

Matching Biological Motion across Viewpoints

Nicola Ballarini^{1,2}, Ian M. Thornton^{*2}

¹Department of General Psychology, University of Padova, Padova, Italy

²Department of Cognitive Science, Faculty of Media and Knowledge Sciences, University of Malta, Msida, Malta

Abstract. There has been much debate as to how objects can be recognized across viewpoint changes. Here we ask whether viewpoint changes affect performance when participants make judgements about human actions depicted as point-light stimuli. Previous research has suggested that bodies may be “special” objects and may thus be immune to such viewpoint costs. We used a concurrent matching task in which 3 dynamic point-light figures performed familiar actions taken from a standard biological motion database. On each trial the action performed by the central “target” figure was also performed by one of the two flanking figures. The participants task was to make a speeded left/right response to indicate which flanker was copying the target. Separate, random depth orientations were assigned to the two flanking figures and the target could either have the same orientation or appear with an offset of 45° or 90° relative to the matching flanker. The starting animation frame was randomly chosen for each of the three figures. We found that viewpoint differences between the target and matching flanker affected both speed and accuracy. This indicates that the recognition of human bodies depicted as biological motion stimuli is viewpoint-dependent, as with many other types of object. We also suggest that concurrent matching is a flexible tool for exploring biological motion as decisions can be made on a variety of actions without the need for explicit action-naming or training.

Keywords: Biological Motion, Action Understanding, Object Recognition, Viewpoint Dependence, Concurrent Matching, Object Constancy

1 Introduction

Helping us perceive and understand the actions of other people is a primary function of the human visual sys-

tem. Vision allows us to adaptively interact with others in our own social environment and to comprehend the meaning, goals and intentions behind behaviours we see from afar (Johnson & Shiffrar, 2012; Knoblich, Thornton, Grosjean & Shiffrar, 2006). The importance of action understanding is reflected in the wide range of brain areas and brain networks that become activated whenever we watch other people behave (Downing & Peelen, 2011; Giese & Poggio, 2003; Grossman & Blake, 2002; Peuskens, Vanrie, Verfaillie & Orban, 2005; Rizzolatti, Fogassi & Gallese, 2001; Saygin, 2007; Thompson & Parasuraman, 2012).

One particular research area has emphasised how *movement* of our bodies, rather than their form (e.g., size or shape), is important for action understanding. This field of “biological motion” research dates back to the classic work of Gunnar Johansson, who first demonstrated the way in which dynamic point-light figures could be used in an experimental setting (Johansson, 1973, 1976; Marey, 1895, see Fig. 1). Point-light stimuli – where movement is conveyed by the relative motions of a small number of bright markers located on the head and the principal joints (i.e., shoulders, elbows, wrists, hips, knees, ankles) – remains the most popular technique for isolating the dynamic aspects of action and is the method we also use in the current work (Blake & Shiffrar, 2007; for a review of the point-light technique, see Thornton, 2006). What is most salient about these stimuli is the fact that when they are viewed as static images, they appear to the naïve observer only as a collection of random dots. However, when set in motion the underlying human behaviour is immediately revealed.

Johansson’s original stimuli were created by filming actors in low-lighting conditions while they were wearing light sources attached to their joints. Later techniques included computer simulations (Cutting, 1978),

*Correspondence to: Ian M. Thornton (ian.thornton@um.edu.mt)

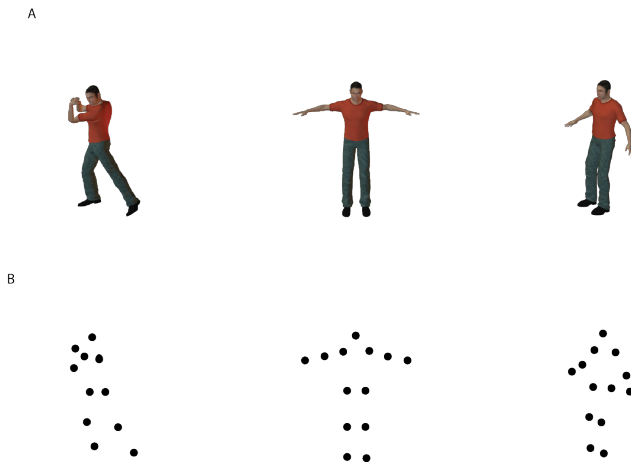


Figure 1: Concurrent Matching Task. A) Three example actions. B) Point-light depictions of the same actions, as shown in the current experiment. On each trial, the task was to decide if the left or right flanking figure performed the same action as the central target, irrespective of difference in depth orientation and action cycle.

identifying joints within each frame of a video sequence (Ahlström, Blake & Ahlström, 1997) and now most commonly purpose-built motion capture systems (Dekeyser, Verfaillie & Vanrie, 2002; Giese & Poggio, 2003; Ilg, Bakir, Mezger & Giese, 2004; Manera, Schouten, Becchio, Bara & Verfaillie, 2010; Troje, 2002; Vanrie & Verfaillie, 2004). This latter technique – made popular by movies such as *The Lord of the Rings* and *Beowulf* – has a number of advantages, but in particular records the position of each point as a 3D coordinate. This means that actions can be replayed from any viewpoint, a feature we exploit in the current work, as described shortly.

