

# Visual search by action category

Jeroen J. A. van Boxtel

Psychology Department, University of California,  
Los Angeles (UCLA), USA



Hongjing Lu

Psychology Department, University of California,  
Los Angeles (UCLA), USA



Humans are sensitive to different categories of actions due to their importance in social interactions. However, biological motion research has been heavily tilted toward the use of walking figures. Employing point-light animations (PLAs) derived from motion capture data, we investigated how different activities (boxing, dancing, running, and walking) related to each other during action perception, using a visual search task. We found that differentiating between actions requires attention in general. However, a search asymmetry was revealed between boxers and walkers, i.e., searching for a boxer among walkers is more efficient than searching for a walker among boxers, suggesting the existence of a critical feature for categorizing these two actions. The similarities among the various actions were derived from hierarchical clustering of search slopes. Walking and running proved to be most related, followed by dancing and then boxing. Signal detection theory was used to conduct a non-parametric ROC analysis, revealing that human performance in visual search is not fully explained by low-level motion information.

Keywords: biological motion, activities, point-light animations, motion capture, action perception, visual search, attention

Citation: van Boxtel, J. J. A., & Lu, H. (2011). Visual search by action category. *Journal of Vision*, 11(7):19, 1–14, <http://www.journalofvision.org/content/11/7/19>, doi:10.1167/11.7.19.

## Introduction

Ever since Johansson (1973) developed point-light displays, numerous studies have reported that humans from a very young age exhibit an exquisite sensitivity to biological motion (see, e.g., Blake & Shiffrar, 2007). The exceptional ability of the human visual system in perceiving actions is revealed by the three major findings. (1) Good recognition performance is obtained for sparse visual stimuli. In particular, point-light displays, which eliminate most form/structural information by reducing the actor to a set of 11 moving dots, still allow for meaningful interpretation (Johansson, 1973). In addition to the point-light display, researchers also found that observers are able to recognize actions with partial information due to occlusions, for example, when actors are viewed through a set of apertures (Lu, 2010; Shiffrar, Lichty, & Heptulla Chatterjee, 1997). (2) Recognition performance is robust with noisy displays. For example, humans can still recognize point-light actions embedded in a noisy background (Bertenthal & Pinto, 1994; Cutting, Moore, & Morrison, 1988; Neri, Morrone, & Burr, 1998; Thompson, Hansen, Hess, & Troje, 2007), assigned with random contrasts (Ahlstrom, Blake, & Ahlstrom, 1997) or jittered positions (Beintema & Lappe, 2002), associated with scrambled depth (Bulthoff, Bulthoff, & Sinha, 1998; Lu, Tjan, & Liu, 2006), or assigned with a general transformation such as inversion (Bertenthal & Pinto, 1994; Bertenthal, Proffitt, & Cutting, 1984; Fox & McDaniel,

1982; Pavlova & Sokolov, 2000; Pinto & Shiffrar, 1999; Simion, Regolin, & Bulf, 2008; Sumi, 1984), and even when spatially scrambled (Chang & Troje, 2009; Troje & Westhoff, 2006). (3) Humans exhibit fine discrimination ability in perceiving actor characteristics such as gender (Cutting & Kozlowski, 1977b; Mather & Murdoch, 1994), identity (Cutting & Kozlowski, 1977a), emotion (Dittrich, Troscianko, Lea, & Morgan, 1996; Roether, Omlor, Christensen, & Giese, 2009), and sign language meaning (Poizner, Bellugi, & Lutes-Driscoll, 1981). These data have been incorporated into various theories that aim to explain why the visual system is so adept at interpreting biological motion (e.g., Giese & Poggio, 2003; Troje & Westhoff, 2006).

An important limitation of work on biological motion, however, is that research has been heavily tilted toward the use of walking figures. This bias may be due to historical accident (walking was the first action investigated using point-light stimuli; Johansson, 1973), as well as to the mathematical tractability of the combined joint movements in spatial-temporal space (Cutting, 1978; Troje, 2002). However, humans are able to perceive a wide range of action categories other than walking when presented as point-light displays (e.g., boxing, dancing, jumping jacks; Brown et al., 2005; Dittrich, 1993; Dittrich et al., 1996; Giese & Lappe, 2002; Ma, Paterson, & Pollick, 2006; Norman, Payton, Long, & Hawkes, 2004; Thurman & Grossman, 2008) and communicative interactions (Manera, Schouten, Becchio, Bara, & Verfaillie, 2010; Neri, Luu, & Levi, 2006). Comparing human

performance for these different actions may help us understand how the visual system conducts an intelligent spatial–temporal analysis for action recognition (Giese & Lappe, 2002; Thurman & Grossman, 2008).

Various actions are not equally easy to identify (although differences are generally small: between 87% and 95% correct in Dittrich, 1993 for actions that we employed in our study). However, there have been few systematic investigations into how the human visual system categorizes different actions and assesses the similarities among them (Giese & Lappe, 2002; Manera et al., 2010; Vanrie & Verfaillie, 2004). Furthermore, actions other than walking, such as running, boxing, and dancing, may be far more informative, especially as they relate to social and threatening situations. For example, a punching figure is likely to be of more interest to a video surveillance system than a walking figure would be.

To address these issues, we employed a visual search approach to study perception of action categories. This method allows us to measure the efficiency of search for a certain action among other actions and, thereby, infer what information the visual system uses to assess the similarity between the different actions. Visual search has long been used as a paradigm to investigate relationships among different types of stimuli (Treisman & Souther, 1985). Typically, visual search involves one target that the observer needs to find, within a field of distractors that differ from the target in certain features. The number of distractors is varied, creating different numbers of total items (i.e., set sizes). The dependence of the reaction time on the number of items (the “search slope”) is an indication of search efficiency: the steeper the slope, the less efficient is the search (Treisman & Souther, 1985; Wolfe, 1998).

Visual search is mainly used for two reasons. First, the forced search for one item among many distractors involves the use of divided attention, and therefore, visual search provides a means to investigate attentional demands of a task. Second, some experimental conditions have revealed search asymmetries, such that the search for target A among distractors of type B is more efficient than the reverse search for target B among distractors of type A. According to Treisman and Souther (1985), finding “basic preattentive features” is faster than recognizing the absence of these preattentive features. Accordingly, visual search asymmetries are indicative of the potential presence of critical features in the target stimulus.

Search asymmetries have been observed in low-level visual processing, such as color and orientation (Treisman & Gormican, 1988; Wolfe, 2001), showing that low-level features can aid in rapid detection when they distinguish between target and distractors. Asymmetry effects have also been found in high-level visual processing, such as letter and object recognition (Wang, Cavanagh, & Green, 1994; Wolfe, 2001). However, here one should be careful when interpreting these findings, as search asymmetries in

a high-level visual task could be due to low-level or middle-level visual features. For example, Levin, Takarae, Miner, and Keil (2001) reported an asymmetry effect in a search task by object category. The researchers used two heterogeneous and complex object categories, artifacts and animals in the study. Participants were asked to search for a line drawing of an object randomly selected from one of the categories among distractors from another category. They found a search asymmetry favoring detection of artifacts among animals rather than finding animals among artifacts. Their follow-up experiments further revealed that this search asymmetry effect was largely due to two critical features, global contour shape (e.g., rectilinearity/curvilinearity) and visual typicality of objects, which distinguishes artifact and animal categories. It is, therefore, important to check whether search asymmetries in high-level tasks are not due to unintended low-level influences.

