



Head-mounted mobile eye-tracking in the domestic dog: A new method

Madeline H. Pelgrim^{1,2} · Julia Espinosa^{2,3} · Daphna Buchsbaum^{1,2}

Accepted: 14 June 2022
© The Psychonomic Society, Inc. 2022

Abstract

Humans rely on dogs for countless tasks, ranging from companionship to highly specialized detection work. In their daily lives, dogs must navigate a human-built visual world, yet comparatively little is known about what dogs visually attend to as they move through their environment. Real-world eye-tracking, or head-mounted eye-tracking, allows participants to freely move through their environment, providing more naturalistic results about visual attention while interacting with objects and agents. In dogs, real-world eye-tracking has the potential to inform our understanding of cross-species cognitive abilities as well as working dog training; however, a robust and easily deployed head-mounted eye-tracking method for dogs has not previously been developed and tested. We present a novel method for real-world eye-tracking in dogs, using a simple head-mounted mobile apparatus mounted onto goggles designed for dogs. This new method, adapted from systems that are widely used in humans, allows for eye-tracking during more naturalistic behaviors, namely walking around and interacting with real-world stimuli, as well as reduced training time as compared to traditional stationary eye-tracking methods. We found that while completing a simple forced-choice treat finding task, dogs look primarily to the treat, and we demonstrated the accuracy of this method using alternative gaze-tracking methods. Additionally, eye-tracking revealed more fine-grained time course information and individual differences in looking patterns.

Keywords Eye-tracking · Comparative cognition · Domestic dog · Visual attention

What is your dog really looking at on a walk? What are young infants attending to when they interact with a parent? While it is impossible to ask nonverbal individuals what they are or are not attending to, we can gather other metrics of attention and perception in order to answer questions about the cognitive processes underlying their behavior. Eye-tracking provides a quantifiable metric of visual attention, providing information about what the participant is seeing and focusing on. Over the past two decades, the development of lightweight head-mounted eye-trackers has expanded the environments in which eye-tracking can be used. By affixing cameras to appropriate

items such as glasses, soft caps, or goggles, head-mounted eye-trackers can capture gaze at real-world 3D objects and scenes (versus gaze during conventional screen-based eye-tracking tasks). This approach has been used particularly extensively with infants and has provided insight into how infants visually respond to their mothers' voices (Franchak et al., 2011), highlighted differences in how crawling and walking toddlers see the world (Kretch et al., 2014), and helped capture the temporal dynamics of how infants and parents coordinate social interactions such as joint visual attention to objects (Yu & Smith, 2013). Despite the growth and success of real-world head-mounted eye-tracking methods in infants and children, this approach has remained comparatively rare in nonhuman species in general, and more specifically dogs.

Due to their unique evolutionary history with humans and their responsiveness to human social-communicative gestures, dogs provide an interesting opportunity for the exploration of social learning abilities in particular (Agnetta et al., 2000; Hare et al., 2002), and nonhuman cognition more broadly (Byosiere et al., 2017; Tecwyn & Buchsbaum, 2019). Dogs are also an integral part of many people's daily lives, either as pets or in a working role. Despite our reliance

✉ Madeline H. Pelgrim
mpelgrim@gmail.com

¹ Department of Cognitive, Linguistic, & Psychological Sciences, Brown University, 190 Thayer St, Providence, RI 02912, USA

² Department of Psychology, University of Toronto, 100 St. George St, Toronto, ON M5S 3G3, Canada

³ Department of Human Evolutionary Biology, Harvard University, 11 Divinity Ave, Cambridge, MA 02138, USA

on dogs to complete vital tasks (e.g., search and rescue dogs, assistance dogs, detection dogs, etc.) and for companionship and emotional support (e.g., emotional support animals), little is known about the specific features of their environment or of their human companions that dogs visually attend to.

In the following sections we will first review existing non-invasive eye-tracking systems used with dogs, including both screen-based and real-world eye-tracking systems, and discuss their advantages and limitations. We will next introduce our novel, easy to train, head-mounted eye-tracking system for dogs, inspired by systems previously used with human infants and toddlers, and discuss how it addresses many of the limitations associated with screen-based systems. Next, we will provide results validating the system's performance using alternative gaze-tracking methods (recorded head movements, verifying the system's accuracy at identifying the target of gaze) as well as internal reliability calculations (specifying the system's degree of spatial precision). We will show results from eye-tracking data using the system during an object choice task, and finally discuss possible future applications of the system.

Screen-based eye-tracking

Conventional screen-based stationary eye-tracking methodology, hereafter screen-based eye-tracking, requires the participant to remain in a stationary position for the duration of the session, while viewing stimuli presented on a screen. This approach can provide highly accurate and precise results for tasks where the participant can remain still and the stimuli can be presented as images or short videos. However, screen-based methods can often be limited by the requirement for a precise setup (i.e., darkened room, remain at an exact distance from the screen, rest head on a custom-built apparatus, etc.), that may be challenging for some species.

Screen-based eye-tracking has frequently been used to study visual attention in nonverbal populations, particularly human infants (e.g., Gredebäck et al., 2009; Laing, 2017) and primates (e.g., Kano & Tomonaga, 2009; Kotani et al., 2017), but also including dogs (e.g., Karl, Boch, Virányi, et al., 2020a; Téglás et al., 2012), as well as other taxa such as birds and rodents (Tyrrell et al., 2015; Zoccolan et al., 2010). Comparisons of gaze patterns and visual attention across species are also possible using screen-based eye-tracking, though these have typically been limited to comparisons between humans and nonhuman primates, such as comparing patterns of visual attention to scenes and faces between infants, chimpanzees, and macaques (e.g., Damon et al., 2017; Kano & Tomonaga, 2009).

In human infants, screen-based eye-tracking has been widely used to explore phenomena such as word learning

(Laing, 2017), causal reasoning (Sobel & Kirkham, 2006), and social cognition (Dean et al., 2021). Infants and toddlers typically do not require any habituation or familiarization to the screen-based eye-tracking apparatus itself as they simply have to sit in front of a screen. Like children, great apes can participate in screen-based eye-tracking without apparatus-specific training. This is thanks to a relatively new method of encouraging apes to keep their heads stationary by providing juice out of a straw at the precise location where their heads should be positioned (Kano et al., 2017). These methods have been used to explore phenomena such as false beliefs (Kano et al., 2017) or response to emotional scenes (Kano & Tomonaga, 2010)

In dogs, screen-based eye-tracking has been used to explore visual attention to familiar and unfamiliar human and dog faces, as well as their attention to upright versus inverted faces (Somppi et al., 2014), and to different emotional facial expressions (Karl, Boch, Zamansky, et al., 2020b; Kis et al., 2017; Somppi et al., 2016). This approach has also been used to determine the order, timing, and general pattern of how dogs visually attend to videos of human social-communicative gestures like pointing (Ogura et al., 2020). Outside of the social domain, screen-based eye-tracking has led to significant advances in understanding of dogs' physical reasoning through the use of 3D animated objects. More specifically, by observing dogs looking patterns and pupil sizes, we now have converging evidence that dogs do not have a gravity bias (something also seen in behavioral studies, Tecwyn & Buchsbaum, 2019) and that dogs, like human infants, understand contact causality (Völter & Huber, 2021). Screen-based eye-tracking has also been used to explore gaze bias—the tendency to look more at a section of a stimulus (i.e., the left side) regardless of what the stimulus is (Guo et al., 2009).

Unlike apes and children, screen-based measures with dogs have historically typically required extensive shaping and training in order to ensure that the dogs remain motionless with their head on a platform for an extended period of time (Karl, Boch, Virányi, et al., 2020a; Somppi et al., 2012). This training may impact task performance and requires extensive time commitments from researchers. In the absence of training, screen-based measures require dog owners to hold their dogs' bodies in position, similar to parents holding children on their laps, but when using this method, many dogs had to be excluded due to extensive head movements (Kis et al., 2017). Recent advances in screen-based eye-tracking, however, have reduced the immobilization/training requirement in many situations, by using a new method involving a combination of treat lures and a forehead marker worn by the dog, to account for some amount of head movement (Correia-Caeiro et al., 2020). This novel approach facilitated a comparison between humans' and dogs' perceptions of facial expressions under

comparable experimental conditions (specifically untrained and unrestrained).

In sum, screen-based eye-tracking is particularly well suited for questions involving visual attention to stimuli that can be presented as images and/or videos, and allows for a great deal of control over the precise timings and features of the images/videos. Additionally, a major strength of screen-based eye-tracking is the high degree of spatial accuracy this method provides for such stimuli (e.g., Kano & Tomonaga, 2009; Morgante et al., 2012; Somppi et al., 2012). Screen-based eye-tracking is, however, not applicable to all experimental situations. Screen-based eye-tracking collects eye movements in isolation when naturalistically participants would be moving around and physically interacting with objects and agents presented as stimuli. There remains a need for an eye-tracking methodology which can capture freely moving participants interacting with the 3D real-world environment.

Head-mounted eye-tracking

To study gaze behavior in a 3D real-world environment, researchers working with a variety of species, and particularly those working with human infants, have begun using mobile, head-mounted eye-trackers. These noninvasive systems have particularly grown in popularity when working with children as they allow for the capture of first-person views and eye movements in a variety of every-day contexts, allowing participants to move through the world and interact with real physical stimuli and with other people (Smith et al., 2015). For instance, mobile head-mounted eye-tracking has been used to explore how children navigate obstacles in their environment and play in an unstructured context (Franchak et al., 2011), and to understand how the daily visual lives of infants change as they transition from crawling to walking, such as what they pay attention to and what visual information is available to them (Kretch et al., 2014; Smith et al., 2015). While calibration procedures for head-mounted eye-trackers require the participant to stay relatively still for a brief period of time, they allow the eye-tracking to be robust without compromising the participants mobility, as the eye-tracker moves with the head (Franchak et al., 2011; Kretch et al., 2014; Yu & Smith, 2013).

Head-mounted eye-tracking has been used to explore the natural behaviors of a variety of nonhuman species, with a variety of different head morphologies, by creatively mounting the eye-tracking cameras (e.g., onto caps, glasses, or goggles). For instance, head-mounted eye-trackers have been used to explore gaze behavior during social interactions with conspecifics in both peacocks and lemurs (Shepherd & Platt, 2006; Yorzinski et al., 2013). Researchers have also explored

human-chimpanzee interactions using an adapted goggle setup (Kano & Tomonaga, 2013).

