

Journal of Visualized Experiments

Testing ideas about visual perception from real-world observations using behavioral and eye tracking measures --Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE59004R1
Full Title:	Testing ideas about visual perception from real-world observations using behavioral and eye tracking measures
Keywords:	Visual perception; Abductive designs; behavior; eye tracking; Multiple object tracking; Gestalt grouping
Corresponding Author:	Jennifer Corbett UNITED KINGDOM
Corresponding Author's Institution:	
Corresponding Author E-Mail:	jennifer.e.corbett@gmail.com
Order of Authors:	Jennifer Corbett Jaap Munneke
Additional Information:	
Question	Response
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US\$4,200)
Please indicate the city, state/province, and country where this article will be filmed . Please do not use abbreviations.	Uxbridge London UK

28 August, 2018

Dear Editor,

Please find attached our manuscript, "Testing ideas about visual perception from real-world observations using behavioral and eye tracking measures," which was invited by *JoVE*'s Dr. Aaron Berard.

We present a protocol for conducting abductive research, observing a visual perception phenomenon in the real-world that cannot be accounted for by existing theories, and then using a range of different behavioral and eye tracking measures to provide the most parsimonious explanation for these observations.

Thank you for considering this manuscript for publication in *JoVE*.

Sincerely,



Jennifer E. Corbett
Brunel University London
College of Health and Life Sciences
Centre for Cognitive Neuroscience
jennifer.e.corbett@gmail.com



Jaap Munneke
Brunel University London
College of Health and Life Sciences
Centre for Cognitive Neuroscience
jaap.munneke@gmail.com

TITLE:

Testing Ideas about Visual Perception from Real-world Observations using Behavioral and Eye Tracking Measurements

AUTHORS & AFFILIATIONS:

Jennifer E. Corbett^{1,2}, Jaap Munneke^{1,2}

¹Brunel University London, College of Health and Life Sciences, Division of Psychology, London, Uxbridge, UK

²Brunel University London, Centre for Cognitive Neuroscience, London, Uxbridge, UK

Corresponding Author:

Jennifer E. Corbett

Email Address: jennifer.e.corbett@gmail.com

Tel: +44 737 984 6994

Email Addresses of Co-authors:

Jaap Munneke (jacobalex.munneke@brunel.ac.uk)

KEYWORDS:

Visual perception; Abductive designs; Behavior; Eye tracking; Multiple object tracking; Gestalt grouping

SUMMARY:

We present a protocol for using convergent behavioral and eye tracking measures to develop an explanation for a real-world observation that cannot fully be accounted for by existing theories.

ABSTRACT:

The majority of scientific inquiry relies on deducing specific conclusions from general rules. However, when existing theories fail to fully account for a given observation, an abductive approach can allow for the formulation of new theories that can then be tested and gradually expanded to account for a wider range of similar observations. Here, we present an abductive approach using convergent behavioral and eye tracking measures to explain a real-world observation that grouping by color similarity facilitates observers in tracking the interactions between players on different sports teams.

INTRODUCTION:

For over a century, investigations of visual perception have relied predominantly on deductive methods. Predictions of an existing theory are tested and the resultant data call for the theory to be supported, modified, or abandoned in favor of an alternative or newly developed theory. However, it is not always possible to deduce specific ideas from general principles when the goal is to understand these general principles. For example, our current understanding of how the limited capacity visual system is able to construct the illusion of stable and complete perception stems largely from abductive work pioneered by Gestalt Psychologists, who developed a set of

heuristics for perceptual organization based on observations of inherent regularities in the surrounding environment¹⁻³. Yet, this sort of abductive research that forms the bedrock of current theories has fallen out of fashion in the scientific community. Despite the inherently cyclical nature of inductive and deductive reasoning, our modern discussion of abductive research has become restricted to illusion contests and demo nights. In an attempt to promote the reintegration of this fundamental scientific link between observation and explanation, we provide an example of a method for observing a real-world phenomenon, recreating it in the lab, and generating testable hypotheses based on these observations. Given the long-standing scientific trend to move away from abductive reasoning, we promote using a range of converging behavioral and eye tracking methods to assess real-world phenomenon aimed to help gather as much information as possible when testing potential explanations.

Our particular investigation began with the observation that the different-colored jerseys worn by teams of players in a soccer match somehow seemed to help keep track of how the players were interacting, both in a collective sense of monitoring the movements of all the players on the field, and in terms of the relationships between the players on a given team. Importantly, studies of multiple object tracking suggest the opposite effects, such that observers are less able to track objects which are part of the same Gestalt group compared to tracking performance for ungrouped objects (*e.g.*, see references⁴⁻⁵). In an effort to explain our observations, we hypothesized that grouping information by similarity of color (the Gestalt principle of grouping by similarity) might allow observers to track more information than the typical “four-object limit” that has been widely accepted in the visual perception literature⁶⁻⁷. Along these lines, organizing players into fewer teams with more players on each team compared to more teams with fewer players on each team may facilitate observers in tracking players’ interactions, as indexed by their performance in counting the number of times a given player gained possession of the ball. Furthermore, the effect of this grouping by color similarity might affect observers’ tracking performance differently depending on whether they were attempting to monitor the interactions of all the players or only the interactions of a specific team. To explore these proposals, we began by abstracting the soccer match scenario to a set-up that could be experimentally manipulated in the laboratory.

Given that existing theories could not explain our observations, we set-out to gather as much data as possible to provide the most complete assessment of the phenomenon in question. Specifically, we had access to psychophysical and eye tracking testing facilities that allowed us to concurrently measure observers’ behavioral accuracy and pupil size while they tracked the interactions of numbered, colored-circle “players” passing a soccer ball. This convergent approach allowed us to measure not only the effects of grouping that could be directly observed from the participants’ response behavior, but also cognitive factors that could be derived from eye movements not contaminated with the extra factors required to plan and execute a manual button press response. Pupil diameter has been demonstrated to index the amount of mental effort observers invest, such that pupil size increases with difficulty and set size in visual search and counting tasks⁸. Along these lines, more effortful top-down task demands tend to elicit larger pupil diameters, and performance in similar tasks; for example, visual search is typically facilitated by top-down attention⁹. Therefore, pupil diameter can index potential differences in

the amount of mental effort participants invested to track the ball target as a function of the top-down set elicited by instructing them to count the possessions by players on a specific color team when there were fewer teams of more players each and when there were more teams of fewer players each, as compared to when they were instructed to simply count the possessions of all the players.

PROTOCOL:

All aspects of the protocol were approved by the local research ethics committee.

1. Participant Recruitment

1.1. Recruit participants aged 18-35 years to ensure that their eyes are fully developed but likely not subject to the effects of aging (*e.g.*, presbyopia).

1.2. Stress during recruitment that participants must have normal or corrected-to-normal vision.

2. Experimental Design

NOTE: Using a stimulus presentation software, construct stimuli that can be manipulated along the dimensions that are key to answering the experimental question while holding all other parameters constant or randomized. This protocol focuses on the effects of grouping by color similarity on observers' abilities to track the interactions between multiple objects. Therefore, we used numbered, colored circles and a JPEG image of a soccer ball as visual stimuli.

2.1. Conditions: Construction of a design matrix of the experimental factors

NOTE: See **Figure 1**.

2.1.1. To examine the effects of the number of different colored teams (NTeams) and the particular task the participant was attempting to perform (CountTask), design a scenario where the same number of circle stimuli (*e.g.*, 20) can be divided equally into a small number and a large number of individual groups (*e.g.*, two groups of 10 circles each, and five groups of four circles each).

2.1.2. Importantly, include a baseline condition to obtain a measure of the observers' performance when there is no grouping (*e.g.*, 20 circles that are all the same color).

2.2. Randomization in space and time

2.2.1. Randomize an equal number of trials for each of the (five) combinations of the experimental factors (CountAll2Teams, CountAll5Teams, CountTeam2Teams, CountTeam5Teams, and 1Team/Baseline) separately for each participant.

NOTE: Ensure that there are enough trials for proper statistical analysis given potential data loss due to human and mechanical error. Typically, 25-50 trials per condition are sufficient in behavioral and eye tracking experiments.

2.2.2. Randomly select the colors of the different (two or five) teams on each trial from a base set of RGB values to ensure that any differences in luminance akin to differences in real-world team jerseys will average out over trials.

NOTE: We used Red = (255 0 0); Green = (0 255 0); Blue = (0 0 255); Yellow = (255 255 0); Magenta = (255 0 255).

2.2.3. Make a base set of screen locations with more locations than the total number of stimuli that cover the central 3/4 of the visible screen. Randomly select where each circle will start for each trial, and then add a small x- and y-jitter to each location so that the stimuli do not appear to be grouped into any geometric formations that might otherwise confound the results.

2.2.4. Randomly select the number of each circle on each team from one to the total number of stimuli so that there is an equal probability of the same number (*e.g.*, 15) appearing in any of the teams presented on the trial.

2.2.5. Randomize the starting location/player who has possession of the soccer ball.

2.2.6. Randomly select the length of a given trial from an interval that allows time for a sufficient number of possessions for participants to count (*e.g.*, 12-14 s allowing for an average of 11 possessions per team).

2.2.7. Cut the trial into shorter intervals (*e.g.*, randomly selected intervals from 400-600 ms in 50 ms steps) and assign each player a 50% chance of changing location at the start of each interval.