Regardless of how the stimuli are created, point-light figures share the characteristic of rapidly conveying a range of information about the underlying action. Not only can naïve participants quickly report that the figure represents a human actor, they can usually identify the action (Dittrich, 1993; Hemeren, 2008), and also extract a range of other characteristics, such as the gender (Kozlowski & Cutting, 1977; Pollick, Kay, Heim & Stringer, 2005) and emotional state (Atkinson, Dittrich, Gemmell & Young, 2004; Dittrich, Troscianko, Lea & Morgan, 1996; Pollick, Paterson, Bruderlin & Sanford, 2001) of the actor. The ability to perceive biological motion with point-light figures appears to be robust against a number of manipulations, such as masking with additional “noise” dots (Bertenthal & Pinto, 1994; Cutting, Moore & Morrison, 1988), spatial and temporal degradation (Beintema & Lappe, 2002; Thornton, Pinto & Shiffrar, 1998; Thurman & Grossman, 2008) and large changes in apparent distance (Thornton, Wootton & Pedmanson, 2014). The rapid and robust processing of

point-light stimuli has led to suggestions that both passive, bottom-up mechanisms (Bosbach, Prinz & Kerzel, 2004; Johansson, 1973, 1976; Mather, Radford & West, 1992; Thornton & Vuong, 2004) and active, top-down mechanism (Bertenthal & Pinto, 1994; Chandrasekaran, Turner, Bühlhoff & Thornton, 2010; Thompson & Parasuraman, 2012; Thornton, Rensink & Shiffrar, 2002) are brought to bear by the visual system in order to solve the problem of action understanding.

In the current paper, there were two main objectives. First, we wanted to further develop a novel matching task for biological motion that we had previously introduced as a new method to study the perception of action at extreme distances (Thornton et al., 2014). On each trial of this task, three point-light figures are presented, one central target and two flanking figures. The two flanking figures always perform different actions, and the target figure copies one of these actions (see Fig. 1). The participant’s task is simply to report whether the left or right flanker matches the target. The task is flexible because the nature of actions displayed (e.g., novel or familiar) and the characteristics of the figures (e.g., relative size, step-cycle, intact or scrambled) can all be manipulated independently without affecting the basic response demands. Participants always simply have to match the left or right flanker.

Here, we show how the task can be easily adapted to answer other biological motion research questions. Specifically, our second objective was to explore the effect that *viewpoint* has on the recognition speed and accuracy of a range of human actions. In the current study, we focus on viewpoint change involving a rotation in depth around the vertical axis. To take walking as an example, rotations in depth would vary whether the figure appears to be moving to the left, right or towards an observer. Within the context of our matching task, in the current study we systematically varied the viewpoint difference between the target and flanking figures.

There have been many previous studies of biological motion that have examined picture plane rotation – turning the figures upside down (e.g., Bertenthal & Pinto, 1994; Hemeren, 2008; Pavlova & Sokolov, 2000, 2003; Sumi, 1984; Troje, 2003). Indeed, perceptual difficulty in processing such “inverted” point-light stimuli – as well as similar manipulations using other forms of body representations (Reed, Stone, Bozova & Tanaka, 2003; Slaughter, Stone & Reed, 2004) – is often cited as one of the main lines of evidence that the human body is “special”, similar to claims for holistic processing of faces (Diamond & Carey, 1986; Rhodes, Brake, Taylor & Tan, 1989; Richler, Mack, Gauthier & Palmeri, 2009; Rossion, 2008; Van Belle, De Graef, Verfaillie, Busigny & Rossion, 2010; Yin, 1969; Young, Hellawell & Hay, 1987). Although some caution has been urged in us-

ing picture plane rotation as a “gold standard” for global processing of biological motion (Pinto & Shiffrar, 2009; Troje & Westhoff, 2006), it remains by far the most common form of viewpoint manipulation. As we note shortly, far fewer studies have examined viewpoint changes in depth, and here we try to link this manipulation more directly to the object recognition literature. Specifically we explore the nature of *object constancy* in the context of human point-light actions other than walking.

Object constancy refers to the ability to recognise an object despite spatial transformations (i.e. changes in orientation, position in the space, size) that give rise to large variations in the image that is projected onto the retina (Graf, 2006; Lawson, 1999). Such variations naturally occur in our everyday experience as we move around the world or objects move relative to us. It is clear that we are able to solve this problem of object constancy as object identity is not usually lost during everyday dynamic interactions. However, there has been considerable debate within the object recognition literature as to how such constancy is achieved.

One class of theories has suggested that recognition is achieved by matching current visual experience to 3D, object-centred representations that are inherently *viewpoint independent* (Biederman, 1987; Marr & Nishihara, 1978). In contrast, another class of theories suggests that our object knowledge consists of stored templates related to previous experience with specific 2D views (Bülthoff & Edelman, 1992; Tarr & Bülthoff, 1995). In order to recognise an object from a specific viewpoint, some sort of internal transformation process is required in order to align the current viewpoint with previously stored *viewpoint dependent* representations. Empirically, the crucial difference between these two theoretical approaches is that viewpoint dependent theories predict there will be a perceptual/behavioural cost when tasks involve viewpoint changes, while *viewpoint independent* theories do not.

