



# Does optic flow provide information about actions?

Hannah L. Masoner<sup>1</sup> · Alen Hajnal<sup>1</sup>

Accepted: 9 February 2023 / Published online: 14 March 2023  
© The Psychonomic Society, Inc. 2023

## Abstract

Optic flow, the pattern of light generated in the visual field by motion of objects and the observer's body, serves as information that underwrites perception of events, actions, and affordances. This visual pattern informs the observer about their own actions in relation to their surroundings, as well as those of others. This study explored the limits of action detection for others as well as the role of optic flow. First-person videos were created using camera recordings of the actor's perspective as they performed various movements (jumping jacks, jumping, squatting, sitting, etc.). In three experiments participants attempted to detect the action from first-person video footage using open ended responses (Experiment 1), forced-choice responses (Experiment 2), and a match-to-sample paradigm (Experiment 3). It was discovered that some actions are more difficult to detect than others. When the task was challenging (Experiment 1) athletes were more accurate, but this was not the case in Experiments 2 and 3. All actions were identified above chance level across viewpoints, suggesting that invariant information was detected and used to perform the task.

**Keywords** Motion: Biological · Perception and action · Visual perception

## Introduction

Have you ever sat in a driving simulator that you were not controlling, yet you knew which direction the virtual car was moving? Many studies have demonstrated that the dynamically changing pattern of scattered, reflected, and refracted light, i.e. optic flow, serves as information for guiding ambulatory activities such as steering and avoiding obstacles (Matthis et al., 2022; Warren et al., 2001). A related, but understudied role of optic flow is to specify the activity of the agent in non-translatory activities. For example, the egocentric optic flow generated when a person does jumping jacks contains information that is unique to jumping jacks. By the same token, the egocentric optic flow pattern is different when a person is simply jumping. Can observers identify the activity and differentiate it from other similar activities by attending to the egocentric optic flow pattern? If you ever played a first-person video game and were required to understand your avatar's actions while your own body was stationary, you were successfully attending to the

optic flow patterns that allowed you to recognize the actions undertaken by the agent. Another example of action detection from egocentric viewpoints is the use of body worn cameras (BWCs) in law enforcement. The footage from these cameras not only reveals information about the visual field ahead but also reveals the movement of the officer wearing the camera. There is an ongoing debate about the ethics and efficiency of these devices, yet they are commonly used as evidence in court (Laming, 2019; Lum et al., 2020) to infer what was the person wearing the camera doing during law enforcement activity.

Furthermore, the detection of actions based on optic flow is not exclusive to humans. Artificial agents such as computers use algorithms that detect and analyze optic flow patterns in order to control self-driving vehicles such as cars and airplanes (Fan et al., 2019; Ruffier & Franceschini, 2005). This line of research traces its lineage to Gibson (1947), who first posed landing of airplanes as a perceptual problem for which optic flow can serve as information. Danafar and Gheissari (2007) explored the application of optic flow algorithms in computer vision when assessing surveillance footage from security cameras. Actions such as walking, jogging, clapping, and boxing were evaluated. The success rate in determining the actions performed was around 85% even though the videos were taken from different viewpoints

---

✉ Hannah L. Masoner  
hannah.masoner@usm.edu

<sup>1</sup> School of Psychology, The University of Southern Mississippi, Hattiesburg, MS, USA

and under different levels of illumination, suggesting that the pattern of motion of an action must be invariant across vantage points (Holte et al., 2010). Optic flow is also used in creating models that allow robots to more effectively interact with humans. Vignolo et al. (2017) created a computational model to allow robots to distinguish biomotion from non-biomotion so that they may advance the robots' social behaviors. Noceti et al. (2019) have utilized optic flow patterns created by the performance of an actor to enhance the coupling of robot and human interaction when performing an action-timing task (e.g., walking in synchrony), much like humans learn to work together.

What is the nature of the information contained in optic flow that specifies actions? Relatedly, what aspects of optic flow do perceivers attend to detect actions? The theory of kinematic specification of dynamics (KSD; Runeson & Frykholm, 1981) suggests that observers attend to and are able to detect the pattern of kinematic variables such as displacement and speed of visual elements of the optic array, and that these optic patterns specify actions and events. Runeson and Frykholm showed that the patterning of optic flow is detectable even in situations in which only a few visual elements are visible, such as in point-light displays (PLDs) of the joints of the body in motion (Johansson, 1973). Importantly, the information in egocentric optic flow patterns should be the same as the information in point-light displays of a moving body. Even though in the egocentric view the body is not visible, the consequences of the action are present in the optic flow pattern of the changing light intensities of visual elements of the ambient optic array as the egocentric viewpoint continuously changes location during body movements. Importantly, we can recognize actions without acting ourselves and without seeing the body of the actor (or an avatar of the body). The goal of the current study was to demonstrate that (1) perception of human activity is possible based on viewing egocentric optic flow alone, and that (2) this perceptual skill is a function of experience. Successful demonstration of the perceptual skill will serve as a preliminary step to future investigations of the nature of the invariant pattern regardless of whether the optic flow is experienced from a first-person (egocentric) viewpoint or from a third-person (allocentric) viewpoint.

## Athletes and action detection

The ability to recognize current and future action possibilities (i.e., affordances) for others is especially relevant in sports. It is a key component of skillful timely decisions during a game to better the play or overthrow the competition. Competitive athletes must read the play scenario, considering information from their own movements as well as those of the opponent and their teammates. They identify the action capabilities of all parties and then attune their

own actions to the information (Hacques et al., 2021; Vickers, 2007). For instance, in volleyball, the typical pattern of play on one side of the net is: pass, set, attack. A defensive player (on the opposing side of the net) must recognize an attacker's affordances based on the location of the ball during the second contact, the attacker's location in relation to the ball, their physical capabilities (e.g., jumping height), and their hand and shoulder positions (Klostermann et al., 2015). In beach volleyball, players must be very skilled in identifying the action possibilities of their partner because they must make the appropriate subsequent move based on their partner's play. At a high level the speed of the game is so fast that a player does not have time to react after their teammate's contact but must be able to anticipate the path of the ball to some degree beforehand so that they can act ahead of the play.

Weast et al. (2011) discovered that basketball players were significantly better at judging a person's ability to jump and reach when compared to non-basketball players. However, there were no differences in judging ability to sit or reach without jumping. It seems that athletes are more sensitive to affordances directly influenced by kinematic information as opposed to static measurements alone.

In cases where biological motion is the only information (i.e., physical details about shape are not available) athletes have demonstrated impressive skills in perceiving actions from PLDs, including whether the actor was a teammate or stranger (Steel et al., 2015). Weast et al. (2014) found that body motion alone provided enough information for athletes to detect affordances for another person when related kinematic information was observed (e.g., watching the motion of an actor squat and then estimate their reaching height while jumping). Athletes are more attuned to these tasks than non-athletes (Fajen et al., 2009).

In summary, athletes have a keen ability to judge action possibilities by observing another player's body movements. Likewise, competitive athletes are more accurate in action detection based on the amount of time they spend intentionally studying actions and making visual observations while performing. For these reasons the goal of the present contribution is to compare perception of athletes and non-athletes in action detection tasks.

## Purpose and hypotheses

We sought to determine if it is possible to detect another person's actions from a video sample of their first-person (egocentric) perspective view during the activity. To wit, can perceivers detect actions when the body of the actor is not visible, and the video footage only contains what the actor sees in front of her during the activity? Furthermore, does extensive physical training provide athletes with a superior ability to perceive actions?

First, we predicted that observers could perceive actions from a video sample of a first-person view recording that only shows the consequences of the motion of the body, but not the body itself. Second, athletes should have an advantage in determining these actions compared to non-athletes. Athletes are expected to be more accurate and faster than non-athletes. We tested these hypotheses using three different empirical methods: open-ended responses (Experiment 1), forced-choice responses (Experiment 2), and a match-to-sample paradigm (Experiment 3).

## Experiment 1

The goal of the experiment was to determine if human observers can perceive an activity based on video footage recorded from the point of view of the actor who was engaged in the action. We created first-person videos of an actor performing six separate actions. These included jumping jacks, jumping, sitting, squatting, skipping, and jogging. The videos showed the actor's perspective during movement, but not their body. The key component of this manipulation is to demonstrate whether observers can recognize the activity based on the head-mounted camera's movements without seeing the body of the actor. We hypothesized that the optic flow pattern generated by the camera movement contains information that specifies the action, and that this information can be detected by observers.

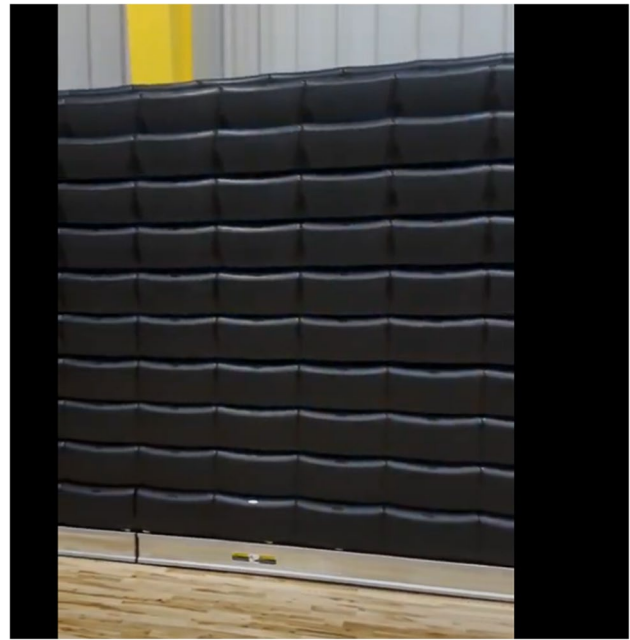
## Method

### Participants

This experiment utilized an online platform and was available to several groups of participants. The first group consisted of participants recruited via the Psychology Department's SONA participant pool who received course credit in their psychology classes for their contribution. The second group was made up of students who competed for one of the varsity sports teams at the university. Participants were categorized in two groups: Non-Athletes ( $n = 50$ ) and Athletes ( $n = 19$ ).