2.2.8. If a player is assigned to change location, immediately represent this stimulus at a new location randomly selected from any of the locations that do not currently have a player assigned (offset the new location by small, randomly selected amounts of jitter in the x- and y-directions to avoid any grid configurations).

2.2.9. Ensure that the ball has a probability of changing players on a large enough proportion of the intervals in a given trial (*e.g.*, 90%).

2.2.10. Randomly select one of the team colors at the start of the trial, and give players on this team approximately half of the possessions to ensure conditions are approximately equal in difficulty in terms of the number of possessions observers must count.

2.2.11. To ensure that individual elements are not truncated, construct an imaginary boundary from the top, bottom, and sides of the screen and restrict the players and the ball from moving within this region.

2.3. Ecological validity: Keep it as real world as possible

2.3.1. Use numbers to individuate different players.

2.3.2. Present the numbers centered on the top half of each circle like numbers on actual soccer jerseys.

2.3.3. To ensure that the ball never obscures the possessing player's number, center the ball horizontally with the player, and position it vertically from the center to the bottom of the player (at the player's "feet").

2.3.4. Allow players to overlap as they would in a real-world match.

2.4. Other design considerations

2.4.1. Use a gray background to minimize the likelihood that visual afterimages or transients confound any observed effects.

2.4.2. Divide the trials into several blocks to ensure that participants take regular breaks every 3-5 min and rest their eyes.

3. Instructing the Participants

NOTE: See **Figure 2**.

3.1. Ensure that the participant fully understands the task and how it relates to the real-world scenario before beginning the experimental blocks.

3.2. Present each participant with written, illustrated instructions and read these instructions aloud with the participant at the start of the experimental session.

3.3. Instruct participants to estimate the number of possessions on each trial as accurately as possible.

3.4. Ensure that the participant understands that the players might overlap at times like in a real soccer match and to make their best guess if this happens.

3.5. Ensure that the participant understands what does count as a possession.

NOTE: Possessions are defined as receiving the ball from a player on the same team, or from a player on a different team.

3.6. Ensure that the participant understands what does not count as a possession.

NOTE: The starting position/player of the ball does not count as a possession and the same player moving with the ball does not count as a possession.

3.7. Before setting up the eye tracker, give participants a practice block with a few repetitions of each condition (in randomized order) and monitor their performance during this practice block to ensure that they are performing as instructed.

3.7.1. Impose (recommended) some performance criteria during the practice block to ensure participants are performing as instructed, such as requiring them to complete the practice block with an average error below a certain threshold.

3.7.2. Make sure to ask whether participants have any questions and ensure that they understand the answers before beginning the experimental blocks.

4. Eye Tracking Set Up

4.1. Determine the sampling rate of the eye tracking system.

NOTE: The measures available from the eye data are largely dependent on the sampling rate of the eye tracking hardware. For example, if the eye tracking system can only sample up 60 times per second (60 Hz), then a single sample is taken approximately every 16.67 ms. This translates to the potential for up to 16.67 ms of noise to be introduced into individual samples, which usually allows for the measurement of pupil size, but often precludes measuring eye movements like saccades and fixations that depend on more precise sampling rates. If the tracking system can sample at a higher rate (*e.g.*, 500 Hz), then such additional measures can also be included to the extent that they can provide information relevant to answering the experimental question. The eye tracker used in the original study by Akyuz and colleagues¹⁰ tracked participants' right eye positions at 60 Hz. For best results, always follow the instructions of the specific eye tracker used.

4.2. Synchronize the eye tracker with the stimulus presentation software and eye tracking recording software according to manufacturer's instructions.

4.3. Position the participant at a constant distance from the center of the monitor.

4.4. Ask the participant to sit with their chin in a chinrest and adjust the chair height and chinrest position as needed to ensure that the participant's head is centered horizontally and vertically with the monitor.

4.5. Before beginning each experimental block, calibrate the eye tracker to ensure it is accurately tracking participants' gaze within 1° of accuracy by having each participant complete a calibration task.

4.5.1. Ask participants to follow the white center of a black dot with their eyes as it moves in random order through nine different points arranged in an imaginary 3 x 3 square grid covering the central 90% of the screen.

4.5.2. Remind the participant to move as little as possible for the duration of each block.

4.6. After successfully calibrating the eye tracker, ensure the participant is ready to begin the task, remind them to sit as still as possible, and begin recording the eye data.

4.7. If the eye tracking software indicates that the eye position has been lost, stop recording, recalibrate, and redo the block.

5. Trial Sequence

5.1. Begin each trial by instructing each participant to press the space bar on the keyboard when ready.

5.2. Present the specific task instructions on a gray background (e.g., “ALL” in the CountAll conditions, “GREEN” in the count team conditions, and “YELLOW ALL” in the 1Team/Baseline conditions). Ask the participant to press the space bar again after reading the trial instructions to initiate the trial.

5.3. Present the players passing the ball as described in the “Experimental Design” section (see section 2). Keep the task instructions at the corner of the screen throughout the entire duration of the trial in case the participant forgets what task to perform.

5.4. At the end of the trial, present a gray screen prompting the participant to respond “How many possessions?” using the number keys on the computer’s keyboard.

6. Data Analysis

6.1. Co-register the eye tracking and behavioral data.

6.2. For each participant, exclude from analysis any trials with absolute errors greater than a pre-determined criterion, trials without responses, and trials where the eye tracker lost the pupil for more than a given number of samples.

6.3. Calculate the behavioral percent correct as follows.

6.3.1. Calculate the percent correct on each trial using the following formula:

$$\%Correct = \frac{NPos - AbsError}{NPos} \times 100$$

where *NPos* = the actual number of possessions, and *AbsError* = the absolute error on each trial.

6.3.2. Subtract the resultant average value in the baseline condition from the average of each of the four main conditions, such that negative values indicate lower accuracy than baseline.

6.4. Measure the pupil diameter.

NOTE: Typically, pupil diameter is measured in arbitrary units specific to the particular eye tracking software used.

6.4.1. Divide each participant's average pupil diameter in a given condition by their pupil diameter averaged over all trials to calculate the given participant's average relative pupil diameter for that condition.

6.4.2. Subtract the resultant average value in the baseline condition from the average of each of the four main conditions, such that negative values indicate smaller pupil diameters than baseline.

REPRESENTATIVE RESULTS:

The current protocol was previously used to test the effects of grouping by color similarity on participants' abilities to track the interactions between individual objects¹⁰. In that study, we tested 29 students from Bilkent University (mean age = 19.8 years, 20 females), who voluntarily participated in exchange for course credit or monetary compensation. All aspects of the protocol were approved by Bilkent University's Ethics Committee. We present statistical analyses based on these previously obtained results, and illustrations of representative results based on hypothetical data to best illustrate the protocol's possible outcomes.

The experiment consisted of a 2 x 2 design with an additional baseline condition. The experimental factors consisted of "CountTask" (count either the ball possessions within one team or within all teams) and "Nteams" (either two teams of 10 players, or five teams of four players were present in the display on a given trial). The baseline condition consisted of one group of 20 homogenously colored circles. The dependent measures were accuracy in counting the number of possessions and the average pupil diameter throughout a given trial. Crucially, all statistical analyses were performed on baseline corrected data: For each participant, the average performance obtained in the baseline condition was subtracted from the participant's average performance on each of the four experimental conditions. Using baseline-corrected data provides a better measurement of the influence of the experimental factors on the participant's tracking performance as it negates the large amount of noise that is the result of using dynamic displays rendered over multiple seconds.

Throughout the experiment, participants were tasked with counting the number of ball possessions of players of either the same team or of all players present in the display, independent of team membership. **Figure 2a** provides an overview of what the dynamic displays in this experiment looked like, and **Figure 2b** instructs what does and does not constitute a possession.

The applied methodology provided a number of quantitative results. First, using the difference in accuracy relative to the baseline as a dependent measure, a repeated measures ANOVA with CountTask (CountAll vs CountTeam) and Nteams (2Teams vs 5Teams) as factors indicated that participants were significantly more accurate when counting the possessions of all players present in the display, as compared to counting the possessions of players within one specified team ($F(1, 28) = 26.759, p < 0.001, \eta_p^2 = 0.489$). Similarly, participants were more accurate when trials consisted of two teams of 10 players as compared to trials on which five teams of four players were presented ($F(1,28) = 6.735, p = 0.015, \eta_p^2 = 0.194$; see **Figure 3a**). Subsequent analyses suggested that participants were more accurate when counting the number of possessions of a specific team when there were two teams of 10 players, as compared to five teams of four players ($t(28) = 2.732, p = 0.011, \text{Cohen's } d = 0.41$). This difference between two and five teams was absent when participants were counting possessions of all players regardless of team ($p = 0.66$). In addition, when tracking the number of possessions of a specified team, participants performed significantly better than in the baseline condition regardless of whether two or five teams were presented in the display. This was confirmed by a one sampled t-test against 0 (baseline; both $ts(28) > 5.981$, both $ps < 0.001$, both Cohen's $ds \geq 1.1$). When players were counting possessions of players in all teams, performance was marginally better than in the baseline condition (both $ts(28) > 1.91$, both $ps < 0.066$, both Cohen's $ds \geq 0.355$).