A few studies have employed a visual search paradigm to investigate biological motion using walking stimuli (Cavanagh, Labianca, & Thornton, 2001; Hirai & Hiraki, 2006; Wang, Zhang, He, & Jiang, 2010). Cavanagh et al. asked subjects to detect a target (a point-light actor walking toward either the right or the left) among the distractors with the opposite gait. They found a significant increment of reaction time with increase in the number of display items. An event-related potential study by Hirai and Hiraki reported a strong association between behavioral reaction time and enhanced negativity in the ERP signal in a task of searching for a point-light walker among scrambled walkers. These findings suggest an attention-based visual routine for processing biological motion information (Cavanagh et al., 2001; Thornton, Rensink, & Shiffrar, 2002). However, Thornton and Vuong (2004) used a peripheral flanker paradigm to show that “to-be-ignored” surrounding walkers influence the discrimination of a central walker, suggesting incidental processing of biological motion. These data suggest that different tasks could lead to different influences of attention on biological motion processing. It is, therefore, very well possible that attentional processing differs between different action categories. This possibility was investigated in the present study.

Wang et al. (2010) performed a series of informative experiments in which observers were asked to search for an intact walker among inverted walkers, and vice versa. No search asymmetry was apparent in reaction time (RT) for these two conditions, suggesting the absence of preattentive features in processing point-light walking actions. Interestingly, when point-light walker stimuli were spatially scrambled or the stimuli only included feet points, observers did show a search asymmetry effect, searching more efficiently for an upright scrambled walker among inverted scrambled walkers than vice versa. The search asymmetry obtained for the scrambled and feet-only conditions suggests the important role of low-level

features, such as local biological motion signals, in perception of actions (see also Troje & Westhoff, 2006). Since the perception of action category was out of the scope of the work by Wang et al., these authors only used walking figures in their study. We will extend the experimental paradigm to visual search by action category.

In summary, a visual search paradigm was used in the present study to assess attentional demands for the detection of a certain action among other actions, to investigate whether there exist any preattentive features that distinguish among the different actions, and to determine whether human performance in the visual search task can be used to measure similarities between different action categories.

## Methods

### Subjects

Sixteen subjects participated in this experiment. One was excluded because of overall low performance. Of the remaining subjects, 8 were male and 7 were female (mean age of  $22.5 \pm 4.5$  years). Subjects were paid for their participation, and all were naive to the purpose of the experiment. All procedures received prior approval from the UCLA IRB.

### Stimuli

#### *Point-light actions*

We drew our action stimuli from a free online motion capture database (<http://mocap.cs.cmu.edu>), which provides the three-dimensional coordinates for all joints of the actors over time. We used 13 joints to display the point-light walkers (the two feet, knees, hips, wrists, elbows, shoulder joints, and the head).

The 3D positions of the joints were projected on the screen using orthogonal projection. The height of the figures was scaled down to be  $3.5^\circ$ . Point-light displays were created by displaying black dots (diameter of  $0.25^\circ$ ) on a white background (Figure 1 shows two sample stimuli). The individual actions rotated in depth at  $150^\circ/\text{s}$ . The rotation was added for two reasons: (1) some dancing actions included self-rotation around the vertical axis, which we wished to make less salient (without completely erasing this important component of dancing) by adding rotation in depth on each action; (2) the rotation increased the complexity of dot movements and brakes the periodicity of the 2D movements. Accordingly, this manipulation enhanced the heterogeneity of actions within each category and decreased low-level motion informativeness. All actions were still easily recognized in isolation, although occasional depth reversals occurred. Our pilot data showed that the occasional depth reversals did not affect the identification of each action. The initial rotation angle was randomly selected for each item. Each item started at a random frame within the movie sequence.

In order to allow for looping of the movie sequences (required when subjects' response times were longer than the movie sequences), we devised an in-house software to find two frames that allowed for the smoothest transition without a clear discontinuity. Briefly, we performed an exhaustive search for the minimal mean of all squared distances between matching joints in the two frames. We chose the frames with the minimum mean squared distance as start and end frames, with the restriction that the length of the movie was at least 200 frames (80 for walking and running actions), all dots continued moving in the same direction, and the action was clearly recognizable. This procedure may leave small movement discontinuities at the time of looping; however, these were not easily observable in the periphery and, moreover, did not differ among categories. For example, mean of squared errors between start and end frames did not differ between walking and boxing categories (the categories



Figure 1. Stimulus illustration. (Left) A frame from a trial with 6 items, without a target (distractors are walkers in this case). (Right) A boxing target among five walking distractors.

that are most distinct; see [Results](#) section),  $t(5.01) = 2.33$ ,  $p > 0.05$ , two-tailed  $t$ -test.

We used four different action categories: boxing, dancing, running, and walking. Note that even though walking and running seem very similar, dynamic models of gait production show that they are, in fact, different categories (Golubitsky, Stewart, Buono, & Collins, 1998, 1999). We used different numbers of movies per action category (6 movies for boxing, 8 movies for dancing, 10 movies for running, and 14 movies for walking). The larger set of movies for running and walking was chosen in order to introduce more movies from different actors so as to increase the variability within these relatively low heterogeneous categories. Boxing and dancing categories are composed of more rich and varied subactions, which we describe here. Two boxing movies consisted mainly of regular punches as in boxing matches. In between punches, the arms were sometimes kept in front of the face. Each movie had one instance in which the boxer ducked away from a punch from an invisible opponent. A third movie contained 3 regular punches and 3 uppercuts. The fourth movie contained 2 regular punches and 1 uppercut. The fifth movie had 5 regular punches. The sixth movie contained several punches that might be made toward a speed bag. No kicks were performed in any movie, and a step was made in only one boxing movie. Several movies contained some rebalancing actions (with small foot movements). The dancing category consisted of a movie with a ballet movement with slow expressive arm movements and a pirouette, a movie with a slow turning movement with the hands slightly up and extended, and 6 movies with male salsa dancers, which consisted of a few dance steps, with one or two leads to a turn for the invisible partner and one or two turns of the dancer. Care was taken to make both dancing and boxing movies also appear rather cyclic by taking parts that showed several reoccurring

actions, e.g., punches, with more or less regular timing. From the sample movement traces in [Figure 2](#), one can see that the amplitude of the dot motions is comparable between boxing and walking for most joints, except for the wrist joints (blue lines).

### Display

A black fixation mark (size: 0.4 deg) was drawn at the center of the screen. Around the fixation mark at a distance of 6.7 deg, 9 equally spaced positions were assigned on an invisible circle, at which positions the PLAs could appear. Subjects used a chin rest to maintain a distance of 57 cm from the screen. Screen resolution was 1024 pixels  $\times$  768 pixels and the refresh rate was 75 Hz.

### Procedure

Subjects were asked to search for a specified target action among different distractor actions. Within each trial, the different distractor items were randomly sampled from one action category. Participants were asked to indicate whether the target was absent or present by pressing one of two buttons.

Each trial started with a 1-s pause during which on-screen instructions informed observers which target action to search for on that trial and the completed/total number of trials. Subjects were asked to fixate the center point during this period. Once the PLAs were displayed, the subject was allowed to move his/her gaze around.