### Materials

For all experiments we created a set of video stimuli using a GoPro (Hero8) sports camera. The videos for Experiment 1 provided a first-person world view and did not give any information about the actor's physicality such as body shape and size. The backdrop for the videos was a set of black retractable bleachers that were withdrawn so that they create a vertical wall-like structure (Fig. 1). The intention for using this background was to provide enough disparity and texture



**Fig. 1** Picture of the set of bleachers used as the scene of the optic flow sample videos

to give rich visual information, but not to give a surplus of detail to make the task too easy. Videos were recorded for six actions. The actions were grouped as three action pairs:

1. Jumping–Jumping Jacks
2. Squatting–Sitting
3. Skipping–Jogging

These actions were chosen because they should be somewhat familiar to most people and are commonly incorporated in exercise, sports, and everyday behavior. The actions were paired with the intent of being similar, so that the task was not too easy, yet different enough to be distinguishable. Specifically, the movement patterns of each action pair had similar cycles, directions and ranges of motion to make the perceptual discrimination hard, but not impossible.

### Experimental design

In Experiment 1 we employed a 2 Athletic Status (athlete versus non-athlete)  $\times$  6 Action mixed-design ANOVA to observe the differences in athlete status and all six actions. Additionally, a 2 Athletic Status (athlete versus non-athlete)  $\times$  2 Action pair mixed design was performed so that athlete status was a between-subject variable, and Action pair was a within-subjects variable. Three mixed 2  $\times$  2 ANOVAs were conducted for the following Action pairs, respectively: jumping and jumping jacks, squatting and sitting, skipping and jogging. All participants underwent the same experimental

procedures with stimuli being presented in a randomized order.

## Procedure

Online experiments were programmed using the Collector data collection software (Garcia et al., 2015) to randomize stimuli for each participant. An online link for the experiment was distributed to both target populations simultaneously so that data for both groups was collected over the same window of time. Participants accessed the online link by using their laptop or desktop computer. A demographic questionnaire was initially presented that inquired about the person's athletic status. This allowed us to determine if they met the qualifications for being included in the athlete group. Any participant who was currently rostered on a university sports team or had been rostered within the past year was included as an athlete.

For the experiment each video was presented randomly four times for a total of 24 trials. Each video was presented one time per trial and lasted about 5 s. For actions such as jumping and squatting the movement was repeated for the 5-s time frame until the participant responded. For actions that require covering ground such as jogging and skipping, a consistent distance was set, and the movement was recorded for the duration of the distance.

To assess people's ability to detect the action we began by asking the general question: "What is the person doing in this video?" Instructions read: "Be as specific as possible but describe the action in no more than two words." In the case that the video did not play appropriately due to technological issues like internet connection, the participant was instructed to enter the word "ERROR" into the text box. Response time for each trial was measured from the moment the response text box appeared and ended when the participant submitted their response. Figure 2 depicts the trial sequence.

## Analyses

A coding scheme was created to categorize participant responses. Categories were determined based on the data collected. For instance, one-word responses such as "jump," "hop," and "bounce" were coded as a jump. After categorizing responses, we determined the accuracy for each trial

and labeled them based on correctness (1 for correct, 0 for incorrect). Trials that resulted in an error response due to malfunction or glitch were removed, as well as trials where the participant clearly did not follow the instructions. This resulted in the removal of 11.8% of trials.

A repeated-measures analysis of variance (ANOVA) was performed to observe both dependent variables: accuracy and response time. It was expected that all participants could decipher the type of action in the videos to some degree. It was also anticipated that athletes would perform more accurately and take less time in responding.

## Results

### Accuracy

A 2 Athletic status  $\times$  6 Action repeated-measures ANOVA on accuracy revealed a main effect of Action,  $F(5,315) = 40.72$ ,  $p < .001$ ,  $\eta_p^2 = 0.39$ . Jumps were perceived most accurately ( $M = 0.63$ ,  $SD = 0.37$ ), whereas sitting was perceived least accurately ( $M = 0.01$ ,  $SD = 0.05$ ). There was also a main effect of Athletic status,  $F(1,63) = 11.29$ ,  $p = .001$ ,  $\eta_p^2 = 0.15$ . Athletes were more accurate ( $M = 0.49$ ,  $SD = 0.44$ ) than non-athletes ( $M = 0.35$ ,  $SD = 0.44$ ). There was no significant interaction.

In order to get a more detailed look at the data we followed up the omnibus analysis with separate 2 Athletic status  $\times$  2 Action ANOVAs for each action pair: jog versus skip, jump versus jumping jacks, and sit versus squat. The 2 Athletic Status  $\times$  2 Action pair (jog vs. skip) ANOVA on accuracy revealed a significant effect of Athletic Status,  $F(1,66) = 5.99$ ,  $p = .017$ ,  $\eta_p^2 = 0.08$ . Specifically, athletes ( $M = 0.65$ ,  $SD = 0.35$ ) were more accurate than non-athletes ( $M = 0.44$ ,  $SD = 0.49$ ). No other effects were significant. The same ANOVA comparing jumps and jumping jacks revealed a significant difference between actions,  $F(1,66) = 113.7$ ,  $p < .001$ ,  $\eta_p^2 = 0.63$ . Specifically, jumps ( $M = 0.64$ ,  $SD = 0.38$ ) were detected more accurately than jumping jacks ( $M = 0.07$ ,  $SD = 0.18$ ). The Athletic Status  $\times$  Action pair interaction was also significant,  $F(1,66) = 5.01$ ,  $p = .03$ ,  $\eta_p^2 = 0.07$ . Athletic Status was not significant. The ANOVA comparing accuracy of perceiving sitting and squatting returned a significant difference between actions,  $F(1,64) = 121.57$ ,  $p < .001$ ,  $\eta_p^2 = 0.66$ . Specifically, squats ( $M = 0.60$ ,  $SD = 0.44$ ) were detected more accurately



**Fig. 2** Example of trial for Experiment 1. Responses were collected by typing into a textbox that appeared on the computer screen after the video was presented

than sitting down ( $M = 0.01$ ,  $SD = 0.05$ ). The Athletic Status  $\times$  Action pair interaction was also significant,  $F(1,64) = 5.2$ ,  $p = .03$ ,  $\eta_p^2 = 0.08$ . Athletic Status was also significant,  $F(1,64) = 5.37$ ,  $p = .03$ ,  $\eta_p^2 = 0.08$ . Specifically, athletes ( $M = 0.4$ ,  $SD = 0.46$ ) were more accurate than non-athletes ( $M = 0.27$ ,  $SD = 0.41$ ). The average accuracy rates for each action pair and group are shown in Fig. 3.

Degrees of freedom varied for the  $2 \times 2$  ANOVAs because due to technical glitches data were not recorded from some actions for four participants. This was because some videos failed to load because of poor internet connection. In the jog-skip and jump-jumping jack analyses we had to drop one participant per analysis; in the sit-squat analysis we had to drop three participants.

### Response time

In order to remove the skewness of the response time distribution, responses that were 3 standard deviations above the mean were removed. This resulted in the removal of 1.6% of trials.

The initial omnibus ANOVA showed a main effect of Action,  $F(5,315) = 5.38$ ,  $p < .001$ ,  $\eta_p^2 = 0.08$ . Responses to jogging actions were the fastest ( $M = 5,946$  ms,  $SD = 2,075$  ms), whereas responses to sitting were the slowest ( $M = 7,527$  ms,  $SD = 2,516$  ms). There was no main effect of Athletic status and no interaction.

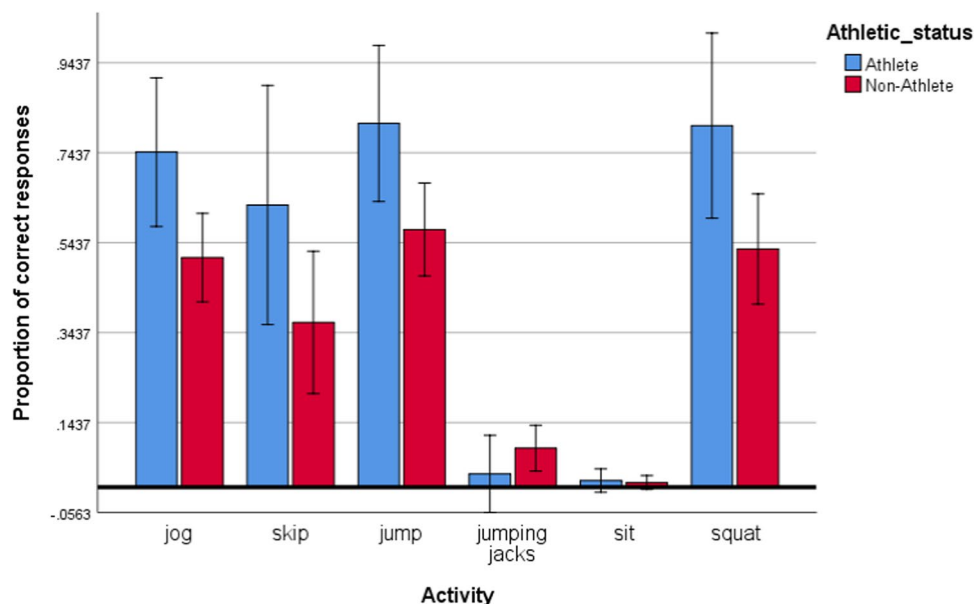
The 2 Athletic Status  $\times$  2 Action pair (jog vs. skip) ANOVA on response time revealed no significant effects. The same ANOVA comparing response times for jumps and jumping jacks revealed no significant effects or interactions. The ANOVA comparing response times of perceiving sitting and

squatting returned a significant difference between actions,  $F(1,64) = 10.65$ ,  $p = .002$ ,  $\eta_p^2 = 0.14$ . Specifically, average response time for squats ( $M = 6,180$  ms,  $SD = 2,121$  ms) was shorter than for sitting down ( $M = 7,507$  ms,  $SD = 2,502$  ms). No other effects were significant. The average response times for each action pair and group are shown in Fig. 4.