Second, a similar ANOVA on the average relative pupil diameter (calculated as the mean conditional diameter divided by the overall pupil diameter), showed that participants had a smaller relative pupil size when counting possessions of all players independent of team, compared to counting possessions by players from one team (independent of number of teams; $F(1, 28) = 15.14, p = 0.001, \eta_p^2 = 0.351$; see **Figure 3b**). In addition, participants displayed a smaller relative pupil size when two teams of 10 players were presented as compared to five teams of four players ($F(1, 28) = 11.223, MSE = 0.001, p = 0.002, \eta_p^2 = 0.286$). Post-hoc comparisons showed a smaller relative pupil size when participants were counting possessions by players of a specific team, when only two, as compared to five teams were presented ($t(28) = 4.429, SEM = 0.005, p < 0.001, \text{Cohen's } d = 0.521$ See **Figure 3b**). Finally, one-sampled t-tests showed that in all conditions, pupil size was smaller than in the baseline condition (all $ts(28) \geq 2.54$, all $ps \leq 0.018$, all Cohen's $ds \geq 0.474$).

FIGURE AND TABLE LEGENDS:

Figure 1: Design matrix. A 2 x 2 design was used to investigate the role of grouping by color similarity on observers' ability to track a soccer ball in a dynamic display. Critical factors were the number of teams in the display (*i.e.*, the number of different colors; NTeams) and whether the participants were instructed to track possessions within one specified team or to track the possessions of all players in the display (CountTask). A baseline condition was introduced consisting of 20 players from the same team. To ensure measurements with the least amount of noise, we subtracted performance in the baseline condition (the condition with one team/color) from the average performance in each of the four experimental conditions, separately for each participant.

Figure 2: Written, illustrated instructions for participants. (a) Players of teams with differently colored “jerseys” moved in small steps through the display. Individual players could overlap, reflecting real-world soccer conditions (where one player can occlude another player). However, to ensure accurate task performance both the soccer ball and the jersey-number were always clearly visible on top of each player. (b) Example displays of what does and does not constitute a possession. A possession is defined as a player receiving the ball from another player, regardless of whether the player that “played” the ball was from the same or a different team. Crucially, the first possession when the trial starts does not count as a possession, nor is it a possession when the player moves with the ball to a new location.

Figure 3: Measurements. (a) Relative accuracy in counting the number of possessions in the experimental conditions minus the performance in the baseline condition. Negative values signify an increase in accuracy as compared to baseline. (b) Relative pupil size during the experimental conditions, divided by the overall pupil size over all trials. Negative values indicate that pupil size was smaller compared to baseline conditions. All data is hypothetical data that follow numerical and statistical patterns presented in Akyuz *et al.*¹⁰. Error bars reflect the standard error of the mean.

DISCUSSION:

The protocol presented here promotes a form of abductive research, beginning with a real-world observation and then using behavioral and eye tracking measures to seek out the most parsimonious explanation. Although the current research climate almost exclusively promotes deductive research, there are numerous circumstances in which an observation does not fit within current theories and an alternative explanation must be formed. The present protocol focusses on a situation where a benefit of being able to pay attention to more objects organized into fewer groups compared to fewer objects arranged in more groups was abstracted after observing a soccer match with players in different colored jerseys. As this increase in information processing could not be explained by current theories which posit that the human visual system can keep track of about four individual objects at one time⁴⁻⁵, it was necessary to develop a new account for these observations. Importantly, the present protocol is aimed to provide researchers an avenue for testing hypotheses when observations cannot be interpreted within an already established theoretical framework.

The most critical steps in this protocol were using convergent measures obtained from available behavioral and eye tracking recording equipment, keeping the experimental stimuli as close to the real-world situation as possible while holding all other parameters constant or randomizing parameters so that only the key variables of interest (in this case, color similarity, grouping, and the observer’s task) could account for any differences in the observed results.

There are several parts of the current protocol that may potentially require some troubleshooting. First, depending on the eye tracking equipment and setup, and the participants, there may be a need to recalibrate frequently. For example, it is crucial that the experimenter monitors the eye tracking software during recording to ensure that the pupil is not lost. If the eye tracker loses the pupil, the block should be stopped and re-done to ensure proper data collection.

For this reason, it is essential both to include enough trials per condition and to organize trials into short blocks that allow participants a break every 3 to 5 min. However, there are some situations where the particular participant has great difficulty in performing the eye tracking task over the entire duration of the experiment, due to fatigue, dry eyes, contact lens problems, *etc.* In such cases, it may be necessary to dismiss the participant from the experiment. In addition, participants will occasionally make mistakes in their behavioral responses, for example, by accidentally typing “99” instead of “09.” Depending on the task and the experimental question, it will be necessary to determine whether to allow participants the opportunity to correct their responses on each trial, or whether extreme responses are trimmed at the data analysis stage. If the experimental question is more concerned with perceptual processes, as in the present scenario, filtering out extreme responses at the analysis stage can help to avoid situations in which participants spend more time thinking about their responses than they normally would observing the parallel real-world event. However, if the experiment is designed to investigate longer-duration processes such as memory or information retention, allowing participants the opportunity to correct their responses during the experiment may be more beneficial.

The present method is limited in two main ways. First, as with all abductive designs, the goal is to find the simplest explanation of the observed phenomenon. Obviously, there are plenty of opportunities for the resultant explanation to fail to extend to similar circumstances. Therefore, it is absolutely essential that any findings from abductive studies be used to form further, more specific hypotheses that can then be tested using more traditional deductive methods. In addition, our study was limited by the relatively low sampling rate of the available eye tracking equipment. As we could sample every 16.67 ms, we were able to record a relative measure of observers’ pupil size. Using equipment with a sampling rate of 500 Hz or greater would allow future investigations to measure more temporally fine-grained aspects such as ballistic eye movements (saccades), which can provide further insight into how the experimental parameters affect perceptual fluency and speed.

Abductive research is meant to find the simplest explanation when none currently exist. Then, this explanation can be posed as a hypothesis with predictions that can be tested in a more traditional, systematic, deductive framework. For example, building on our initial findings that observers are more accurate and exert less effort when tracking the interactions of more objects organized into fewer groups *versus* fewer objects organized into more groups, future studies can extend the external validity of this paradigm by investigating the contributions of social cues such as body language to tracking the interactions of players in real-world sports matches, perhaps by superimposing simple colored-circles on players from real-world footage to retain their actual movement patterns.

ACKNOWLEDGMENTS:

The authors have no acknowledgements.

DISCLOSURES:

The authors have nothing to disclose.

REFERENCES:

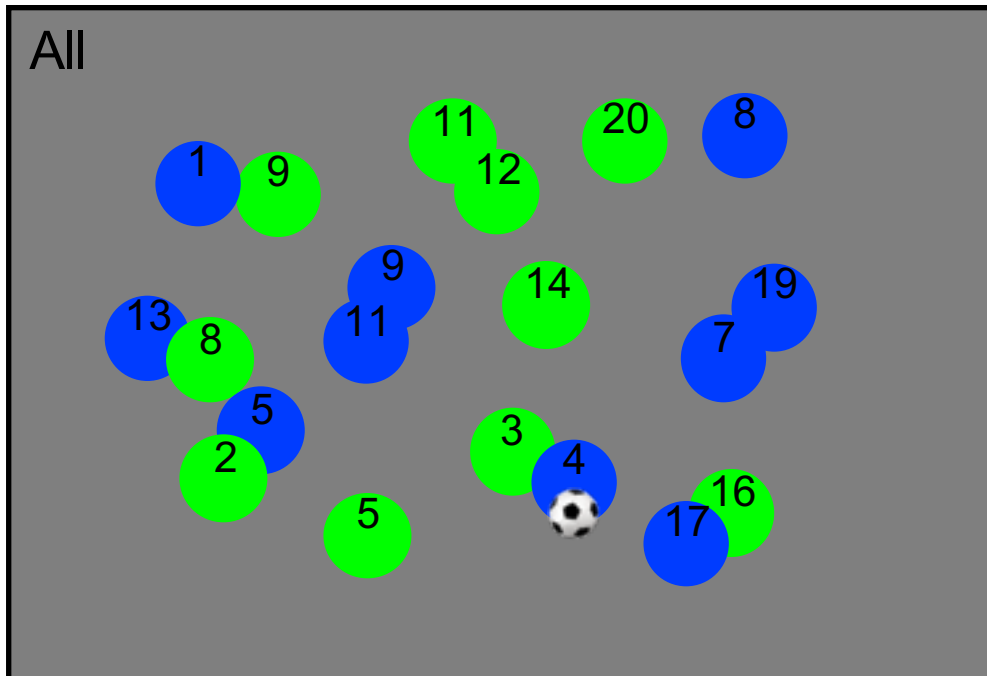
1. Koffka, K. Principles of Gestalt psychology. Harcourt, Brace. Oxford, England. (1935).
2. Kanizsa, G. Subjective Contours. *Scientific American*. **234**(4), 48-53 (1976).
3. Wertheimer, M., Riezler, K. GESTALT THEORY. *Social Research*. **11**(1), 78-99 (1944).
4. Erlikhman, G., Keane, B. P., Mettler, E., Horowitz, T. S., Kellman, P. J. Automatic feature-based grouping during multiple object tracking. *Journal of Experimental Psychology: Human Perception and Performance*. **39**(6), 1625 (2013).
5. Scholl, B. J., Pylyshyn, Z. W., Feldman, J. What is a visual object? Evidence from target merging in multiple object tracking. *Cognition*. **80**(1-2), 159-177.4 (2001).
6. Luck, S. J., Vogel, E. K. The capacity of visual working memory for features and conjunctions. *Nature*. **390**, 279-281 (1997).
7. Pylyshyn, Z.W., Storm, R.W. Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*. **3**, 179–197 (1988).
8. Porter, G., Troscianko, T., Gilchrist, I. D. Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology*. **60**(2), 211-229 (2007).
9. Chun, M. M., Jiang, Y. Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*. **10**, 360–365 (1999).
10. Akyuz, S., Munneke, J., Corbett, J. E. Set similarity modulates object tracking in dynamic environments. *Attention, Perception & Psychophysics*. **80**(7), 1744-1751 (2018).