There has been relatively little previous research directly addressing the question of object constancy in relation to biological motion processing. Several studies have examined viewpoint dependency in specific relation to walking. Karl Verfaillie for example, used a sequential priming paradigm and found reliable RT advantages in both object (human or non-human decision Verfaillie, 1993) and action (left or right facing decision Verfaillie, 2000) related tasks. Stimuli were always profile views of walking figures and speed of response was found to be faster when the primed and target view were facing in the same direction. Mark Bradshaw and colleagues used a detection task with normal and spatially scrambled walkers embedded in noise masks (Bradshaw, Leach, Hibbard, van der Willigen & Rushton, 1999). They found

detection thresholds were reliably lower when the walker was facing the observer than when oriented at 30°, 60° or 90° away from the line of sight. Troje, Westhoff and Lavrov (2005) used a person identification task with walking patterns and found that performance consistently dropped when different study and test views were used. Interestingly, the same group found some evidence for viewpoint independent processing of walking patterns when they used a more basic person identification task but included the condition where one of the target figures belonged to the observer (Jokisch, Daum & Troje, 2006). They found that the recognition of one’s own walking patterns was viewpoint independent, whereas the recognition of all other walking patterns was more accurate from a frontal and half-profile view than a profile view, consistent with findings of Bradshaw et al. (1999).

Moving beyond point-light walkers, Daems and Verfaillie (1999) used photographic depictions of possible and impossible actions in a priming task. They only found facilitation when prime and target stimuli were physically plausible and had the same in-depth orientation, consistent with the idea of viewpoint dependent action recognition. Two more recent studies, however, have produced somewhat conflicting results with respect to question of viewpoint dependency. de la Rosa, Mieskes, Bülthoff and Curio (2013) examined the ability to recognise dyadic interactions (e.g., handshake, hug) using stick-figure stimuli. They found both reaction time and accuracy costs such that each type of interaction had a preferential viewpoint. In contrast Platonov and Orban (2016) used video sequences of observed manipulative actions (e.g., rolling or rotating an object) and found no viewpoint costs. They specifically suggest that action observation may differ from object recognition in being “for the most part viewpoint-independent” (p. 10).

From the above brief review, it is clear that there is still fairly limited – and somewhat conflicting – evidence concerning the question of viewpoint dependency and action recognition in general (as opposed to walking specifically). In the current paper we restrict ourselves to full-body, point-light stimuli and ask the question of whether there are viewpoint costs – the signature of viewpoint dependent representations – when matching biological patterns other than walking.

2 Method

2.1 Participants

Sixteen observers from the University of Malta took part in the current experiment. All participants had normal or corrected to normal vision and none reported any history of motor impairments. Prior to data collection all participants gave written informed consent although

the precise research question was only explained during debriefing. They were paid €5 for taking part in the single experimental session which lasted approximately 15 minutes. All aspects of current study were reviewed and approved by the Research Ethics Committee of the Faculty of Media and Knowledge Sciences, University of Malta.

2.2 Equipment

Stimuli were displayed and data collected on a Macintosh Mini Computer connected to Fujitsu B24T-7 LCD display. The display had a visible area of 54×30 cm, a resolution of 1280×1024 pixels and a refresh rate of 75 Hz. Custom written code was developed in Matlab, using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). The viewing distance was approximately 90 cm.

2.3 Stimuli

In the current study we used the stimulus display shown in Fig. 1B. On each trial there were always three dynamic point-light figures. The two flanking figures always performed two different actions. The central target copied the action of one of the two flankers. Each figure was composed of 13 white dots drawn on a uniform black background. Note that for illustration purposes, the contrast of figure dots and background have been reversed in Fig. 1B. The 13 dots represent head, shoulders, elbows, wrists, hip, knees and ankles. All dots were always visible, even when they would have ordinarily been occluded by other parts of the body. Each dot subtended approximately 0.22° visual angle. The figures were orthographically projected and aligned so that their centres were at the vertical midline of the screen. Each figure was the same size and subtended approximately 5° visual angle in height. The target figure was positioned at the very centre of the display area and the flankers were offset approximately 8° to the left and right of the target.

Each figure performed a familiar, periodic human action. The 11 actions performed by the figures were chosen from those within the database created by Vanrie and Verfaillie (2004). Specifically, the actions used were: chop, jump, mow, paint, pump, saw, shoot, spade, sweep, tap and wave. Note that walking was specifically not chosen for the current study as previous work had already explored this action in detail. We used all other actions from the database that were periodic and had a basic upright posture. The actions were randomly selected from this set on a trial-by-trial basis. The starting frame of each figure was always determined randomly. The “yaw” rotation around the y -axis (360° orientation range) of the two flanking figures was independently randomised. The orientation of the target was constrained to be offset 0° , 45° or 90° in respect to the orientation of

the matching flanker. Trial order was determined randomly for each participant. Animations were displayed at 30 frames/s, and all of the three actions continued to play until a response was made.

