snow or grass terrains within a trial. Each session consisted of 32 trials with 8 trials testing each of the 4 combinations of objects and terrains. The order of trials was randomized within each session. Within each trial, there were six repeated presentations of

the same object on the same terrain. Thus, there were a total of 192 object presentations per session. Each trial maintained the same randomly selected configuration of object, surface material, and camera speed for its duration. These properties were counterbalanced and equated across trials. Finally, the paths of the objects were randomized within a trial with each path tested twice. Two sessions of this type of training were conducted.

Shape discrimination. Next, the pigeons were switched to a discrimination in which only the object's shape was relevant. Thus, both red and (now introduced) green cubes were reinforced when selected on the grass terrain and green and (now introduced) red pyramids were reinforced on the snow terrain. Otherwise, the organization of the trials and sessions remained the same as in the previous phase, with each trial having a randomly selected configuration of object, surface material, and camera speed that was counterbalanced across a session. Initially, pigeon #1D often failed to peck at the red pyramid at the beginning of this phase. To promote pecking behavior, this bird was given several sessions in which access to the mixed grain was automatically provided following presentations of the red pyramid in its context regardless of whether it was selected or not. This quickly fixed the issue and was stopped. Shape discrimination training continued for 25 sessions.

Within-trial object discrimination. Finally, we introduced trials containing mixtures of both objects with both colors. Each trial now consisted of 12 successive object presentations. Within a trial each of the objects was tested in each of the three object paths ($2 \text{ objects} \times 2 \text{ colors} \times 3 \text{ paths} = 12 \text{ presentations total}$). Across trials, the terrain features were counterbalanced in the same manner as above. Sessions now consisted of 16 trials (192 total object presentations per session). Eight sessions of this type of within-trial discrimination training were conducted.

Results

Overall, the pigeons exhibited little difficulty in learning the object identification task and its relation to the contextual properties of the digital terrain over the three phases of Experiment 1. Figure 2 shows this performance measured as d' across the three phases. This was calculated from the rate of *hits* (percent selection of correct objects) and *false alarms* (percent selection of incorrect objects) in each session. The d' statistic is calculated as $d' = \Phi^{-1}(\text{hit rate}) - \Phi^{-1}(\text{false alarm rate})$ where Φ is the cumulative distribution function (cdf) for the standard normal probability distribution (Macmillan & Creelman, 1991). Because the objects were repeated within a trial for the first two phases, and thus subsequent responses could have been based on recent reinforcement history within the terrain, d' was calculated based on only the first object presentation within a trial for those two phases. To avoid redundancy, detailed analyses of within-trial dynamics are described only for the third phase.

During redundant cue discrimination, the pigeons quickly appeared to distinguish the green pyramid and red cube on their correct terrains relative to incorrect backgrounds (leftmost panel of Figure 2). Across the two sessions, object selection accuracy increased quickly. This discrimination was not solely due to a neophobic reaction on trials containing the newly introduced incorrect terrain and object configurations, which might partially account for this accuracy. Both pigeons started off pecking at the objects during the first few trials that were presented on these new backgrounds, but they quickly stopped within the very first session.

When we started training with only the object's shape as the discriminative cue, object selection accuracy quickly dropped (see middle panel of Figure 2). This drop for each pigeon indicates that they used the redundant color cue to discriminate in the first phase. Nevertheless, within a few sessions both pigeons' object selection accuracy significantly increased as they began to use object shape to determine their choices. Evaluating these data in 4-session blocks

revealed that by the second block (sessions 5-8) accuracy was significantly above a chance d' value of 0 (single mean $ts(3) > 4.0$, $ps < .03$; an alpha level of $p < .05$ was used for all analyses).

The introduction of the seemingly more demanding within-trial discrimination, where any object could appear within a single trial, had little impact on the birds' ongoing object selection accuracy (see right panel of Figure 2). There was little or no drop in accuracy with this shift, despite the requirement for each of the six objects to be assessed during a trial. Over these 8 sessions, each pigeon's object selection accuracy was significantly above chance, single mean $ts(7) > 5.8$, $ps < .001$.

During this final phase, the average trial duration for the 12 presentations across the birds was 244.7 s (bird #1D = 244.1 s, bird #2L = 245.3 s). Despite this extended time, accuracy did not change over the time course of a trial. For the first three object presentations, the d' was 3.9 for bird #1D and 3.8 for bird #2L, whereas for the last three object presentations the d' was 4.3 for bird #1D and 3.6 for bird #2L. There was no significant effect of object serial position for either bird as determined by an ANOVA that compared each session's d' across presentation order, $Fs(11,77) < 1.4$.

We next examined performance in this third phase to determine if there were any effects of terrain, object path, or camera speed on the pigeons' abilities to correctly identify objects or terrains. For this, we separately computed d' for each level of a factor and conducted repeated measures ANOVAs to examine if these properties affected the pigeons' performance. These analyses revealed that both birds were very good and above chance at object identification on each terrain, all $ts(7) > 4.2$. Pigeon #2L exhibited no significant difference in d' accuracy with either surface material (grass mean = 3.2; snow = 4.4; $F(1,7) < 1$), whereas pigeon #1D was

significantly more discriminating with the grass surface material (grass = 5.5; snow = 3.9; $F(1,7) = 14.7, p = .006$).

An object's speed or trajectory also had little impact on performance. Analyses revealed that the speed of the approaching object (mean d' fast = 4.6; slow = 4.8) had no significant influence on pigeons' object selection accuracy, $F_s(1,7) < 3.2, p_s > .12$. The objects' path had a small, but significant, effect, $F(2,14) > 3.8, p_s < .048$. The birds were slightly better with the lateral paths (mean d' left = 5.0; right = 5.0) than with the central one (center = 3.9). This difference stemmed from their making slightly more false alarms on presentations of central objects rather than lateral ones. A repeated measures ANOVA on accuracy on the negative trials showed decreased accuracy (i.e., increased false alarms) when the object was centrally placed as compared to on the sides $F_s(2,14) > 11.0, p_s < .001$.