In dogs, relatively few previous efforts have been made to develop a working head-mounted eye-tracking system. One previous study piloted a muzzle-mounted mirror-based prototype. One dog was successfully trained and calibrated on the setup, but no future studies were conducted using this method (Williams et al., 2011), perhaps because the muzzle-mount was relatively difficult to train. A second exploratory study used a modified pair of dog goggles with processing and recording equipment carried in a backpack by the dog. This study explored dogs' gazing patterns at human gestures made by a familiar versus an unfamiliar person (Rossi et al., 2014). While this exploratory work was promising, no further studies were conducted using this method, possibly due to cited challenges with the prototype's camera and recording quality which resulted in a limited subset of dogs actually providing usable data (Rossi et al., 2014).

As discussed above, dogs' real-world visual attention is a particularly rich research area because of their responsiveness to human social-communicative cues (Agnetta et al., 2000) and their history of partnership and companionship with humans. Using eye-tracking to investigate this relationship can shed light on similarities and differences in the role of features of visual interaction in cooperation and relationship building in this unique cross-species context. In addition, using eye-tracking to understand the real-world visual environment of dogs would allow for the exploration of their perceptual and cognitive abilities more generally, contributing to our understanding of dogs' social, physical and causal reasoning abilities.

The present study

Head-mounted eye-tracking has proved to be a promising methodology with human infants and other nonhuman animals, but thus far work with dogs has not moved past proof of concept. Dogs provide an interesting opportunity for the use of real-world eye-tracking due to living in a human-built visual environment, and their close social and working relationships with humans. In addition, many behavioral studies in canine cognition currently rely on dogs visually searching for food or other items (see review by Bensky et al., 2013), yet we know very little about how they actually complete these tasks.

The present study develops and validates a new method for real-world head-mounted eye-tracking in dogs using 3D stimuli. At the time of writing, this is the first robust, validated, and widely implementable method for eye-tracking in dogs. It is based on a commercially available system in wide use with infants and children, allowing for cross-species comparisons and providing existing calibration and validation benchmarks.

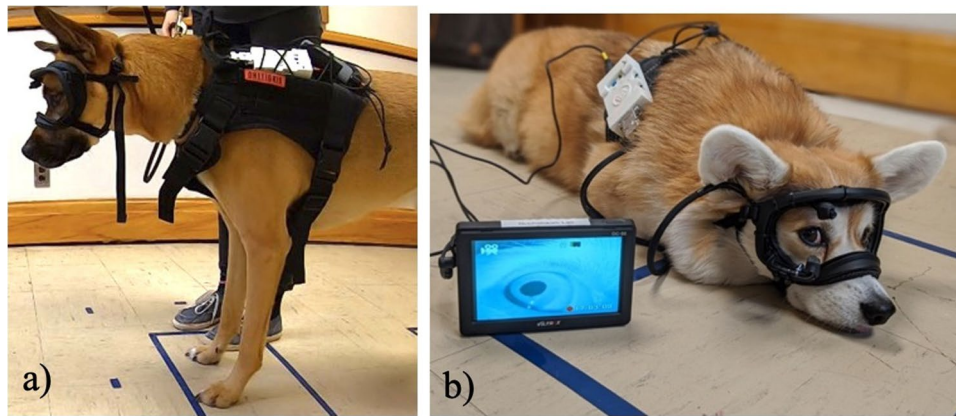


Fig. 1 **a** A dog wearing size XL goggles, in full eye-tracking equipment during a trial. **b** A dog wearing size M goggles, in full the eye-tracking gear with the LCD screen that was removed during participation

Training for this new method does not need to be completed in the lab environment, but instead can be largely completed by pet dog owners at home, reducing the researcher time commitment typically required with canine eye-tracking studies (i.e., Karl, Boch, Virányi, et al., 2020a). Further, the eye-tracking setup in the present study provides a variety of size and position adjustments which enhances the stability of the system and allows for it to be reliably used on moving dogs with a variety body and facial morphologies. Calibration of the system is able to occur offline after a given session, reducing time wearing the apparatus for dogs.

While wearing the apparatus, dogs completed a two-alternative forced-choice task between a plate with a treat on it and an empty plate. We chose this task because we wanted a task that dogs generally choose the plate with the treat on it with a high degree of accuracy (Espinosa et al., 2021), to see whether wearing the apparatus impacted dogs' performance. Further, we had strong expectations about where dogs would look during the task on a coarse time scale, due to previous work using this same task, which measured dogs' gaze using looks recorded from a room camera (Espinosa et al., 2021). To validate the eye-tracking data collected during the task, dogs' head orientation and visible eye movements were recorded from an additional camera in the testing room (room camera), as in Espinosa et al. (2021), allowing us to compare our results to an alternative measure. While this task took place in our lab space, we expect this same eye-tracking methodology to be applicable for use in a variety of environments, such as outdoors or in the dog's home.

Materials and methods

Participants

Participants were five pet dogs (male = 2; mean age = 48 months; age range = 21–71 months). All participants were

recruited on a volunteer basis from the Greater Toronto Area after already completing other cognitive studies (average 2.43 studies), none of which involved eye-tracking. Requirements for participation included owner willingness and availability, and appropriate head size to fit the three pairs of eye-tracking goggles available in our lab (we excluded brachycephalic and toy dogs from this validation study, but the goggle sizes could be extended to encompass them). Three additional dogs began training but did not contribute data, due to leaving the study ($n = 1$) and the early cessation of data collection due to the COVID-19 pandemic ($n = 2$). See Supplementary Materials for a detailed breakdown of all trained participants (https://osf.io/v8tdj/?view_only=2513d2cafc164586be7cf2fb580c5c0f).

Materials

The headgear consisted of two cameras mounted to dog goggles (Fig. 1). Headgear was custom developed by Positive Science, Inc., who have previously developed head-mounted eye-trackers for infants and adults (e.g., Franchak et al., 2011), as well as peacocks (Yorzinski et al., 2013). One camera (the eye camera) is an infrared eye camera with an adjacent infrared emitting diode which recorded the participant's right eye, similar to hardware developed for humans by Positive Science, Inc. The other camera (the scene camera) was placed off-center to be approximately over the right eye and recorded a field of view of 101.55° horizontal and 73.60° vertical. Videos from both cameras were digitized at 29.96 frames per second. The entire headgear weighed approximately 112 grams (g).

Dogs wore a commercially available harness appropriately sized for the dog with Velcro patches which allowed a battery pack and video processor to be attached (see Figs. 1 and 2). The video processor recorded both the eye and scene

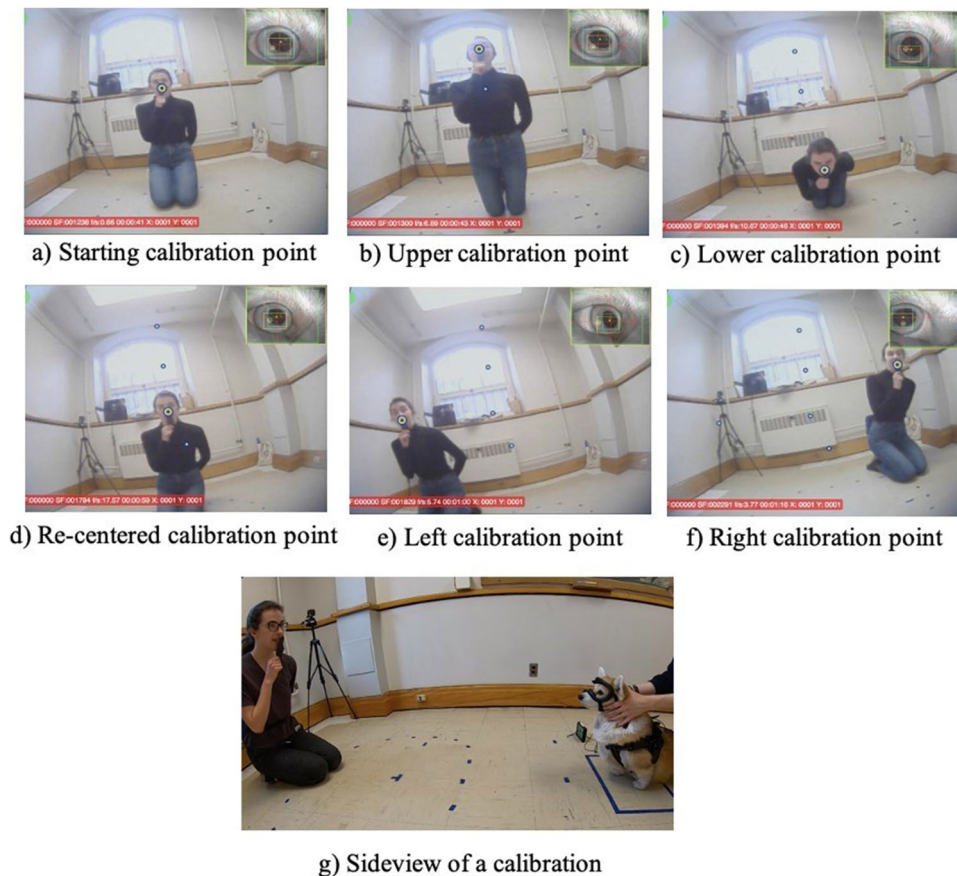


Fig. 2 This full calibration procedure (a–f, seen here in the Yarbus software) is repeated at the start and end of each session. Each point provides known points where the gaze from the eye camera (top

right) lines up to a specific known location from the scene camera's field of view (as indicated by the yellow dots). The dog's owner (g) immobilizes the head

camera footage to memory cards, and the files recorded during a given session were uploaded after the fact to be calibrated and coded offline. The total weight of the harness and associated materials is 612 g. By adhering all materials to the harness, dogs were able to freely move through the room and interact with stimuli. A handheld screen was plugged into the white video processor pack during calibrations to allow for the Experimenter to verify the eye image (see Fig. 1b) but this screen was detached from the video processor prior to trials starting.