CountTask

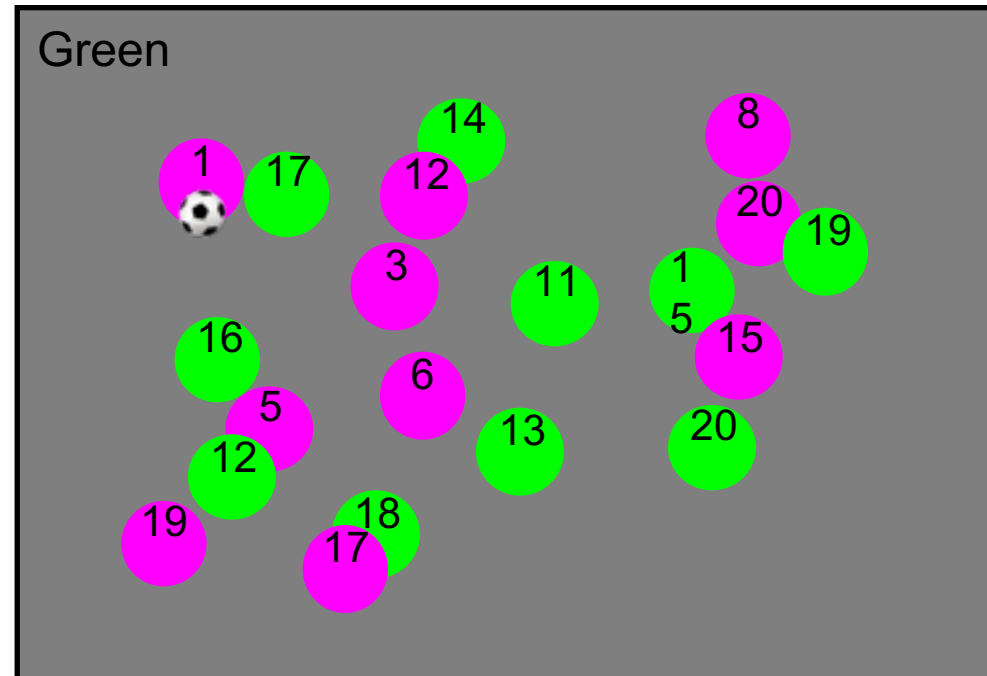
NTeams

2

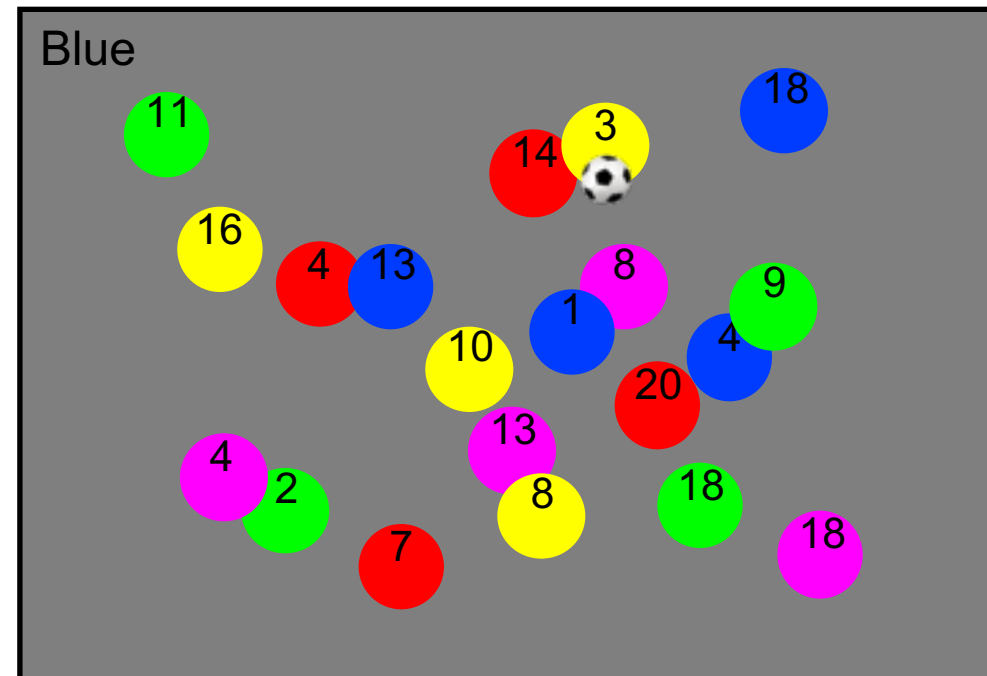
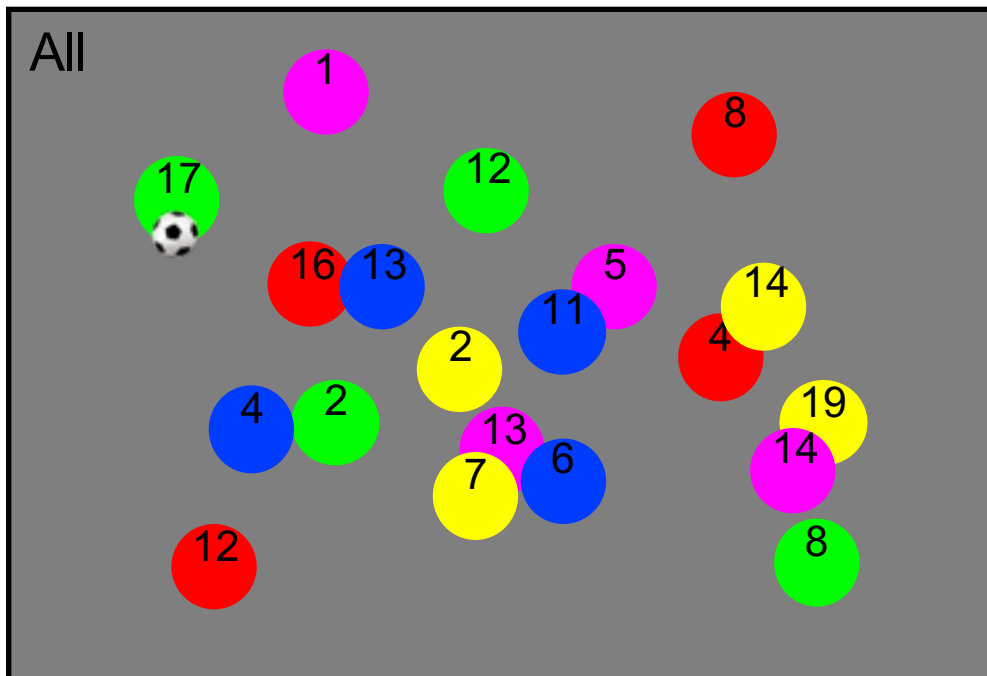
All



Group



5



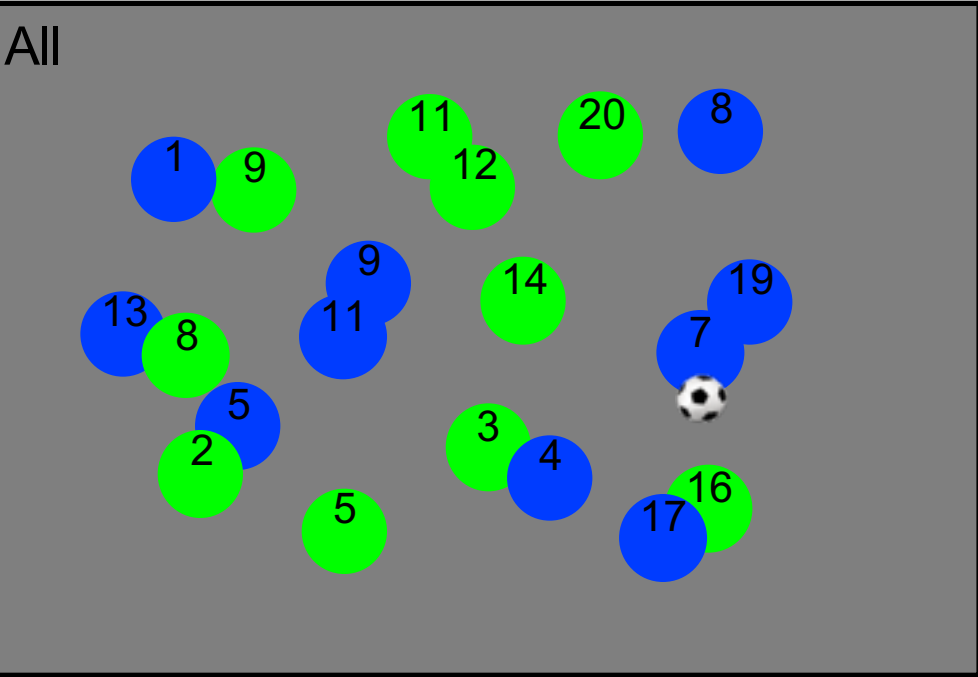
Subtract

Baseline

1Team/Baseline



a.