Finally, we examined the time course of pecking behavior and object selection within the presentation of a single object. Shown in Figure 3 are the probability distributions for all pecks across time relative to the onset of the object's appearance during this phase. Regardless of its identity, when a correct target object appeared on the horizon, it generally generated immediate and considerable pecking behavior. The pigeons did not appear to have to wait any appreciable length of time for it to get closer. This is reflected in the high percentage of *hit* pecks within the first few seconds in Figure 3. The pecks that were made during correct rejections and false alarms were far more diffuse over time for both animals and clustered later in time. Of the two, bird #1D showed more loss of stimulus control with time for both types of responses as reflected in the greater number of pecks to these types. A similar increase in later responding can also be seen in #2L's *false alarm* pecks. The loss of inhibition reflected in these later occurring pecks are likely the result of the increasing size and salience of the objects as they neared the camera and

potentially the hopper's location. Finally, #2L shows a peak in its peck distribution for correct rejections that is near the main peak for hits or correct object selections. This suggests that #2L may have first localized the target and distractor objects, and then stopped pecking the latter after a few "observation pecks" that did not meet our FR 5 object selection criterion.

Discussion

Experiment 1 found that both pigeons had little difficulty identifying and tracking the correct objects within their assigned digital terrain or context. This was true whether the objects were redundantly defined by color and shape or distinguished only by the latter. In both cases, the pigeons readily learned to peck at correct objects when they first appeared on the horizon and suppress pecking at the incorrect objects as they progressed along and out of the display. Overall object selection accuracy was high on both terrains. This was true most importantly when the objects were mixed for extended minutes within the same terrain and trial. The speed of the moving terrain had no impact on this accuracy, and the object's path only had a small effect. In sum, the pigeons virtually had no challenges in sustaining attention and accurately performing their object identification task as visually configured by the features of the current virtual environment.

The only minor difficulty may have been that both pigeons showed a loss of stimulus control as negative objects continued toward them on the screen. Specifically, as a distractor object grew nearer, the likelihood of pecking it increased for both animals. As mentioned, this loss of inhibition might have been caused by the increased size and salience of the objects as they neared the camera and also neared the location where food was delivered (i.e., the food hopper). In differential reinforcement of other behavior (DRO) paradigms, pigeons have been known to disengage from the screen, including turning around and engaging in grooming behavior (i.e., "other behaviors"). We have similarly observed other pigeons moving or looking away from S-

displays on the negative trials of go/no-go discriminations. Had this been a closed-loop system, we might have observed the pigeons navigating away from such negative objects in addition to not selecting them (cf., Wasserman, Franklin, & Hearst, 1974). The continuous and sustained need in the current search task to keep looking for the target objects may have made this kind of avoidance behavior difficult for the pigeons to employ. The peak of late pecking may also reflect a growing expectation of a potentially better object appearing in the immediate future.

Of course, the central and pivotal question is whether any of this realistic virtual simulation really mattered to the birds. This question will be engaged again more completely in the general discussion. For the moment, an appropriately skeptical view is that it did not. Despite the high-end graphics and extravagant object/terrain nomenclature, the critical feature of the terrain provided a consistent and large static color cue to which the birds could have attended in order to perform the discrimination (i.e., is the lower portion of the screen currently mostly green or mostly white). From this perspective, we simply established that pigeons can use a shape-based cue of some type to respond to and track a moving stimulus contingent on a large surrounding color cue, and nothing more. If the screen's background had simply been solid white or solid green and if during the trials a simple 2D square or triangle had moved down the center of the screen, then few would be surprised that the pigeons could learn the discrimination.

Is there more here? The ease and rapidity of the learning and the birds' high accuracy does suggest that all the extra and potentially distracting "realistic" features (i.e., dynamic paths, changing object size, moving terrain features, different rates of transit) did not appear to greatly hurt their performance. Given the consistency of such perspective and motion cues with what the birds naturally experience as they move in reality, perhaps the birds' ready accommodation of the features in this virtual setting is unsurprising. Moving over terrains constantly looking for

correct objects (i.e., food) is what they naturally do. It is not clear how similar variability would impact the simpler 2D discrimination outlined above. Of course, most of this is just conjecture. Experiment 2 tries to address some of these concerns.

Experiment 2

In Experiment 2, we added a new dynamic cue that we hoped would require more complete processing of the virtual environment than might be demanded by the colored surface material. For this purpose, we introduced terrain topography as a conditional cue for the pigeons' object identifications. We did this because our previous research had found that birds are capable of seeing convex and concave surfaces that have shapes similar to changes in the topography of a terrain (Cook, Qadri, Kieres, & Commons-Miller, 2012; Qadri, Romero, & Cook, 2014).

To create this cue, we constructed a terrain of endlessly streaming "hills" and "valleys." We rendered the scenes with the camera appropriately moving up and down these virtual hills in a dynamic fashion, with the objects located intermittently in the valleys between them. This required integrating the varying visual cues to determine whether a terrain was flat or not. This we suspected would challenge any 2D processes the pigeons might have used to recognize objects by placing them with novel 3D-like perspectives. On the flat topography, the objects were always visible when they appeared on the horizon. In the "hills" topography, however, objects could be surprisingly close and lower when first encountered after the crest of a ridge or appear to be on the next slope. Seeing these familiar objects under such conditions would be most coherent if the pigeons perceived the 3D nature of the environment and objects.

With the addition of dynamic topography (flat vs. hills) and the continued use of surface material (grass vs. snow), we could now create four continuous terrains or contexts in which the pigeons could encounter different objects. Because of this, we also added two new distinctively shaped objects to be identified in addition to the two old ones. Using distinct objects that

minimized inter-object confusions allowed us to infer the attention the pigeons paid to the properties of the surrounding context by examining their different object choices. Accordingly, for any particular context, the birds' errors in mistakenly selecting one of the three distractor objects should reveal which aspects of the context (terrain, topography, or their combination) were most influential at any one moment. Finally, this discrimination was now a conjunctive conditional relation between the two different contextual cues, because it required that two context features (surface material and topography) be simultaneously discriminated to successfully perform this more complex object identification task.

Method

Animals & Apparatus

The same animals and apparatus were used as in Experiment 1.