2.4 Task

The task was the same as that used by Thornton et al. (2014). On each trial the participant had to decide which of the two flanking figures was performing the same action as the central target figure. Responses were indicated by pressing designated keys on a standard USB keyboard. They used the right hand to press the “l” key if the right flanker matched the target and the “s” key with the left hand if the left flanker matched the central target.

2.5 Design

Participants worked through two experimental blocks consisting of 90 trials in each block. The first block was preceded by approximately 20 familiarization trials. Each block was composed of 15 repetitions \times 2 matching flanker positions (flanker matching on the right or on the left of the target) \times 3 orientations offsets (0° , 45° or 90° with respect to the target). The 11 actions were not parametrically varied within the design. Rather, on each trial, two actions were selected at random. Trial order was determined separately for each block and participant.

2.6 Procedure

The experiment took place in a silent and dimly lit room. Participants were familiarised with the task and method of responding. They were instructed to respond as quickly and accurately as possible. Trials were separated by an inter-trial interval of 0.5 seconds. Errors were indicated by a visible “Error” message on the screen and a short additional pause of 0.5 seconds. The opportunity to rest was given to participants after the first experimental block. The entire experiment lasted about 15 minutes.

2.7 Data Analysis

Response time and error rates were analysed separately using the same One-Way repeated measures ANOVA. The independent variable, Viewpoint, had three levels (0° , 45° and 90°) reflecting the angular difference between target and flanker. We used planned linear contrasts to explore whether performance degraded in a systematic manner with increasing viewpoint differences. When data violated the assumption of sphericity, as assessed with Mauchly’s test, we applied the Greenhouse-Geisser correction to the degrees of freedom.

3 Results

Fig. 2A shows the pattern of median reaction times for correct responses. There are two points to note. First,

overall responses times are relatively slow in this task. The time needed to compare at least two actions that evolve over time means that participants are not able to make the rapid reactions typical of simple detection or direction discrimination tasks. Second, there is a very clear pattern of increasing reaction time as a function of the variation in viewpoint. Specifically, the time needed to recognize two identical actions increases in proportion to the angular rotation difference between target and flanker. Participants were fastest when target and flanker had the same viewpoint (0°) and increased in the 45° and 90° conditions.

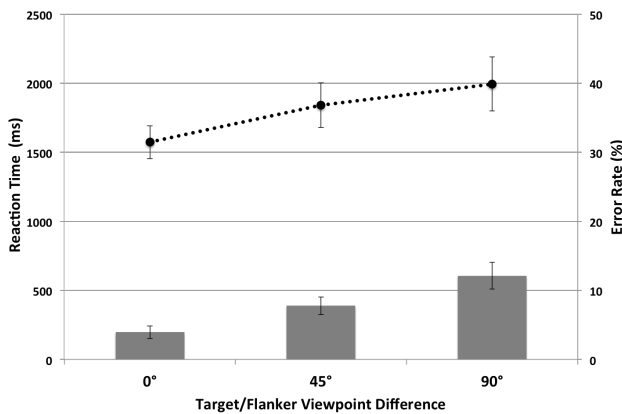


Figure 2: The left axis shows Median Reaction time data as a function of the viewpoint difference between target and flanker in a concurrent matching task. The right axis shows percentage error rates as a function of viewpoint difference. Error bars indicate one standard error of the between-participant means.

Consistent with these patterns there was clear main effect of Viewpoint, $F_{(2,30)} = 10.5$, $MSE = 69345$, $p < 0.001$, $\eta^2_p = 0.41$. Within subjects contrasts indicated that a linear model was the best fitting prediction of the overall pattern seen in Fig. 2A, $F_{(1,15)} = 14.6$, $MSE = 97425$, $p < 0.01$, $\eta^2_p = 0.49$.

Fig. 2B shows the pattern of error data. Again, there are two points to note. First, the overall level of performance was very good, with error rates remaining below 15% in all conditions. This indicates that, as with other familiar objects, participants are able to generalise quite well across viewpoints changes. Second, despite this general level of performance, there were clear costs associated with viewpoint change that mirror those seen with response times. Participants were most accurate when making judgements about actions shown from the same orientation (0°), while performance dropped in the other two conditions (45° and 90°).

There was again a main effect of Viewpoint, $F_{(1.4,20.8)} = 15.18$, $MSE = 0.003$, $p < 0.001$, $\eta^2_p = 0.5$. As with the RT data, within subjects contrasts indicated that a linear increase in error rates was the best fitting

model for the overall pattern of data, $F_{(1,15)} = 23.1$, $MSE = 0.002$, $p < 0.001$, $\eta^2_p = 0.61$.