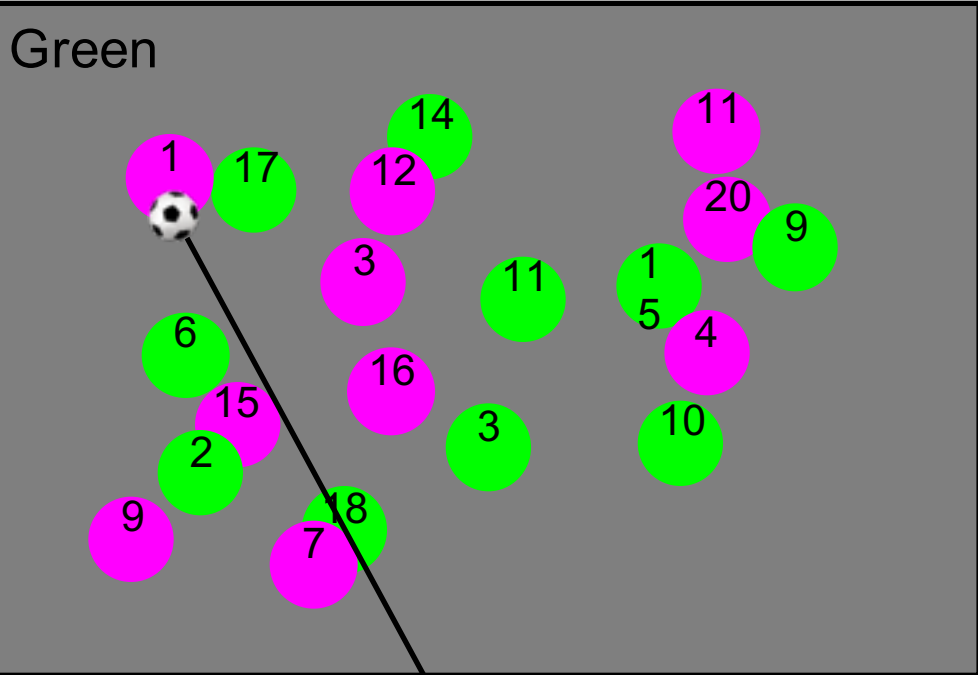


Sometimes the players might overlap (you might have to guess like in a real soccer match)
The ball will be in the center of the player’s “feet”

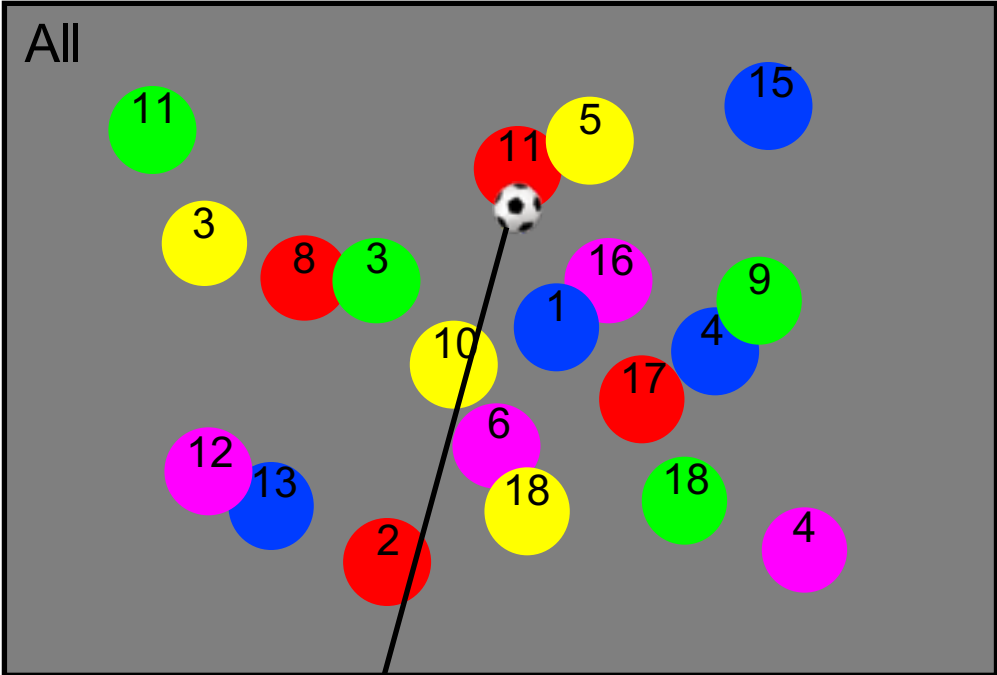
b.

What is a possession?

Getting the ball from another player on a different team

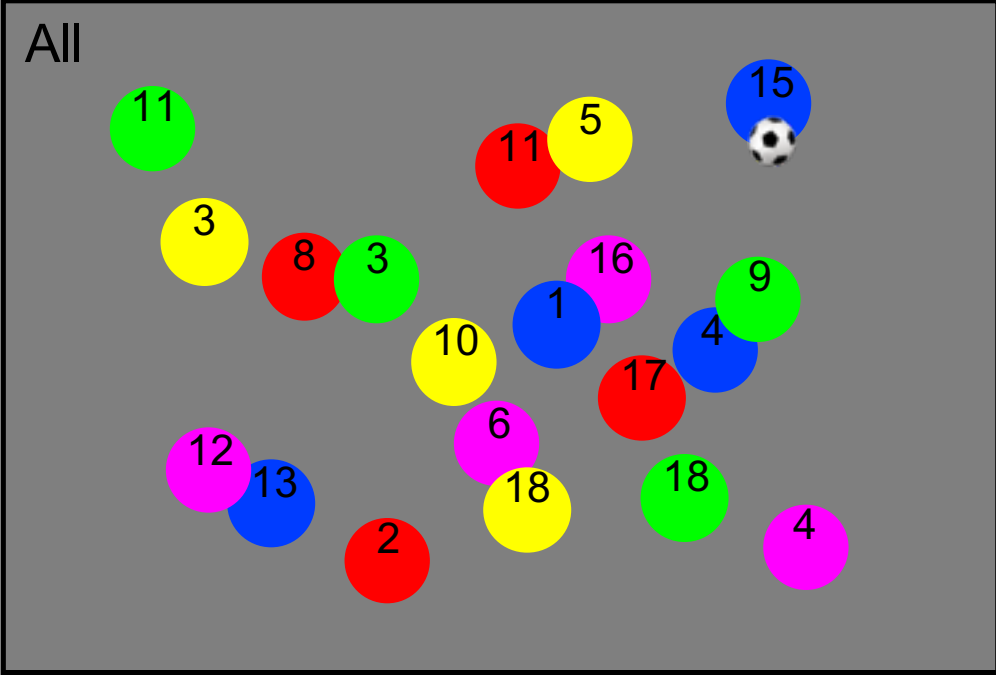


Getting the ball from another player on a the same team

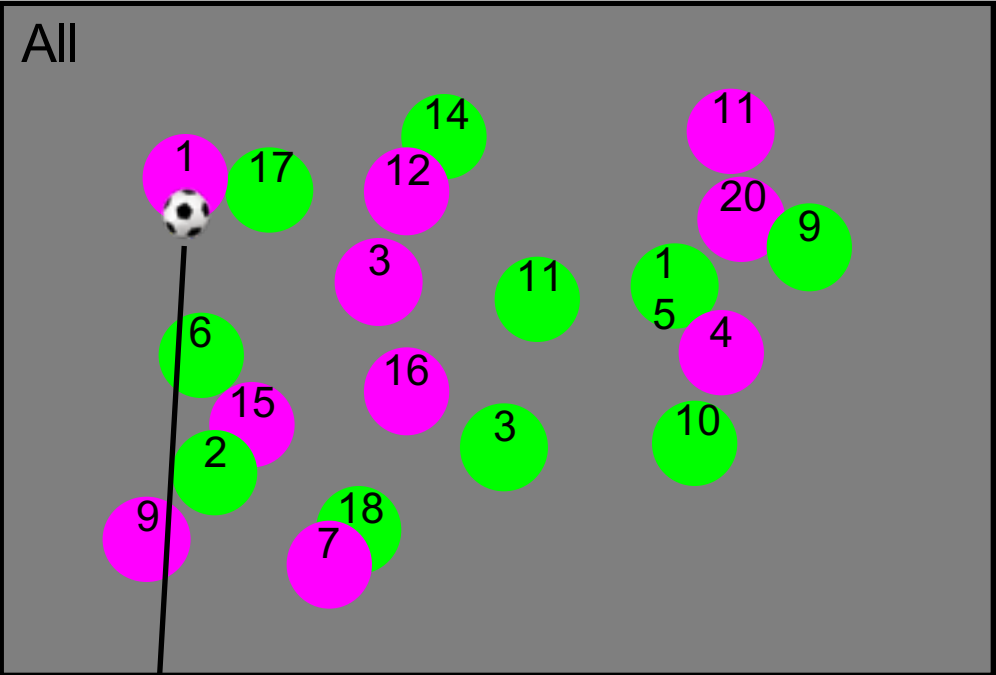


What is not a possession?