Virtual Environment and Objects

A new "hills" topography and two new objects were added. The hills topography varied sinusoidally, creating a series of alternating virtual hills and valleys (see Figure 4; also see supplemental material). Thus, as the camera transited forward, it appeared to ascend and descend over the terrain. The camera's perspective turned upward at the beginning of a hill until it reached the top, at which point it pivoted and pointed down the slope. At the fast speed, this terrain would be traversed (hilltop to hilltop) in 10.9 s, and at the slow speed, it would be traversed in 21.8 s. In these terrains, the objects always appeared in the valleys between the peaks and closer to the far hill than the one underneath the camera (to ensure that objects could be pecked in time). As a result, the objects' visible durations on the displays were shorter than in the flat terrain (lateral placements -5.0 s; central placement -5.6 s). Objects were appropriately tilted so that their base was parallel to the ground and appeared to sit properly within the terrain.

Both the flat and hills topographies could be created with either the grass or snow surface materials.

The two new objects were a sphere and a complex prism with eight extended star-like outward points (see examples in Figure 4). They were sized similarly to the other two objects (1 virtual unit). The sphere was correct and reinforced in the context of “grassy hills” and the pointed object was correct and reinforced when it appeared on the “snowy hills.” The terrain assignments for the two old objects remained the same. Finally, because we were concerned about the similarity of the green object color and the green surface material, we replaced this color before this experiment for the objects. So, in addition to the continued use of the red base color, we introduced, trained, and tested a blue base color for all of the objects (RGB: 0, 0, 255).

Procedure

Initial training. Training with the new objects was conducted in a manner similar to the later stages of previous training in Experiment 1. The pigeons first received two sessions with the two new objects presented within only their positive “hills” terrains. These objects were randomly colored either blue or red. The pigeons quickly learned to select the two introduced objects. They then received four sessions of shape discrimination training in the “hills” contexts by adding the negative surface material for each object. This phase was also quickly learned. Thus, by the end of this training, the pigeons were correctly selecting the new objects on their corresponding positive terrain (“grassy hills” or “snowy hills”) and not doing so on the selected portion of the negative terrains (respectively, “snowy hills” or “grassy hills”).

Conjunctive cue discrimination. At this point, we reintroduced trials with flat terrains to make both topography and surface material relevant cues. In addition, the green objects tested in Experiment 1 were switched to be blue. Thus, the pigeons were trained to conditionally discriminate the shape of four objects based on the correct conjunction of topography and surface

material features in order to receive a food reward. Each session consisted of 96 total object encounters made from eight trials. Each trial consisted of 12 randomly intermixed object encounters (3 target and 9 distractor presentations for that context). Each object appeared once in each path within a trial. Within a session, all terrains were tested twice and all objects were scheduled to appear once in every path and every color. Half the trials were conducted at the fast speed and half at the slow speed. This training and testing lasted for 24 sessions.

Within-trial terrain discrimination. Finally, the pigeons were trained in a situation where the four terrains themselves changed as a trial progressed. Each session now contained 64 total object encounters presented across four trials. Each trial contained 16 object encounters (4 target and 12 distractor presentations) divided among the four contexts. Within a trial, shifting terrains started at the horizon and progressed as the camera approached the dividing line. Terrain shifts occurred only after every object was seen on that terrain (thus, after each block of four different objects). Object presentations were delayed until the terrain had fully changed, so only the fastest speed was used. The order of appearance for the four terrains was randomly determined for each trial. Object color and path was randomly determined for each encounter. Other timing and trial features remained the same as in Experiment 1. This mixed terrain testing lasted for 12 sessions.

Results

Both pigeons learned the relations between the four objects and the four terrains. This can be seen in the upper and lower panels of Figure 5 showing the session-wise d' values for each pigeon. The rapid upward trend over the first few sessions and subsequent high d' values indicate how quickly and accurately the pigeons came to perform the discrimination. Thus, despite reward being dependent on the complex conjunction of object, topography, and surface material cues, the pigeons were very good at the task.

The three lines for each pigeon reveal the types of errors they made from their particular object choices. For any particular context (e.g., “snowy hills” as a working example), each of the three distractor objects represented a different possible type of error, depending on the context in which it was the correct object. For example, one distractor object would be correct on a terrain that was similar to the current one in neither topography nor surface material (i.e., “flat grass”), a second distractor object would be correct on terrain sharing topography cues (i.e., “grassy hills”), and the third would share surface material (i.e., “flat snow”).

Using the percentage of hits to correct objects adjusted for each terrain, we calculated the frequency of *false alarms* to each of the three distractor objects. These separate *false alarm* rates were used to compute separate d' values for each distractor type. The higher d' curve for the redundant condition in Figure 5 reflects the fact that across the terrains, the pigeons were best at not selecting the distractor object associated with terrains that differed in both features from the currently visible terrain. Thus, at the beginning of conjunctive cue testing, both pigeons rejected best the distractor objects whose correct terrain shared no features with the currently visible one. This redundant facilitation in both birds suggests they were sensitive to differences in both surface material and topography from early in testing. Next, both pigeons rejected best those distractor objects associated with a different surface material than the currently visible one. Given the explicit training to discriminate among surface materials in Experiment 1 and the initial training of the current experiment, this is not surprising. Finally, the most frequently selected incorrect object (i.e., lowest d') was the one associated with a different topography, but sharing the same surface material. Thus, the newly introduced topography cue took several sessions to be incorporated into the birds' ongoing discrimination.

A regression analysis of each bird's first eight sessions confirmed an effect of these three conditions ($F_s(2,20) > 4.2$, $p_s < .031$), as well as a linear effect of session ($F_s(1,20) > 5.1$, $p_s < .036$) for each bird (the session values were centered prior to the regression). By the end of conjunctive cue training, however, the differences between terrain conditions were absent and difficult to see or isolate as both pigeons became very good at selecting only the correct object within each of the four terrain contexts.