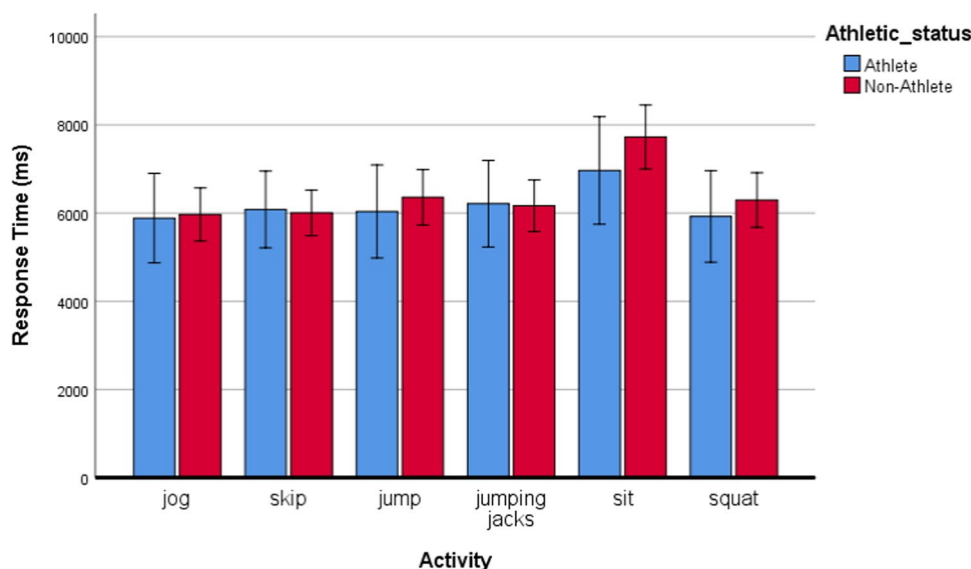
### Discussion

Some actions were more difficult to detect than others. Participants struggled to recognize jumping jacks and sitting. This could be because the natural optic flow patterns for these actions are not as unique as others, and are therefore easily confused with other actions. Another possibility is that jumping jacks and sitting may generate optic flow patterns that are more complex than for other actions, rendering them hard to detect. Sitting might have proven difficult because it is typically not a repetitive movement, however our video sample captured it as such (with the actor sitting down and standing up several times). Jumps were detected more accurately than jumping jacks, perhaps due to the relative simplicity of jumping motions. Athletes were more accurate than non-athletes, consistent with our predictions. This is most likely due to their trained eye and a lot of experience with physical activity with extensive focus and awareness of body movements. It is also possible that in the open-ended response type design, athletes were better equipped to report answers within the constraints of the task than non-athletes because of their familiarity with exercise names and types of movement.

Response times for the sitting activity were the longest of all actions. This is consistent with the difficulty in



**Fig. 3** Accuracy expressed as proportion of correct responses as a function of athletic status and action type in Experiment 1. Error bars represent 95% confidence intervals



**Fig. 4** Response time as a function of athletic status and action type in Experiment 1. Error bars represent 95% confidence intervals

detecting sitting action and shows that perhaps it was not the optimal choice for this task due to it not being a cyclical action. There were no differences in the speed of responding between athletes and non-athletes, contrary to our prediction. This may have been because participants were not prompted in any way to respond as quickly as possible.

We also must consider the limits of the open-ended response method, which was utilized to increase the external validity of the task. At the same time, the open-ended nature of task invited a variety of responses, which decreased experimental control and resulted in low internal validity. The absence of clear differences between groups and activities may have been the result of passive responses (lack of inherent motivation to answer accurately), variations in participants' typing speeds, and uncertainty about the exact labels for the various categories of activities. In some cases, participants were able to report the general movement but did not give a concise enough response to be considered correct (e.g., “up and down,” “moving forward”). In the second experiment we chose to use a forced choice response paradigm to reduce variability due to the open-ended responses. We predicted that the forced choice paradigm would make the task easier and result in less variable responses.

## Experiment 2

The second experiment was conducted to refine and verify the results of Experiment 1. Participants were asked to determine the action presented in the first-person videos by means of a forced-choice task.

## Method

### Participants

Participants for the second experiment were recruited in the same manner as Experiment 1. All participants were new individuals with no pre-existing knowledge of the study. Two groups were recruited and formed via Sona and email: Non-athletes ( $n = 29$ ) and Athletes ( $n = 29$ ). For one non-athlete participant we could not record any responses to jump videos due to technical difficulties, therefore this person's data were not included in the analyses.

### Materials

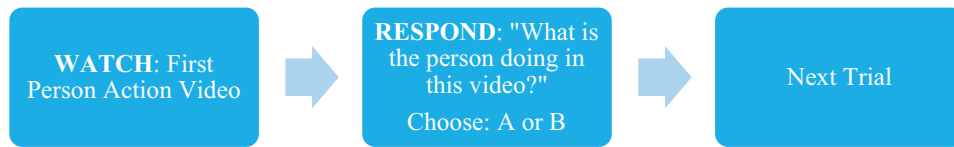
Materials were the same as Experiment 1. We utilized the six first-person action videos as visual stimuli and conducted the experiment online with the Collector software.

### Experimental design

Experiment 2 employed the same experimental design as Experiment 1. The only difference was the manner in which the dependent variable was measured: instead of an open-ended response, a binary forced-choice response mechanism was used.

### Procedure

Using Collector, the same demographic questionnaire as in Experiment 1 was administered. Each target video



**Fig. 5** Example of trial for Experiment 2. Options A and B were presented as the following pairs: jogging or skipping; jumping or jumping jacks; sitting or squatting

was presented randomly four times for a total of 24 trials. After each target video participants answered the question: “Which action is depicted in the video?” The relevant options for each target video appeared as a binary forced-choice answer with two choices: jumping or jumping jacks, squatting or sitting, skipping or jogging. The participant chose which of the two actions they thought was depicted in the target video. Figure 5 depicts an example trial sequence.

It was anticipated that the added context information provided in the multiple-choice format would help guide participants in their decisions and assist them in distinguishing between actions, resulting in overall better accuracy compared to the first experiment. Additionally, athletes were expected to perform more accurately and take less time responding. Participants were allowed to take as much time as they wish to respond.

**Data processing**

Due to technical errors with the internet connection and software 66 trials (4.84%) had to be dropped from the statistical

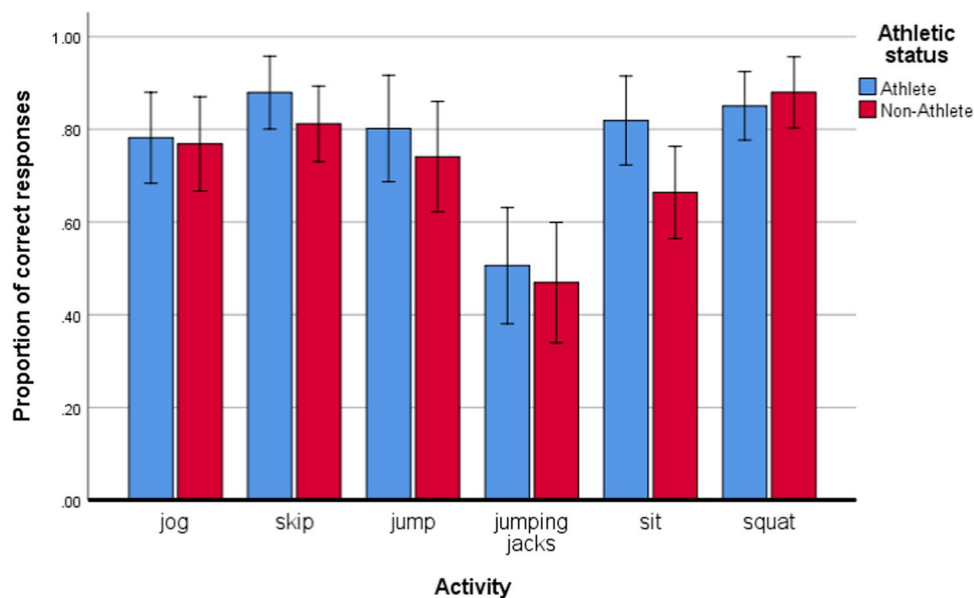
analyses. Another 13 trials (0.95%) were removed because the response time was more than 3 standard deviations above the mean.

**Results**

**Accuracy**

A 2 Athletic status × 6 Action repeated-measures ANOVA on accuracy revealed a main effect of Action,  $F(5,270) = 17.30, p < .001, \eta_p^2 = 0.24$ . Accuracy was highest for squatting ( $M = .86, SD = .2$ ) and lowest for jumping-jacks ( $M = .49, SD = .33$ ). There was no significant effect of athletic status on accuracy. The Athletic status × Action interaction was not significant. The average accuracy rates were reported in Fig. 6.