4 Discussion

In this paper, we examined the perception of point-light actions across viewpoint change using a concurrent matching task. Consistent with previous studies that had used only point-light walkers (Bradshaw et al., 1999; Jokisch et al., 2006; Troje et al., 2005; Verfaillie, 1993, 2000) we found that for a range of actions both reaction time and accuracy performance grew consistently worse as the angular difference between target and matching flanker increased. While the magnitude of these viewpoint costs were relatively small – consistent with the general notion of object constancy for human actions – their presence supports the notion that the representations underlying action understanding – as with both static (e.g., Bühlhoff & Edelman, 1992; Lawson, 1999; Tarr & Bühlhoff, 1995) and dynamic (e.g., Chuang, Vuong & Bühlhoff, 2012; Friedman, Vuong & Spetch, 2010; Vuong, Friedman & Plante, 2009) object recognition – are viewpoint dependent.

The present study also demonstrates how the simple concurrent matching task we had previously developed (Thornton et al., 2014) could be used as a tool for exploring other biological motion questions and, more generally, help shed light on action understanding. As mentioned in the Introduction, we believe this task has advantages over other methods used to examine biological motion – such as direction discrimination and action naming – in that the action content can be changed quite dramatically without the need for familiarization and/or training, as the basic response demands always stay the same.

In general terms, the current task would be classified as an ABX design (Kingdom & Prins, 2016; Macmillan & Creelman, 1991), where A & B are samples, and X is the to-be-matched target. This is a standard psychophysical task, originally introduced for comparing auditory samples (Munson & Gardner, 1950), but one that can be applied to any sensory dimension. Presenting two sample stimuli concurrently, with random key assignments on a trial-by-trial basis reduces the likelihood that systematic response bias will affect the outcome, for example, a tendency to always respond “same” in a Same/Different design. Typically, ABX designs have sequential presentation of each item, but concurrent presentation also has the advantage of reducing memory load and the tendency to prefer the first or second presented sample.

Within the context of biological motion research, we have already used the task to explore distance perception (Thornton et al., 2014). In addition to exploring viewpoint costs, as we do in the current pa-

per, the task could also be used whenever an explicit discrimination is required. For example, it could be used to assess how well people are able to distinguish male and female actors, the speed with which actions are being performed, whether actions are synchronised/desynchronised or even which of the two flanking actions is being performed by the same actor as the target. While many of the questions have been addressed with other methods, we stress again the advantages of the basic ABX design mentioned above, and also note that because the same left/right decision is used regardless of the underlying decision, performance across different types of discrimination could also be compared. Of course, the task may not be appropriate for exploring some issues, for example, in studies that test the ability to process biological motion incidentally (Bosbach et al., 2004; Thornton & Vuong, 2004; Veto, Einhäuser & Troje, 2017, 2013).

Three further points are worth noting about the observed viewpoint costs in the context of current matching. First, as all stimuli were always visible, it suggests that the viewpoint dependent processing we observe with this task relate to immediate perceptual or very short-term/working memory representations. It has long been known that perceptual transformations – such as mental rotation (Shepard & Metzler, 1971) – do occur when participants have to make judgements about concurrently visible stimuli. It is our suggestion that such transformations underlie the current performance deficits. As mentioned in the Introduction, the need to make such transformations is the hallmark of viewpoint dependent processing.

Second, we did not constrain the absolute orientation of the figures. That is, the orientation of the two flanking figures was independently selected at random on each trial from the full 360° options. The target figure orientation was then constrained to be either 0°, 45° or 90° relative to the matching flanker orientation. While this does not allow us to examine possible preferential viewpoints for processing point-light stimuli (Bradshaw et al., 1999; Jokisch et al., 2006), it does suggest that transformation costs generalise and are unlikely to arise just because of problems perceiving point-light figures from specific angles.

Third, our design, was also random with respect to the action that was selected on a given trial. One of our main goals was to extend previous findings beyond simple walking patterns and our design included 11 actions from the Vanrie and Verfaillie (2004) database. In future studies, it would be interesting to systematically examine the costs associated with particular actions, as this might shed further light on the nature of the perceptual transformations involved. Here, we can simply note that an item analysis – collapsing across participants –

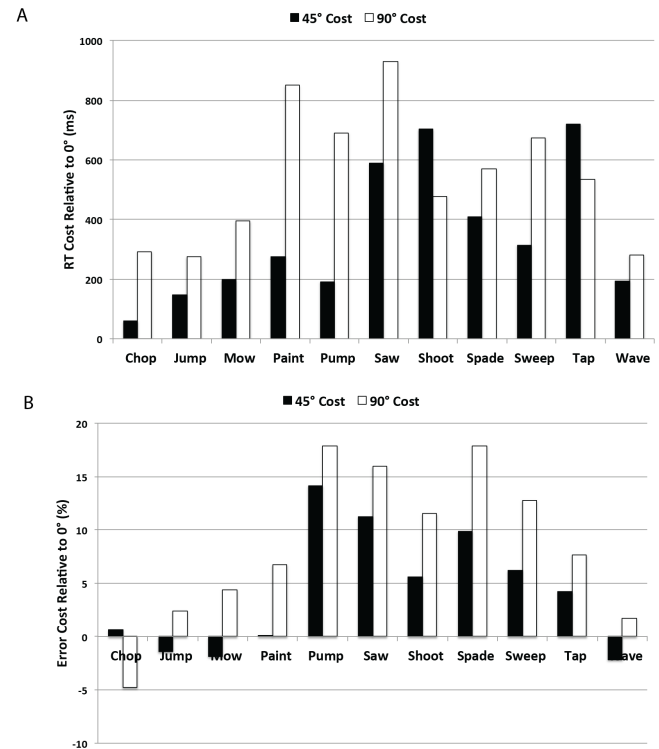


Figure 3: A) Reaction time costs in milliseconds; B) Percentage error rate costs for both the 45° and 90° trials relative to 0° difference. In both panels data have been collapsed across participants and are shown separately for each of the 11 actions.