When we subsequently introduced the phase where the terrains shifted frequently within a trial, both pigeons were largely unaffected. This is seen by the continued excellence in discrimination by both pigeons in the right panels of Figure 5 showing the results of this mixed terrain testing. A regression analysis of the each bird's 12 sessions found no significant effects of session, ($F_s(1,32) < 1.3$), or condition, ($F_s(2,32) < 2.1$), for each bird (again, the session values were centered prior to the regression). To examine the effect of terrain changes within a trial, we computed d' separately for the presentations that immediately followed changes in surface material, versus changes in topography, versus changes in both.

For bird #1D, changing both surface material and topography ($d' = 4.6$) and changing only the topography ($d' = 4.6$) yielded similar performance, but changing just the surface material resulted in slightly worse performance ($d' = 3.4$). For bird #2L, changing both aspects ($d' = 3.1$) yielded better performance than changing only surface material ($d' = 2.9$), which was still better than changing topography alone ($d' = 2.4$). We found little change in performance over the course of the multiple object presentations within a terrain. Whereas bird #1D had a d' of 3.4 on the first object of a terrain, he ended with a d' of 3.0 on the last object, and bird #2L started with a d' of 2.4 and ended with a d' of 2.6.

We next examined the birds' performance over successive object presentations within a trial to determine whether the extended search time (> 14 min.) created any difficulty. Shown in Figure 6 are the d' values across presentation time within a trial averaged over all sessions. Although the first minute reliably had at least 96 object presentations, the remaining minutes contained from 75 to 24 presentations depending on how the birds pecked and the composition of terrains within a trial. We truncated the analysis at 14 minutes, because there were so few observations beyond that time. Both pigeons exhibited good discrimination across the extended trials. Bird #1D seemed to maintain a steady discrimination throughout a trial, whereas bird #2L seemed, if anything, to improve over the course of a trial.

Lastly, we tried to capture whether the pigeons recognized the difference in the terrains upon their appearance. For this, we collected those pecks that occurred to the display before objects were scheduled to appear. As the birds were trained to inhibit pecking during this time, the number of these "anticipatory" pecks was limited (on average 4 pecks per inter-object period for #1D, 1 peck for #1L). We then examined where on the display these pecks were located. From visual observation, we knew that pigeons on the flat terrain often pecked near the horizon, likely because new objects appeared at that location. Further, during the hilly terrains, the pigeons often appeared to peck lower on the displays, again perhaps because objects regularly first became visible to the birds on the lower parts of the screen, corresponding to the valleys of these terrains.

Displayed in Figure 7 is the mean vertical position of each bird's pecking behavior to the flat and hilly terrains. In order to ensure that these were anticipatory pecks, only pecks that occurred more than 1 s after the previous object (or warning signal) had disappeared are included. Across their serial encounters in the display, both birds pecked significantly lower (#1D, $Y = 318.1$ pixels; #1L, $Y = 286.4$ px) on hilly terrains than on flat terrains (#1D, $Y = 401.5$ px; #1L, $Y =$

355.0 px; independent sample $t(30) > 4.3$, $ps < .001$ for each bird). Similar examinations of the horizontal position found no differences between these two terrains for #2L, $t(30) < 1$, $p = .4$, and a significant rightward shift for #1D, $t(30) = 5.9$, $p < .001$, on the hilly terrain.

Discussion

Experiment 2 revealed that pigeons could conjunctively combine multiple context cues to control their search behavior for four different objects. Across all four terrains varying in surface material and topography, the pigeons were able to accurately identify and quickly peck the correct object for each context. They showed considerable flexibility in being able to do so whether the terrain varied between trials or smoothly shifted within an extended trial. Anticipatory pecks made to the different topographical terrains prior to object appearance suggested the pigeons knew where objects were likely to appear for each of the contexts. Finally, the pigeons were able to sustain attention in this task for an extended duration of time, showing little variation in accuracy even when faced with changing environments portrayed for more than 14 minutes at a time.

The types of *false alarms* made to the different distractor objects at the beginning of the experiment help to illuminate what the pigeons were attending to in these contexts. Based on their previous training, it appears the pigeons were often attending to the nature of the surface material as a source of information early in the experiment. They most frequently pecked at the distractor object that shared its surface material assignment with the current one. The topography cue took a session or so to be incorporated, and with its processing came a corresponding rise in accuracy. However, the pigeons incorporated this cue quickly, a bit to our surprise. They may have benefited by being better at not selecting the distractor objects that failed to share either cue with the current context. This redundant benefit shown by both birds suggests that both cues were salient even at their introduction, even with surface material being more dominant early on.

Despite the ease by which topography came to control object selection behavior, it is unclear how the pigeons processed and recognized this new cue. One possibility is that they recognized the contexts using visual or perspective cues, seeing them as flat or hilly terrains (Cavoto & Cook, 2006; Cook et al., 2012). If this is the case, then a combination of linear perspective, object interposition, occlusion, size changes, and textured gradient cues could all have contributed to context recognition. As suggested from visual observation and confirmed by differences in the location of anticipatory pecks, the pigeons did appear to be able to recognize where objects were likely to appear based on the visual or dynamic appearance of the context. Static visual cues have been demonstrated to support scene-based spatial cuing in the pigeon (Wasserman, Teng, & Brooks, 2014), so perhaps this anticipatory behavior reflects simple visual feature recognition.

A second and more speculative possibility is that the pigeons encoded the different terrains as the experience of the camera motion through the context (Cook, Shaw, & Blaisdell, 2001; Frost & Sun, 1997). Powerful sensations of self-motion can be induced in humans by camera movement in theatrical presentations and in virtual realities (i.e.,vection). For flat terrains, this experience would be a straight advancing motion sensation over a plane, as opposed to the up and down motion sensation within the hills context. As the primary means to tell which terrain was being presented, our own experience was that visual cue was more important than any encoding of camera motion or self-action. That being said, it is intriguing to think about whether animals could encode and distinguish experiences by the context of their ongoing actions (e.g., walking vs. flying vs. diving/swimming). Experiments which put self-motion cues and perspective visual cues in conflict would be interesting to evaluate this notion – and are primarily made possible using virtual reality.