Training

Dogs first completed a fitting appointment in the lab, where they were sized for a pair of training goggles to be used for training at home. Commercially available dog goggles (Rex Specs) were used for training as they were very similar to the pairs with mounted cameras, and come in a variety of sizes (M, L, and XL were used in this study). At the fitting appointment, owners were instructed on how to adjust the goggles and acclimate their dog to

the wearing of the goggles. Specifically, dog owners were instructed to train their dog in short sessions daily, starting with the goggles on the loosest setting and leaving them on for less than 30 seconds while constantly giving treats. They were instructed to gradually adjust the fit of the goggles to be snug on their dogs' face and slowly progress to longer wear periods with larger intervals between treats and increased activities with their dogs as they become comfortable. Dog owners were encouraged to incorporate goggle training into their dogs' daily lives, and suggested activities to engage in while the dog wore the goggles included walks, mealtimes, unrelated training sessions, and various games. Similar positive-reinforcement-based training procedures have been used by some screen-based eye-tracking designs, where dogs have been trained to place their heads on chin rests by owners (Somppi et al., 2012).

Approximately one week after the fitting, owners completed a virtual check-in to review their dogs' progress. At this virtual check-in (usually completed via video call), dog owners were encouraged to ask any questions and

discuss issues they encountered during training. Two weeks after the fitting (schedules permitting), dogs came into the lab for an in-person check-in where they were evaluated for their comfort level wearing the goggles while completing practice trials and a practice calibration procedure. If dogs were able to wear the goggles comfortably in the lab environment for 10 minutes, they were invited back to complete their test session and contribute eye-tracking data. If the dog tried to remove the goggles (typically by using a forepaw or rubbing their head against furniture) or seemed uncomfortable (i.e., reluctant to move or engage in normal behaviors, showing stress signals), they were given an additional one to two weeks' training time (schedules permitting), after which we repeated the in-person check-in to observe progress. If after a training extension the dog exhibited repeated stress signals or discomfort wearing the goggles, they left the study. See Supplementary Materials for detailed by-dog training breakdown.

Procedure

During their test session, dogs completed a two-alternative forced choice task, by repeatedly choosing between two plates, only one of which had a treat on it, while wearing the eye-tracking goggles. Prior to this task, dogs first completed a short warm up task and an eye-tracker calibration task, which was also repeated after the choice task. Throughout the session, dogs' behavior was recorded independent of eye-tracking cameras using a wide-angle lens camera in the corner of the room (the room camera). This captured both their behavior and their head movements. Head orientation, as coded from the room camera, was later used to further validate the eye-tracking data (see "Data scoring and analysis" for more details).

Warm-up

At the start of their test session, dogs completed a brief warm-up activity wearing their training goggles and the harness (without battery pack or video processing adaptor) that would be used during the eye-tracking session. This ensured that the dog was comfortable moving through the room and taking treats while wearing the harness and goggles. The experimenter sat centered approximately 1.2 m away from where the dog waited on-leash with a trained research assistant handler. The experimenter presented a single treat between the thumb and index finger of her right hand. The experimenter then placed the treat down on a plate directly in front of her and verbally released the dog, cueing the handler to physically release the dog to come get the

treat. This was repeated for a minimum of three trials or until the dog was comfortable approaching and eating the treat off the plate.

Calibration

After the warm-up, dogs completed a calibration procedure, used to map the movement of the eye (from the camera recording the dog's eye or the eye camera) onto the real-world environment (what the dog was looking at in the world, as recorded from the first-person view camera or scene camera). In this procedure, dogs visually tracked a large treat lure to a predetermined constellation of points spread widely across their field of view. Each of these points provided a known point—a point where the dog was looking at a known location in their first-person view. Offline after the session, we identified the position of the eye (from the eye camera) during the known point from the scene camera, resulting in a mapping of how the location of dogs' looks in the real world (from the scene camera) change as a function of their eye movements. To aid in calibration, a small LCD screen was plugged into the video processing pack for this phase, which allowed the handler and the experimenter to view the eye camera and verify a good and clear eye image.

We used a five-point calibration procedure (Fig. 2), adapted from the infant head-mounted eye-tracking literature (Franchak et al., 2011; Kretch et al., 2014), and similar to procedures used to calibrate existing screen-based eye-tracking systems (Somppi et al., 2012). More specifically, we used a treat lure to get dogs to look at five unique locations, with the lure held in front of the experimenter's mouth so that auditory attention-getters could also be used (more details included below). These known points were chosen to span the dogs' first-person perspective, covering the edges and the center of their field of view which provided a wide range of eye positions. The alignment of the known points (how the eye was positioned when the dog was looking at a given known location in the real world) were then used to extrapolate where the dog was looking for the remainder of the session using eye-tracking software (Yarbus by Positive Science Inc.). The calibration procedure was completed both at the start and at the end of each session to maximize the calibration points available, as well as to allow us to recalibrate following any major headgear movements, if needed (not observed in our sample).

During the calibration procedure, dog owners restrained their dogs' heads (Fig. 2g) while the dogs were sitting in the starting box (a procedure practiced during the check-in). This was done to ensure that dogs looked at each of the five points by moving their eyes, rather than by moving their heads. The procedure took approximately 30–45 seconds and was the only time the dog was physically restrained. The experimenter knelt 1.2 m away from the dog, with one arm

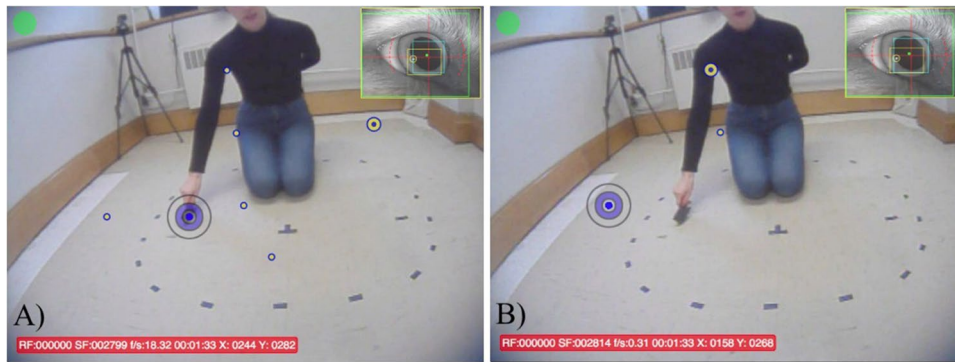


Fig. 3 The test-calibration and verifier points allow for the identified center of gaze, shown here as a blue cursor, to be checked for spatial accuracy. **A** An example of a good calibration: the center of gaze is in the correct place, there are enough calibration points (indicated here by small yellow and blue dots) spread across the field of view. The

next test-calibration point can now be verified. **B** An example of a poor and incomplete calibration: the center of gaze is offset from the known true gaze location (the treat), and the calibration does not have enough identified points (small yellow and blue dots). More calibration points must be included before verification can begin

behind her back and the other holding a large treat in front of her mouth. The experimenter then rose up on her knees and lifted the treat, then bent over to get calibration points at the upper and lower limits of the scene camera's field of view (see Fig. 2b and c). The experimenter then re-centered to the starting kneeling position and moved approximately 30 cm to her right and left sides respectively (Fig. 2d–f). Throughout the calibration procedure, the experimenter called the dog's name and waved the treat as needed before each calibration point, to make sure the dog was looking at the treat for each point. In between calibration points the experimenter broke off a small piece of the treat and gave it to the dog to maintain their motivation.

Verification

To ensure that the calibrated mapping of the dog's gaze to the scene was accurate, we collected additional verification points following the start and end of session calibration procedures (two locations on the ground, start and end verifications), and at the midpoint of the trials (three points, positioned at the experimenter's face and the same ground locations, mid-session verification), which allowed us to confirm that the mapping completed by the software generalized accurately (Fig. 3). These points were used to assess the accuracy of the extrapolated mapping completed using points from the calibration procedure. The dogs' head was not held or in any way immobilized during the verifications, so the calculated accuracy was representative of the rest of the test session. To collect the start and end session verifications following each calibration procedure the owner released the dog's head and the experimenter sequentially tapped two locations on the ground with a treat to encourage the dog to orient their head towards each location. For the mid-session verification, the experimenter called the dog's

name and held a treat in front of her mouth, then tapped the same two ground locations individually with the treat, providing three known points. During all calibration and verification procedures, the experimenter visually verified that the dog was looking at the treat at each location and that the eye image (as displaced on the LCD screen) was clear. If the dog did not track the treat to a given point or if the eye image was unclear, the experimenter repeated that location and verified that the dog was looking.

Choice task

To validate and test the eye-tracking system, dogs wore the eye-tracking apparatus while completing a two-alternative forced choice task between a treat and an empty plate (Espinoza et al., 2021). The two plates were placed 0.75 m apart on marked locations on the floor, the same locations used for the start, end, and mid-session verifications. At the start of a trial, the experimenter placed both plates onto the marked locations. The experimenter then held out both hands at the dog's eye level, holding a treat between her thumb and index finger, with her other hand open. She then moved her hands simultaneously, placing the treat on one plate while her empty hand hovered over the other. After placing the treat, the experimenter returned both hands to her knees and verbally released the dog, signaling the handler to release any tension in the leash. The experimenter's movements were timed so that the regions of interest (ROIs) were visually available for comparable amounts of time. The dog was then allowed to approach and make one choice (defined as making physical contact with the plate) between the two search locations. If dogs chose the plate with the treat on it, they were allowed to eat the treat, and if they chose the empty plate they were **not** rewarded. Dogs were then recalled by the handler to begin the next trial. Placement of the treats was pseudo-randomized across right

and left, and the treat was placed with equal numbers to both sides. After five trials or in the event of any disturbance to the goggles (observed only once in our sample, when a dog ran into the legs of the handler while being recalled between trials), the eye image was once again verified using the LCD screen and a test-calibration was conducted as described in the Calibration section. The procedure was repeated for a total of 10 choice trials.

Data scoring and analysis

Calibration

A subset of the calibration points (5–9 total points used) from both the start and end of the session were used offline to map the dog's eye movements onto the recording of their first-person view. From these calibration points, the dog's point of regard (where in space they were looking as measured in cartesian coordinates) and the timing of each look (when the look started and ended) were then extrapolated for the entirety of the session. This resulted in a video of each dog's first-person perspective with their point of regard overlaid, from which the dog's visual attention to specific ROIs was coded (discussed below).