To get a more detailed picture of the results we followed up the omnibus analysis with separate 2 Athletic status × 2 Action ANOVAs for each action pair: jog versus skip, jump versus jumping jacks, and sit versus squat. For accuracy, we found significant differences in all three pairs: jog-skip,  $F(1,55) = 4.82, p = .032, \eta_p^2 = 0.08$ ,



**Fig. 6** Accuracy expressed as proportion of correct responses as a function of athletic status and action type in Experiment 2. Error bars represent 95% confidence intervals

jump-jumping jacks,  $F(1,54) = 43.01$ ,  $p < .001$ ,  $\eta_p^2 = 0.44$  sit-squat,  $F(1,55) = 12.72$ ,  $p = .001$ ,  $\eta_p^2 = 0.19$ . Specifically, skipping ( $M = .85$ ,  $SD = .21$ ) was more readily detected than jogging ( $M = .76$ ,  $SD = .28$ ), jumping ( $M = .77$ ,  $SD = .31$ ) more so than jumping-jacks ( $M = .49$ ,  $SD = .33$ ), and squatting ( $M = .87$ ,  $SD = .2$ ) more than sitting ( $M = .74$ ,  $SD = .27$ ). There was also a significant interaction between action and athletic-status for sit-squat,  $F(1,55) = 7.41$ ,  $p = .009$ ,  $\eta_p^2 = 0.12$ . Athletes ( $M = .82$ ,  $SD = .22$ ) were significantly more accurate than non-athletes ( $M = .65$ ,  $SD = .3$ ) when perceiving sitting but there was no significant difference when perceiving squatting.

A sensitivity analysis for Experiment 2 suggested that a sample size of  $N = 57$  across two groups (which is what we had) would be sensitive to effects of  $\eta_p^2 = 0.1$  (medium effect size) with 80% power (alpha = 0.05) for a between-subjects effect of Athletic status. This means the study would not be able to reliably detect effects smaller than  $\eta_p^2 = 0.1$ . Our actual obtained  $\eta_p^2$  was 0.03, which was much smaller. This means we had enough power to detect a medium-sized effect, but did not; therefore, we conclude that the non-significance was not due to lack of statistical power.

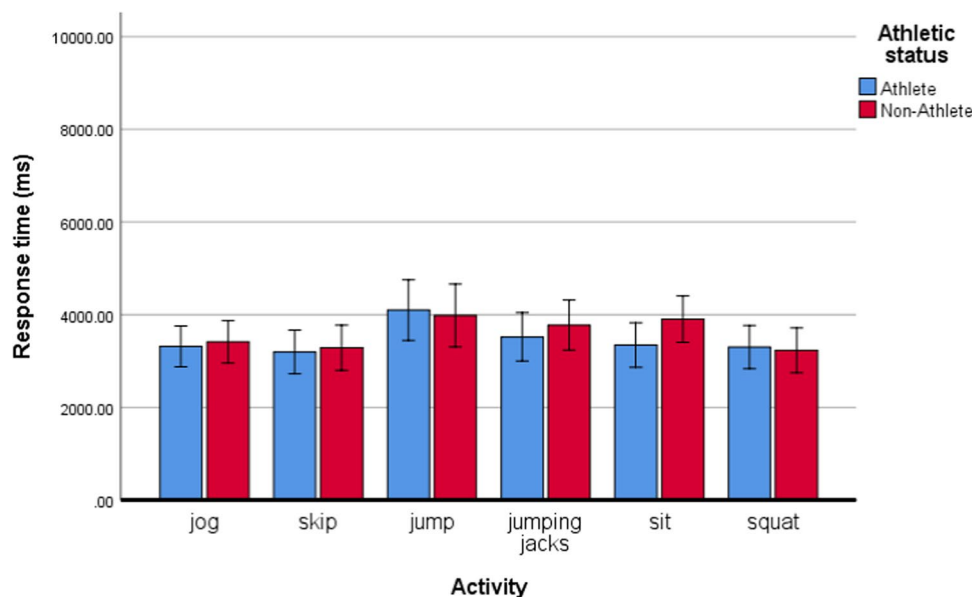
### Response time

There was a main effect of action on response time,  $F(5,270) = 3.52$ ,  $p = .004$ ,  $\eta_p^2 = 0.06$ . Overall, response times were longest for jumping ( $M = 4,043$  ms,  $SD = 1,736$  ms). There were no significant results for athletic status. The 2 Athletic status  $\times$  2 Action ANOVAs for each action pair revealed no

significant findings. The average response times are shown in Fig. 7.

### Discussion

In the first round of analyses, we discovered a main effect of action such that squatting was most accurately detected and jumping-jacks the least. This may speak to the nature of each of these activities and how the optic patterns associated with them differ in complexity. Although we have not performed any type of video differencing or computed any complexity measures on these videos, observation by the naked eye suggests jumping-jacks has a very complex and compact pattern associated with the movement, whereas squatting has a smoother and less complex pattern. Most likely this has something to do with the fact that while performing jumping-jacks the person must leave the ground and then land repeatedly, whereas during squats the person is stable on the ground during the entire motion. Nevertheless, direct comparison between squats and jumping jacks was not sought in the present study due to the fact that these two activities are not in the same category of actions and therefore trivially distinguishable. The second round of analyses (focusing on pairs of actions separately) showed distinguishable differences between the actions in all pairs, but it may be the case that some are more difficult to tell apart than others. Skipping, jumping, and squatting were recognized more than jogging, jumping-jacks, and sitting. It could be that skipping, jumping, and squatting have very distinct optic patterns that are more familiar and easier to detect than the other three.



**Fig. 7** Response time as a function of athletic status and action type in Experiment 2. Error bars represent 95% confidence intervals



The absence of general differences between athletes and non-athletes suggests that the added context in this task (i.e., the provided choice responses) made the task easier and “levels the playing field” so that athletic experience does not provide much of an advantage. Athletes outperformed non-athletes when judging sitting. Since sitting may probably be one of the most difficult optic flow patterns to recognize, athletes’ ability to do so better supports the notion that they have an upper hand in action detection. The lack of notable findings for response time might be attributed to the fact that there was no encouragement to perform quickly and that the relatively easy forced choice paradigm compared to the task in Experiment 1 made all groups respond at a similar rate.

### Comparison of Experiments 1 and 2: Do athletes perform better when the task is hard?

Since Experiment 2 was designed as a follow-up to the first experiment we decided to compare them in a 2 Experiment  $\times$  2 Athletic Status  $\times$  6 Action mixed ANOVA on accuracy and response time, respectively.

#### Accuracy

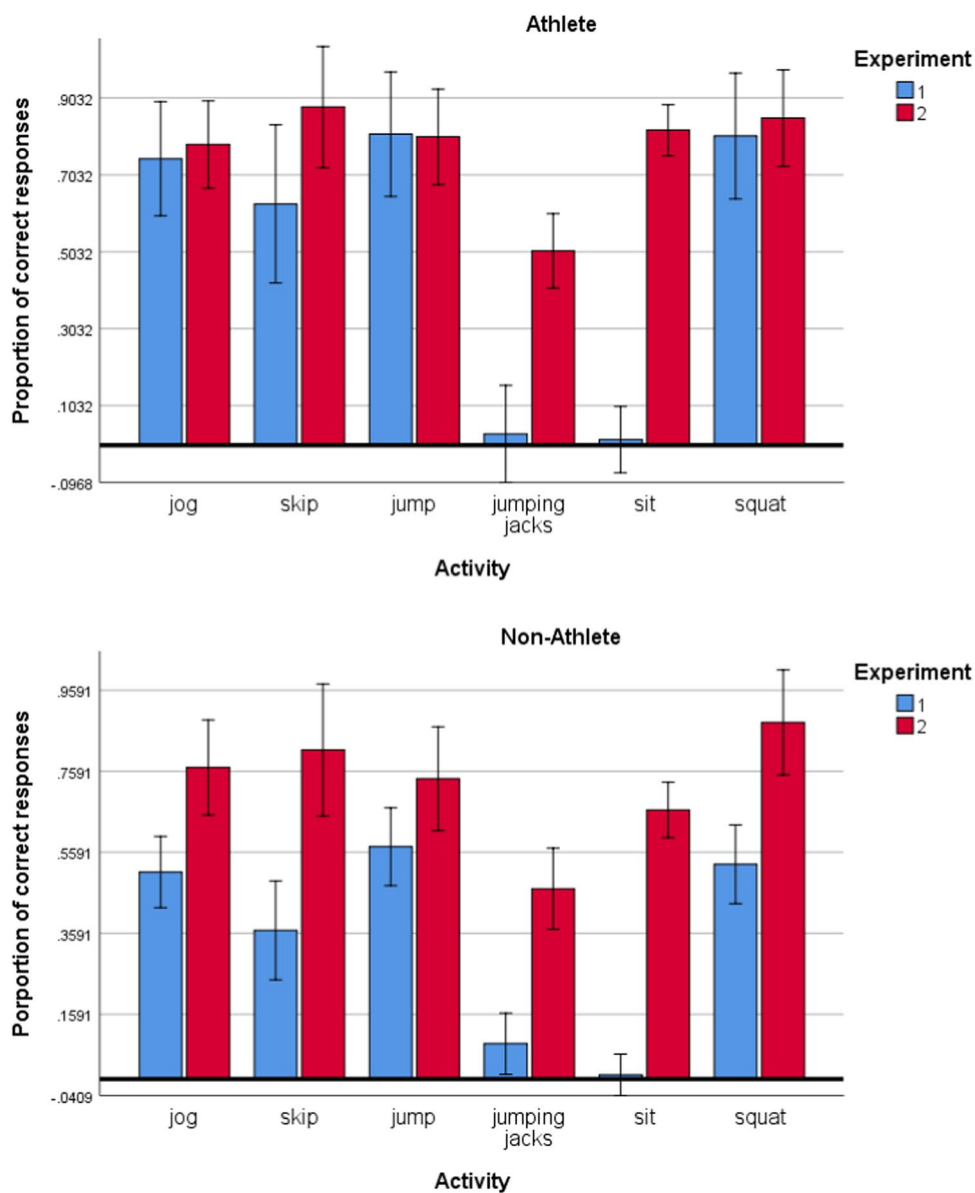
Evidence that the task (forced-choice) in Experiment 2 ( $M = .74$ ,  $SD = .3$ ) was easier than Experiment 1 ( $M = .38$ ,  $SD = .44$ ) was found in a main effect,  $F(1,117) = 109.26$ ,  $p < .001$ ,  $\eta_p^2 = 0.48$ , such that accuracy rates were much higher for the second experiment. A main effect of Action,  $F(5,585) = 50.01$ ,  $p < .001$ ,  $\eta_p^2 = 0.30$ , revealed squatting ( $M = .72$ ,  $SD = .37$ ) to be the most accurately detected action across experiments and jumping-jacks ( $M = .26$ ,  $SD = .33$ ) to be the least accurately detected action. An Action  $\times$  Experiment interaction,  $F(5,585) = 16.33$ ,  $p < .001$ ,  $\eta_p^2 = 0.12$ , revealed that the largest improvement of accuracy from Experiment 1 to Experiment 2 was in the actions most difficult to detect: jumping-jacks (Exp. 1:  $M = .07$ ,  $SD = .18$ , Exp. 2:  $M = .49$ ,  $SD = .33$ ) and sitting (Exp. 1:  $M = .01$ ,  $SD = .05$ , Exp. 2:  $M = .74$ ,  $SD = .27$ ). A three-way interaction between Action  $\times$  Experiment  $\times$  Athletic Status,  $F(5,585) = 2.47$ ,  $p = .03$ ,  $\eta_p^2 = 0.02$ , shows, overall, Athletes ( $M = .48$ ,  $SD = .44$ ) were better than Non-athletes ( $M = .35$ ,  $SD = .44$ ) in Experiment 1; however, in Experiment 2, Athletes were only better at detecting the difficult action of sitting (Athletes:  $M = .82$ ,  $SD = .22$ , Non-Athletes:  $M = .66$ ,  $SD = .30$ ). Lastly, a main effect of Athletic Status,  $F(1,117) = 11.57$ ,  $p = .001$ ,  $\eta_p^2 = 0.09$ , suggests that athletes perform better than non-athletes across both experiments but the three-way interaction mentioned above gives us a clearer picture that Athletes’ better performance is carried in Experiment 1 (except for sitting in Experiment 2). Visualization of the accuracy data across Experiments 1 and 2 is displayed in Fig. 8.