Finally, the pigeons had little difficulty with sustaining attention over the course of these long trials. There was no decline in accuracy during a trial. Although sustained attention has received considerable discussion in the human literature, its investigation in animals has not been as great. That the pigeons had little trouble is again unsurprising given that the current task is little different from what they naturally do for extended periods of time – forage. Instead of simple attentional strain from extended duration trials, however, a more interesting avenue for future research would be to see what demands on sustained attentional capacity could be induced by varying stress (Dukas & Clark, 1995) or information load (Riley & Leith, 1976).

General Discussion

These two experiments demonstrate that pigeons can readily process object information in a continuously changing and streaming virtual environment. The pigeons conditionally selected different objects as they successively encountered them across four terrains as composed from different surface materials and topography. During Experiment 2, the birds seemed to be engaged in different scanning behaviors in anticipation of objects' appearance depending on topographical context cues. Finally, by the end of the experiments, the pigeons were attending to these environments and searching for objects for extended time periods (> 14 minutes). These outcomes are consistent with the possibility that the pigeons were processing this streaming visual information much like an object-filled natural environment. With that said, because we are testing nonhuman animals, we must never lose sight of the illusion created from the easy use of natural language to describe these virtual methods. Evidence and caution are required before making strong claims as to the real experience of the animal in such settings (Morgan, 1896).

To perform accurately in the last phase of the experiment, the pigeons certainly needed to make at least three discriminations simultaneously. The first was to discriminate among the shapes of four different foreground objects. The second required discriminating the background

surface material. The third required discriminating the nature of the topography. It was only the correct conjunction of all three values (the correct object in the correct topography and correct surface material) that could lead to consistent food reward and avoid delaying opportunities for future rewards by pecking at incorrect objects.

As a visual search task, these discriminations within the digital environment were easy to learn and perform by both pigeons. In fact, the nature of the visual configuration itself may have made it so easy and apparently “natural.” Of course, this is the pivotal question. Did the pigeons truly experience the shifting, complex, continuous, virtual stimulus environments as anything like the real world and was that illusion responsible for the excellent discrimination?

A skeptic of such a dynamic 3D interpretation hypothesis would notice that all three conditional discriminations of objects, terrain, and topography can be reduced to simpler 2D features. The surface material cue can be easily reduced to a color discrimination. Likewise, the dynamic topography cue, although not perfectly reducible to a simple static cue, may be identified by simple 2D patterns or movement cues that might not demand a 3D interpretation. The objects were designed to be easily discriminated from each other and so plenty of 2D features also reside in them. Thus, there is not one conclusive piece of evidence that would dissuade such a skeptic. Similar kinds of concerns have been raised with regard to simpler picture perception of static images and their possible correspondence to the real world (Bovet & Vauclair, 2000; Weisman & Spetch, 2010).

That caution acknowledged, the really intriguing possibility is that the pigeons were experiencing the display as designed and intended – a changing 3D virtual environment in which objects were successively encountered (i.e., a virtual walk in the park). The visual organization of this environment strongly induced that impression in humans.

Many of the elements needed to support such an impression have been shown to operate in and to be perceived by pigeons. Strong optic flow perception would be needed to induce the apparent forward movement created by the expanding field created from the terrain and the expanding size of the approaching objects. Given the visual demands of flight, it is not surprising that pigeons have this capacity (Frost & Sun, 1997; Wylie & Frost, 1990). They can also discriminate the relative coherence of separated texture-like elements (Nguyen et al., 2004).

We have found that pigeons also have the capacity to recognize concave and convex patterns as portrayed in receding 3D shaded planes, even when in motion (Cook et al., 2012). Additionally, there is good evidence that pigeons can use the many monocular cues present in such rich environments to perceive depth and perspective (Cavoto & Cook, 2006; Reid & Spetch, 1998). Finally, the pigeons also seem quite capable of tracking a variety of moving objects (McVean & Davieson, 1989; Wilkinson & Kirkpatrick, 2009). Each of these previous outcomes is consistent with the possibility that the pigeons in this experiment could have experienced the shape and perspective of these terrains as intended.

After all, avian visual mechanisms did not develop in the ecological vacuum of the operant chamber. Rather, they evolved in a visually rich environment enabling the extraction of exactly the kind of complex and dynamic information portrayed here (Cook et al., 2015). One hypothesized source for the number of apparent divergences between avian and mammalian testing is the impoverished and abstract visual nature of the experimental operant setting (Qadri & Cook, 2015). From this perspective, experiments with more full-featured or realistic displays, similar to the current situation, best capture the true visual abilities of birds.

One line of evidence that might persuade a skeptic that our pigeons truly saw a virtual world would be transfer of the discrimination to the real world. This approach pivots around the same

issues and limitations of whether animals share a correspondence between their experience of 2D pictures and 3D objects (Weisman & Spetch, 2010).

For humans, correspondence has been shown in many designs, such as exhibiting similarities in acquisition and mechanisms (e.g., Sturz, Bodily, & Katz, 2006; Sturz, Bodily, Katz, & Kelly, 2009). Previous investigations on correspondence with pigeons expose the animals to a stimulus (e.g., a conspecific, 3D object, or location) and examine subsequent responses to pictorial representations of that stimulus (e.g., a picture or video) or vice versa (Dawkins, Guilford, Braithwaite, & Krebs, 1996; Shimizu, 1998; Spetch & Friedman, 2006). For the present experiment, such a correspondence might be found in several ways.

One way would be to see real physical compensations by the animals to actions in the virtual world, such as reacting appropriately to different types of camera motions or taking a defensive reaction to a looming object. We saw none of those in the different times we watched these pigeons perform this task, but such instances are surely rare and likely best observed only on the first few encounters. Another approach would be to look for explicit or implicit transfer of experiences in either type of world to the other one. Ideally, any such virtual to reality transfer would be more than just responding appropriately to otherwise “identical” visual stimulation. If the pigeon cannot differentiate between the virtual and real situations, then what really is being tested or transferred? The best evidence would stem from experiences in one world being the indirect origins of actions taken in the other – e.g., being able to navigate in reality to a novel location or along a novel path based exclusively on relations learned virtually (cf. Cole & Honig, 1994; Tolman, 1948; Tolman, Ritchie, & Kalish, 1946).