Temporal data

Start and end times for each trial, as well as events within a trial, were coded using Datavyu software (Datavyu Team, 2014). To create a consistent data set across dogs, we only considered fixations occurring during the observation phase of trials, which began when the trial started (when the experimenter began lifting the treat) and ended when the dog was released to move forward and make their choice. This phase was used, in keeping with past work, because unlike eye-tracking cameras, which continue to effectively track looks while the dog is in motion, estimating where dogs were looking from their head orientation as viewed by the room camera was very difficult while the dog was moving forward to make their choice (Espinosa et al., 2021). All analyses presented in the results section were completed for the observation phase¹. The presentation phase began when the trial started and ended when the experimenter placed the treat onto the plate. This subset of the overall trial was used in separating some ROIs, discussed below.

¹ To be comparable to prior work and maintain consistency throughout, all analyses presented here in the main text cover only the observation phase (start of trial until dog is released). Analyses conducted using only eye-tracking data were also run for the full trial (start of trial until dog made choice), and these results were consistent with those presented below. Full trial analyses along with all data can be found in the Supplementary Materials https://osf.io/v8tdj/?view_only=2513d2cafc164586be7cf2fb580c5c0f

Regions of interest

Fixations, defined as stationary pauses in eye movement lasting longer than 100 ms (Franchak et al., 2011; Patla & Vickers, 1997) were automatically identified from the complete video file and then manually coded as falling into one of six defined ROIs using GazeTag software (Positive Science, Inc.). These ROIs were chosen to capture dogs' attention to social features (e.g., the experimenter's face and hands) and to target objects and locations, namely attention to the treat and to the two search locations. The ROIs we analyzed are expanded from past room camera-based gaze recordings (Espinosa et al., 2021), since eye-tracking is able to accurately differentiate gaze location at a higher degree of spatial accuracy. Specifically, the ROIs included (1) the experimenter's face/body, (2) the experimenter's hand holding and placing the treat, (3) the experimenter's empty hand placing nothing, (4) the experimenter's empty hands coming back up from placement after the treat presentation, (5) the treat on the plate/baited location, and (6) the empty plate/unbaited location. An additional ROI, looking backward to the owner or handler, was considered a priori; however, looking to this ROI was not observed within our sample. These ROIs were also categorized either as being explicitly social (two ROIs: the experimenter's face/body and the experimenter's empty hand after the treat presentation), or as not explicitly social (the remaining four ROIs). See Fig. 4 for a visual breakdown of the locations of the ROIs. Any fixations falling outside of these locations were coded as off-target.

All fixations were initially coded by the first author, and 100% of ROI categorizations were also verified by a second coder (percent similarity of coding was 97.5%). In the event of disagreement on a coded fixation (seen in 78 of the 3095 fixations), both coders reviewed the fixation together and discussed the two choices until they reached an agreement. For each fixation, the timing (onset and offset), pupil height and width, and the point of regard (measured by x and y values as mapped onto the linear plane of the scene camera) were recorded. The manual coding of which ROI the dog was looking at for each fixation is a limitation of the current system, in particular compared to screen-based systems which automatically categorize ROI data.

Room camera gaze coding

In prior work (Espinosa et al., 2021), dogs show consistent relative gaze patterns across the ROI on this task, such as gazing more often at the treat versus empty locations. These patterns can be accurately coded at a coarse spatial and temporal granularity from the room camera (Espinosa et al., 2021). In order to verify that the eye-tracking data from our new approach are consistent and accurate, we examined whether the pattern of looking across ROIs was consistent

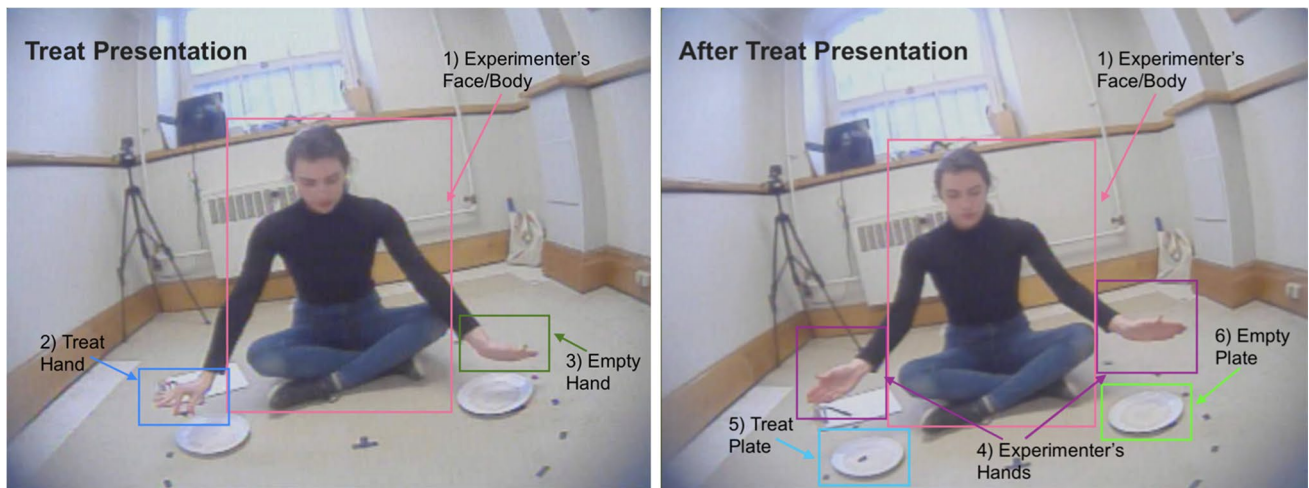


Fig. 4 The six regions of interest coded during a trial from eye-tracking data. Room camera data were coded as falling into one of four ROIs, where ROIs 2 and 5 (treat-hand and treat-plate) and ROIs 3 and

6 (empty-hand and empty-plate) were combined to be coded as Treat or Empty, respectively

between the manual room coding and the eye-tracking data. Room data were collected from a camera providing a view of the room and dog's face, and the dog's gaze was coded by observation of head orientation and visible movements of the eyes and surrounding muscles (see Supplementary Materials video for side-by-side comparison of eye-tracking versus room cameras for a single trial). Gaze was coded as falling into one of five ROIs, a subset of the six used in the eye-tracking data described above (Fig. 4). In keeping with past work (Espinosa et al., 2021) the ROIs included (1) the experimenter's face/body, (2) the treat, either in the experimenter's hand during placement or on the plate, (3) the empty location, either the empty hand during placement or the plate, and (4) the experimenter's empty hands coming back up from placement, meaning after the treat presentation. As with the larger set of ROIs, these ROIs were categorized as explicitly social (as above, the experimenter face/body and empty hands after presentation) and not explicitly social (the treat and empty location ROIs). All room gaze coding was completed by a researcher assistant with extensive experience with this method from Espinosa et al. (2021) using Datavyu software.

Comparison of eye-tracking and room data

We predicted the same overall pattern of looking to the ROIs across the eye-tracking and room cameras, as measured by the proportion of observation phase time spent looking to each of the ROIs. We predicted that across camera types, dogs would proportionally look more to the treat (vs. the empty location). We tested this using a one-sample t-test comparing the proportion of time during the observation phase averaged across

trials that dogs looked to the treat (vs. empty location) relative to chance (chance would be looking equally to both locations or 0.5). See looks during trials section below for more detailed analysis of dogs' looking behaviors. To validate the location of looks generated by the eye-tracking data statistically, we explored whether individual dog, trial, and the proportion of time looking to the treat (vs. empty) from the room camera predicted the proportion of time looking to the treat from the eye-tracking cameras. We also completed a mixed-effects logistic regression examining the effects of trial number, camera type (eye-tracking vs. room), and individual dog on the proportion of each trial spent looking to the treat versus empty search location. This would provide confirmation that the eye-tracking data are recording dogs' looks accurately as measured by a previously accepted alternative method of recording where dogs are looking.

Given that we have a coding that is consistent and accurate according to what we already know about dogs' gaze behavior, we wanted to examine additional information that we gain from eye-tracking vs the room. Specifically, we expected the eye-tracking data to provide a more detailed data set, due to the higher spatial resolution of the data, allowing additional ROIs to be coded, and because of the greater temporal specificity offered by eye-tracking. In order to qualitatively validate the eye-tracking data and explore the advantages over recording from a room camera, we compared the counts and length of recorded fixations from both data sets.

Choice task performance

We predicted that dogs would succeed at the task, choosing the plate with the treat on it at above chance levels. We also expected that wearing the eye-tracking goggles would

not have a significant impact on dogs' performance. All participating dogs had previously completed a comparable two-alternative forced choice task while on-site for non-eye-tracking studies, with the exception of one dog who completed the task on one of her goggle training check-in visits without goggles on². We used this past performance to evaluate whether the eye-tracking goggles had an impact on dogs' choices, and conducted a mixed-effects logistic regression exploring dogs' choices (correct or incorrect) across trials and visits (with or without the eye-tracker).

Looks during trials

We did not have strong expectations about the specifics of how dogs would look at the ROIs during the trials. We conducted a series of exploratory analyses exploring different ways of measuring dogs' looking behaviors. We conducted linear regressions exploring the impact of individual dog, trial number, ROI, and the interaction between dog and ROI on both the number of looks (counted fixations) and durations of look (how long in ms the fixations lasted). We evaluated how dogs looked to social, versus non-social targets as a function of both individual dog and trial number. Specifically, we used a multiple linear regression comparing the proportion of trial time looking to social ROIs (looks to the experimenter's face/body and hands) versus looks to the other ROIs. We also qualitatively explored how many trials in which dogs looked to both search locations to investigate their information-gathering strategies.

Results and discussion

Training and calibration

Training on average took 34.83 days, and dogs came in for an average of 1.14 in-person and 1 virtual check-in sessions. A full breakdown of by-dog training times can be found in the Supplementary Materials.³ All dogs who contributed data were able to be successfully calibrated, meaning that the extrapolated gaze (generated by the eye-tracking software using the calibration points) was on-target for the verifier points and test-calibration point. All dogs also contributed data on each trial. Spatial accuracy was assessed for each dog using the distance between the extrapolated gaze and the known point (the treat during the verifiers

and test-calibration). More specifically, when the dog was known to be looking at the treat, which was 4.5 cm across on average, as the experimenter tapped it on the ground during the verifiers and test-calibration, approximately 125 cm away from the dog, we calculated the offset distance in degrees. A minimum of 20 frames, an averaged number from previous work (Watalingam et al., 2017; Yorzinski et al., 2013), were used for each dog across individual dogs' test-calibrations and verifiers. The average calibration accuracy in degrees (*min.* = 2.41°, *max.* = 5.15°, *mean* = 3.56°) was lower than screen-based eye-tracking systems, which typically have calibration accuracy of less than 1° (i.e., Correia-Caeiro et al., 2020 [0.25°–0.5°]; Karl, Boch, Virányi, et al., 2020a [$< 0.5^\circ$]; Somppi et al., 2012 [1°]; Téglás et al., 2012 [0.5°–0.7°]), but consistent with results from head-mounted systems in other species [Franchak et al., 2011 [2°]; Watalingam et al., 2017 [3.8°]; Yorzinski et al., 2013 [3.9°]).