showed reaction time viewpoint costs for all 11 actions. However, there were considerable differences in the magnitude of such costs as a function of action. Similar variation was seen in an accuracy item analysis, although error rates were higher for the majority (7 out of 11) of actions when the target and flanker had different viewpoints. These cost patterns are illustrated in Fig. 3.

To conclude, it seems highly likely that the perception of full-bodied point light actions is achieved via mechanisms that are viewpoint dependent. Previous studies using walking patterns with a variety of tasks have generally come to the same conclusion. Here, we have generalised this finding across a broader range of actions using concurrent matching. This does not necessarily mean that all forms of action understanding rely on viewpoint dependent representations. Indeed, above we have noted exceptions to this rule in previous studies using both point-light (Jokisch et al., 2006) and video (Platonov & Orban, 2016) stimuli. Although the viewpoint dependent and viewpoint invariant views of object recognition are often thought of as being mutually exclusive a number of authors have pointed out that the visual system is more likely to adopt a flexible approach and use whatever form of representation suits the current stimuli and task (Foster & Gilson, 2002; Vanrie,

Willems & Wagemans, 2001). Nevertheless, our current reading of the literature, together with the results of this experiment lead us to speculate that viewpoint dependent action understanding is the rule, rather than the exception.

References

- Ahlström, V., Blake, R. & Ahlström, U. (1997). Perception of Biological Motion. *Perception*, 26(12), 1539–1548.
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J. & Young, A. W. (2004). Emotion Perception from Dynamic and Static Body Expressions in Point-Light and Full-Light Displays. *Perception*, 33(6), 717–746.
- Beintema, J. A. & Lappe, M. (2002). Perception of biological motion without local image motion. *Proc. Natl. Acad. Sci.* 99(8), 5661–5663.
- Bertenthal, B. I. & Pinto, J. (1994). Global Processing of Biological Motions. *Psychol. Sci.* 5(4), 221–225.
- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* 94(2), 115–147.
- Blake, R. & Shiffrar, M. (2007). Perception of human motion. *Annu. Rev. Psychol.* 58, 47–73.
- Bosbach, S., Prinz, W. & Kerzel, D. (2004). A Simon Effect With Stationary Moving Stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 30(1), 39–55.
- Bradshaw, M. F., Leach, R., Hibbard, P. B., van der Willigen, R. & Rushton, S. K. (1999). The walker's direction affects the perception of biological motion. In M. A. Grealy & J. A. Thomson (Eds.), *Stud. Percept. action V/Tenth International Conference Perception Action Aug. 8–13, 1999, Edinburgh, Scotland* (pp. 3–6). Mahwah, NJ: L. Erlbaum Associates.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spat. Vis.* 10, 433–436.
- Bülthoff, H. H. & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proc. Natl. Acad. Sci.* 89(1), 60–64.
- Chandrasekaran, C., Turner, L., Bülthoff, H. H. & Thornton, I. M. (2010). Attentional networks and biological motion. *Psihologija*, 43(1), 5–20.
- Chuang, L. L., Vuong, Q. C. & Bülthoff, H. H. (2012). Learned Non-Rigid Object Motion is a View-Invariant Cue to Recognizing Novel Objects. *Front. Comput. Neurosci.* 6.
- Cutting, J. E. (1978). A program to generate synthetic walkers as dynamic point-light displays. *Behav. Res. Methods Instrum.* 10(1), 91–94.
- Cutting, J. E., Moore, C. & Morrison, R. (1988). Masking the motions of human gait. *Percept. Psychophys.* 44(4), 339–347.
- Daems, A. & Verfaillie, K. (1999). Viewpoint-dependent Priming Effects in the Perception of Human Actions and Body Postures. *Vis. cogn.* 6(6), 665–693.
- de la Rosa, S., Mieskes, S., Bülthoff, H. & Curio, C. (2013). View dependencies in the visual recognition of social interactions. *Front. Psychol.* 4, 752.
- Dekeyser, M., Verfaillie, K. & Vanrie, J. (2002). Creating stimuli for the study of biological-motion perception. *Behav. Res. Methods, Instruments, Comput.* 34(3), 375–382.
- Diamond, R. & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *J. Exp. Psychol. Gen.* 115(2), 107–117.
- Dittrich, W. H. (1993). Action Categories and the Perception of Biological Motion. *Perception*, 22(1), 15–22.
- Dittrich, W. H., Troscianko, T., Lea, S. E. G. & Morgan, D. (1996). Perception of Emotion from Dynamic Point-Light Displays Represented in Dance. *Perception*, 25(6), 727–738.
- Downing, P. E. & Peelen, M. V. (2011). The role of occipitotemporal body-selective regions in person perception. *Cogn. Neurosci.* 2(3-4), 186–203.
- Foster, D. H. & Gilson, S. J. (2002). Recognizing novel three-dimensional objects by summing signals from parts and views. *Proc. R. Soc. B Biol. Sci.* 269(1503), 1939–1947.
- Friedman, A., Vuong, Q. C. & Spetch, M. (2010). Facilitation by view combination and coherent motion in dynamic object recognition. *Vision Res.* 50(2), 202–210.
- Giese, M. A. & Poggio, T. (2003). Cognitive neuroscience: Neural mechanisms for the recognition of biological movements. *Nat. Rev. Neurosci.* 4(3), 179–192.
- Graf, M. (2006). Coordinate transformations in object recognition. *Psychol. Bull.* 132(6), 920–945.
- Grossman, E. D. & Blake, R. (2002). Brain Areas Active during Visual Perception of Biological Motion. *Neuron*, 35(6), 1167–1175.
- Hemeren, P. E. (2008). *Mind in action: action representation and the perception of biological motion*. Lund University cognitive studies. Lund: Univ.
- Ilg, W., Bakir, G. H., Mezger, J. & Giese, M. A. (2004). On the representation, learning and transfer of spatio-temporal movement characteristics. *Int. J. Humanoid Robot.* 1(04), 613–636.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Percept. Psychophys.* 14(2), 201–211.
- Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychol. Res.* 38(4), 379–393.