Unlike the majority of previous attempts to examine virtual realities with animals that have focused on navigation as the main measure of simulation, we focused in these experiments on

determining whether the pigeons were capable of using features of the surrounding context to modulate their behavior. In Pavlovian conditioning settings, the analysis of associative control by context has been extensively researched. One of the persistent problems within this study has been the challenge of controlling and manipulating such contexts. Testing rendered virtual environments would help considerably on the latter front.

One view is that a conditioning context is a form of stimulus that as a compound gains and loses strength as an independent unit from the conditioned stimulus (Rescorla & Wagner, 1972). Although the objects certainly acted as direct signals for reinforcement, it is not so clear that the nature of the different virtual environments tested here operated in this fashion. It makes more sense to us to consider these contexts as serving a modulatory role, effectively telling the animal what is now the appropriate object for which to search. In this sense, the virtual contexts seem to us to be acting more like occasion setters that activate the correct object representation to search for in each terrain (cf. Holland, 1992).

There are many avenues to pursue using this new approach. One of the more important directions to add is a direct coordination between the pigeons' actions and the virtual environment. As mentioned, the current approach is an example of an open-loop system in which the animal's behavior had very minimal impact on the displays. In the present case, they could only make the objects disappear when pecked or prolong the delay until the next object appeared. Closed-loop systems tie the perspective's motion to the animal's movement. The latter has much appeal, but is less practical for many investigations given the difficulties of creating appropriate responses and sensors for each species. Given the smaller size of pigeons, however, this is much more practical than for primates, at least for walking (virtual flight is an altogether different complication). Without being completely immersive or reactive, we doubt the pigeons ever

became confused and thought they were in the “natural” world. The operant chamber was always right there, with only a virtual portal into another perspective. With that said, however, we think the intended realism of the current displays likely strongly activated the same visual mechanisms that birds use to perceive the event structure of the natural world. Finally, there is no need to limit digital environments to a post-apocalyptic world of lifeless objects – digital animals such as conspecifics, prey, or predators could be animated and behave within digital environment to great interest and advantage (Asen & Cook, 2012; Qadri, Asen, et al., 2014; Qadri, Sayde, et al., 2014; Watanabe & Troje, 2006).

References

- Ahrens, M. B., Li, J. M., Orger, M. B., Robson, D. N., Schier, A. F., Engert, F., & Portugues, R. (2012). Brain-wide neuronal dynamics during motor adaptation in zebrafish. *Nature*, 485(7399), 471-477. doi: 10.1038/nature11057
- Asen, Y., & Cook, R. G. (2012). Discrimination and categorization of actions by pigeons. *Psychological Science*, 23, 617-624. doi: 10.1177/0956797611433333
- Biocca, F., & Delaney, B. (1995). Immersive virtual reality technology. In F. Biocca & B. Delaney (Eds.), *Communication in the age of virtual reality* (pp. 57-124): Lawrence Erlbaum Associates, Inc.
- Blough, D. S. (1977). Visual search in the pigeon: Hunt and peck method. *Science*, 196(4293), 1013-1014. doi: 10.1126/science.860129
- Bohil, C. J., Alicea, B., & Biocca, F. A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews Neuroscience*, 12(12), 752-762.
- Bovet, D., & Vauclair, J. (2000). Picture recognition in animals and in humans: A review. *Behavioral Brain Research*, 109, 143-165. doi: 10.1016/S0166-4328(00)00146-7
- Cavoto, B. R., & Cook, R. G. (2006). The contribution of monocular depth cues to scene perception by pigeons. *Psychological Science*, 17, 628-634. doi: 10.1111/j.1467-9280.2006.01755.x
- Cole, P. D., & Honig, W. K. (1994). Transfer of a discrimination by pigeons (*Columbia livia*) between pictured locations and the represented environments. *Journal of Comparative Psychology*, 108(2), 189-198. doi: 10.1037/0735-7036.108.2.189

- Cook, R. G., Beale, K., & Koban, A. (2011). Velocity-based motion categorization by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 37(2), 175-188. doi: 10.1037/A0022105
- Cook, R. G., Levison, D. G., Gillett, S. R., & Blaisdell, A. P. (2005). Capacity and limits of associative memory in pigeons. *Psychonomic Bulletin & Review*, 12(2), 350-358. doi: 10.3758/BF03196384
- Cook, R. G., & Murphy, M. S. (2012). Motion processing in birds. In O. F. Lazareva, T. Shimizu, & E. A. Wasserman (Eds.), *How Animals See the World: Behavior, Biology, and Evolution of Vision* (pp. 271-288). London: Oxford University Press.
- Cook, R. G., Qadri, M. A. J., & Keller, A. M. (2015). The analysis of visual cognition in birds: Implications for evolution, mechanism, and representation. *Psychology of Learning and Motivation*, 63, 173-210. doi: 10.1016/bs.plm.2015.03.002
- Cook, R. G., Qadri, M. A. J., Kieres, A., & Commons-Miller, N. (2012). Shape from shading in pigeons. *Cognition*, 124, 284-303. doi: 10.1016/j.cognition.2012.05.007
- Cook, R. G., Shaw, R., & Blaisdell, A. P. (2001). Dynamic object perception by pigeons: Discrimination of action in video presentations. *Animal Cognition*, 4, 137-146. doi: 10.1007/s100710100097
- Dawkins, M. S., Guilford, T., Braithwaite, V. A., & Krebs, J. R. (1996). Discrimination and recognition of photographs of places by homing pigeons. *Behavioral Processes*, 36, 27-38.
- Dombeck, D. A., Harvey, C. D., Tian, L., Looger, L. L., & Tank, D. W. (2010). Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nature Neuroscience*, 13(11), 1433-1440. doi: 10.1038/nn.2648