Choice task performance

Overall, dogs succeeded at the choice task, choosing the plate where the treat had been placed over the empty plate (total = 45/50 trials correct, *mean* = 9, *SD* = 1). Individually, all dogs chose correctly on eight or more trials. A mixed-effects logistic regression exploring dogs' choice (correct or incorrect) as a function of trial number and visit (with or without the eye-tracker) with a random intercept for dog, found that wearing the eye-trackers did not alter the dogs' performance on the task. Neither trial, $z = 1.64$, $p = .101$, nor visit, $z = -1.76$, $p = .079$ was a significant predictor of choice, with a trend towards better performance on the second visit when dogs were wearing the eye-tracker (mean proportion correct without eye-tracker = 0.78, mean proportion correct with eye-tracker = 0.9). Dogs succeeded at this task, and wearing the eye-tracking goggles did not appear to cause any significant behavioral changes, or impact choice performance. While it is possible that there were subtler impacts of goggle training, the results presented above suggest that, once trained at home, dogs wearing the eye-trackers are likely to perform similarly on cognitive tasks to how they would without wearing goggles.

Validation

To validate that the eye-tracking data were recording the target and approximate timing of looks appropriately, we conducted a multiple linear regression exploring whether proportional looks to the treat (vs. empty location) as coded from the room camera, trial number, and individual dog predicted proportional looking to the treat from the eye-tracking data. We found that the proportion of time dogs spent looking to the treat as recorded from the room camera significantly predicted the proportional looking from

² Link to all data including data from past performances: https://www.osf.io/v8tdj/?view_only=2513d2cafc164586be7cf2fb580c5c0f. Prior data were collected as part of Espinosa and Buchsbaum (manuscript in preparation) and Pelgrim et al. (2021).

³ Link to all data and Supplementary Materials: https://osf.io/v8tdj/?view_only=2513d2cafc164586be7cf2fb580c5c0f

the eye-tracking data, $F(1) = 5.50$, $p = .025$. There was no significant effect for trial number, $F(9) = 1.83$, $p = .1$, or for individual dog, $F(4) = 1.04$, $p = .4$.

To further validate the eye-tracking data, we conducted a multiple linear regression to compare the proportion of looks to the treat as a function of camera type (eye-tracking vs. room camera), as well as individual dog and trial number. There was no significant difference in proportion of looks to treat by camera type—eye-tracking ($M_{\text{proportion}} = 0.78$, $SD = 0.1$) versus room ($M_{\text{proportion}} = 0.77$, $SD = 0.06$)—when controlling for trial and dog, $F(1) = .0005$, $p = .98$. The same model revealed a difference in proportional looks to treat by trial, independent of camera type, with dogs looking more to the treat-location over the course of trials, $F(9) = 3.14$, $p = .003$. Finally, there was no significant difference in proportional looks to the treat across individual dogs, $F(4) = .8$, $p = .53$.

Taken together, this suggests that the eye-tracking data and the room-view data are highly proportionally aligned, and that the proportion of time dogs spent looking at the treat (vs. the empty location) did not differ between the two methods of capturing dogs' looks. As a final validation measure, we conducted two one-sample t -tests comparing dogs' proportional looks to the treat from both camera types against chance (chance being 0.5 or looking equally to both search locations).

When comparing looks to the treat versus empty locations, averaged across their trials, dogs proportionally looked to the treat at above chance levels for both the eye-tracking, $t(4) = 6.58$, $p = .003$, and the room cameras, $t(4) = 10.37$, $p < .001$. This is consistent with previous room-view work on a comparable task (Espinosa et al., 2021), and suggests that the eye-tracking cameras are recording dogs' point of regard accurately. Finally, dogs near ceiling performance on the choice task prevented analysis of what looking behaviors predict success, as there were very few error trials. Nonetheless, dogs looked longer at the treat than the empty location, consistent with previous findings that proportional gaze predicts choice.

Eye-tracking versus room data

As mentioned above, eye-tracking and room-view cameras recorded similar proportional looks to the key search locations; however, eye-tracking data provided a more detailed data set than the data collected from the room-view cameras. Across dogs and trials ($M_{\text{observation phase length}} = 5041.52$ ms, $SD = 356.8$ ms) gaze coding from the room-view cameras generated a total of 170 distinct looks ($M_{\text{per trial}} = 3.4$, $SD = 1.16$), averaging 1478.57 ms in length ($SD = 1123.42$ ms), whereas data from the eye-trackers generated 309 distinct fixations ($M_{\text{per trial}} = 6.18$, $SD = 2.67$), each averaging 582.81 ms in length ($SD = 676.29$ ms). In other words, the eye-tracking data contain more individual looks that are shorter in length. The observed differences in timing and counts of looks across the two camera types can be explained

by the finer grain with which the eye-tracking data capture looking behaviors (see Fig. 5 for an example trial).

More specifically, as reported above, the eye-tracking and room-view cameras comparably capture dogs' point of regard when examining proportional data for search locations, like looks to the treat. However, relative to room-view camera data, eye-tracking data are characterized by greater spatial and temporal accuracy. This allows for short-duration changes in point of regard to be captured in a way that estimating looks from head orientation recording from a room camera does not (i.e., off-target looks in Figs. 5 and 6, glances to experimenter's hands after presentation in Fig. 5), allowing for rapid, shorter-duration looks to be accounted for. It further provides greater spatial accuracy (relative to room-view cameras) allowing more accurate distinctions to be made between spatially adjacent ROIs, such as the experimenter's hands after their presentation, versus their face/body than can be made from the room-view cameras (see Figs. 5 and 6). Further, eye-tracking data capture looks off-target ($n = 30$) that could not be determined from the room-view camera (see Fig. 6).

Exploring differences in individual patterns of looking

We next examined other aspects of dogs' visual attention, such as characteristics of looks including frequency and length, and how they changed as a function of ROI and trial number. Overall, dogs averaged $M = 6.18$, $SD = 2.67$ fixations per trial, with an average fixation length of $M = 582.81$ ms, $SD = 676.29$ ms.

Duration of looks

To explore dogs' visual attention during trials, we conducted a multiple linear regression exploring fixation duration as a function of ROI, trial number, and individual dog, as well as a possible interaction between individual dog and ROI. We found a significant effect of ROI $F(6) = 11.84$, $p < .001$, but found that average duration did not change over trial, $F(1) = .25$, $p = .62$ or by dog $F(4) = .56$, $p = .69$. This suggests that dogs' patterns of looks, as defined by how long each look is, did not change over the course of their sessions and did not vary significantly by dog identity in our sample. We also found a significant interaction between ROI and individual dog, $F(24) = 2.11$, $p = .002$. This suggests that there were individual differences in what dogs preferred to look at (see Fig. 7 for an overview of individual differences). More specifically, while there were ROIs that were universally more interesting (such as the treat in the hand or on the plate, see Table 1), different dogs also found different ROIs more or less interesting, resulting in longer durations, on average, for certain ROIs, which we will return to below.

Averaged across dogs and trials the longest fixations occurred when looking at the experimenter's hands, not

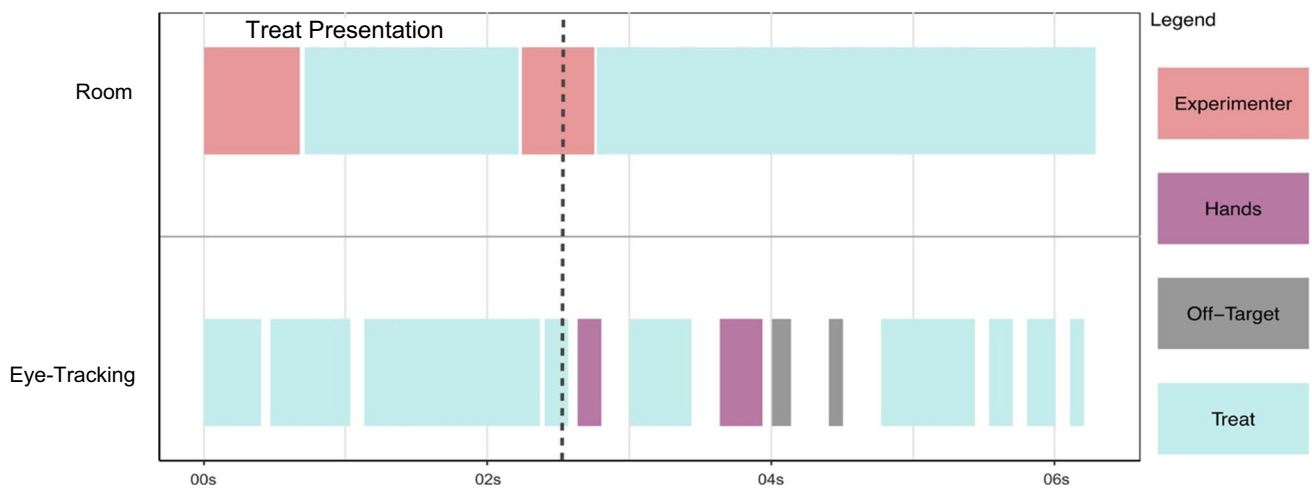


Fig. 5 A comparison of the data collected from the room camera (top) and the eye-tracking data (bottom) for a single trial. This demonstrates the finer spatial and temporal resolution of the eye-tracking data. Eye-tracking is able to capture rapid, shorter-duration looks. In addition, eye-tracking captures moments where the dog's point of

regard (where they are actually looking) differs from their head orientation, as observed from the room camera, like the glances off-target seen here in gray. See Supplementary Materials for video of this trial (https://osf.io/v8tdj/?view_only=2513d2cafc164586be7cf2fb580c5c0f)



Fig. 6 A comparison of the same moment in a trial from the two camera types. **Left:** Mid-trial data from the room camera. Tracking the head orientation from this angle suggests the dog is looking at the

treat. **Right:** Mid-trial data from the eye-tracker. Using eye-tracking shows the dog was actually looking out the window

during the treat presentation ($M=951.52$ ms, $SD=814.14$ ms). Dogs also had the fewest total fixations to this location ($n=22$). In other words, dogs rarely looked to the experimenter's hands after the presentation, but when they did it was typically for a comparatively long time. Averaged across dogs and trials, the shortest individual fixations were looks to off-target locations ($M=181.17$ ms, $SD=182.08$ ms), and they occurred a moderate number of times ($n=30$). This provides good support that dogs were generally engaged with the task as they were not distracted for long periods of time during trials. See Table 1 for a summary of the looks to each ROI.

