Orientation Specific Effects of Automatic Access to Categorical Information in Biological Motion Perception

Paul E. Hemeren (paul.hemeren@his.se)
University of Skövde, School of Humanities and Informatics
Box 408, 541 28, Skövde, Sweden, and
Lund University Cognitive Science
Kungshuset, Lundagård, 222 22, Lund, Sweden

Abstract

Previous findings from studies of biological motion perception suggest that access to stored high-level knowledge about action categories contributes to the fast identification of actions depicted in point-light displays of biological motion. Three priming experiments were conducted to investigate the automatic access to stored categorical level information in the visual processing of biological motion and the extent to which this access varies as a function of action orientation. The results show that activation of categorical level information occurs even when participants are given a task that does not require access to the categorical nature of the actions depicted in point-light displays. The results suggest that the visual processing of upright actions is indicative of Hochstein and Ahissar's notion of vision at a glance, whereas inverted actions indicate vision with scrutiny.

[Key words: visual processing, biological motion, view-dependence, conceptual knowledge]

Introduction

Gunnar Johansson's (1973) point-light technique reveals the sensitivity of human vision to biological motion. Depicted actions in the point-light displays seem to pop out as soon as motion is perceived among the point-light elements. When presented in an upright orientation, we appear to have phenomenally direct access to the high-level categorical nature of the motion pattern. When presented with point-light displays, we see actions, not just local motion patterns. An effective method of disturbing action recognition is to show the actions upside-down, i.e., to invert them. Despite having the same structural hierarchy among elements as in upright actions, inversion leads to severely reduced performance in tasks of identification, detection and recognition, as well as long-term priming (Dittrich, 1993; Bertenthal & Pinto, 1994; Pavlova & Sokolov, 2000; Pinto & Shiffrar, 1999; Ahlström, Blake & Ahlström, 1997 to name a few).

The difference in behavioral results for the visual processing of upright and inverted displays suggests differential access to stored high-level representations. More specifically, findings from experiments on biological motion perception indicate the following differences in the visual processing of upright and inverted displays. For upright displays, visual processing:

- is fast and "automatic" (indicates pop-out) (Jokisch, Daum, Suchan & Troje, 2005; Giese & Poggio, 2003);
- involves high-level global processing mechanisms (Bertenthal & Pinto, 1994; Shiffrar, Lichtey, Heptulla Chatterjee, 1997);
- involves access to categorical level information (Dittrich, 1993, Pinto & Shiffrar, 1999);
- requires attention (Thornton, Rensink & Shiffrar, 2002). (But see Giese and Poggio (2003) for a different view.)

Concerning the relationship between global processing mechanisms and access to categorical level information, results from object recognition indicate that access to the global shape of objects automatically activates identification (Boucart & Humphreys, 1992). Therefore, to the extent that displays of biological motion represent dynamic objects, information about the categorical nature of the depicted actions may be automatically accessed if visual processing occurs on a global level.

In relation to the factors characterizing the processing of upright actions and for the purpose of the work presented here, there is evidence to suggest that the visual processing of inverted actions:

- is slower and indicates less (if any) pop-out (Dittrich, 1993; Pavlova & Sokolov, 2000);
- impairs accurate high-level global processing and appears to rely more on local motion processing (Pavlova & Sokolov, 2003; Pinto & Shiffrar, 1999);
- impairs access to categorical level information (Pinto & Shiffrar, 1999).
- There is no specific data on the role of attention in the visual processing of inverted displays.

Hochstein and Ahissar's (2002) Reverse Hierarchy Theory (RHT) provides a plausible theoretical framework from which to understand and investigate the relationship between the factors characterizing the processing differences between upright and inverted actions. According to RHT explicit high-level visual processing involves initial feedforward mechanisms that implicitly follow a bottom-up hierarchical pathway. The end product of this processing, and the beginning of explicit visual perception, is conscious access to perceptual content in high-level cortical areas. The further claim is that explicit high-level perception is holistic and is where basic-level category judgments are made. This ability to quickly make basic-level category judgments is termed 'vision at a glance.' Regarding the role of access to global level representations, Hochstein and Ahissar assert, "…the whole is perceived first due to explicit perception initially accessing only high-level representations." (p. 796)

It should also be mentioned that according to Hochstein and Ahissar explicit visual perception requires spread attention.