- Dombeck, D. A., & Reiser, M. B. (2012). Real neuroscience in virtual worlds. *Current Opinion in Neurobiology*, 22(1), 3-10. doi: 10.1016/j.conb.2011.10.015
- Dukas, R., & Clark, C. W. (1995). Sustained vigilance and animal performance. *Animal Behaviour*, 49(5), 1259-1267. doi: 10.1006/anbe.1995.0158
- Frost, B. J., & Sun, H. (1997). Visual motion processing for figure/ground segregation, collision avoidance, and optic flow analysis in the pigeon. In M. V. Srinivasan & S. Venkatesh (Eds.), *From living eyes to seeing machines* (pp. 80-103). Oxford, U.K: Oxford University Press. .
- Fry, S. N., Rohrseitz, N., Straw, A. D., & Dickinson, M. H. (2008). TrackFly: Virtual reality for a behavioral system analysis in free-flying fruit flies. *Journal of Neuroscience Methods*, 171(1), 110-117. doi: 10.1016/j.jneumeth.2008.02.016
- Garbers, C., Henke, J., Leibold, C., Wachtler, T., & Thurley, K. (2015). Contextual processing of brightness and color in Mongolian gerbils. *Journal of Vision*, 15(1), 13. doi: 10.1167/15.1.13
- Gray, J. R., Pawlowski, V., & Willis, M. A. (2002). A method for recording behavior and multineuronal CNS activity from tethered insects flying in virtual space. *Journal of Neuroscience Methods*, 120(2), 211-223. doi: 10.1016/S0165-0270(02)00223-6
- Harvey, C. D., Collman, F., Dombeck, D. A., & Tank, D. W. (2009). Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature*, 461(7266), 941-946. doi: 10.1038/nature08499
- Hoffman, H. G. (2004). Virtual-reality therapy. *Sci Am*, 291, 58-65.
- Holland, P. C. (1992). Occasion setting in Pavlovian conditioning. *The psychology of learning and motivation*, 28, 69-125. doi: 10.1016/S0079-7421(08)60488-0

- Hölscher, C., Schnee, A., Dahmen, H., Setia, L., & Mallot, H. A. (2005). Rats are able to navigate in virtual environments. *The Journal of Experimental Biology*, 208(3), 561-569. doi: 10.1242/jeb.01371
- Katayama, N., Hidaka, K., Karashima, A., & Nakao, M. (2012, 20-23 Aug. 2012). *Development of an immersive virtual reality system for mice*. Paper presented at the SICE Annual Conference (SICE)
- Koban, A., & Cook, R. G. (2009). Rotational object discrimination by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 35, 250-265. doi: 10.1037/a0013874
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press
- McVean, A., & Davieson, R. (1989). Ability of the pigeon (*Columba livia*) to intercept moving targets. *Journal of Comparative Psychology*, 103(1), 95. doi: 10.1037/0735-7036.103.1.95
- Miyata, H., & Fujita, K. (2010). Route selection by pigeons (*Columba livia*) in “traveling salesperson” navigation tasks presented on an LCD screen. *Journal of Comparative Psychology*, 124(4), 433. doi: 10.1037/a0019931
- Morgan, C. L. (1896). *An introduction to comparative psychology*. London, U.K.: Walter Scott Ltd
- Nguyen, A. P., Spetch, M. L., Crowder, N. A., Winship, I. R., Hurd, P. L., & Wylie, D. R. (2004). A dissociation of motion and spatial-pattern vision in the avian telencephalon: implications for the evolution of “visual streams”. *The Journal of Neuroscience*, 24, 4962-4970. doi: 10.1523/JNEUROSCI.0146-04.2004

- Peckmezian, T., & Taylor, P. W. (2015). A virtual reality paradigm for the study of visually mediated behaviour and cognition in spiders. *Animal Behaviour*, 107, 87-95. doi: 10.1016/j.anbehav.2015.06.018
- Qadri, M. A. J., Asen, Y., & Cook, R. G. (2014). Visual control of an action discrimination in pigeons. *Journal of Vision*, 14, 1-19. doi: 10.1167/14.5.16
- Qadri, M. A. J., & Cook, R. G. (2015). Experimental divergences in the visual cognition of birds and mammals. *Comparative Cognition & Behavior Reviews*, 10, 73-105. doi: 10.3819/ccbr.2015.100004
- Qadri, M. A. J., Romero, L. M., & Cook, R. G. (2014). Shape-from-shading in European starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology*, 128(4), 343-356. doi: 10.1037/a0036848
- Qadri, M. A. J., Sayde, J. M., & Cook, R. G. (2014). Discrimination of complex human behavior by pigeons (*Columba livia*) and humans. *PLoS ONE*, 9(11), e112342. doi: 10.1371/journal.pone.0112342
- Reid, S., & Spetch, M. L. (1998). Perception of pictorial depth cues by pigeons. *Psychonomic Bulletin and Review*, 5, 698-704.
- Rescorla, R., & Wagner, A. R. (1972). A theory of Pavlovian conditioning. Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory*. New York, NY: Appleton-Century-Crofts.
- Riley, D. A., & Leith, C. R. (1976). Multidimensional psychophysics and selective attention in animals. *Psychological Bulletin*, 83, 138-160. doi: 10.1037/0033-2909.83.1.138

- Sakuma, M. (2002). Virtual reality experiments on a digital servosphere: guiding male silkworm moths to a virtual odour source. *Computers and Electronics in Agriculture*, 35(2–3), 243-254. doi: 10.1016/S0168-1699(02)00021-2
- Saleem, A. B., Ayaz, A., Jeffery, K. J., Harris, K. D., & Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. *Nature Neuroscience*, 16(12), 1864-1869. doi: 10.1038/nn.3567
- Sato, N., Sakata, H., Tanaka, Y. L., & Taira, M. (2004). Navigation in virtual environment by the macaque monkey. *Behavioural Brain Research*, 153(1), 287-291. doi: 10.1016/j.bbr.2003.10.026
- Sato, N., Sakata, H., Tanaka, Y. L., & Taira, M. (2006). Navigation-associated medial parietal neurons in monkeys. *Proceedings of the National Academy of Sciences*, 103(45), 17001-17006. doi: 10.1073/pnas.0604277103
- Sato, N., Sakata, H., Tanaka, Y. L., & Taira, M. (2010). Context-dependent place-selective responses of the neurons in the medial parietal region of macaque monkeys. *Cerebral Cortex*, 20(4), 846-858. doi: 10.1093/cercor/bhp147
- Shimizu, T. (1998). Conspecific recognition in pigeons (*Columba livia*) using dynamic video images. *Behaviour*, 135, 43-53. doi: 10.1163/156853998793066429
- Skinner, B. F. (1938). The behavior of organisms: An experimental analysis.
- Sofroniew, N. J., Cohen, J. D., Lee, A. K., & Svoboda, K. (2014). Natural Whisker-Guided Behavior by Head-Fixed Mice in Tactile Virtual Reality. *The Journal of Neuroscience*, 34(29), 9537-9550. doi: 10.1523/jneurosci.0712-14.2014
- Spetch, M. L., & Friedman, A. (2006). Pigeons see correspondence between objects and their pictures. *Psychological Science*, 17(11), 966. doi: 10.1111/j.1467-9280.2006.01814.x