#### Response time

Response time data for both experiments delivered a significant main effect of Action,  $F(5,585) = 5.85$ ,  $p < .001$ ,  $\eta_p^2 = 0.48$ , where sitting took the most time ( $M = 5,706$  ms,  $SD = 2,800$  ms) and skipping took the least ( $M = 4,765$  ms,  $SD = 2,076$  ms). An Action  $\times$  Experiment interaction,  $F(5,585) = 3.58$ ,  $p = .003$ ,  $\eta_p^2 = 0.03$ , was found because of a drastic change in response times for the difficult action of sitting across experiments. The forced choice versus open-ended paradigm created a stark difference for this action between Experiments 1 ( $M = 7,482$  ms,  $SD = 2,491$  ms) and 2 ( $M = 3,617$  ms,  $SD = 1,298$  ms). Lastly, it became clear via a main effect of Experiment,  $F(1,117) = 125.45$ ,  $p < .001$ ,  $\eta_p^2 = 0.52$ , that trials for Experiment 1 ( $M = 6,334$  ms,  $SD = 2,164$  ms) took participants longer to complete than the trials in Experiment 2 ( $M = 3,535$  ms,  $SD = 1,375$  ms). Visualization of the response time data across Experiments 1 and 2 is displayed in Fig. 9.

### Experiment 3

The third experiment was conducted as an extension of Experiments 1 and 2. We sought to discover if the information provided from a sample of optic flow is invariant across viewpoints, and to find out if this invariance is equally easy to detect for all actions. How easy or difficult is it to recognize the same action based on a first-person perspective viewpoint as compared to a third-person perspective? Could actions of an actor performing the motion from a third person viewpoint be matched to the first-person video of the same action? We used a nonverbal matching to sample method (Barth et al., 1995) in the present experiment. The target stimuli were presented one at a time as either third-person or first-person videos, and the two available options for response were the relevant paired action videos recorded from the opposite perspective of the target. We hypothesized that participants would successfully match videos from different viewpoints, demonstrating the invariance of the information pattern across viewpoints. Second, participants would be more accurate when introduced to a third-person view of the action than when working with only first-person videos. Third, athletes will outperform non-athletes. Fourth, different action pairs will result in different levels of recognition accuracy. The experimental design including both first- and third-person perspectives as target videos allowed us to examine the relationship between information type and to determine which is more readily utilized to perceive actions.



**Fig. 8** Proportion of correct responses as a function of experiment, action type, and athletic status in Experiments 1 and 2. Error bars represent 95% confidence intervals

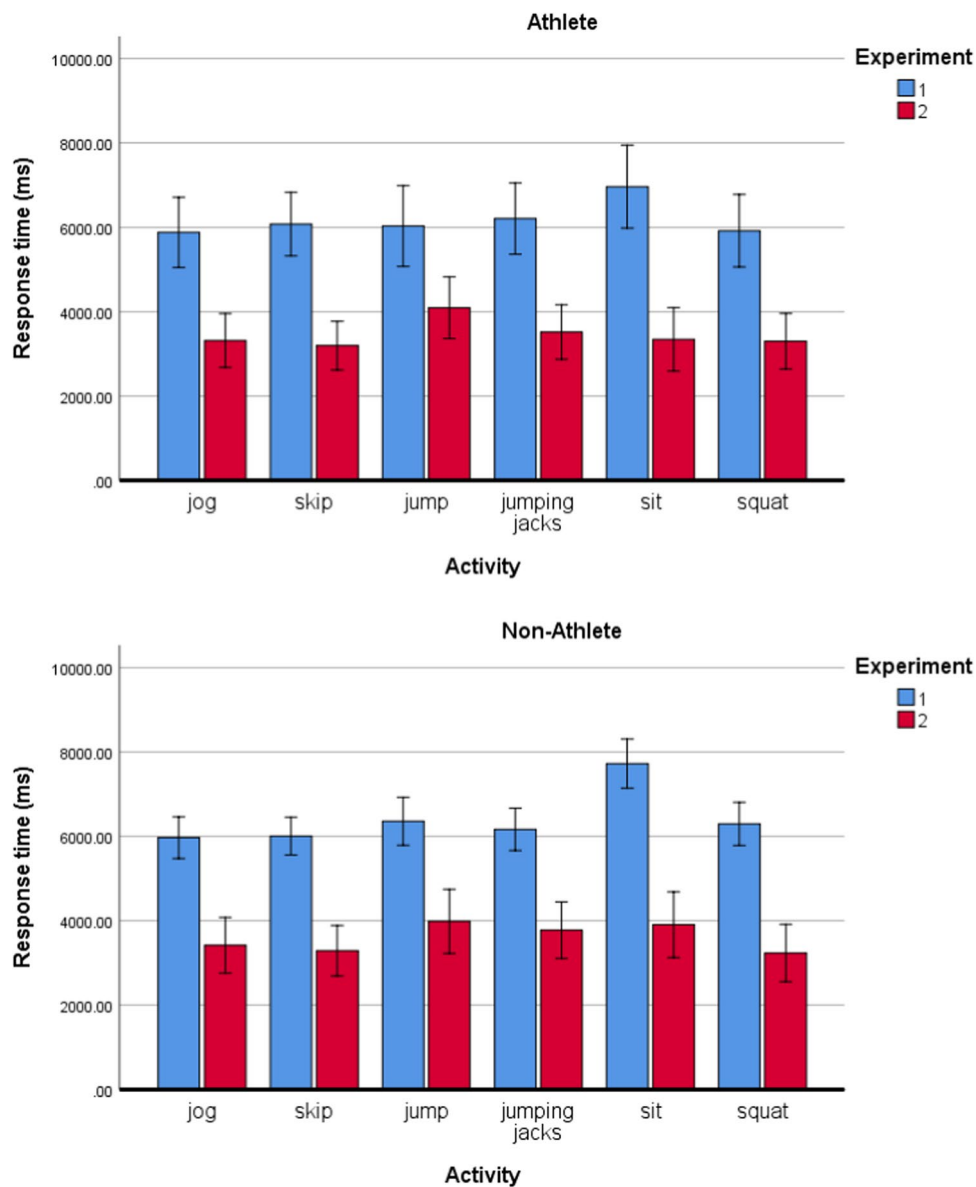
## Methods

### Participants

Participants for this experiment were recruited in the same manner as in Experiments 1 and 2. There was an athlete group ( $n = 30$ ) and a non-athlete group ( $n = 35$ ). Individuals who participated in Experiment 1 or Experiment 2 were not allowed to partake in the present study.

### Materials

For this experiment we used the set of first-person action videos from the previous experiments as well as videos of the same actor performing the action from a third-person view. Third-person videos were filmed in the same manner as the previous first-person videos using a Go Pro sports camera. The videos were filmed on the same day using the same actor (the author of the paper) and the actions were



**Fig. 9** Response time as a function of experiment, action type, and athletic status in Experiments 1 and 2. Error bars represent 95% confidence intervals

made to be as similar and consistent with the first-person videos as possible. The videos were also trimmed to be five seconds long.