As noted above, there were also individual differences between dogs in the average length of their fixations to the different ROIs, suggesting differences in either which of the ROIs individual dogs found visually interesting or which ROIs they found helpful or informative for completing the

task. As an example, individual dogs' looks to the empty hand during the presentation ranged from very short in duration (i.e., $M_{CCL268}=100$ ms, the minimum definition of a fixation) to averaging half a second (i.e., $M_{CCL345}=500.83$ ms). Dogs also varied in their average fixation length to the treat on the plate, with one dog averaging moderate-length fixations ($M_{CCL345}=352.74$ ms) and another averaging nearly a second for each look ($M_{CCL485}=907.41$ ms).

Frequency of looks

We next explored what factors predicted the number of fixations, measured by counting individual fixations rather than the length of those fixations as presented above. The ROIs dogs looked at most frequently were not necessarily those that had the longest individual looks (see Table 1). To

explore this other facet of looking behavior, we conducted a multiple linear regression exploring number of fixations to each of the seven eye-tracking ROIs as a function of trial, ROI, individual dog, and the interaction between dog and ROI. There was a significant main effect of dog, $F(4) = 3.00$, $p = .019$, meaning that overall dogs differed in their number of fixations, even when controlling for which ROI they were looking at or what trial they were on. This provides very strong evidence for individual differences in visual searching strategy and visual attention between dogs.

Combining across their session, dogs ranged in total fixation counts from 45 to 81, and per trial they ranged from 1 to 13 fixations. Controlling for the effects of individual dog and trial, there was a significant main effect of ROI on number of fixations, $F(6) = 10.16$, $p < .001$, with the most individual fixations occurring to experimenter ($n = 77$) and the fewest occurring to the experimenter's hands after the presentation phase ($n = 22$). There was no significant effect of trial on fixation count, $F(1) = .099$, $p = .75$, meaning that the pattern of dogs looks was not changing over the course of their session. Finally, there was also a significant interaction between individual dog and ROI on number of fixation $F(24) = 2.00$, $p = .004$, meaning that dogs showed individual differences in their pattern of fixations across ROIs (see Fig. 7 for an overview of individual differences in looking patterns between dogs).

Taken together the individual differences observed in fixation counts may be reflective of differences in visual search strategy. For example, one dog (CCL485) never tracked the experimenter's hands after the treat presentation, focusing instead mostly on the treat, while another dog (CCL220) looked at the experimenter's empty hands after she had placed the treat on 6/10 trials. It is possible that such individual differences may be driven by seeking help or guidance (Passalacqua et al., 2011), and dogs with fewer looks to the experimenter could be categorized as more independent problem solvers. Alternatively, these patterns could be indicative of more general differences in social drive or motivation (Cook et al., 2016). The looking patterns observed here may have also been impacted by other factors, such as individual dogs' previous training history (e.g., being previously rewarded for eye contact with their owner).

Social analyses

As a final exploration into the looking behaviors between dogs, we compared looks to the explicitly social ROIs (the experimenter's face/body and her hands after presentation) with the non-social or semi-social targets (the treat in the hand or the empty hand during the presentation, and the two plates). Off-target looks were removed for the purposes of these analyses, and we collapsed across social and non-social looks as described to calculate the average duration

of looks. Averaged across trials, dogs varied in their average durations to social versus non-social targets. We conducted a multiple linear regression examining average duration of looks as a function of ROI (here, social vs. non-social), dog, trial number, and an interaction of ROI and dog. As in previous analyses, we found no significant effect of trial number, $F(1) = .11$, $p = .75$, indicating the mean duration of looks does not change over the course of the session. We also found no significant main effect of dog, $F(4) = 1.27$, $p = .29$, meaning that when controlling for the target they were looking at, dogs did not differ in their average duration for each look. Finally, we found no significant effect of ROI, $F(1) = .79$, $p = .38$, indicating that across dogs and trials, dogs did not consistently have longer looks to social or non-social targets. This may be due to a balancing effect caused by individual differences—there was a significant interaction between ROI and dog, $F(4) = 2.89$, $p = .03$, meaning that individual dogs differed in the degree to which they preferentially looked to social versus non-social targets.

Some dogs, such as CCL485, averaged very short durations of looks to social targets ($M_{\text{social}} = 131.94$ ms), and tended to look much longer to non-social targets (on average, $M_{\text{non-social}} = 536.58$ ms), whereas other dogs (i.e., CCL220) showed the opposite trend ($M_{\text{social}} = 677.78$ ms, $M_{\text{non-social}} = 396.50$ ms). Finally, dogs like CCL345 were more evenly split between social and non-social targets ($M_{\text{social}} = 410.14$ ms, $M_{\text{non-social}} = 401.47$ ms). As observed in previous analyses, dogs may differ in how they prioritize independent information-gathering versus looking at the human experimenter. These patterns may indicate differences in motivation, help-seeking, or training history, and will be discussed in more detail below.

We can also explore social and non-social targets by comparing the number of looks that occurred to these locations. As in the duration analyses, we removed off-target looks and collapsed across social ROIs (the experimenter's face/body and her hands after presentation) and the non-social targets (the treat in the hand or the empty hand during the presentation, and the two plates). After averaging the number of looks for each dog and trial across these two domains, we conducted a multiple linear regression exploring the number of looks as a function of ROI (social vs. non-social), dog, trial number, and an interaction between ROI and dog. As observed in the analyses, exploring duration of looks to each of social versus non-social targets, there was no significant main effect of dog, $F(4) = 1.82$, $p = .13$, meaning that dogs did not differ in their average overall number of fixations. Also, in keeping with the results exploring all ROIs, there was no main effect of trial, $F(1) = .16$, $p = .69$. There was no main effect of ROI, $F(1) = .64$, $p = .43$, meaning that, averaged across dogs, there was no difference in the number of looks to social versus non-social targets, and there was also no significant interaction between ROI and dog $F(4) = .19$,

Proportion looking time by ROI & Dog for the Observation Phase

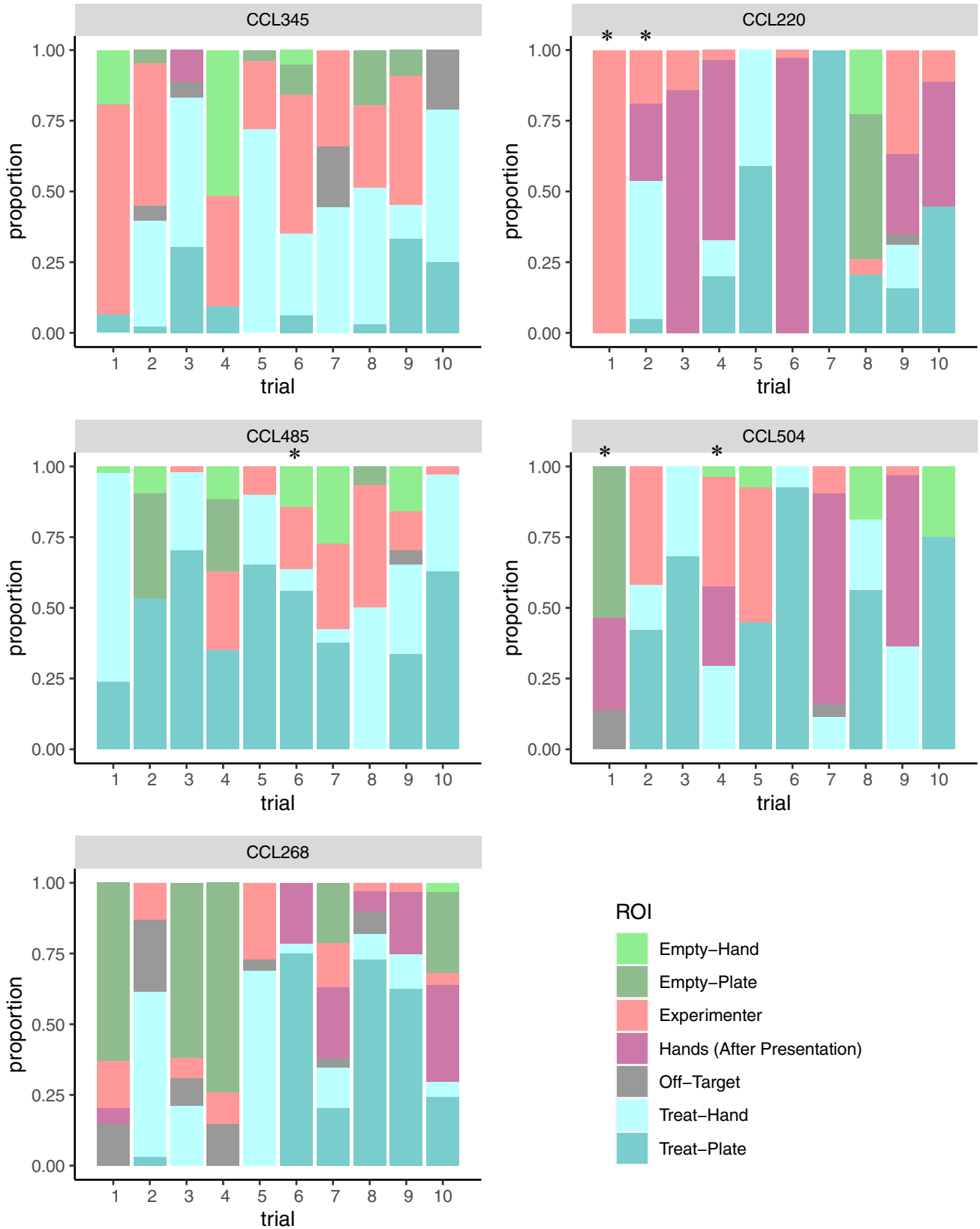


Fig. 7 The proportion of time during the observation phase (start of trial until dogs were released to make their choice) that dogs spent looking to the various ROIs. The five dogs in this sample displayed varying search strategies, with some dogs not looking at a given ROI at all (i.e., CCL485, who never looked at the experimenter's hands after the treat presentation). Incorrect trials, indicated with overhead asterisks (*), do not display a consistent pattern of looks across dogs. Proportions graphs for the full trial (start of trial until choice) for each dog are available in the Supplementary Materials

$p = .12$. This suggests that the observed differences in dogs' attention to social targets comes predominately from the duration of their individual looks, rather than the overall number of those looks.