Findings from the perception of inverted actions, on the other hand, suggest a greater role for 'vision with scrutiny.' According to RHT, vision with scrutiny involves the operation of top-down feedback mechanisms in making fine perceptual discriminations and in the veridical binding of local features. In line with RHT, if the initial explicit processing of inverted biological motion fails to access a stored global representation, then top-down feedback mechanisms will be needed to re-bind local motion elements in order to create a global pattern that more veridically matches a stored representation. This may be one important reason why inverted displays take longer to process.

In relation to biological motion perception previous results suggest that perception of upright actions is indicative of vision at a glance (Jokisch et al., 2005). Visual processing of upright actions of biological motion indicates access to high-level representations that automatically activate categorical information about the depicted action.

If categorical level information is automatically activated as a result of access to global level representations, then we should see categorical level effects during a task that does not explicitly require categorical processing.

Three priming experiments were performed to investigate the potential differential access to categorical level information in upright and inverted actions. Since the activation of categorical level information for upright actions appears to be greater compared to inverted actions, viewing upright actions should lead to greater facilitation for the categorical processing of actions viewed later.

Experiment 1

This experiment used short-term (repetition) priming to assess the extent to which automatic access to categorical level information facilitates the visual processing of upright and inverted actions and the extent to which potential facilitation varies as a function of orientation congruence.

To assess automatic access to categorical level information, an orientation decision task was used. This task simply required subjects to indicate the orientation of the presented point-light action. It seems quite unlikely that categorical information about action categories is needed to make an orientation decision. Three different actions were also used to assess potential categorical effects on priming.

For orientation congruent transitions, categorical level effects will be demonstrated if priming varies as function of type of action. If there is greater priming when an action primes itself compared to when it is primed by a different action (action congruency), this would indicate that subjects are able to distinguish between the different actions. Based on the reasoning presented previously, the prediction here is that facilitation due to categorical level information should be greater for upright than for inverted actions that are orientation congruent. The effect of categorical level information for orientation incongruent transitions may also differ depending on the orientation of the priming action as previously mentioned.

Method

Participants: Eight students (4 f, 4 m) from the University of Skövde participated in the experiment. (Mean age: 22yrs.)

Materials and procedure: Point-light displays of a person walking, doing jumping jacks and climbing a rope were used in the experiment. Each motion sequence contained 26 frames. Frame display rate was set to 20 fps, resulting in a display duration of 1.3 seconds for each action.

Point-light displays were presented upright and inverted. Two neutral stimuli of nonhuman form, based on the previous work of Verfaillie (1993), were used to establish a neutral baseline for potential priming effects. To prevent participants from recognizing the actions by only perceiving the first frames, each action started randomly at one of three different points in the action cycle.

Each action type served as both priming and primed stimulus for each of the other actions, resulting in 9 different action transitions (action congruence transitions). There were also four orientation transitions for the priming and primed stimuli: upright-upright, inverted-inverted, upright-inverted and inverted-upright. The combination of these transition types results in 36 total priming transitions.

The stimuli were displayed on a Macintosh 17" monitor set to black and white color with a resolution of 832 x 624 pixels and a refresh rate of 75Hz. Stimulus presentation was controlled by a Macintosh 7100/66AV (66 MHz) and specially written software (DotPlayer).

Participants participated individually in 5 sessions distributed over a period of 6 days. Each of the 5 sessions was divided into two sub sessions, and each sub session contained 6 blocks of trials. One block contained 144 trials where each action, orientation and neutral display occurred 18 times. Each subsession therefore contained 864 trials, and each session contained 1728 trials. A total of 8640 trials was completed by each participant during the 5 sessions. Trials within blocks and block order within each subsession were randomized for each participant.

Participants were told they would view 5 different patch-light sequences in random order and that 3 of the sequences represented actions performed by a human actor. Each action would be presented upright and inverted. In addition to the 3 human actions, 2 abstract, or neutral, patch-light sequences would also be presented. These sequences would be presented many times throughout the five sessions. When presented with a patch-light action, participants were instructed to simply indicate whether they thought the action was upright or inverted by pressing assigned keys (counter-balanced across participants). Both speed and accuracy were emphasized and no response feedback was given to the participants.

Participants sat in a dimly lit room with a viewing distance of 70cm to the computer screen. The stimuli were presented in the center of the computer screen. Following a response indicating the orientation of the motion, a response-stimulus interval (RSI) of 500 occurred. The time (ms) to make an orientation decision was recorded for each action.

Stimulus configurations appeared one at a time in a random order. Each motion was viewable for up to a maximum of 1.3 seconds, after which followed a blank screen. A participant response that occurred before the end of a sequence terminated the sequence and started the RSI.