There were four target and distractor actions, three set sizes (3, 6, or 9 items) and an equal number of target-absent and target-present trials. All combinations were shown, except for same target and distractor conditions

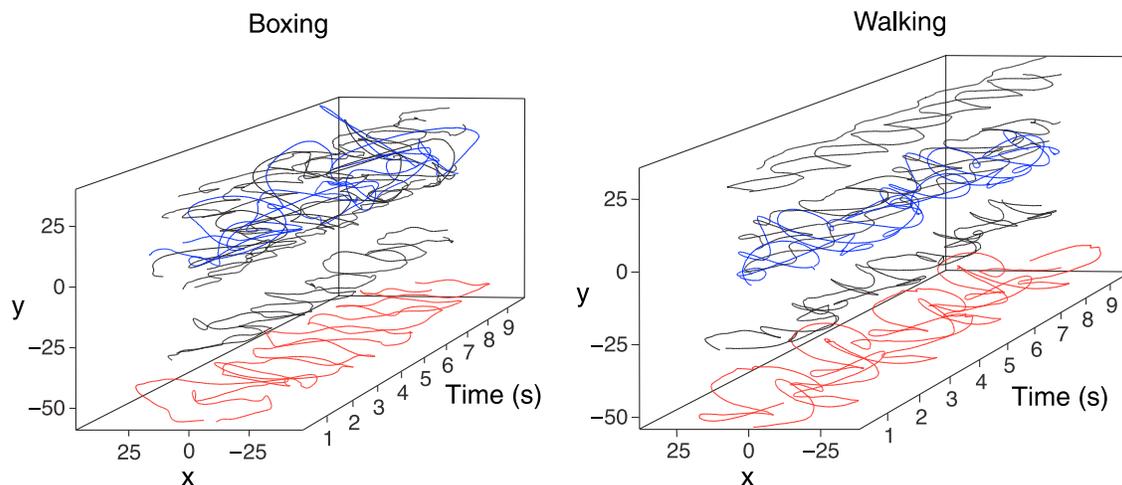


Figure 2. Motion paths. The  $x$  and  $y$  positions of all the dots over the first 10 s of a typical trial of (left) a boxing item and (right) a walking item. The red lines are the two ankle joints, the blue lines are the two wrist joints, and the black lines are the other point-light joints.

(e.g., no search for a boxer among boxers). Each condition was repeated 10 times, yielding a total of 720 trials. Target actions were blocked into different sessions of 180 trials, while all other variations were displayed in random order within each block. The order of the blocks was counterbalanced across subjects.

## Results

### Task performance

First, we analyzed task performance using signal detection theory, in order to determine whether searching for actions is generally a well-performed task and whether there are any inherent biases in indicating whether or not actions are present in a search display.

Performance was analyzed using sensitivity index,  $d'$ . We also analyzed the response bias, criterion, in order to determine whether or not people were more or less likely to report the presence or absence of an action. Values of  $d'$  were generally high (consistent with the instruction to be as accurate as possible) but decreased with increasing number of items (Figure 3, left), as revealed with a repeated-measures ANOVA ( $F(2,18) = 14.96, p < 0.0001, \eta^2 = 0.52$ ).

The criterion, the measure of response bias, was significantly above zero, indicating that the observers were biased to report “absent” in the visual search task, even in 3-item displays. This bias leads to elevated miss rates when targets are rare (Wolfe, Horowitz, & Kenner, 2005; Wolfe & Van Wert, 2010), suggesting that the observers anticipated the targets to be absent. Furthermore, the criterion increased with the number of items ( $F(2,18) = 4.27, p = 0.03, \eta^2 = 0.23$ ). This result (Figure 3, right) indicates that with a more complex action background (i.e., an increasing number of items), subjects were more likely to report the absence of the target.

The  $d'$  and criterion did not differ among the different target–distractor combinations, with the exceptions of a lower  $d'$  when searching for a dancer among boxers (compared to all other  $d'$ -means,  $Z = 2.77, p < 0.01$ ; see Figure 4) and a lower criterion level for a boxer among dancers ( $Z = 5.94, p < 0.001$ ).

### Visual search efficiency

Visual search efficiency is generally analyzed in terms of the linear relationship between RT and set size. Figure 4 shows the RT  $\times$  set size plots for all combinations of target actions and distractor actions. Clearly, a wide range of patterns is evident: some slopes are nearly flat (indicating efficient search), whereas others are relatively steep, up to about 250 ms/item (indicating less efficient search). Mean slope values are depicted in Figure 4. Mean intercepts (at set size 1) for present and absent trials are reported in Table 1.

As shown in Figure 4, all slopes were significantly different from zero (all  $p < 0.02$ ), indicating that there was no pop-out of actions in the displays we used (slopes ranged from 90 ms/item to 251 ms/item for present trials and from 225 ms/item to 502 ms/item for absent trials).

All but one slope in the absent condition were significantly larger than the corresponding slopes in the present condition, replicating the well-known absent/present search asymmetry (Chun & Wolfe, 1996; Treisman & Souther, 1985). (The exception was the runner among boxers,  $p > 0.4$ , paired  $t$ -test.) A near-significant correlation was obtained between the size of the slope in the present and absent conditions ( $r = 0.57, p = 0.054$ ), with the absent slopes being, on average, 2.15 times larger than the present slopes (not significantly larger than 2,  $t$ -test; cf. Wolfe, 1998), suggesting a self-terminating search (Sternberg, 1966). It appears that the visual system conducts search among actions by contrasting expected (template) target and distractor actions, regardless of whether the target item is, in fact, present.

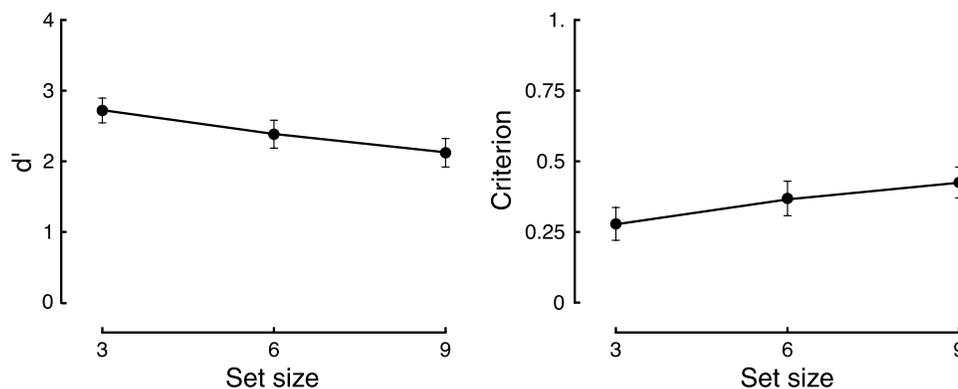


Figure 3. Signal detection analysis, showing (left)  $d'$  and (right) criterion values. With increase in set size,  $d'$  decreases, and the criterion increases.