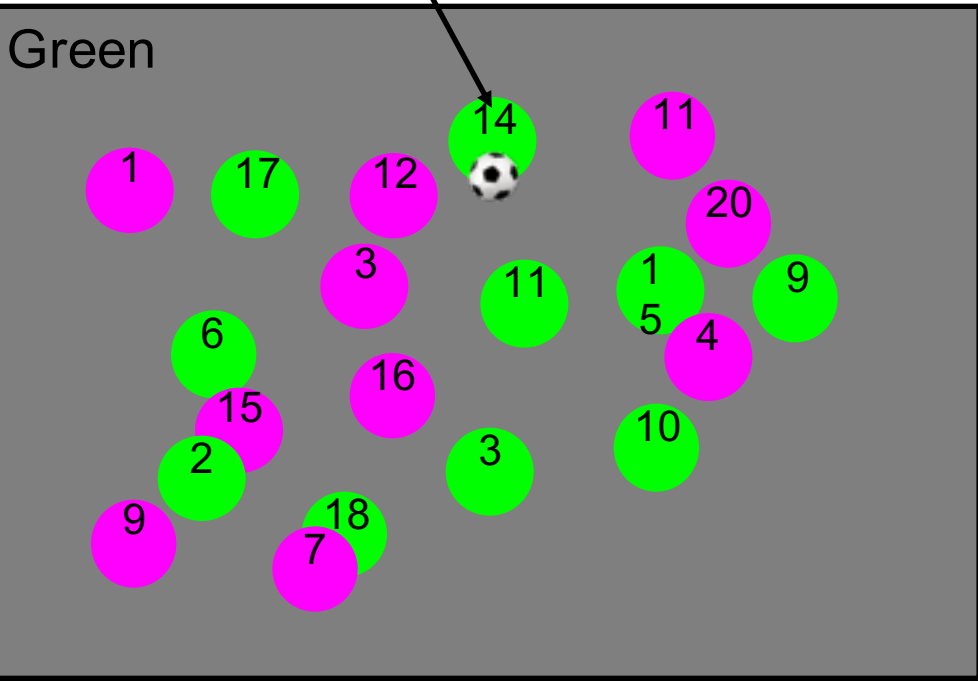
NOT when a player starts the trial with the ball



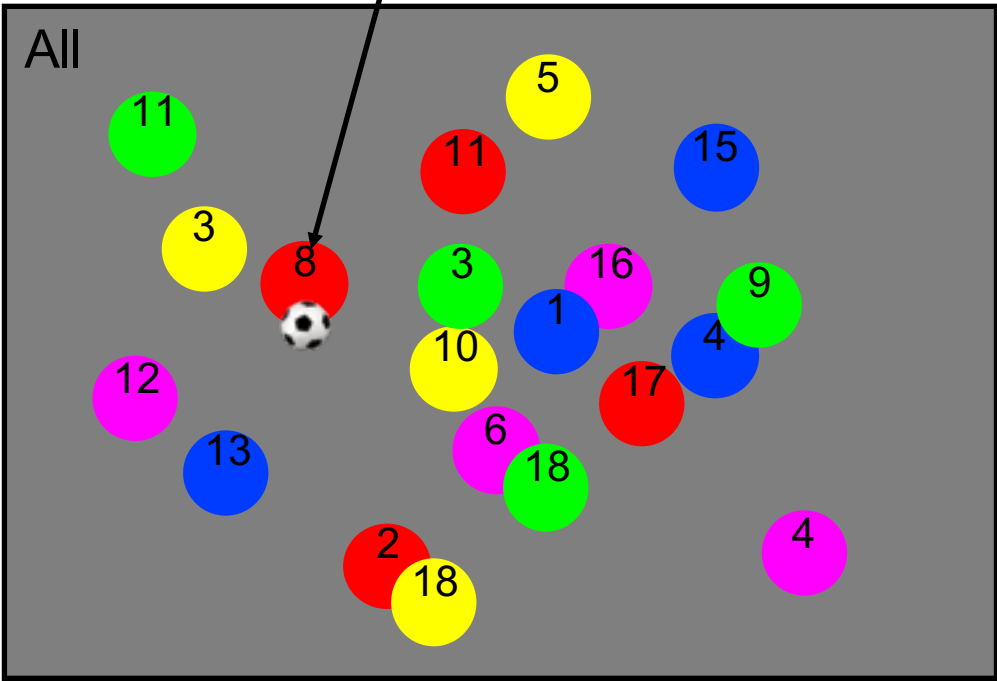
NOT when the same player holds the ball and moves