- Johnson, K. & Shiffrar, M. (2012). *People watching: Social, perceptual, and neurophysiological studies of body perception*. Oxford University Press.
- Jokisch, D., Daum, I. & Troje, N. F. (2006). Self Recognition versus Recognition of others by Biological Motion: Viewpoint-Dependent Effects. *Perception*, *35*(7), 911–920.
- Kingdom, F. A. A. & Prins, N. (2016). *Psychophysics: a practical introduction* (Second edi). Amsterdam: Elsevier/Academic Press.
- Kleiner, M., Brainard, D. H., Pelli, D. G., Ingling, A., Murray, R., Broussard, C. et al. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.
- Knoblich, G., Thornton, I. M., Grosjean, M. & Shiffrar, M. (Eds.). (2006). *Human body perception from the inside out*. Advances in visual cognition. Oxford ; New York: Oxford University Press.
- Kozlowski, L. T. & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Percept. Psychophys.* *21*(6), 575–580.
- Lawson, R. (1999). Achieving visual object constancy across plane rotation and depth rotation. *Acta Psychol. (Amst)*. *102*(2–3), 221–245.
- Macmillan, N. A. & Creelman, C. D. (1991). *Detection theory: a user's guide*. Cambridge [England]; New York: Cambridge University Press.
- 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. *Behav. Res. Methods*, *42*(1), 168–178.
- Marey, É.-J. (1895). *Movement: The Results and Possibilities of Photography* (tr. E. Pritchard). London: William Heinemann. [French original: 1894. *Le Mouvement*. Paris: Masson].
- Marr, D. & Nishihara, H. K. (1978). Representation and Recognition of the Spatial Organization of Three-Dimensional Shapes. *Proc. R. Soc. B Biol. Sci.* *200*(1140), 269–294.
- Mather, G., Radford, K. & West, S. (1992). Low-Level Visual Processing of Biological Motion. *Proc. R. Soc. B Biol. Sci.* *249*(1325), 149–155.
- Munson, W. A. & Gardner, M. B. (1950). Standardizing Auditory Tests. *J. Acoust. Soc. Am.* *22*(5), 675.
- Pavlova, M. & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Percept. Psychophys.* *62*(5), 889–899.
- Pavlova, M. & Sokolov, A. (2003). Prior Knowledge about Display Inversion in Biological Motion Perception. *Perception*, *32*(8), 937–946.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* *10*(4), 437–442.
- Peuskens, H., Vanrie, J., Verfaillie, K. & Orban, G. A. (2005). Specificity of regions processing biological motion. *Eur. J. Neurosci.* *21*(10), 2864–2875.
- Pinto, J. & Shiffrar, M. (2009). The visual perception of human and animal motion in point-light displays. *Soc. Neurosci.* *4*(4), 332–346.
- Platonov, A. & Orban, G. A. (2016). Action observation: the less-explored part of higher-order vision. *Sci. Rep.* *6*(1).
- Pollick, F. E., Kay, J. W., Heim, K. & Stringer, R. (2005). Gender Recognition From Point-Light Walkers. *J. Exp. Psychol. Hum. Percept. Perform.* *31*(6), 1247–1265.
- Pollick, F. E., Paterson, H. M., Bruderlin, A. & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*(2), B51–B61.
- Reed, C. L., Stone, V. E., Bozova, S. & Tanaka, J. (2003). The Body-Inversion Effect. *Psychol. Sci.* *14*(4), 302–308.
- Rhodes, G., Brake, S., Taylor, K. & Tan, S. (1989). Expertise and configural coding in face recognition. *Br. J. Psychol.* *80*(3), 313–331.
- Richler, J. J., Mack, M. L., Gauthier, I. & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. *Vision Res.* *49*(23), 2856–2861.
- Rizzolatti, G., Fogassi, L. & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci*, *2*(9), 661–670.
- Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychol. (Amst)*. *128*(2), 274–289.
- Saygin, A. P. (2007). Superior temporal and premotor brain areas necessary for biological motion perception. *Brain*, *130*(9), 2452.
- Shepard, R. N. & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. *Science* (80-). *171*(3972), 701–703.
- Slaughter, V., Stone, V. E. & Reed, C. (2004). Perception of Faces and Bodies: Similar or Different? *Curr. Dir. Psychol. Sci.* *13*(6), 219–223.
- Sumi, S. (1984). Upside-down Presentation of the Johansson Moving Light-Spot Pattern. *Perception*, *13*(3), 283–286.
- Tarr, M. J. & Bühlhoff, H. H. (1995). Is human object recognition better described by geon structural descriptions or by multiple views? Comment on Biederman and Gerhardstein (1993). *J. Exp. Psychol. Hum. Percept. Perform.* *21*(6), 1494–1505.
- Thompson, J. & Parasuraman, R. (2012). Attention, biological motion, and action recognition. *Neuroimage*, *59*(1), 4–13.