Information-gathering strategies

One potential explanation for dogs' differences in looking patterns, like those described above, relate to different information-gathering strategies. Dogs are choosing the treat at comparable levels, but they differ in the kinds of information that they sought out prior to making that choice. We explored this further by examining whether dogs evaluated both search locations before making their choice. On half of the trials (25/50) dogs looked only to one item, and of these, 22/25 dogs tracked only the treat. Considering only the treat did not guarantee success, as on one of the five incorrect trials the dog looked exclusively at the treat. On the other 3/25 trials where dogs only considered one search location, they looked only to the empty location. Interestingly, only on one of these did the dog choose incorrectly. Despite never looking at the treat location prior to being released on the other two trials, dogs still chose the treat location, suggesting that they were perhaps avoiding the empty location. On 22/50 trials, dogs considered both search locations before making their choice, and two of the five incorrect trials occurred after the dog considered both search locations. On the final 3/50 trials, dogs did not look at either of the search locations. These trials showed increased levels of social engagement, and looks were at the experimenter's face/body or at her hands after the treat was placed onto the plate. Of these 3/50 trials, one was incorrect, and on the other two, the dog successfully chose the treat despite not looking at it prior to being released to make her choice.

In each trial, the two choices (treat vs. empty) were presented simultaneously. To further examine how dogs were gathering their information before making their choice, we explore looking behavior during the presentation phase (from the start of the trial to when the treat was placed onto the plate). The majority of dogs only looked to one location (only 8/50 dogs looked to both locations during this phase). In other words, on 8/50 total trials, dogs switched the search location they were looking at,

looking at both the empty hand and the hand with the treat. Overall, this suggests that dogs largely just track the one moving item at a time (in our case mostly the treat in the hand) and that consideration of the second object largely occurs after both are stationary.

The extent to which dogs looked to both search locations within trials also varied by dog, ranging from looking to both locations on just 1/10 trials to looking at both locations on 7/10 trials. Further, all three trials where a dog did not look at either search location were from the same dog. Overall, despite having comparable choice behavior on the task, individual dogs differed in their information-gathering strategies. Some dogs took a more social approach and spent much of their trial looking at the experimenter and not at the search locations. Others favored a comprehensive independent approach, reliably checking both of the search locations before making their choice.

General discussion

We have presented a novel method for real-world mobile eye-tracking, developed specifically for dogs that can be directly compared to systems successfully in use with other species. We have shown that this method is easy to use and that it calibrates with a high degree of accuracy, comparable to that found in other species. We were able to collect data for all participating dogs on all trials, gathering information on the timing and focus of visual attention. We were able to explore individual differences in visual attention on a simple object choice task and demonstrate changes in dogs' visual attention patterns over the course of trials. We have validated the accuracy of the recorded data using an alternative gaze-tracking technique, demonstrating both that the eye-tracking system is capturing data accurately as well as highlighting the additional information it provides. We chose a simple task for dogs to complete while wearing the eye-trackers as we had strong predictions about how they would perform, and we found that wearing the eye-trackers did not have a negative impact on their task performance.

One limit of the present study, however, is that dogs were so successful on the task that it was not possible to complete any analysis exploring gaze patterns as predicting correct choices. When considering further applications of this task across cognitive domains, future work could explore whether features of gaze are predictive of success on tasks with greater performance variability. Future work can also further explore differences in individual dogs' information-gathering strategies. Despite similar choice performance, dogs differed in their looking behaviors, looking to different ROIs for different amounts of time. Dogs also differed in the extent to which they explored both search locations, with

Table 1 Statistics of the regions of interest from eye-tracking data for the observation phase

ROI	Total number of looks	Avg. number of looks per trial	Mean look duration	SD look duration
Empty-hand	24	2.67	337.71	314.86
Empty-plate	29	2.9	509.31	441.71
Experimenter	77	7.7	452.64	592.73
Hands (post-presentation)	22	2.44	951.52	814.14
Off-target	30	3.33	181.17	182.08
Treat-hand	71	7.1	608.73	613.15
Treat-plate	56	5.6	942.35	926.44

some dogs searching both consistently and others focusing on only one location. Finally, dogs also differed in how socially oriented they were, as indicated by different durations to social and non-social targets by dog. This suggests that some dogs sought to solve the problem at hand by focusing on the human, whereas others focused on independently solving the problem and focused on the treat.

We believe that the head-mounted eye-tracking method we have presented here will facilitate the exploration of research questions about dogs' gaze behaviors in a 3D real-world environment that are not currently addressable with existing methods. For instance, future work can take inspiration from the developmental psychology literature (i.e., Yu & Smith, 2013), such as evaluating how dogs visually interact with their owners in an everyday context, such as during play. Hallmarks of human–human visual interaction, like joint attention during cooperative tasks, can now be explored in a cross-species context, contributing to a broader understanding of key features of visual communication and their role in components of human–dog partnerships like cooperation and bonding. This new method can also be used to evaluate complex social dynamics such as re-engagement, a process facilitated by joint intentionality and something recently demonstrated in dogs (Horschler et al., 2022). In an applied context, this method could clarify how, for example, guide dogs are able to safely navigate city streets to assist visually impaired individuals, or how search and rescue dogs are able to navigate rubble to locate missing people, facilitating improved training regimes.

While this system opens up additional research avenues, it has both advantages and limitations relative to existing screen-based systems. In particular, screen-based systems provide greater spatial accuracy than the method presented here, making them more suitable for fine-grained analyses of dogs looks to smaller ROIs. Screen-based systems also provide increased control over precisely what the participant sees. Although training for the method presented here can be conducted by dog owners at home, newer screen-based eye-tracking methods

(e.g., Correia-Caeiro et al., 2020) require no training for dogs, reducing time to data collection and also eliminating concern over potential training effects not detectable through behavioral measures. Finally, screen-based eye-tracking systems typically allow for automatic detection of the ROI the participant is looking at, facilitating faster and less labor-intensive data processing than the manual coding currently implemented in our system (something future work can improve on).

In sum, we have presented a reliable and validated new method allowing for eye-tracking of dogs as they interact with real-world stimuli. This method allows for more naturalistic data collection, and expands the types of tasks that can be completed while collecting eye-tracking data, facilitating future research in both a theoretical and applied context. Using this new method, we can explore the information available to dogs in their environment, as well as what they visually attend to, learning more about a unique cross-species relationship, and about nonhuman cognitive abilities more generally.

Acknowledgements Thank you to the dogs and their owners for their support and participation in this project. Thanks to the members of the Canine Cognition Lab at the University of Toronto for their assistance in data collection and training. Further thanks to Liyuzhi Dong for completing the gaze coding from the room cameras. Thanks to Jason Babcock and the team at Positive Science for their ongoing support and assistance in getting this project running. This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC), [2016-05552], and the Canada Foundation for Innovation [32653].

Author contributions Madeline H. Pelgrim: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Visualizations, Writing – Original Draft, Writing – Editing. Julia Espinosa: Conceptualization, Investigation, Methodology, Writing – Editing. Daphna Buchsbaum: Supervision, Conceptualization, Methodology, Formal Analysis, Writing – Review & Editing, Funding Acquisition

Data availability All data from the study and code used for analysis are public and can be accessed on the Open Science Framework: https://www.osf.io/v8tdj/?view_only=2513d2cafc164586be7cf2fb580c5c0f. This study was not pre-registered.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest. The experiment was approved by the University of Toronto's University Animal Care Committee (UACC). Procedures were in accordance with Ontario's Animals for Research Act, the federal Canadian Council on Animal Care and fully complied with the APA Ethical Standards for Use of Animals in Research.