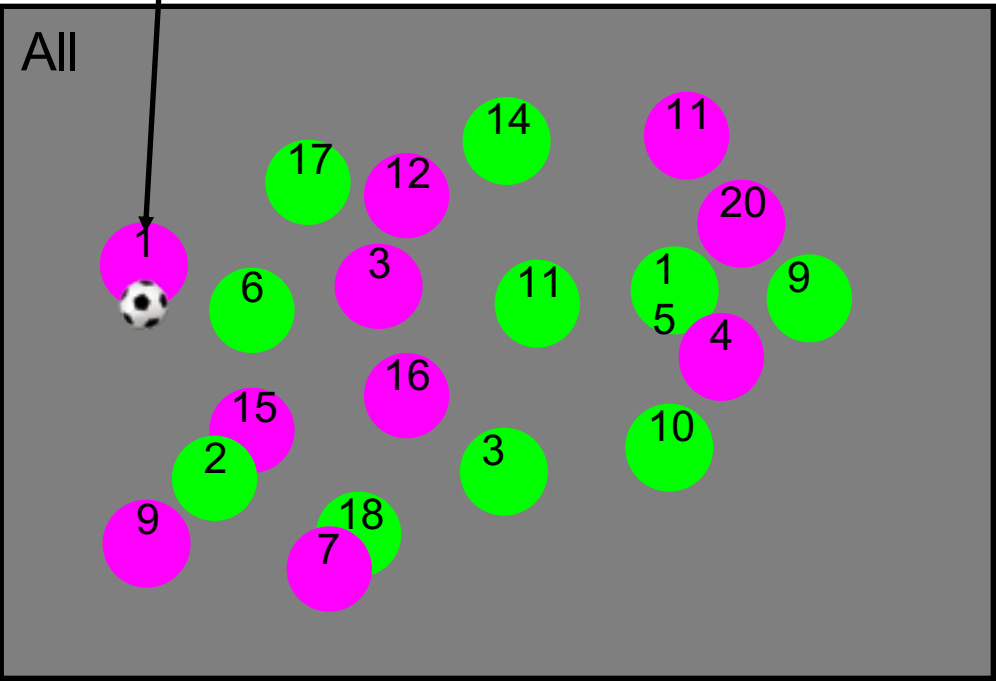
Green



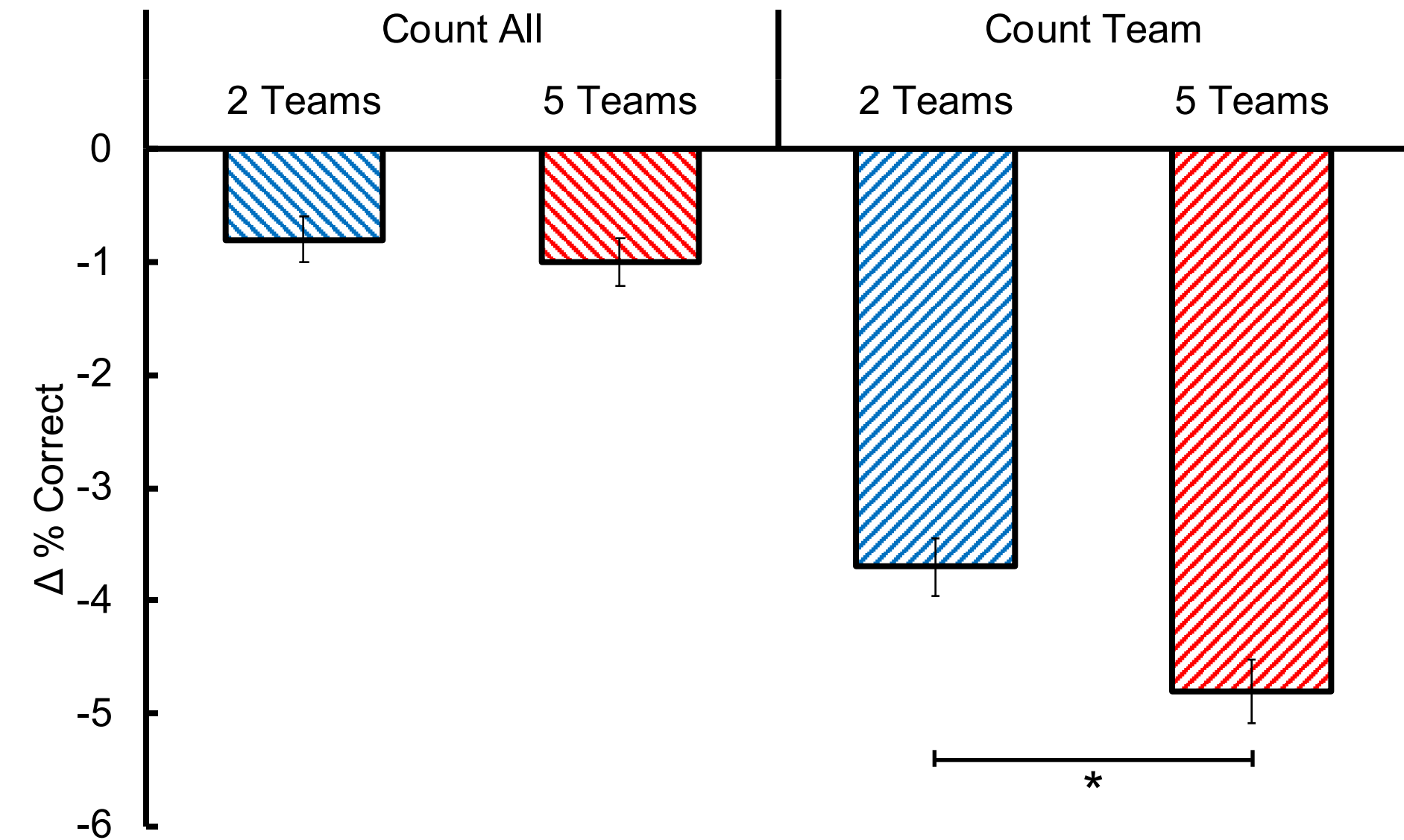
All



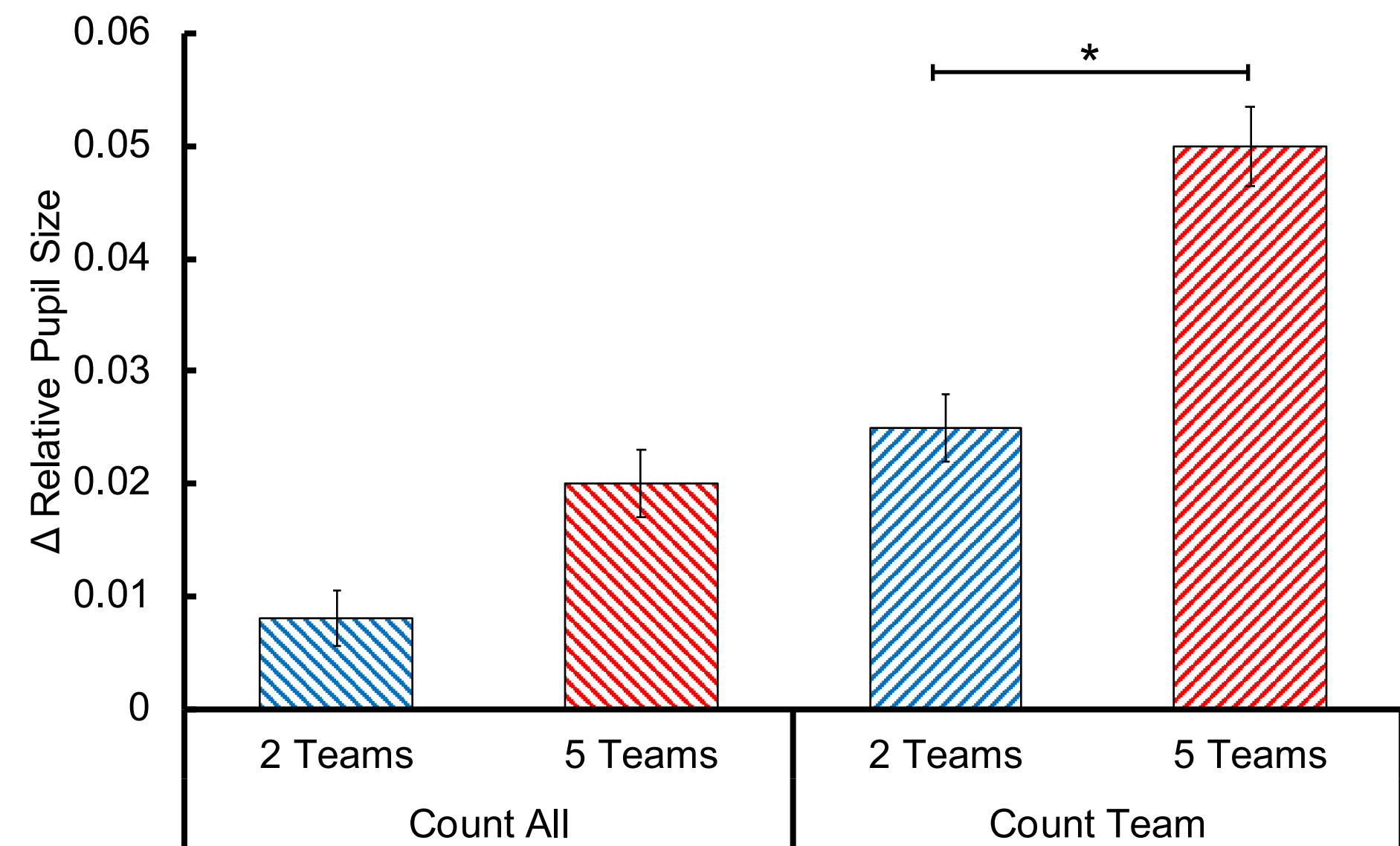
All



a.



b.



Name of Material/ Equipment	Company	Catalog Number	Comments/Description
None until we actually film			

ARTICLE AND VIDEO LICENSE AGREEMENT - UK

Title of Article:	jTesting ideas about visual perception from real-world observations using behavioral and eye tracking measures
Author(s):	Jennifer E. Corbett & Jaap Munneke

Item 1: The Author elects to have the Materials be made available (as described at <http://www.jove.com/publish>) via:

☐ Standard Access ☒ Open Access

Item 2: Please select one of the following items:

- ☒ The Author is **NOT** a United States government employee.
- ☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.
- ☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution 3.0 Agreement (also known as CC-BY), the terms and conditions of which can be found at: <http://creativecommons.org/licenses/by/3.0/us/legalcode>; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion

of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4 and 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video - Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video - Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be

deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole

ARTICLE AND VIDEO LICENSE AGREEMENT - UK

discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. Indemnification. The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or

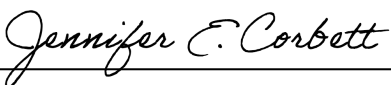
decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. Fees. To cover the cost incurred for publication, JoVE must receive payment before production and publication of the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US\$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. Transfer, Governing Law. This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to be one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Jennifer E. Corbett	
Department:	College of Health and Life Sciences	
Institution:	Brunel University London	
Title:	PhD	
Signature:		Date: 28 August, 2018

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

Editorial Comments:

Comment 1: “Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammatical errors.”

[Response 1: We have proofread the article. We have flagged all changes noted in this letter in blue ink in the text of the revised document.](#)

Comment 2: “Textual Overlap: Significant portions show significant overlap with previously published work. Please re-write the text indicated in red in the attached document to avoid this overlap.”

[Response 2: We have rewritten these sections of red text.](#)

Comment 3: “Please include an ethics statement before your numbered protocol steps indicating that the protocol follows the guidelines of your institutions human research ethics committee”.

[Response 3: We now include an ethics statement from Bilkent University, where the research was originally conducted. Please note that this is not our current research institute, Brunel University London, where the video will be filmed.](#)

Comment 4: “Protocol Language:

1) Please ensure that ALL text in the protocol section is written in the imperative voice/tense as if you are telling someone how to do the technique (i.e. “Do this”, “Measure that” etc.) Any text that cannot be written in the imperative tense may be added as a “Note”, however, notes should be used sparingly and actions should be described in the imperative tense wherever possible. Some examples that are NOT in imperative voice: 2.1.1., 2.1.2, 7.1, 8.1.

2) Every main step should have at least 1-2 substeps under it (please edit steps 5, 6).“

[Response 4: We now ensure all text is in imperative tense, and that there are at least 1-2 sub-steps under each main step.](#)

Comment 5: “Protocol Detail: Please note that your protocol will be used to generate the script for the video, and must contain everything that you would like shown in the video. Please add more specific details (e.g. button clicks for software actions, numerical values for settings, etc) to your protocol steps. There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol.”

[Response 5: We have given all details of button presses \(numbers on the keyboard and space bar presses\), and detailed all general actions regarding the stimulus presentation and eye tracking software, equipment, and numeric settings. We note that the particular settings for the eye tracking recording system will be dependent on the manufacturer’s recommendations.](#)

Comment 6: *“Protocol Highlight: Please highlight ~2.5 pages or less of text (which includes headings and spaces) in yellow, to identify which steps should be visualized to tell the most cohesive story of your protocol steps. Please see JoVE’s instructions for authors for more clarification. Remember that the non-highlighted protocol steps will remain in the manuscript and therefore will still be available to the reader.*

- 1) *The highlighted steps should form a cohesive narrative, that is, there must be a logical flow from one highlighted step to the next.*
- 2) *Please highlight complete sentences (not parts of sentences). Include sub-headings and spaces when calculating the final highlighted length.*
- 3) *Notes cannot be filmed and should be excluded from highlighting.*
- 4) *Please bear in mind that software steps without a graphical user interface/calculations/ command line scripting (such as steps 2-2.1.2, 3-6, 10-10.4.2) cannot be filmed.*
- 5) *Please ensure that the manuscript title best reflects the filmable content (i.e. the portions you highlight).”*

[Response 6:](#) *We have now highlighted the steps of the protocol to be filmed.*

Comment 7: *“Discussion: JoVE articles are focused on the methods and the protocol, thus the discussion should be similarly focused. Please ensure that the discussion covers the following in detail and in paragraph form (3-6 paragraphs):*

- 1) *modifications and troubleshooting,*
- 2) *limitations of the technique,*
- 3) *significance with respect to existing methods,*
- 4) *future applications and*
- 5) *critical steps within the protocol. “*

[Response 7:](#) *The discussion covers each of these points.*

Comment 8: *“Figure/Table Legends: Please expand the legends to adequately describe the figures/tables. Each figure or table must have an accompanying legend including a short title, followed by a short description of each panel and/or a general description.”*

[Response 8:](#) *The figures each have a short title and description.*

Comment 9: *“Commercial Language: JoVE is unable to publish manuscripts containing commercial sounding language, including trademark or registered trademark symbols (TM/R) and the mention of company brand names before an instrument or reagent. Examples of commercial sounding language in your manuscript are Matlab®.*

- 1) *Please use MS Word’s find function (Ctrl+F), to locate and replace all commercial sounding language in your manuscript with generic names that are not company-specific. All commercial products should be*

sufficiently referenced in the table of materials/reagents. You may use the generic term followed by “(see table of materials)” to draw the readers’ attention to specific commercial names.”