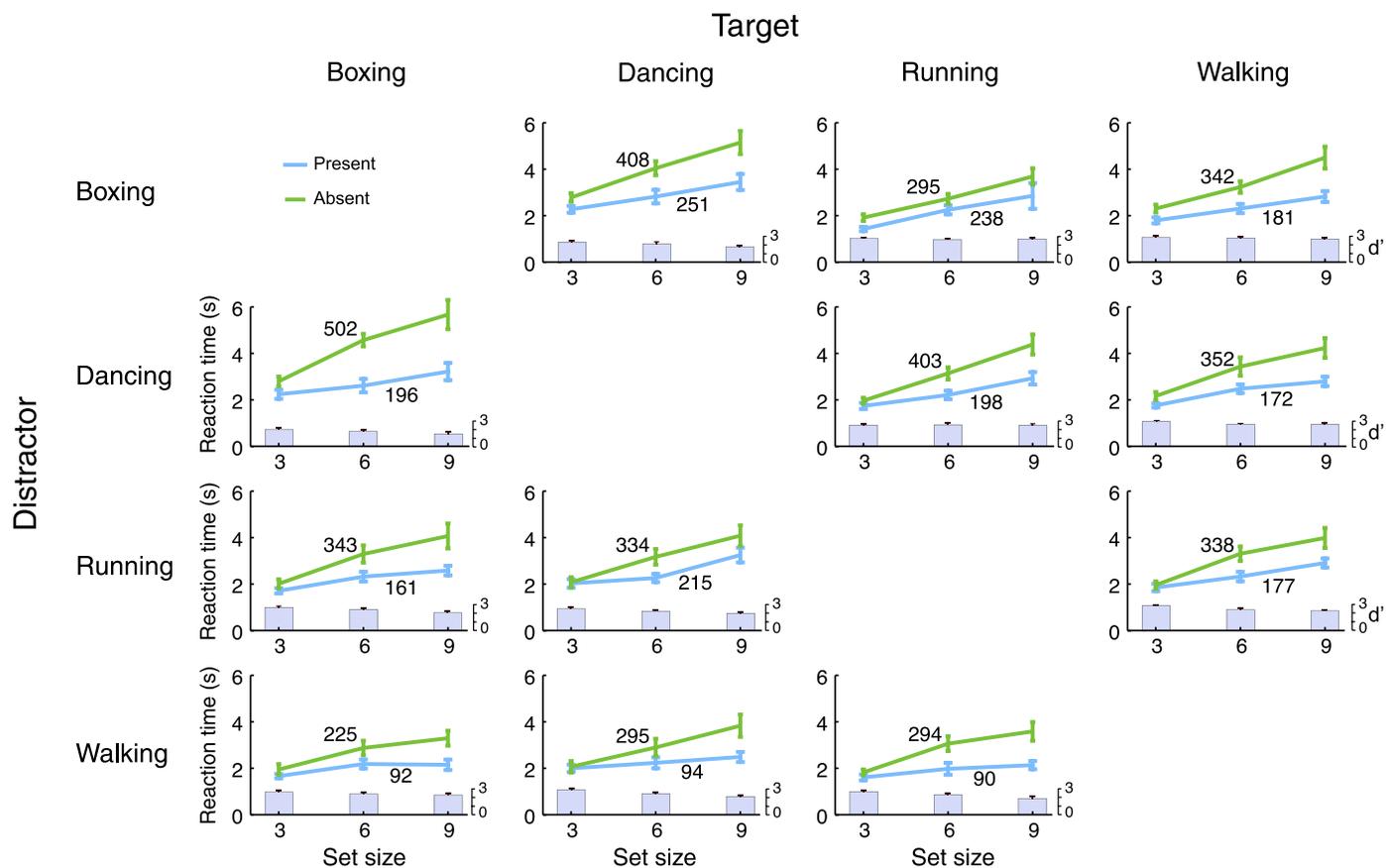


Figure 4. Reaction time and  $d'$  as a function of set size. Search time varied depending on different combinations of target action and distractor action. The green and blue lines indicate search performance when the target was present and absent, respectively. The blue bars represent the  $d'$  values. The  $d'$  values use the secondary abscissa. Data points depict mean  $\pm$ SEM over subjects. Values represent the slopes in ms/item obtained with a linear fit.

### Visual research asymmetries

To more clearly illustrate the differences among different target–distractor combinations, we calculated the slopes in each condition for each subject separately and devised a “slope cube.” The slope cube displays the

		Target			
		Boxing	Dance	Run	Walk
Distractor	Boxing		1.87	1.15	1.45
	Dance	1.91		1.30	1.48
Run	Boxing	1.99		1.15	1.55
	Dance	1.48	1.49		1.48
Walk	Boxing	1.45	1.44		1.39
	Dance	1.59	1.83	1.47	
		1.58	1.46	1.35	

Table 1. Mean slope intercept at set size 1 (in seconds) in the trials of target present (top value in each cell) and target absent (bottom value in each cell) per target–distractor combination.

relative magnitudes of the slopes for all measured conditions (Figure 5). The thickness of the arrows scales with the magnitude of the slope, and the arrows point from the target action to the distractor action. The left plot shows the slopes for the target-present conditions, and the right plot shows the slopes for target-absent conditions.

Clearly, different combinations of actions led to different search slopes. For example, when a target walker was presented in a display of runners, it proved relatively easy to find, revealed by a flat slope as shown in Figure 4 (the plot at row 3 and column 1). However, when the same action was displayed among boxers, the steeper RT slope indicates the greater difficulty of finding a walker among boxers (the plot at row 1 and column 1). Such different relationships were found among different action combinations.

Some searches were relatively fast and efficient either way around, e.g., finding a walker among runners and finding a runner among walkers were both relatively easy searches. Other searches were slow and less efficient both ways around, e.g., finding a dancer among boxers and finding a boxer among dancers were both accompanied by large search slopes. Finally, some conditions led to

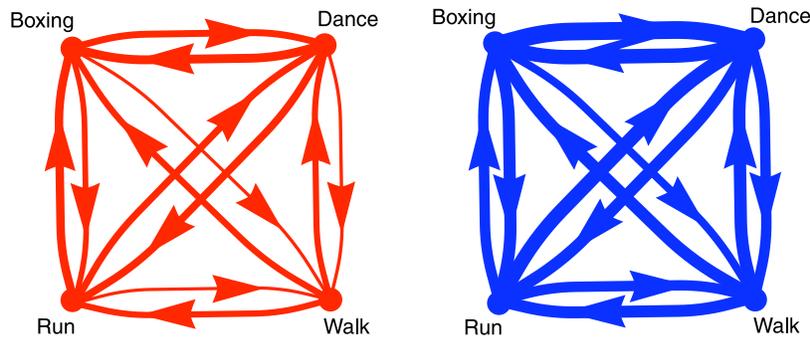


Figure 5. Graphical presentation of the different search slopes. All the search slopes between the different conditions are shown with the arrow pointing from the target to the distractor type. The thicker is the arrow, the larger is the slope. The left panel shows the target-present conditions, whereas the right panel shows the target-absent conditions (same scaling).

asymmetrical relationships in search slopes, e.g., a boxer is easy to find among walkers, whereas a walker is difficult to find among boxers. Such a search asymmetry has been assumed to be diagnostic of the presence of preattentive features in the fast-search target (Treisman & Souther, 1985).

We analyzed search asymmetries by subtracting for each subject the slopes of one search direction (e.g., boxer among walkers) from the other direction (walker among boxers). The difference, averaged over subjects, is plotted in Figure 6 for both present and absent searches. There was only one condition that led to significant search asymmetries: the boxer–walker combination ( $p < 0.02$  in present (Cohen’s  $d = 0.68$ ),  $p < 0.05$  in absent (Cohen’s  $d = 0.55$ ), paired  $t$ -tests between boxer-in-walkers and walker-in-boxers conditions).

## Target recognition versus distractor rejection

For slow search performance, as found in our experiment, it has been suggested that search is serial, with each item being scrutinized individually and accepted or rejected as a target. In this type of search, it has been proposed that the search slope depends on the speed at which distractors can be rejected (Duncan & Humphreys, 1989; Treisman & Souther, 1985). To test this hypothesis, we performed a repeated-measures ANOVA on the search slopes, with the target and distractor types as independent variables (setting the slopes to zero for identical target and distractor combinations because in such conditions the first gaze would always fall on a target, and therefore, there should be no dependence on the number of distractors). The main effect of target was not significant ( $F(3,42) < 1$ ), whereas the main effect of distractor was significant ( $F(3,42) = 7.70$ ,  $p < 0.0005$ ,  $\eta^2 = 0.36$ ). This result is in agreement with the previous findings concerning the important role of distractor rejection in determining the search slope. The interaction of target  $\times$  distractor was also significant (Greenhouse–Geisser correction:  $F(3.79,53.08) = 12.04$ ,  $p < 0.0001$ ,  $\eta^2 = 0.46$ ), indicating

that similarity between the target and distractors can also influence search efficiency.