References

- Agnetta, B., Hare, B., & Tomasello, M. (2000). Cues to food location that domestic dogs (*Canis familiaris*) of different ages do and do not use. *Animal Cognition*, 3(2), 107–112. <https://doi.org/10.1007/s100710000070>
- Bensky, M. K., Gosling, S. D., & Sinn, D. L. (2013). The World from a Dog's Point of View. In: *Advances in the Study of Behavior* (Vol. 45, pp. 209–406). <https://doi.org/10.1016/B978-0-12-407186-5.00005-7>
- Byosiere, S.-E., Feng, L. C., Chouinard, P. A., Howell, T. J., & Bennett, P. C. (2017). Relational concept learning in domestic dogs: Performance on a two-choice size discrimination task generalises to novel stimuli. *Behavioural Processes*, 145, 93–101. <https://doi.org/10.1016/j.beproc.2017.10.009>
- Cook, P. F., Prichard, A., Spivak, M., & Berns, G. S. (2016). Awake canine fMRI predicts dogs' preference for praise vs food. *Social Cognitive and Affective Neuroscience*, 11(12), 1853–1862. <https://doi.org/10.1093/scan/nsw102>
- Correia-Caeiro, C., Guo, K., & Mills, D. S. (2020). Perception of dynamic facial expressions of emotion between dogs and humans. *Animal Cognition*, 23(3), 465–476. <https://doi.org/10.1007/s10071-020-01348-5>
- Damon, F., Méary, D., Quinn, P. C., Lee, K., Simpson, E. A., Paukner, A., ... Pascalis, O. (2017). Preference for facial averageness: Evidence for a common mechanism in human and macaque infants. *Scientific Reports*, 7(1), 46303. <https://doi.org/10.1038/srep46303>
- Datavyu Team. (2014). Datavyu: A Video Coding Tool. Databrary Project, New York University. <http://datavyu.org>
- Dean, B., Ginnell, L., Ledsham, V., Tsanas, A., Telford, E., Sparrow, S., ... Boardman, J. P. (2021). Eye-tracking for longitudinal assessment of social cognition in children born preterm. *Journal of Child Psychology and Psychiatry*, 62(4), 470–480. <https://doi.org/10.1111/jcpp.13304>
- Espinosa, J., Dong, L., & Buchsbaum, D. (2021). Domestic dogs' gaze and behavior association in 2-alternative choice tasks. In: *Proceedings of the 43rd Annual Conference of the Cognitive Science Society*. <https://doi.org/10.31234/osf.io/pt3vh>
- Franchak, J. M., Kretch, K. S., Soska, K. C., & Adolph, K. E. (2011). Head-Mounted Eye Tracking: A New Method to Describe Infant Looking: Head-Mounted Eye Tracking. *Child Development*, 82(6), 1738–1750. <https://doi.org/10.1111/j.1467-8624.2011.01670.x>
- Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye Tracking in Infancy Research. *Developmental Neuropsychology*, 35(1), 1–19. <https://doi.org/10.1080/87565640903325758>
- Guo, K., Meints, K., Hall, C., Hall, S., & Mills, D. (2009). Left gaze bias in humans, rhesus monkeys and domestic dogs. *Animal Cognition*, 12(3), 409–418. <https://doi.org/10.1007/s10071-008-0199-3>
- Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The Domestication of Social Cognition in Dogs. *Science*, 298(5598), 1634. <https://doi.org/10.1126/science.1072702>
- Horschler, D. J., Bray, E. E., Gnanadesikan, G. E., Byrne, M., Levy, K. M., Kennedy, B. S., & MacLean, E. L. (2022). Dogs re-engage human partners when joint social play is interrupted: A behavioural signature of shared intentionality? *Animal Behaviour*, 183, 159–168. <https://doi.org/10.1016/j.anbehav.2021.11.007>
- Kano, F., & Tomonaga, M. (2009). How chimpanzees look at pictures: A comparative eye-tracking study. *Proceedings of the Royal Society B: Biological Sciences*, 276(1664), 1949–1955. <https://doi.org/10.1098/rspb.2008.1811>
- Kano, F., & Tomonaga, M. (2010). Attention to emotional scenes including whole-body expressions in chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology*, 124(3), 287–294. <https://doi.org/10.1037/a0019146>
- Kano, F., & Tomonaga, M. (2013). Head-Mounted Eye Tracking of a Chimpanzee under Naturalistic Conditions. *PLoS One*, 8(3), e59785. <https://doi.org/10.1371/journal.pone.0059785>
- Kano, F., Krupenye, C., Hirata, S., & Call, J. (2017). Eye tracking uncovered great apes' ability to anticipate that other individuals will act according to false beliefs. *Communicative & Integrative Biology*, 10(2), e1299836. <https://doi.org/10.1080/19420889.2017.1299836>
- Karl, S., Boch, M., Virányi, Z., Lamm, C., & Huber, L. (2020a). Training pet dogs for eye-tracking and awake fMRI. *Behavior Research Methods*, 52(2), 838–856. <https://doi.org/10.3758/s13428-019-01281-7>
- Karl, S., Boch, M., Zamansky, A., van der Linden, D., Wagner, I. C., Völter, C. J., Lamm, C., & Huber, L. (2020b). Exploring the dog-human relationship by combining fMRI, eye-tracking and behavioural measures. *Scientific Reports*, 10(1), 22273. <https://doi.org/10.1038/s41598-020-79247-5>
- Kis, A., Hernádi, A., Miklósi, B., Kanizsár, O., & Topál, J. (2017). The Way Dogs (*Canis familiaris*) Look at Human Emotional Faces Is Modulated by Oxytocin. An Eye-Tracking Study. *Frontiers in Behavioral Neuroscience*, 11, 210. <https://doi.org/10.3389/fnbeh.2017.00210>
- Kotani, M., Shimono, K., Yoneyama, T., Nakako, T., Matsumoto, K., Ogi, Y., ... Ikeda, K. (2017). An eye tracking system for monitoring face scanning patterns reveals the enhancing effect of oxytocin on eye contact in common marmosets. *Psychoneuroendocrinology*, 83, 42–48. <https://doi.org/10.1016/j.psyneuen.2017.05.009>
- Kretch, K. S., Franchak, J. M., & Adolph, K. E. (2014). Crawling and Walking Infants See the World Differently. *Child Development*, 85(4), 1503–1518. <https://doi.org/10.1111/cdev.12206>
- Laing, C. E. (2017). A perceptual advantage for onomatopoeia in early word learning: Evidence from eye-tracking. *Journal of Experimental Child Psychology*, 161, 32–45. <https://doi.org/10.1016/j.jecp.2017.03.017>
- Morgante, J. D., Zolfaghari, R., & Johnson, S. P. (2012). A Critical Test of Temporal and Spatial Accuracy of the Tobii T60XL Eye Tracker. *Infancy*, 17(1), 9–32. <https://doi.org/10.1111/j.1532-7078.2011.00089.x>
- Ogura, T., Maki, M., Nagata, S., & Nakamura, S. (2020). Dogs (*Canis familiaris*) Gaze at Our Hands: A Preliminary Eye-Tracker Experiment on Selective Attention in Dogs. *Animals*, 10(5), 755. <https://doi.org/10.3390/ani10050755>
- Passalacqua, C., Marshall-Pescini, S., Barnard, S., Lakatos, G., Valsecchi, P., & Prato Previde, E. (2011). Human-directed gazing behaviour in puppies and adult dogs, *Canis lupus familiaris*. *Animal Behaviour*, 82(5), 1043–1050. <https://doi.org/10.1016/j.anbehav.2011.07.039>
- Patla, A. E., & Vickers, J. N. (1997). Where and when do we look as we approach and step over an obstacle in the travel path? *NeuroReport*, 8(17), 3661–3665.
- Pelgrim, M. H., Espinosa, J., Tecwyn, E. C., Marton, S. M., Johnston, A., & Buchsbaum, D. (2021). What's the point? Domestic dogs' sensitivity to the accuracy of human informants. *Animal Cognition*, 24(2), 281–297. <https://doi.org/10.1007/s10071-021-01493-5>

- Rossi, A., Smedema, D., Parada, F. J., & Allen, C. (2014). Visual Attention in Dogs and the Evolution of Non-Verbal Communication. In: A. Horowitz (ed.), *Domestic Dog Cognition and Behavior* (pp. 133–154). <https://doi.org/10.1007/978-3-642-53994-76>
- Shepherd, S. V., & Platt, M. L. (2006). Noninvasive telemetric gaze tracking in freely moving socially housed prosimian primates. *Methods*, 38(3), 185–194. <https://doi.org/10.1016/j.ymeth.2005.12.003>
- Smith, L. B., Yu, C., Yoshida, H., & Fausey, C. M. (2015). Contributions of Head Mounted Cameras to Studying the Visual Environments of Infants and Young Children. *Journal of Cognition and Development*, 16(3), 407–419. <https://doi.org/10.1080/15248372.2014.933430>
- Sobel, D., & Kirkham, N. (2006). Blickets and Babies: The Development of Causal Reasoning in Toddlers and Infants. *Developmental Psychology*, 42, 1103–1115. <https://doi.org/10.1037/0012-1649.42.6.1103>
- Somppi, S., Törnqvist, H., Hänninen, L., Krause, C., & Vainio, O. (2012). Dogs do look at images: Eye tracking in canine cognition research. *Animal Cognition*, 15(2), 163–174. <https://doi.org/10.1007/s10071-011-0442-1>
- Somppi, S., Törnqvist, H., Hänninen, L., Krause, C. M., & Vainio, O. (2014). How dogs scan familiar and inverted faces: an eye movement study. *Animal Cognition*, 17(3), 793–803. <https://doi.org/10.1007/s10071-013-0713-0>
- Somppi, S., Törnqvist, H., Kujala, M. V., Hänninen, L., Krause, C. M., & Vainio, O. (2016). Dogs Evaluate Threatening Facial Expressions by Their Biological Validity – Evidence from Gazing Patterns. *PLoS One*, 11(1), e0143047. <https://doi.org/10.1371/journal.pone.0143047>
- Tecwyn, E. C., & Buchsbaum, D. (2019). What factors really influence domestic dogs' (*Canis familiaris*) search for an item dropped down a diagonal tube? The tubes task revisited. *Journal of Comparative Psychology*, 133(1), 4–19. <https://doi.org/10.1037/com000145>
- Téglás, E., Gergely, A., Kupán, K., Miklósi, Á., & Topál, J. (2012). Dogs' Gaze Following Is Tuned to Human Communicative Signals. *Current Biology*, 22(3), 209–212. <https://doi.org/10.1016/j.cub.2011.12.018>
- Tyrrell, L. P., Butler, S. R., & Fernández-Juricic, E. (2015). Oculomotor strategy of an avian ground forager: tilted and weakly yoked eye saccades. *Journal of Experimental Biology*, 218(16), 2651–2657. <https://doi.org/10.1242/jeb.122820>
- Völter, C. J., & Huber, L. (2021). Dogs' looking times and pupil dilation response reveal expectations about contact causality. *Biology Letters*, 17(12), 20210465. <https://doi.org/10.1098/rsbl.2021.0465>
- Watalingam, R. D., Richetelli, N., Pelz, J. B., & Speir, J. A. (2017). Eye tracking to evaluate evidence recognition in crime scene investigations. *Forensic Science International*, 280, 64–80. <https://doi.org/10.1016/j.forsciint.2017.08.012>
- Williams, F. J., Mills, D. S., & Guo, K. (2011). Development of a head-mounted, eye-tracking system for dogs. *Journal of Neuroscience Methods*, 194(2), 259–265. <https://doi.org/10.1016/j.jneumeth.2010.10.022>
- Yorzinski, J. L., Patricelli, G. L., Babcock, J. S., Pearson, J. M., & Platt, M. L. (2013). Through their eyes: Selective attention in peahens during courtship. *Journal of Experimental Biology*, 216(16), 3035–3046. <https://doi.org/10.1242/jeb.087338>
- Yu, C., & Smith, L. B. (2013). Joint Attention without Gaze Following: Human Infants and Their Parents Coordinate Visual Attention to Objects through Eye-Hand Coordination. *PLoS One*, 8(11), e79659. <https://doi.org/10.1371/journal.pone.0079659>
- Zoccolan, D., Graham, B. J., & Cox, D. D. (2010). A Self-Calibrating, Camera-Based Eye Tracker for the Recording of Rodent Eye Movements. *Frontiers in Neuroscience*, 4, 193. <https://doi.org/10.3389/fnins.2010.00193>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.