[Response 9:](#) We have removed any commercial language.

Comment 10: “Table of Materials: Please revise the table of the essential supplies, reagents, and equipment. The table should include the name, company, and catalog number of all relevant materials/software in separate columns in an xls/xlsx file. Please include items such as software used, eye-tracking equipment.

- Please define all abbreviations at first use.*
- If your figures and tables are original and not published previously or you have already obtained figure permissions, please ignore this comment. If you are re-using figures from a previous publication, you must obtain explicit permission to re-use the figure from the previous publisher (this can be in the form of a letter from an editor or a link to the editorial policies that allows you to re-publish the figure). Please upload the text of the re-print permission (may be copied and pasted from an email/website) as a Word document to the Editorial Manager site in the "Supplemental files (as requested by JoVE)" section. Please also cite the figure appropriately in the figure legend, i.e. "This figure has been modified from [citation]."*

[Response 10:](#) We are unsure how to proceed. The protocol will be filmed in a different laboratory than the where the study was originally conducted. Should we include the specifics of the software and hardware that we will use, or that we previously used? For example, we will use a different eye tracking system and different versions of stimulus presentation software to film the video than we did to conduct the original research. However, until a filming date is known, we will not know such specifics because we need to book one of three eye tracking systems, each using slightly different software, depending on which is available on that date at our current university.

Reviewer #1:

Comment 1: “-2.1.1. "... with the same number of circles per group (e.g., two groups of 10 circles each, and five groups of four circles 110 each)." That doesn't seem right: 10 circles and 4 circles is not 'the same number of circles per group.'"

[Response 1:](#) We have reworded this sentence to clarify that we mean the circles should be equally divided into groups (e.g., 4 groups of 5 circles in each of these 4 groups, instead of 4 groups with 3 circles in one group, 5 circles in the second group, 6 circles in the third group, and 2 circles in the fourth group).

Comment 2: “The authors have not taken into account the difference in luminance between the various colors used, which may introduce error variance in the pupil data. Have they considered using equiluminant colors (e.g. Teufel colors)? Is it

possible that attention to one particular color influences pupil size depending on the luminance of that color? Or are those potential confounds negligible compared to the main effects of task and number of teams?”

[Response 2:](#) We have clarified that randomization of team colors on each trial ensures there are no luminance confounds while maintaining more “real-world” circumstances in section 3.2.

[Comment 3:](#) “Under ‘representative results’ the authors describe inferential statistics, which suggests that they actually tested a number of subjects. But information about those subject is missing.”

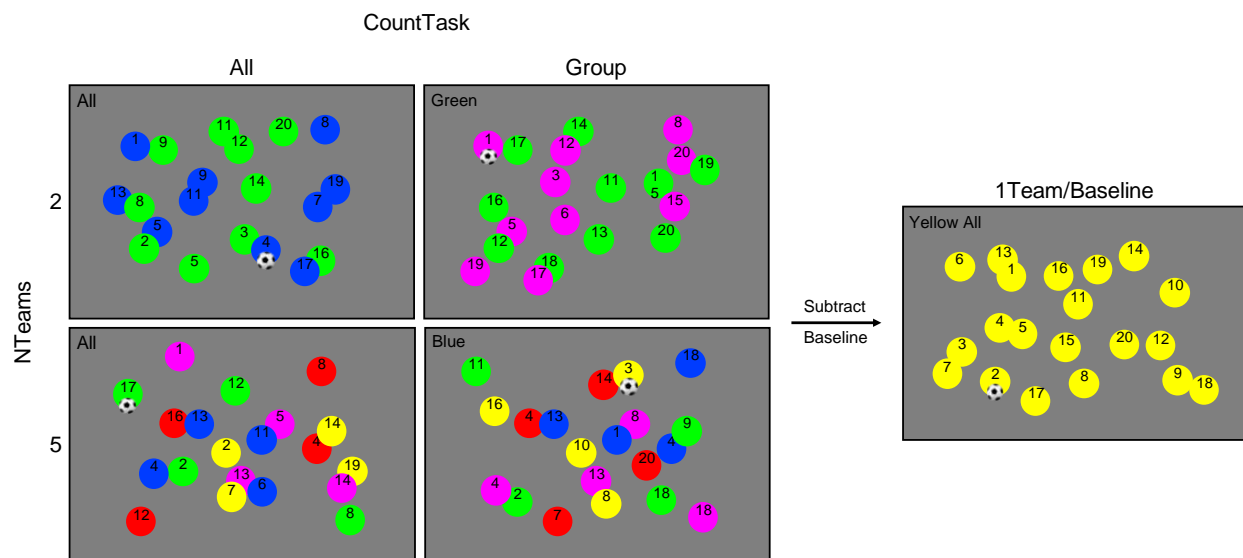
[Response 3:](#) We now include a short description of the participants tested in the original study from which the current protocol is taken.

[Comment 4:](#) “Line 280: “.., as compared the five teams of four 281 players”. ‘the’ should be ‘to’

[Response 4:](#) Corrected

[Comment 5:](#) “I could not find Figure 1, the design matrix”

[Response 5:](#) Perhaps the Figure 1 was not available in the version of the manuscript sent to Reviewers? We have appended it here:



Reviewer #2:

[Concern 1:](#) “It is true that many papers nowadays assume some well-established theory or other as starting points for the investigations they describe - a trend strengthened by the ‘new wave’ of replication-based studies - and that many researchers tend to lose sight of core issues in scientific research (which is

ultimately meant to advance the general body of knowledge). However, the authors should do a better job at explaining exactly what phenomenon they are illustrating by distinguishing the motivation of their approach from its relevance, and by detailing its consequences. A naïve reader might be led to believe that the authors advocate a revolutionary approach to Experimental Psychology.”

Response 1: We reworded portions of the manuscript to better convey the motivation and relevance of our approach. In short, we were motivated by the observation that different colored jerseys in soccer matches seemed to afford the viewer some differential and potentially task-dependent benefits in tracking the interactions between players. However, we could not understand this effect in terms of any available theories of perceptual organization. In fact, as we now discuss, the existing MOT literature suggests that grouping may even *impair* observers’ abilities to track players’ interactions. Therefore, we adopted an abductive approach by forming a more concrete hypothesis based on our observations.

Concern 2. *“Do the authors imply that there is a dearth of new theories available to cover the whole range of observable phenomena and call for new theories to be built based on empirical observations, or do they simply claim that hypothesis-testing need not happen within an already available theoretical framework, thereby advocating a return to the roots of scientific investigation (they mention the work of Gestalt psychologists)? The authors need to explain the difference between theory and hypothesis.”*

Response 2: We have now clarified in the introduction and conclusion of the manuscript that we are advocating for a return to the roots of scientific investigation in the same spirit as Gestalt Psychologists, such that hypothesis testing does not necessarily need to happen within an already established theoretical framework.

Concern 3. *“The authors claim that “existing theories could not explain our observations” (p2, line 71). However, links can still be established between their empirical observation and several already available theoretical principles. In vision research, multiple-object-tracking MOT studies have demonstrated that the number of tracked objects is no longer at issue as long as they are part of the same group. Gestalt principles themselves partly explain the authors’ empirical hypothesis. In memory research, mnemonic strategies include ‘chunking’ and ‘grouping’, which help people better keep track of concepts in working memory. In brief, I don’t feel like there is a scientific void that would need to be filled, but rather dots to be connected and better use to be made of several theoretical principles out there to fully explain the empirical observation the authors discuss.”*

Response 3: Although we had already conducted an intense search of the visual perception literature to find any existing theories that could account for our observations when preparing the manuscript from which the present methodology is taken, we again searched specifically within the MOT literature to find studies

that have demonstrated that the number of tracked objects is not a factor when objects are part of the same group. On the contrary, we could find no such study, but we now reference additional MOT studies suggesting that when objects are part of the same group, tracking is impaired (e.g., Erlikhman, et al., 2013; Scholl, et al., 2001). We do ask that the Reviewer please disclose any particular study or set of studies referencing such MOT effects, as we would need to revise the current manuscript as well as perhaps publish a corrigendum to the already published manuscript from which the present methodology is taken. We agree that Gestalt principles and chunking are somehow a partial explanation for our observations, but it was not clear how either could lead to better performance. As we note in the Introduction, "...it is not always possible to deduce specific ideas from general principles when the goal is to understand these general principles."