**Experimental design**

We implemented a 2 Athletic Status (Athlete, Non-athlete) × 3 Action (Jog, Skip, Sit) × 2 Perspective (First-person, Third-person) mixed design for this experiment. For the purpose of this experiment, we did not assess each action individually but instead evaluated each action pair with the emphasis being on the difference in perspective (first or third). Each participant underwent the same experimental

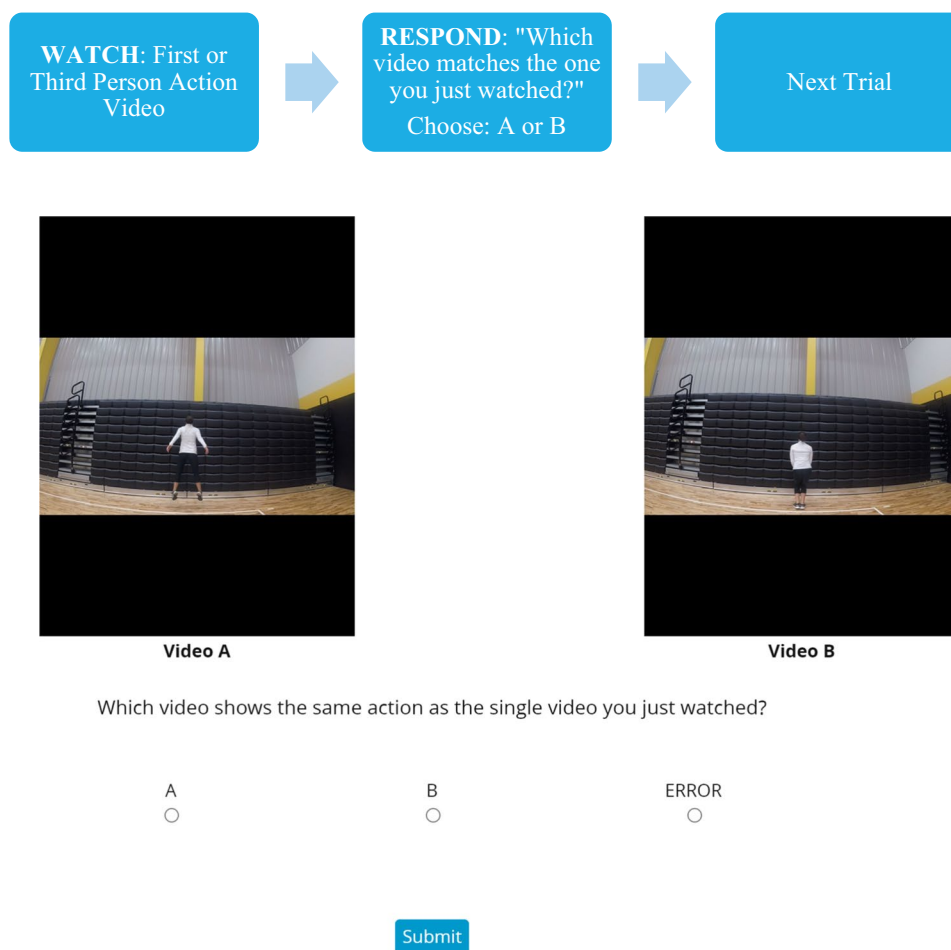
procedures. The video trials were presented in randomized order. Both action and perspective order were randomized.

**Procedure**

Participants accessed the online link by using their personal computer or some type of laptop/ desktop. The first part of the experiment was the demographic questionnaire and then experimental trials began (Fig. 10).

Experimental trials consisted of either a first- or third-person target video followed by two side-by-side videos in the opposite perspective. For example:

Target video: First-person jump



**Fig. 10** Example of a trial sequence in Experiment 3. Options A and B were shown as side-by-side videos (see bottom panel for a screenshot of the response options as presented to participants) depicting

both actions of the relevant action pair (e.g., jog | skip) in the opposite perspective as the previously shown target video (e.g., target = first-person, side-by-side response videos = third-person).

Sample A: Third-person jump | Sample B: Third-person jumping-jacks

This presentation constituted a trial. Each of the six actions (in both perspectives) served as a target stimulus and was repeated four times for a total of 48 trials. The side-by-side sample videos were randomly displayed on the left or right to reduce response bias. All six actions were grouped into their relevant action pairs. The pairs were as follows: (1) jogging and skipping, (2) jumping and jumping-jacks, (3) sitting and squatting. Responses were again assessed for accuracy and response time.

#### Data processing

Due to technical errors with the internet connection and software 87 trials (3.7%) had to be dropped from the statistical analyses. Another 37 outlier trials (1.6%) were removed based on the response time criteria of three standard deviations above the mean.

## Results

### Accuracy

A 2 Athletic Status  $\times$  2 Perspective  $\times$  3 Action Pair ANOVA showed a main effect of Action pair,  $F(2,126) = 7.7, p = .001, \eta_p^2 = 0.11$ , where the jog-skip pair ( $M = .73, SD = .26$ ) was detected significantly less accurately than the jump-jumping-jack pair ( $M = .80, SD = .2$ ), and the sit-squat pair ( $M = .81, SD = .22$ ). This effect was qualified by the interaction of Action pair and Perspective,  $F(2,126) = 8.32, p = .001, \eta_p^2 = .12$ , which revealed that accuracy was consistent across actions when the target video was in the first-person perspective (Jog:  $M = .80, SD = .26$ , Jump:  $M = .79, SD = .2$ , Sit:  $M = .80, SD = .21$ ); however, when the target video was in the third-person perspective, Jog ( $M = .66, SD = .25$ ) was detected significantly less than Jump ( $M = .8, SD = .19$ ) and Sit ( $M = .83, SD = .23$ ). A main effect of Athletic status approached significance,  $F(1,63) = 3.48, p = .067, \eta_p^2 = .06$ , so that Athletes ( $M = .82,$

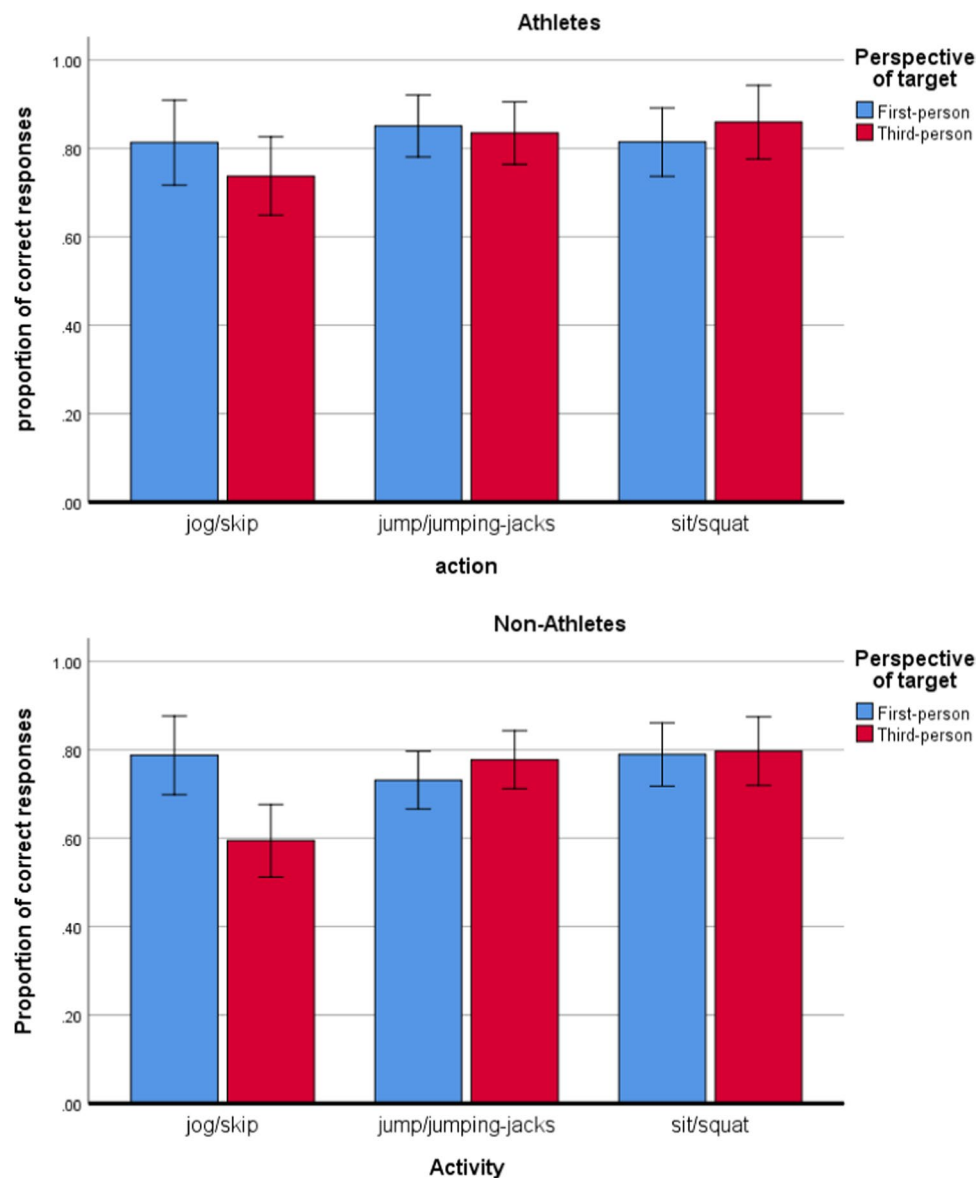
$SD = .2$ ) were more accurate than Non-athletes ( $M = .75$ ,  $SD = .25$ ), but this finding was not significant.

We followed up the omnibus analysis with 2 Athletic Status  $\times$  2 Perspective ANOVAs for each action pair to get a more precise look at the role of perspective and athletic status. There was a main effect of Perspective for the jog-skip pair,  $F(1,63) = 12.5$ ,  $p = .001$ ,  $\eta_p^2 = .17$ . Specifically, the third-person target video ( $M = .66$ ,  $SD = .25$ ) results in less accurate perception than the first person target video ( $M = .8$ ,  $SD = .26$ ). The only difference we see between Athletes and Non-Athletes is for the jump-jumping jacks pair,  $F(1,63) = 5.6$ ,  $p = .02$ ,  $\eta_p^2 = .08$ . Athletes ( $M = .84$ ,  $SD = .18$ ) were more accurate than non-athletes ( $M = .75$ ,  $SD = .2$ ). Mean accuracy rates are displayed in Fig. 11.