Another factor that affects search speed is target recognition time. This time can be best approximated by taking the intercept at set size 1, for the target-present conditions (Tong & Nakayama, 1999). Performing a similar ANOVA as above on these intercept values, we find a significant effect of target ( $F(3,42) = 4.49$ ,  $p < 0.01$ ,  $\eta^2 = 0.24$ ) but a non-significant effect of the distractor ( $F(3,42) < 1$ ). The interaction of target  $\times$  distractor was also significant ( $F(9,126) = 12.36$ ,  $p < 0.001$ ,  $\eta^2 = 0.80$ ).

Overall, the present results revealed that search slopes are more dependent on the type of distractor actions, whereas search intercepts are more dependent on the type of target actions. However, the similarity between the

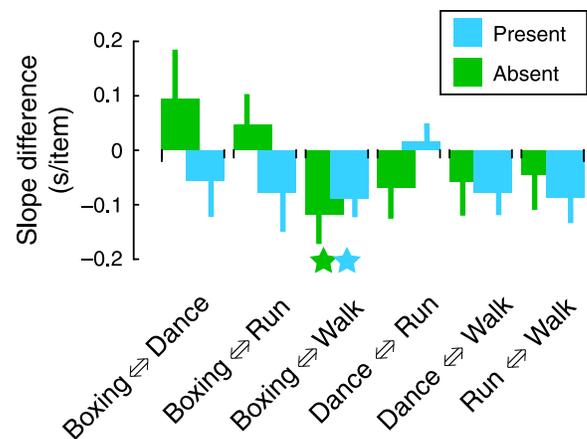


Figure 6. Visual search asymmetries. The differences in slopes (as calculated per subjects and then averaged). Positive values mean that the first-mentioned category (e.g., boxing in “boxing ⇔ dance”) has a higher slope when it is target among distractors of the second-mentioned category (e.g., dance) than the reverse search (e.g., dance target among boxing distractors). Only the combination of walkers and boxers leads to a significant search asymmetry. Data points depict mean  $\pm$  SEM over subjects.

target and distractor actions influences both search efficiency and target recognition time.

## Clustering

Action similarity measures have been an active research topic in the field. Studies have typically involved asking observers to make similarity judgments in psychophysical experiments and applying multidimensional scaling method on human similarity ratings to reveal the low dimensional perceptual space. Giese, Thornton, and Edelman (2008) reconstructed the perceptual space of body movements (including walking, running, and marching) and compared these with the physical space. Their analyses supported the hypothesis that perceptual metrics of actions are characterized by the physical metrics based on the similarity measure of joint trajectories. Pollick, Paterson, Bruderlin, and Sanford (2001) examined the emotion space reconstructed from human performance in categorizing the emotional status by viewing point-light arm movements (knocking and drinking) with different affects. They found that the most important perceptual dimension in the emotion space correlated with kinematics of arm movements.

We applied a hierarchical clustering method on search slopes of all target–distractor combinations to measure similarity of action categories. As can be observed in Figures 4 and 5, there were very obvious differences in the magnitudes of the slopes among the different conditions. In order to assess how the different actions relate to each other in terms of search efficiency, and presumably in terms of the processing mechanisms employed by the brain, we performed a hierarchical clustering analysis based on a nearest neighbor clustering algorithm. We entered the average slopes of each target–distractor combination for both present and absent conditions into the Agglomerate function in Mathematica 7 (Wolfram Research, Champaign, IL), using the standard options (e.g., squared Euclidian distance). Conditions in which the target was identical to the distractor (which were not run in the experiment) were entered as a zero slope. This cluster analysis was based on the measure of slope similarity conditional on action distractors (which was, by necessity, given our experimental design). For example, when a search for walkers and runners yielded similar slopes when shown among dancer or among boxer distractors, then walkers and runners will be clustered close together.

The cluster analysis (Figure 7) revealed boxing as an out-group, followed by the dancer, while the walker and runner proved to be most related. This pattern matches well with our intuition, which would presumably group walking and running as closely related (see also Giese & Lappe, 2002), followed by dancing (as it also involves quite a bit of walking), and treating boxing as quite

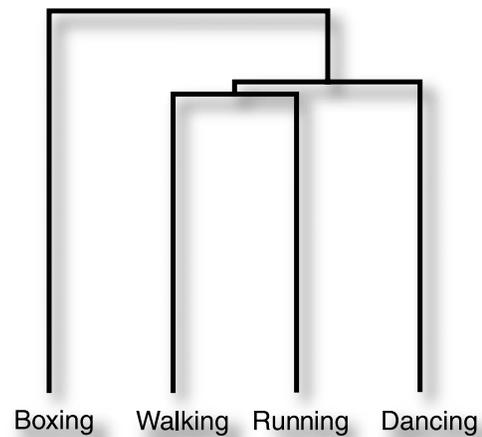


Figure 7. Dendrogram showing the relationship between the different actions. The relationships were derived from a clustering analysis on the target-present slopes.

something different (as it is primarily performed with the arms and not with the feet).

## Is low-level information important? An ROC analysis

To examine the contribution of low-level information in visual search by action categories, we performed non-parametric ROC analyses on the mean and maximum inter-frame velocity of point lights, the mean and maximum acceleration (i.e., derivative of speed) of point lights, as well as on a measure of positional spread (sum of squared distances from the mean position). We calculated these values along both the horizontal and vertical directions independently for each joint of each displayed item within a trial. We then averaged over time and over joints (i.e., average velocity in 2D with  $x$  and  $y$  directions) or took the maximum of all values (i.e., maximum 2D velocity). We did this in each trial for all items. Then, the distance was calculated between each item (now represented by one value (maximum/average speed/acceleration)) from the mean of the remaining items in a trial in velocity/acceleration/position space. Trials in which the maximum distance passed a threshold were identified as “target-present” trials; other trials were “target-absent” trials. These target-present and target-absent markers were compared to the actual presence of a target. These comparisons yielded hits, misses, false alarms, or correct rejections. ROC curves were constructed by varying the threshold, and the area under the curve (AUC) was calculated with the standard trapezoidal numerical integration (trapz function of MatLab).

Figure 8 shows that in no instance did the ROC analysis reach the performance levels of the subjects. In addition, there was no strong evidence for a decrease in performance

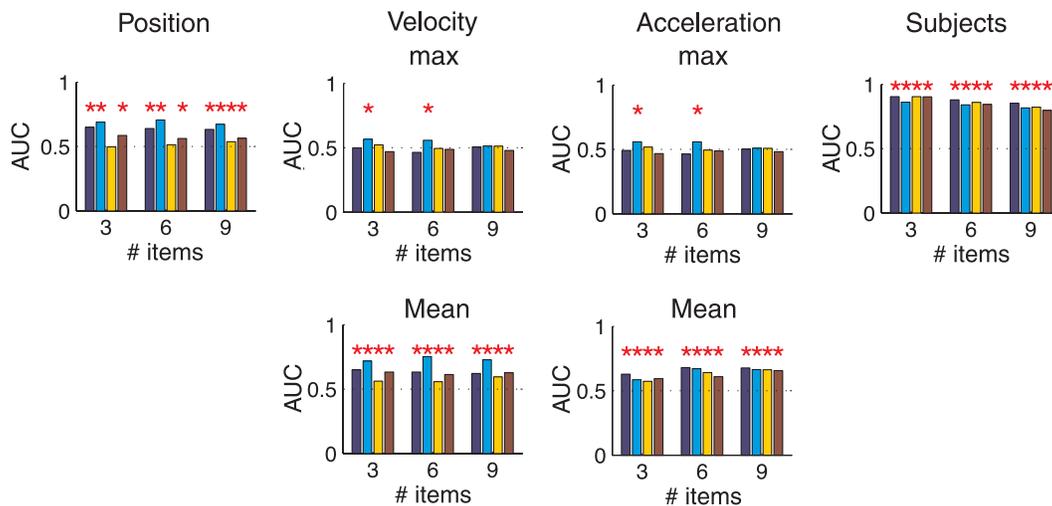


Figure 8. Results of the ROC analysis. AUC values are shown for position, velocity, and acceleration information. For velocity and acceleration, results are shown for both mean and maximum velocities/accelerations. In the rightmost column, the subjects' average results are shown. None of the ROC analyses reached levels comparable to human performance. Values significantly different from chance-level performance of 0.5 are marked by a red star.

with increased amounts of items. Overall, velocity signals seemed to show the best performance, which is consistent with neurophysiological data that showed that “motion” cells signaled different action categories better than “snapshot” cells (cells that were equally sensitive to snapshots from action clips as to the actual movie clip; Vangeneugden, Pollick, & Vogels, 2009). However, performance never reached the average human performance in the most difficult condition (walker among boxers, with an AUC of about 0.85).