- Thornton, I. M. (2006). Biological Motion: Point-Light Walkers and Beyond. In G. Knoblich, I. M. Thornton, M. Grosjean & M. Shiffrar (Eds.), *Hum. body percept. from insid. out* (pp. 271–303). Advances in visual cognition. Oxford ; New York: Oxford University Press.
- Thornton, I. M., Pinto, J. & Shiffrar, M. (1998). The Visual Perception of Human Locomotion. *Cogn. Neuropsychol.* 1998(15), 535–552.
- Thornton, I. M., Rensink, R. A. & Shiffrar, M. (2002). Active versus Passive Processing of Biological Motion. *Perception*, 31(7), 837–853.
- Thornton, I. M. & Vuong, Q. C. (2004). Incidental Processing of Biological Motion. *Curr. Biol.* 14(12), 1084–1089.
- Thornton, I. M., Wootton, Z. & Pedmanson, P. (2014). Matching biological motion at extreme distances. *J. Vis.* 14(3), 13.
- Thurman, S. M. & Grossman, E. D. (2008). Temporal “Bubbles” reveal key features for point-light biological motion perception. *Journal of Vision*, 8(3), 28.
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *J. Vis.* 2(5), 2.
- Troje, N. F. (2003). Reference Frames for Orientation Anisotropies in Face Recognition and Biological-Motion Perception. *Perception*, 32(2), 201–210.
- Troje, N. F. & Westhoff, C. (2006). The Inversion Effect in Biological Motion Perception: Evidence for a “Life Detector”? *Curr. Biol.* 16(8), 821–824.
- Troje, N. F., Westhoff, C. & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Percept. Psychophys.* 67(4), 667–675.
- Van Belle, G., De Graef, P., Verfaillie, K., Busigny, T. & Rossion, B. (2010). Whole not hole: Expert face recognition requires holistic perception. *Neuropsychologia*, 48(9), 2620–2629.
- Vanrie, J. & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behav. Res. Methods, Instruments, Comput.* 36(4), 625–629.
- Vanrie, J., Willems, B. & Wagemans, J. (2001). Multiple Routes to Object Matching from Different Viewpoints: Mental Rotation versus Invariant Features. *Perception*, 30(9), 1047–1056.
- Verfaillie, K. (1993). Orientation-dependent priming effects in the perception of biological motion. *J. Exp. Psychol. Hum. Percept. Perform.* 19(5), 992–1013.
- Verfaillie, K. (2000). Perceiving Human Locomotion: Priming Effects in Direction Discrimination. *Brain Cogn.* 44(2), 192–213.
- Veto, P., Einhäuser, W. & Troje, N. F. (2017). Biological motion distorts size perception. *Sci. Rep.* 7, 42576.
- Veto, P., Thill, S. & Hemeren, P. (2013). Incidental and non-incidental processing of biological motion Orientation, attention and life detection. In *Coop. minds social interaction group dynamics proceedings 35th annual meeting cognitive science society berlin, germany, july 31–august 3, 2013*. Cognitive Science Society, Inc.
- Vuong, Q. C., Friedman, A. & Plante, C. (2009). Modulation of Viewpoint Effects in Object Recognition by Shape and Motion Cues. *Perception*, 38(11), 1628–1648.
- Yin, R. K. (1969). Looking at upside-down faces. *J. Exp. Psychol.* 81(1), 141–145.
- Young, A. W., Hellawell, D. & Hay, D. C. (1987). Configurational Information in Face Perception. *Perception*, 16(6), 747–759.