A sensitivity analysis for Experiment 3 suggested that a sample size of  $n = 65$  across two groups would be sensitive to effects of  $\eta_p^2 = 0.09$  (medium effect size) with 80% power ( $\alpha = 0.05$ ) for a between-subjects effect of Athletic status. This means the study would not be able to reliably detect effects smaller than  $\eta_p^2 = 0.09$ . Our actual obtained  $\eta_p^2$  was 0.05, which was smaller. This means we had enough power to detect a medium sized effect, but did not, therefore we conclude that the non-significance was not due to lack of statistical power.

### Response time

A 2 Athletic Status  $\times$  2 Perspective  $\times$  3 Action Pair ANOVA on response time displayed a main effect of Perspective,

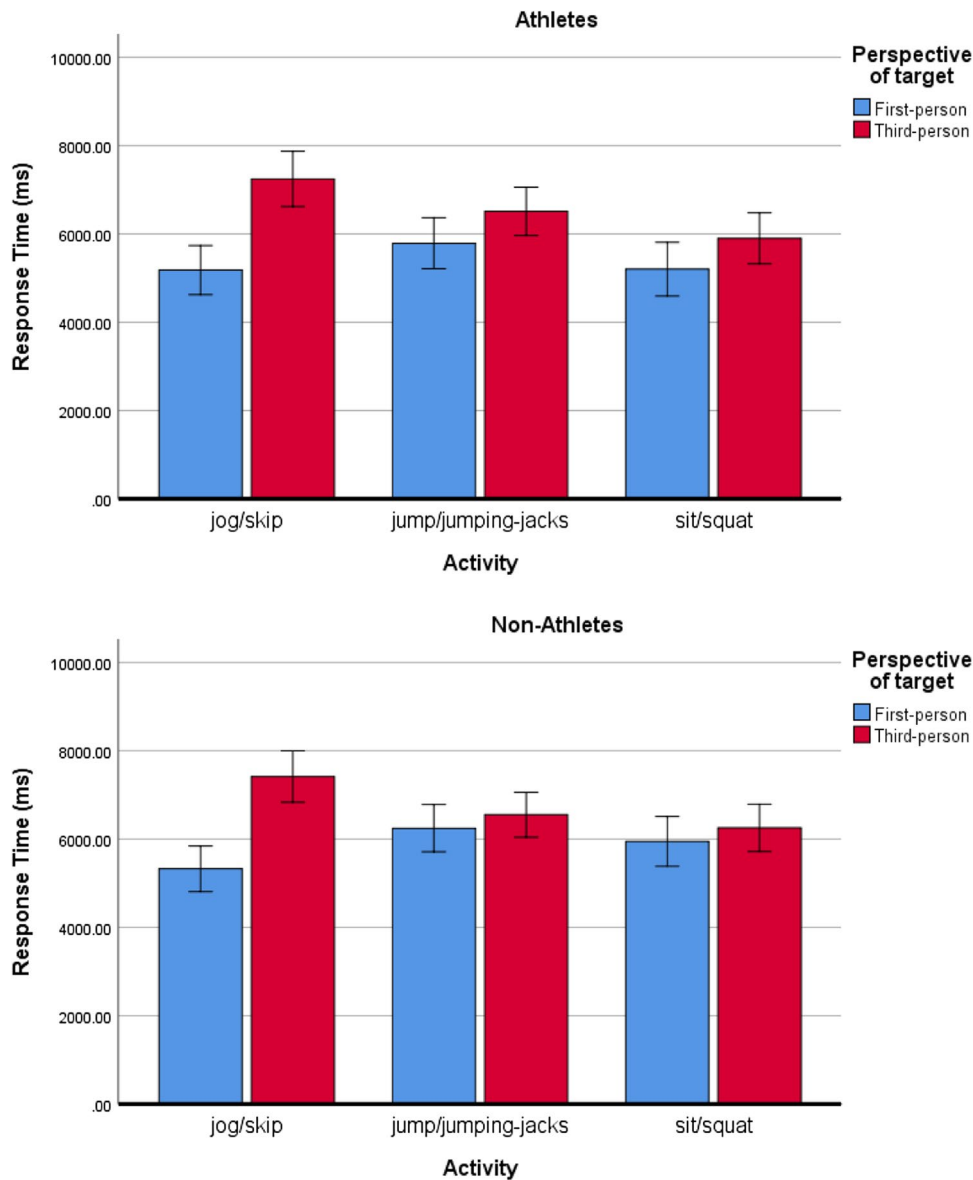


**Fig. 11** Proportion of correct responses as a function of perspective, athletic status, and action pair in Experiment 3. Error bars represent 95% confidence intervals

$F(1,63) = 65.5, p < .001, \eta_p^2 = .51$ , such that participants were significantly faster in responding when the target video was in the first-person ( $M = 5,632$  ms,  $SD = 1,623$  ms) versus the third ( $M = 6,653$  ms,  $SD = 1,668$  ms). The main effect of Action was also significant,  $F(2,126) = 7.41, p = .001, \eta_p^2 = .11$ , where the jog-skip pair ( $M = 6,298$  ms,  $SD = 1,917$  ms) resulted in the longest responses, followed by the jump-jumping-jack pair ( $M = 6,283$  ms,  $SD = 1,553$  ms), and the sit-squat pair ( $M = 5,846$  ms,  $SD = 1,650$  ms). This was qualified by the significant Perspective  $\times$  Action interaction,  $F(2,126) = 17.5, p < .001, \eta_p^2 = .22$ , revealing that the largest difference in response times between perspectives occurred for the jog-skip action pair. No other findings were

significant. The mean response times by perspective and athletic status are presented for each action pair in Fig. 12.

Again, we followed up with 2 Athletic Status  $\times$  2 Perspective ANOVAs. When assessing the jog-skip action pair we found a main effect of Perspective,  $F(1,63) = 78.1, p < .001, \eta_p^2 = .55$ , so that when the target video was presented in the first-person ( $M = 5,259$  ms,  $SD = 1,515$  ms) participants responded faster than in the third ( $M = 7,336$  ms,  $SD = 1,708$  ms). Similarly, we found a significant main effect for Perspective for the jump-jumping jack pair,  $F(1,63) = 5.87, p = .02, \eta_p^2 = .09$ , so that when the target video was presented in the first-person ( $M = 6,034$  ms,  $SD = 1,588$  ms) participants responded faster than in the third ( $M = 6,533$



**Fig. 12** Response times as a function of perspective, athletic status, and action pair in Experiment 3. Error bars represent 95% confidence intervals

ms,  $SD = 1,487$  ms). The main effect of Perspective was also significant for the sit-squat pair,  $F(1,63) = 6.03$ ,  $p = .02$ ,  $\eta_p^2 = .09$ , so that when the target video was presented in the first-person ( $M = 5,602$  ms,  $SD = 1,692$  ms) participants responded faster than in the third ( $M = 6,091$  ms,  $SD = 1,581$  ms). There were no differences between Athletes and Non-Athletes.

## Discussion

In Experiment 3 we tested the hypothesis that visual perception of human activity should be based on invariant information. In particular, the information that specifies a given activity is hypothesized to produce optic flow patterns that are equivalent across viewpoints. To test this, we compared an egocentric (first-person) view to an allocentric (third-person) view of the same activity using video footage of an actor performing various actions. The results showed that participants performed faster and more accurately when the action had to be matched from a first-person video stimulus. The sit/squat action pair was most accurately detected and fastest across conditions, suggesting that the invariance is strongest for this pair. Participants are most efficient in recognizing these actions, probably because the optic flow pattern for standing up and sitting down repeatedly is very unique and different from the optic flow pattern of squatting repeatedly. Sitting is marked by more of a forward and backward movement, with a clear break at the moment when the motion reverses from upward to downward. This is most likely very easy to notice and is at the core of the invariant information across viewing perspectives. On the other hand, all the other actions are cyclical by nature. We do not typically sit down and stand up repeatedly as it was depicted in the video, but it is common to squat repeatedly during exercise or to jump up and down several times in a row. Perhaps cyclical movements' invariant patterns are harder to notice? Future studies will be necessary to investigate this possibility.

What could be the explanation for why participants perform consistently faster as well as more accurately when the target video is in the first-person? The first-person view is “pure optic flow” without any other irrelevant (potentially distracting) information. If this type of information is seen beforehand, it is easy to match it later, because the information being remembered did not co-occur with other, irrelevant, and thus *not* invariant features at the outset. The reverse is not true: if the third-person view is seen first, it contains a lot of distracting info that is not invariant (e.g., details of the actor's body contour or clothes that are not visible from an egocentric viewpoint). This may lead to a lot of unnecessary retention of information that is irrelevant to the task, thus making it more difficult to match it to the same action that is presented from a different viewpoint. It

could be that the longer response times indicate an effect of memory. If there is less information to memorize at the outset (i.e., only the essential invariant without distractors) then it should be easier to match when tested later. However, if the first stimulus is more detailed, then irrelevant features may interfere with matching. This is an example of when less, or simpler information is better for perceiving.

Unexpectedly, Athletes were only marginally better at this task, performing at an 82% average accuracy versus 75% for Non-athletes with large variability ( $\pm 20\%$  standard deviation). This suggests that perceiving invariants is not a function of learned expertise with body movements or consistent physical training, which we assume athletes possess, but rather is an inherent capability of all individuals. A notable exception to this was found in the case of jumps and jumping-jacks where athletic experience seemed to present an advantage. Athletes were better at detecting invariants for these actions (84%) compared to Non-athletes (75%). This may be rooted in the trivial fact that these stereotypical exercises are part of athletic training. Overall, the matching of invariants has been performed above chance (50%) level for all actions for both athletes and non-athletes, shoring up evidence that humans perceive actions by detecting invariants across various viewpoints and in many different circumstances and contexts.