It has been suggested that a target can be found based on it being an outlier among the distractors (e.g., Rosenholtz, 1999). We tested this hypothesis by using Mahalanobis distance (instead of maximum distance) as the decision variable in the ROC analysis (as suggested by Rosenholtz, 1999) and found that performance increased rather than decreased with the number of items (while still not reaching human performance).

Our results, thus, indicate that neither low-level motion information, such as instantaneous speed and acceleration, nor overall positional spread can explain human performance in the action search task. This appears to be contrary to a previous report showing that an ideal observer model can extract biological motion information on walking direction better than human observers can do it (Gold, Tadin, Cook, & Blake, 2008). However, because ideal observer models need to be tailored to each specific task and the given spatial-temporal templates, the two findings are not actually in conflict.

Apparently, a more sophisticated analysis on the kinematic information than velocity or acceleration differences underlies human performance in our task. It is possible that humans may integrate speed or acceleration signals over multiple frames, which might change the results. However, any difference between items is likely to decrease through

this averaging, and therefore, it seems unlikely that simple integration will result in better performance. Alternatively, a more complex analysis of “motion paths” (including measures of motion trajectories, such as curvature) could be involved, or some sort of “life detector” mechanism that is more sensitive to certain movements. Which of these mechanisms is at work is currently unclear.

### Is prior action experience important for visual search among actions?

We also acquired data on whether subjects had any experience with boxing or dancing. Boxing experience was rare (only two subjects). Dancing was more common; 5 subjects reported having experience with dancing. The inclusion of these variables as covariates in the analyses did not lead to any qualitative changes in the reported pattern of data. These results seem to be in conflict with previous studies that have reported that prior experience can influence action recognition (e.g., Calvo-Merino, Ehrenberg, Leung, & Haggard, 2010; Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005); however, the present experiment was not specifically designed to investigate such effects and potentially lacked the statistical power necessary to detect them.

## Discussion

In this study, we have analyzed the relationship between the processing of different actions. Even though it has been shown that different actions can be recognized in point-light displays (Dittrich, 1993; Giese & Lappe, 2002;

Ma et al., 2006; Norman et al., 2004; Thurman & Grossman, 2008), how this feat is achieved is not clear. We addressed this question using a visual search paradigm to investigate the relationships among the different actions (Giese & Lappe, 2002). In this section, we will first compare search efficiency for actions with human performance in other search tasks, followed by a discussion of how search results can help us understand the relations among different action categories. Finally, we will discuss factors influencing action search.

## Visual search efficiency for actions

Our experiment showed that search slopes for actions were always greater than zero, indicating that action search was effortful and required attention (Treisman & Souther, 1985). Search appeared to be serial, as we showed that target-absent trials take about twice as long as target-present trials (e.g., Chun & Wolfe, 1996; Sternberg, 1966).

How do our results compare to other visual search tasks? Our slopes are comparable to earlier reports of biological motion. For example, when searching for leftward (rightward) moving walkers among right (left) walkers, slopes were  $\sim 120$  ms per item (Cavanagh et al., 2001). Likewise, when searching for upright among inverted, the slopes were 50 to 100 ms for intact stimuli (Wang et al., 2010). This comparative similarity between these studies and ours, which involves search for different actions among other actions, suggests that BM search for actions by category is not easier than search by walking directions, which may be considered as a harder task because more refined discriminations are needed.

Overall, our slopes ( $\sim 80$ – $250$  ms/item) are considerably larger than search for simple features and for conjunctions of two (or three) elementary features ( $0$ – $20$  ms/item; e.g., Treisman & Souther, 1985; Wolfe, 1998). Compared to other “complex searches,” such as for faces, search slopes for actions are in the same range, though at the high end. Face search (either of line drawings or pictures) ranges from very rapid ( $<10$  ms/item) to rather slow ( $\sim 100$  ms/item) and may be even  $200$ – $400$  ms for scrambled faces (Kuehn & Jolicoeur, 1994; Nothdurft, 1993).

Action search involves spatial–temporal analyses, which arguably are more complex and time-consuming than processing on static face images. It is, therefore, not surprising that humans achieve slightly lower search efficiency for actions than for faces. However, not all searches for dynamic stimuli are slow; visual search for global motion (translation and expansion) is much faster ( $<20$  ms/item; Wolfe, 1998) than what we found in action search. These data indicate that biological motion is more difficult to discriminate than simple global motions (Neri et al., 1998). However, search efficiency is significantly reduced when searching for unfamiliar and slightly more complex motion patterns. Cavanagh et al. (2001) found that search slopes were greater than  $450$  ms/item when

searching for a tumbling pair of dots (two dots revolving about their center of mass) among orbiting pairs (one dot orbiting another) or vice versa. Importantly, the search slopes with biological motion in our experiment ( $\sim 80$ – $250$  ms/item) indicate higher efficiency than slopes with cycloid motions with two dots ( $>450$  ms/item), even though biological motion is made up of many dots making cycloid motions. This is evidence that the visual system analyzes biological motion in a rapid manner despite the inherent stimulus complexity in dynamics and configuration.

Thus, based on our search slopes ( $\sim 80$ – $250$  ms/item), it does seem that action discrimination is generally an attention-demanding task (even though the size of these slopes may be partly explained by the fact that people were allowed to make eye movements). Still, the slopes also seem shallower than those of complex motion sequences (i.e., tumbling motion pairs; Cavanagh et al., 2001). It is, therefore, possible that high-level influences (e.g., our familiarity with biological motion or phase relationships between joints; Cavanagh et al., 2001; Wang et al., 2010) sped up the search, while it still required some amount of attention.

## Relationships among different actions

Even though the different actions proved to be sufficiently similar to yield significant search slopes, there is an underlying structure in the action category divisions employed by the visual system. We analyzed this structure by performing a cluster analysis on the slopes, and we found that walking and running actions are most similar, followed by dancing and finally boxing (Figure 7). This clustering is consistent with how we would intuitively group these actions. It is also consistent with the finding that the visual system can easily generalize between locomotive actions (walking, running, and marching) but not between, e.g., boxing and walking (Giese & Lappe, 2002). Together, these findings suggest that the brain possesses clear action representations by categories, a suggestion that is supported by the finding the different neurons in the superior temporal sulcus prefer different actions (Vangeneugden et al., 2009).

Intuitively, one may assume that observers need more time to recognize actions occurring less frequent than actions occurring frequently. Arguably, boxing/dancing are less frequent to be observed in everyday life than running/walking. So, we would expect that participants need more time to recognize boxers/dancers than runners/walkers. The expected result is confirmed by our search data on intercepts at set size 1 (Table 1), showing that recognizing boxers/dancers requires  $300$  ms more time than recognizing runners/walkers. However, we acknowledge that other factors other than action frequency may contribute to the difference for search intercept. The heterogeneity of action category and typicality of individual actions may also play an important role in determining

recognition performance. In addition, although boxing/dancing are relatively uncommon actions, they are not more unusual compared to cycling or skipping. When searching is conducted between common actions and rare actions, we expect that high-level information, such as novelty (Wang et al., 2010, 1994), may be a significant factor in contributing how observers conduct action search. More research is needed to examine these possibilities.