- Sturz, B., Bodily, K., & Katz, J. (2006). Evidence against integration of spatial maps in humans. *Animal Cognition*, 9(3), 207-217. doi: 10.1007/s10071-006-0022-y
- Sturz, B., Bodily, K., Katz, J., & Kelly, D. (2009). Evidence against integration of spatial maps in humans: generality across real and virtual environments. *Animal Cognition*, 12(2), 237-247. doi: 10.1007/s10071-008-0182-z
- Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Monographs: General and Applied*, 2(4).
- Thurley, K., Henke, J., Hermann, J., Ludwig, B., Tatarau, C., Wätzig, A., . . . Leibold, C. (2014). Mongolian gerbils learn to navigate in complex virtual spaces. *Behavioural Brain Research*, 266, 161-168. doi: 10.1016/j.bbr.2014.03.007
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189-208.
- Tolman, E. C., Ritchie, B. F., & Kalish, D. (1946). Studies in spatial learning I: Orientation and the short-cut. *Journal of Experimental Psychology*, 36, 13-24.
- Trivedi, C. A., & Bollmann, J. H. (2013). Visually driven chaining of elementary swim patterns into a goal-directed motor sequence: a virtual reality study of zebrafish prey capture. *Frontiers in Neural Circuits*, 7, 86. doi: 10.3389/fncir.2013.00086
- Wasserman, E. A., Franklin, S. R., & Hearst, E. (1974). Pavlovian appetitive contingencies and approach versus withdrawal to conditioned stimuli in pigeons. *Journal of Comparative & Physiological Psychology*, 86, 616-627. doi: 10.1037/h0036171
- Wasserman, E. A., Nagasaka, Y., Castro, L., & Brzykcy, S. J. (2013). Pigeons learn virtual patterned-string problems in a computerized touch screen environment. *Animal Cognition*, 16(5), 737-753. doi: 10.1007/s10071-013-0608-0

- Wasserman, E. A., Teng, Y., & Brooks, D. I. (2014). Scene-based contextual cueing in pigeons. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40(4), 401-418. doi: 10.1037/xan0000028
- Watanabe, S., & Troje, N. F. (2006). Towards a “virtual pigeon”: a new technique for investigating avian social perception. *Animal Cognition*, 9(4), 271-279. doi: 10.1007/s10071-006-0048-1
- Weisman, R. G., & Spetch, M. L. (2010). Determining when birds perceive correspondence between pictures and objects: A critique. *Comparative Cognition & Behavior Reviews*, 5, 117-131. doi: 10.3819/ccbr.2010.50006
- Wilkinson, A., & Kirkpatrick, K. (2009). Visually guided capture of a moving stimulus by the pigeon (*Columba livia*). *Animal Cognition*, 12(1), 127-144.
- Wylie, D. R., & Frost, B. J. (1990). Binocular neurons in the nucleus of the basal optic root (nBOR) of the pigeon are selective for either translational or rotational visual flow. *Visual Neuroscience*, 5, 489-495. doi: 10.1017/S0952523800000614

Author Notes

This research and its preparation was supported by a grant from the National Eye Institute (#RO1EY022655). We would like to thank Clark High School for providing the internship that supported Sean Reid during the conduct of this research. We also thank Ashlynn Keller and Suzanne Gray for their helpful comments on an earlier draft of the manuscript. E-mail: Robert.Cook@tufts.edu. Home Page: www.pigeon.psy.tufts.edu.

Figure Captions

Figure 1. Screen captures of each digital terrain and its correct object as tested in Experiment 1. The left example shows the “grass” surface material and the right shows the “snow” surface material. Each object also appeared in the other terrain (not shown), and its selection was not rewarded in that context. All terrains and the objects within them moved continuously from the top to the bottom of the display.

Figure 2. Object identification accuracy using d' over the three phases of Experiment 1. The top and bottom rows show performance for pigeons #1D and #1L, respectively. The horizontal axis ($d' = 0$) coincides with chance-level performance.

Figure 3. Probability distributions for pecks after object appearance during the within-trial object phase of testing in Experiment 1. Note that the vertical scale represents probability density with respect to all conditions, so the relative areas of the distributions reflect the relative frequency of hits, misses, false alarms, and correct rejections as well as the time that the peck was emitted. All pecks were included, not just pecks registered as being on the object in the display. The distributions were smoothed in R using a Gaussian smoothing kernel with a bandwidth of 1000 ms to generate data for 4096 points between 0 and 40 s.

Figure 4. Screen captures of the two new digital terrains in which topography varied and their correct objects as tested in Experiment 2. The topography of the terrain was sinusoidal and two ridges (“hilltops”) can be seen in these examples. As the camera progressed through these virtual environments, its perspective changed (i.e., pointing upwards or downwards) depending on its location to ensure the visibility of the objects in upcoming portions of the streaming displays.

Figure 5. Object identification accuracy using d' over the two phases of Experiment 2. The d' values for the three different conditions reflect different rates of incorrect object selections (*false*

alarms). These are separated according to the features shared between the incorrect object's assigned terrain and the current terrain.

Figure 6. Object identification accuracy using d' over the time course of a trial during the within-trial terrain discrimination phase of Experiment 2.

Figure 7. Mean vertical peck position in pixels of anticipatory pecking over successive presentations during the within-trial terrain discrimination phase of Experiment 2. Only pecks that occurred after 1 s after the previous object disappeared are included.