## General discussion

Optic flow underwrites the visual capacity to properly navigate in the environment. The optical information manifested during body movements gives context for object location, observer location, observer capabilities within the environment and for the detection of potential possibilities for action. The current project demonstrated that through practice and experience humans possess the ability to extend action detection to optic flow patterns that are not their own, specifically, not observed from their own point of view.

Experiments 1 and 2 showed that there are invariant patterns of optic flow that can be distinguished for certain actions. This was true even when observing the consequences of motion from a first-person perspective in which the body is not visible, thus offering a strong test of the ability to detect optic flow and recognize biological motion patterns. In Experiment 3 we discovered that people are sensitive to optic flow patterns irrespective of point of view. Athletic experience provided added benefit when the task was hard (Experiment 1).

We tested sample groups from two separate populations to investigate the potential advantage that might come with consistent physical training and a more learned eye for action possibilities based on experience in sports. We

found athletic status mattered less and less as we progressed from Experiment 1 (open-ended response to target video) to Experiment 2 (multiple-choice response to target video) and Experiment 3 (matching a target video to sample videos). Experiment 1 forced participants to rely on their own knowledge of the names and descriptions for the actions involved, thus posing a language issue. This could have given athletes an advantage because of their familiarity with the presented actions. Due to this issue it was not clear if we were testing visual perception or language. Experiment 2 provided language labels (words) as response options and helped bring us closer to investigating visual perception. However, participants were still matching perceptual information (video) to abstract concepts (words). Experiment 3 employed a non-verbal response mechanism, thus providing the best test of perception.

We discovered that certain actions are more difficult to recognize than others. Interestingly, the detection of sitting was the worst in Experiment 1, but in Experiment 3 the sit/squat action pair resulted in the best performance. All actions were detected above chance level, but only once measured using the appropriate method (nonverbal match to sample in Experiment 3). Without any context the task to recognize a motion like sitting is extremely difficult. Recognition becomes easier when the task is to distinguish between two very different types of movements for which visual patterns of invariant motion are quite distinct.

There were some limitations to this study. First, online data collection lacked sufficient experimental control and may have resulted in increased variability. An additional difficulty was that we were operating under Covid-19 restrictions at the time of data collection. Secondly, we only considered six action types and all actions were presented in cyclical fashion. This was true even for actions that are not typically cyclical in nature. For example, sitting was videotaped as a sequence of several bouts of sitting down and standing up. It could be that this made the task more difficult because sitting is not typically repeated in sequence. We also did not directly test which action pairs were more similar in nature. It is possible that skipping and jogging are more similar movements than sitting and squatting. We also did not cross-pair any of the actions outside of the originally planned pairings. For instance, it would be interesting to see how jumping and squatting are perceived via direct comparison, however this pairing was not tested.

Additional data processing is planned for future studies to compare the patterns of motion using video analysis. Pixel-by-pixel calculation of mutual information (cross-correlation) would provide a more precise understanding of similarity and disparity for each action and a better understanding for comparing invariance across perspectives. Furthermore, video analysis could be conducted to convert the videos into optic flow footage by only showing those pixels in each

frame that changed from one moment to the next. Once these optic flow patterns are generated a direct comparison could reveal the similarities among all the actions that were tested. Relatedly, future studies should investigate how much visual stimulation is necessary and sufficient to detect the invariant optical pattern. This could be done by visualizing the optic flow pattern with a decreasing number of pixels until we see a breakdown in the ability to detect the action.

Examination of kinematic information and biological motion has a long tradition in behavioral science. Numerous studies of the perception of point-light displays (PLDs) have demonstrated humans' abilities to recognize and make sense of motion even when there is very little information present (Johansson, 1973). Gender, identity, specific details of an action (e.g., lifting something heavy vs. something light), intentions for an action, and whether a person is deceptive are a few perceived events within the context of biomotion stimuli (Runeson & Frykholm, 1983). Even when there are distractions incorporated such as misplaced points of light or added points of light, specific motions can be detected (Neri et al., 1998). There has been very little research about perception of biological motion from a first-person point of view. In one notable study it was demonstrated that both embodiment of a virtual anatomy as well as kinematic illusions are possible after interaction with a PLD-created limb in virtual reality (Giroux et al., 2019). Our study has potential to contribute to this research as it taps into the detection of movements and the understanding of invariants of biological motion perception across viewpoints. Specifically, advances in computer vision could provide a useful application of our current research findings for the purposes of building artificial agents (robots) that are capable of detecting actions based on limited and noisy visual information.

The present study could also set the groundwork for understanding the role of sports training on visual perception. We compared athletes to non-athletes to explore the possibility that athletic training, for the purposes of competition in sports, enhances the ability to detect actions. This prediction is consistent with the fact that athletes spend significantly more time (than non-athletes) focusing on their own physical movements as well as picking up those of others in order to perform better and win games. The general advantage that athletes demonstrated in the present study suggests that training in these areas could potentially lead to better visual perception and more efficient interaction with the environment. However, the lack of differences as the task becomes easier suggests that perception of movement is more inherent than trained even when the actions are specific to athletic activities (e.g., exercise routines).

**Data availability** The datafiles and video stimuli for every experiment are available publicly at <https://osf.io/97ut5/>



## References

- Barth, C., Fein, D., & Waterhouse, L. (1995). Delayed match-to-sample performance in autistic children. *Developmental Neuropsychology*, *11*(1), 53–69.
- Danafar, S., Gheissari, N. (2007). Action recognition for surveillance applications using optic flow and SVM. In: Y. Yagi, S. B. Kang, I. S. Kweon, H. Zha (Eds.), *Computer Vision – ACCV 2007. ACCV 2007. Lecture Notes in Computer Science* (Vol. 4844). Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-540-76390-1\\_45](https://doi.org/10.1007/978-3-540-76390-1_45)
- Fajen, B. R., Riley, M. A., & Turvey, M. T. (2009). Information, affordances, and the control of action in sport. *International Journal of Sport Psychology*, *40*(1), 79–107.
- Fan, R., Jiao, J., Ye, H., Yu, Y., Pitas, I., & Liu, M. (2019). Key ingredients of self-driving cars. *arXiv preprint arXiv:1906.02939*. <https://doi.org/10.48550/arXiv.1906.02939>
- Garcia, M. A., Kerr, T. K., Blake, A. B., & Haffey, A. T. (2015). Collector (Version 2.0.0-alpha) [Software]. Available from <https://github.com/gikeymarcia/Collector/releases>. Accessed 15 May 2020.
- Gibson, J. (1947). *Motion picture testing and research*. AD 651783. Washington DC: US Government Printing Office.
- Giroux, M., Barra, J., Barraud, P. A., Graff, C., & Guerraz, M. (2019). From Embodiment of a Point-Light Display in Virtual Reality to Perception of One's Own Movements. *Neuroscience*, *416*, 30–40.
- Hacques, G., Komar, J., Dicks, M., & Seifert, L. (2021). Exploring to learn and learning to explore. *Psychological Research*, *85*(4), 1367–1379.
- Holte, M. B., Moeslund, T. B., & Fihl, P. (2010). View-invariant gesture recognition using 3D optical flow and harmonic motion context. *Computer Vision and Image Understanding*, *114*(12), 1353–1361.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201–211.
- Klostermann, A., Vater, C., Kredel, R., & Hossner, E. J. (2015). Perceptual training in beach volleyball defence: different effects of gaze-path cueing on gaze and decision-making. *Frontiers in Psychology*, *6*, 1834.
- Laming, E. (2019). Police use of body worn cameras. *Police Practice and Research*, *20*(2), 201–216.
- Lum, C., Koper, C. S., Wilson, D. B., Stoltz, M., Goodier, M., Eggins, E., ... & Mazerolle, L. (2020). Body-worn cameras' effects on police officers and citizen behavior: A systematic review. *Campbell Systematic Reviews*, *16*(3), e1112.
- Matthis, J. S., Muller, K. S., Bonnen, K. L., & Hayhoe, M. M. (2022). Retinal optic flow during natural locomotion. *PLoS Computational Biology*, *18*(2), e1009575.
- Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, *395*(6705), 894–896.
- Noceti, N., Odone, F., Rea, F., Sciutti, A., & Sandini, G. (2019). View-invariant robot adaptation to human action timing. In K. Arai, S. Kapoor, & R. Bhatia (Eds.), *Intelligent Systems and Applications* (IntelliSys 2018. Advances in Intelligent Systems and Computing) (Vol. 868). Springer.
- Ruffier, F., & Franceschini, N. (2005). Optic flow regulation: The key to aircraft automatic guidance. *Robotics and Autonomous Systems*, *50*(4), 177–194.
- Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(4), 733–740.
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*(4), 585–615.
- Steel, K., Ellem, E., & Baxter, D. (2015). The application of biological motion research: biometrics, sport, and the military. *Psychonomic Bulletin & Review*, *22*(1), 78–87.
- Vickers, J. N. (2007). *Perception, cognition, and decision training: The quiet eye in action*. Human Kinetics.
- Vignolo, A., Noceti, N., Rea, F., Sciutti, A., Odone, F., & Sandini, G. (2017). Detecting biological motion for human–Robot interaction: A link between perception and action. *Frontiers in Robotics and AI*, *4*, 14.
- Warren, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience*, *4*(2), 213–216.
- Weast, J. A., Shockley, K., & Riley, M. A. (2011). The influence of athletic experience and kinematic information on skill-relevant affordance perception. *Quarterly Journal of Experimental Psychology*, *64*(4), 689–706.
- Weast, J. A., Walton, A., Chandler, B. C., Shockley, K., & Riley, M. A. (2014). Essential kinematic information, athletic experience, and affordance perception for others. *Psychonomic Bulletin & Review*, *21*(3), 823–829.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.