### Factors that influence search: Distractor rejection and target recognition

Visual search can help characterize the relationships among different actions. However, our paradigm also offers insight into visual search itself. We analyzed the search slopes and the intercepts (at set size 1; Tong & Nakayama, 1999). We found that search slopes depended strongly on the type of distractor action, whereas the intercept depended strongly on the type of target action. This pattern suggests that there are two important processes in the search for actions. The first is distractor rejection. The faster a distractor can be rejected, the shallower the slope (i.e., the more efficient the search). Especially, shallow slopes were found for walking and running as distractor items. The rapid rejection of these distractors suggests that the brain has a very well-defined template for them, which allows for rapid comparison and subsequent rejection. By contrast, boxer and dancer distractors yielded much steeper slopes, indicating less well-defined templates for their action categories. However, it is important to note that because dancer and boxer stimuli are less stereotypical than walker and runner stimuli, they create larger variability when presented as distractors, which may be sufficient to cause augmented slopes (Duncan & Humphreys, 1989; Rosenholtz, 1999).

The second important factor in search is target recognition, i.e., the time it takes to recognize something as a target. Although this process is often strongly correlated with distractor rejection (consistent with the reliable interactions obtained between target and distractor type in our experiment), these processes can be separated (Tong & Nakayama, 1999). Unfortunately, our design does not allow a full separation of target recognition and distractor rejection processes. Nevertheless, we were able to show that the intercept at set size 1 is mainly dependent on the target, consistent with it being an indicator of target recognition ease.

### Visual search asymmetries in action search

Visual search asymmetries have been widely used to assess whether a certain item possesses a critical pre-attentive feature that readily distinguishes it from other items (Treisman & Souther, 1985). Search asymmetries have been found both for low-level features, such as

orientation, color, and motion information (Treisman & Gormican, 1988; Wolfe, 2001), and high-level features, such as letters (Wang et al., 1994), face identity (Tong & Nakayama, 1999), shapes (Kleffner & Ramachandran, 1992), circling/cycloid motion (Cavanagh et al., 2001), and scrambled biological motion (Wang et al., 2010).

We found that search asymmetries generally do not occur in the search for actions by category. Therefore, it seems that there is no universal low-level critical feature for grouping actions into different categories. However, we did find one exception, as there was a search asymmetry between boxers and walkers. What may be the critical difference between these two actions?

It has repeatedly been suggested that local motion information plays an important role in biological motion perception. When local motion information is disrupted by introducing large temporal gaps, or contrast reversals, between frames, biological motion discrimination generally suffers (Mather, Radford, & West, 1992; Thornton, Pinto, & Shiffrar, 1998). However, others have suggested that local motion has only limited, if any, importance in biological motion perception *per se* (Ahlstrom et al., 1997; Beintema & Lappe, 2002) and that local motion is instead important in image segmentation (and thereby, through an indirect route, enhances biological motion perception).

Our results show that simple average or maximum velocity or acceleration signals are insufficient to account for human performance in an action search task (ROC analysis, Figure 8). It is, nonetheless, possible that more complex, novel spatiotemporal motions may have contributed to the asymmetry.

Other factors may contribute to the search asymmetry between boxers and walkers. There exists a potential intermediate level mechanism, the “life detector,” that has been proposed to be based on the orientation and phase information relationship between the two feet (Troje & Westhoff, 2006; Wang et al., 2010). In our boxer stimuli, however, it is unlikely that the feet contributed to the asymmetry, because they are largely immobile. The punch, however, may contain information that can be picked up by some sort of “life detector,” although one that differs in a major way from previous proposals of this nature (Troje & Westhoff, 2006). We also conjecture that higher level configurational information, or other information directly related to the interpretation of actions in point-light displays, may produce the search asymmetry. Future study is needed to tease apart the contributions from different levels of action processing, by inverting or scrambling the point-light actions.

## Conclusion

Search for actions shows many of the hallmarks of other serial and attention-demanding searches (Duncan &

Humphreys, 1989; Sternberg, 1966; Treisman & Souther, 1985; Treisman & Gelade, 1980; Wolfe, 2001). The visual search paradigm allowed us to investigate the relationships among different actions. We showed that the similarity structure of action categories can be derived from human performance in a visual search task. This paradigm provides a means to experimentally assess the similarity between action categories. We found that search for actions is generally not influenced by critical features, although we identified one exception: search for a boxer among walkers. The critical feature is probably not a low-level feature, as evidenced by an ROC analysis but may mainly stem from potentially novel complex spatiotemporal movements patterns in mid- or high-level processes.

## Acknowledgments

The motion capture data used in this project were obtained from <http://mocap.cs.cmu.edu>. The database was created with funding from NSF EIA-0196217. This research was supported by a grant from the National Science Foundation (NSF BCS-0843880). We thank Matthew Weiden, Janice Lau, and Michael Finch for their help in data collection.

Commercial relationships: none.

Corresponding author: Dr. Jeroen J. A. van Boxtel.

Email: [j.j.a.vanboxtel@gmail.com](mailto:j.j.a.vanboxtel@gmail.com).

Address: 6550 Franz Hall, Los Angeles, CA 90095, USA.

## References

- Ahlstrom, V., Blake, R., & Ahlstrom, U. (1997). Perception of biological motion. *Perception, 26*, 1539–1548. [[PubMed](#)]
- Baintema, J. A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 5661–5663. [[PubMed](#)]
- Bertenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science, 5*, 221–224. [[Article](#)]
- Bertenthal, B. I., Proffitt, D. R., & Cutting, J. E. (1984). Infant sensitivity to figural coherence in biomechanical motions. *Journal of Experimental Child Psychology, 37*, 213–230. [[PubMed](#)]
- Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology, 58*, 47–73. [[PubMed](#)]
- Brown, W. M., Cronk, L., Grochow, K., Jacobson, A., Liu, C. K., Popovic, Z., et al. (2005). Dance reveals symmetry especially in young men. *Nature, 438*, 1148–1150. [[PubMed](#)]
- Bulthoff, I., Bulthoff, H., & Sinha, P. (1998). Top-down influences on stereoscopic depth-perception. *Nature Neuroscience, 1*, 254–257. [[PubMed](#)]
- Calvo-Merino, B., Ehrenberg, S., Leung, D., & Haggard, P. (2010). Experts see it all: Configural effects in action observation. *Psychological Research, 74*, 400–406. [[PubMed](#)]
- Calvo-Merino, B., Glaser, D. E., Grezes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249. [[PubMed](#)]
- Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. *Cognition, 80*, 47–60. [[PubMed](#)]
- Chang, D. H., & Troje, N. F. (2009). Characterizing global and local mechanisms in biological motion perception. *Journal of Vision, 9*(5):8, 1–10, <http://www.journalofvision.org/content/9/5/8>, doi:10.1167/9.5.8. [[PubMed](#)] [[Article](#)]
- Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30*, 39–78. [[PubMed](#)]
- Cutting, J. E. (1978). Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception, 7*, 393–405. [[PubMed](#)]
- Cutting, J. E., & Kozlowski, L. (1977a). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society, 9*, 353–356.
- Cutting, J. E., & Kozlowski, L. (1977b). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics, 21*, 575–580. [[Article](#)]
- Cutting, J. E., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Perception & Psychophysics, 44*, 339–347. [[PubMed](#)]
- Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception, 22*, 15–22. [[PubMed](#)]
- Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25*, 727–738. [[PubMed](#)]
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433–458. [[PubMed](#)]
- Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science, 218*, 486–487. [[PubMed](#)]
- Giese, M. A., & Lappe, M. (2002). Measurement of generalization fields for the recognition of biological motion. *Vision Research, 42*, 1847–1858. [[PubMed](#)]

- Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, *4*, 179–192. [PubMed]
- Giese, M. A., Thornton, I., & Edelman, S. (2008). Metrics of the perception of body movement. *Journal of Vision*, *8*(9):13, 1–18, <http://www.journalofvision.org/content/8/9/13>, doi:10.1167/8.9.13. [PubMed] [Article]
- Gold, J. M., Tadin, D., Cook, S. C., & Blake, R. (2008). The efficiency of biological motion perception. *Perception & Psychophysics*, *70*, 88–95. [PubMed]
- Golubitsky, M., Stewart, I., Buono, P. L., & Collins, J. J. (1998). A modular network for legged locomotion. *Physica D*, *115*, 56–72.
- Golubitsky, M., Stewart, I., Buono, P. L., & Collins, J. J. (1999). Symmetry in locomotor central pattern generators and animal gaits. *Nature*, *401*, 693–695. [PubMed]
- Hirai, M., & Hiraki, K. (2006). Visual search for biological motion: An event-related potential study. *Neuroscience Letters*, *403*, 299–304. [PubMed]
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*, 201–211. [Article]
- Kleffner, D. A., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, *52*, 18–36. [PubMed]
- Kuehn, S. M., & Jolicoeur, P. (1994). Impact of quality of the image, orientation, and similarity of the stimuli on visual search for faces. *Perception*, *23*, 95–122. [PubMed]
- Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception & Psychophysics*, *63*, 676–697. [PubMed]
- Lu, H. (2010). Structural processing in biological motion perception. *Journal of Vision*, *10*(12):13, 1–13, <http://www.journalofvision.org/content/10/12/13>, doi:10.1167/10.12.13. [PubMed] [Article]
- Lu, H., Tjan, B. S., & Liu, Z. (2006). Shape recognition alters sensitivity in stereoscopic depth discrimination. *Journal of Vision*, *6*(1):7, 75–86, <http://www.journalofvision.org/content/6/1/7>, doi:10.1167/6.1.7. [PubMed] [Article]
- Ma, Y., Paterson, H. M., & Pollick, F. E. (2006). A motion capture library for the study of identity, gender, and emotion perception from biological motion. *Behavior Research Methods*, *38*, 134–141. [PubMed]
- Manera, V., Schouten, B., Becchio, C., Bara, B. G., & Verfaillie, K. (2010). Inferring intentions from biological motion: A stimulus set of point-light communicative interactions. *Behavior Research Methods*, *42*, 168–178. [PubMed]
- Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society of London B: Biological Sciences*, *258*, 273–279.
- Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings, Biological Sciences*, *249*, 149–155. [PubMed]
- Neri, P., Luu, J. Y., & Levi, D. M. (2006). Meaningful interactions can enhance visual discrimination of human agents. *Nature Neuroscience*, *9*, 1186–1192. [PubMed]
- Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, *395*, 894–896. [PubMed]
- Norman, J. F., Payton, S. M., Long, J. R., & Hawkes, L. M. (2004). Aging and the perception of biological motion. *Psychology and Aging*, *19*, 219–225. [PubMed]
- Nothdurft, H. C. (1993). Faces and facial expressions do not pop out. *Perception*, *22*, 1287–1298. [PubMed]
- Pavlova, M., & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, *62*, 889–899. [PubMed]
- Pinto, J., & Shiffrar, M. (1999). Subconfigurations of the human form in the perception of biological motion displays. *Acta Psychologica*, *102*, 293–318. [PubMed]
- Poizner, H., Bellugi, U., & Lutes-Driscoll, V. (1981). Perception of American sign language in dynamic point-light displays. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 430–440. [PubMed]
- Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*, B51–B61. [PubMed]
- Roether, C. L., Omlor, L., Christensen, A., & Giese, M. A. (2009). Critical features for the perception of emotion from gait. *Journal of Vision*, *9*(6):15, 1–32, <http://www.journalofvision.org/content/9/6/15>, doi:10.1167/9.6.15. [PubMed] [Article]
- Rosenholtz, R. (1999). A simple saliency model predicts a number of motion popout phenomena. *Vision Research*, *39*, 3157–3163. [PubMed]
- Shiffrar, M., Lichtey, L., & Heptulla Chatterjee, S. (1997). The perception of biological motion across apertures. *Perception & Psychophysics*, *59*, 51–59. [PubMed]
- Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 809–813. [PubMed]

- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654. [[PubMed](#)]
- Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, *13*, 283–286. [[PubMed](#)]
- Thompson, B., Hansen, B. C., Hess, R. F., & Troje, N. F. (2007). Peripheral vision: Good for biological motion, bad for signal noise segregation? *Journal of Vision*, *7*(10):12, 1–7, <http://www.journalofvision.org/content/7/10/12>, doi:10.1167/7.10.12. [[PubMed](#)] [[Article](#)]
- Thornton, I. A., Pinto, J., & Shiffrar, M. (1998). The visual perception of human locomotion. *Cognitive Neuropsychology*, *15*, 535–552.
- Thornton, I. M., Rensink, R. A., & Shiffrar, M. (2002). Active versus passive processing of biological motion. *Perception*, *31*, 837–853. [[PubMed](#)]
- Thornton, I. M., & Vuong, Q. C. (2004). Incidental processing of biological motion. *Current Biology*, *14*, 1084–1089. [[PubMed](#)]
- Thurman, S. M., & Grossman, E. D. (2008). Temporal “Bubbles” reveal key features for point-light biological motion perception. *Journal of Vision*, *8*(3):28, 1–11, <http://www.journalofvision.org/content/8/3/28>, doi:10.1167/8.3.28. [[PubMed](#)] [[Article](#)]
- Tong, F., & Nakayama, K. (1999). Robust representations for faces: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1016–1035. [[PubMed](#)]
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*, 15–48. [[PubMed](#)]
- Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, *114*, 285–310. [[PubMed](#)]
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136. [[PubMed](#)]
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, *2*(5):2, 371–387, <http://www.journalofvision.org/content/2/5/2>, doi:10.1167/2.5.2. [[PubMed](#)] [[Article](#)]
- Troje, N. F., & Westhoff, C. (2006). The inversion effect in biological motion perception: Evidence for a “life detector”? *Current Biology*, *16*, 821–824. [[PubMed](#)]
- Vangeneugden, J., Pollick, F., & Vogels, R. (2009). Functional differentiation of macaque visual temporal cortical neurons using a parametric action space. *Cerebral Cortex*, *19*, 593–611. [[PubMed](#)]
- Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods, Instruments & Computers: A Journal of the Psychonomic Society, Inc.*, *36*, 625–629. [[PubMed](#)]
- Wang, L., Zhang, K., He, S., & Jiang, Y. (2010). Searching for life motion signals: Visual search asymmetry in local but not global biological-motion processing. *Psychological Science*, *21*, 1083–1089. [[PubMed](#)]
- Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, *56*, 495–500. [[PubMed](#)]
- Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*, 33–39.
- Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics*, *63*, 381–389. [[PubMed](#)]
- Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Cognitive psychology: Rare items often missed in visual searches. *Nature*, *435*, 439–440. [[PubMed](#)]
- Wolfe, J. M., & Van Wert, M. J. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Current Biology*, *20*, 121–124. [[PubMed](#)]