Results and discussion

Mean reaction time for correct orientation decisions for upright actions was 409ms. For inverted actions, the mean was 436ms. The difference was significant, $F(1,7) = 16.67$, $p < .01$.

Twelve neutral baselines were calculated to reflect the independent effects of different key presses, orientation and the three different actions. A sorting program identified relevant transitions. The neutral baselines were then used to calculate priming effects. The priming results are presented in Figure 1.

Panel A shows the results for upright-upright transitions. In relation to the neutral baseline, all priming actions led to facilitation for the orientation decision in all 9 transitions. Facilitation due to priming was significant as indicated by the 95% confidence intervals. More importantly, in relation to the prediction, there was greater facilitation when an action primed itself compared to when it was primed by other actions, i.e., action congruency effect. This was the case for 5 of the 6 action incongruent priming transitions, indicated by the star (*). This suggests automatic access to the categorical nature of the different actions.

Panel B shows the results for inverted-inverted transitions. The results appear to be similar to the results in Panel A. In relation to the neutral baseline, viewing any of the three actions as an inverted prime led to priming for the orientation decision. There was significant priming in all 9 conditions. A planned comparison between priming levels for Panel A and B revealed no significant difference, $F(1,7) = 2.93$, $p > .10$. This appears to contradict the prediction that access to categorical level information should be greatly reduced for the inverted actions. (This will be discussed below.) However, the effect of action congruency was less pronounced for the inverted-inverted transitions in Panel B. A significant difference between levels of priming when an action primed itself and when it was primed by a different action was found only for 1 of the 6 action incongruent transitions, indicating reduced categorical processing.

Panel C displays the priming effects for orientation incongruent transitions where upright actions act as primes for inverted actions. In relation to the neutral primes, viewing an upright prime facilitated an orientation decision response for inverted actions in 7 of the 9 transitions, indicating that participants have access to categorical information that distinguishes human motion from the neutral displays. Looking at the bars in Panel C shows however that there are no differences between priming levels when an action primed itself and when it was primed by a different action.

The results in Panel D show no facilitation for the orientation decision task when an inverted action precedes an upright action. There is no significant priming for any of the transitions. This stands in stark contrast to the results in Panel A (upright-upright transitions) where an upright action prime leads to a level of activation that facilitates the response to the orientation decision, which also leads to effects of action congruence. In contrast, results in Panel D show that an inverted action prime is no more effective at facilitating a response to the orientation decision than the neutral displays. This clearly indicates a difference in access to information conveyed by upright and inverted displays. This difference is also strengthened by the significant priming effects when upright actions acted as primes for inverted displays. This clearly shows that effects of orientation congruence depend on the orientation of the priming action. This interaction was significant, $F(1,7) = 10.17$, $p < .05$.

In summary, the data presented in Figure 1 show two significant results that support the prediction of different levels of automatic activation of categorical information for upright and inverted actions. The first result was the significant main effect of action congruence, $F(2,14) = 12.73$, $p < .01$. There was greater facilitation when an action primed itself than when it was primed by a different action. The second result is seen in the significant interaction between the orientation of the priming action and action congruence, $F(2,14) = 6.58$, $p < .05$. When a priming action was presented upright there was significantly more priming when an action primed itself than when it was primed by different actions. This however was not the case for inverted actions.

Contrary to predictions, the significant priming effects in Panel B indicate that inverted actions can prime inverted actions and even automatically activate categorical level information to some extent. Since participants were exposed to hundreds of trials of inverted actions, they may have developed representational and processing resources that facilitated the orientation decisions for inverted actions. Although this learning aided orientation decisions for inverted actions, it did not lead to any advantages when inverted actions preceded upright actions.

In this experiment, effects of automatic access to categorical information were evaluated by a task that did not explicitly require access to the categorical nature of the actions. Evidence for categorical effects is admittedly based on a processing assumption of implicit access to categorical level information. The results suggest that access to this information plays a role in biological motion perception. A more direct test of this assumption would be to instruct participants to identify the action, which is a task that explicitly requires access to categorical level information. This was done in the following two experiments.

Experiments 2 and 3

The purpose of these experiments was to directly assess the role of categorical information in biological motion perception by using an action identification task. If action identification is facilitated by previous exposure to the actions, this would indicate that the priming action activates categorical level information. Consistent with previous predictions, upright-upright transitions should lead to significant priming. Furthermore, if priming in upright-inverted transitions is due to access to categorical level

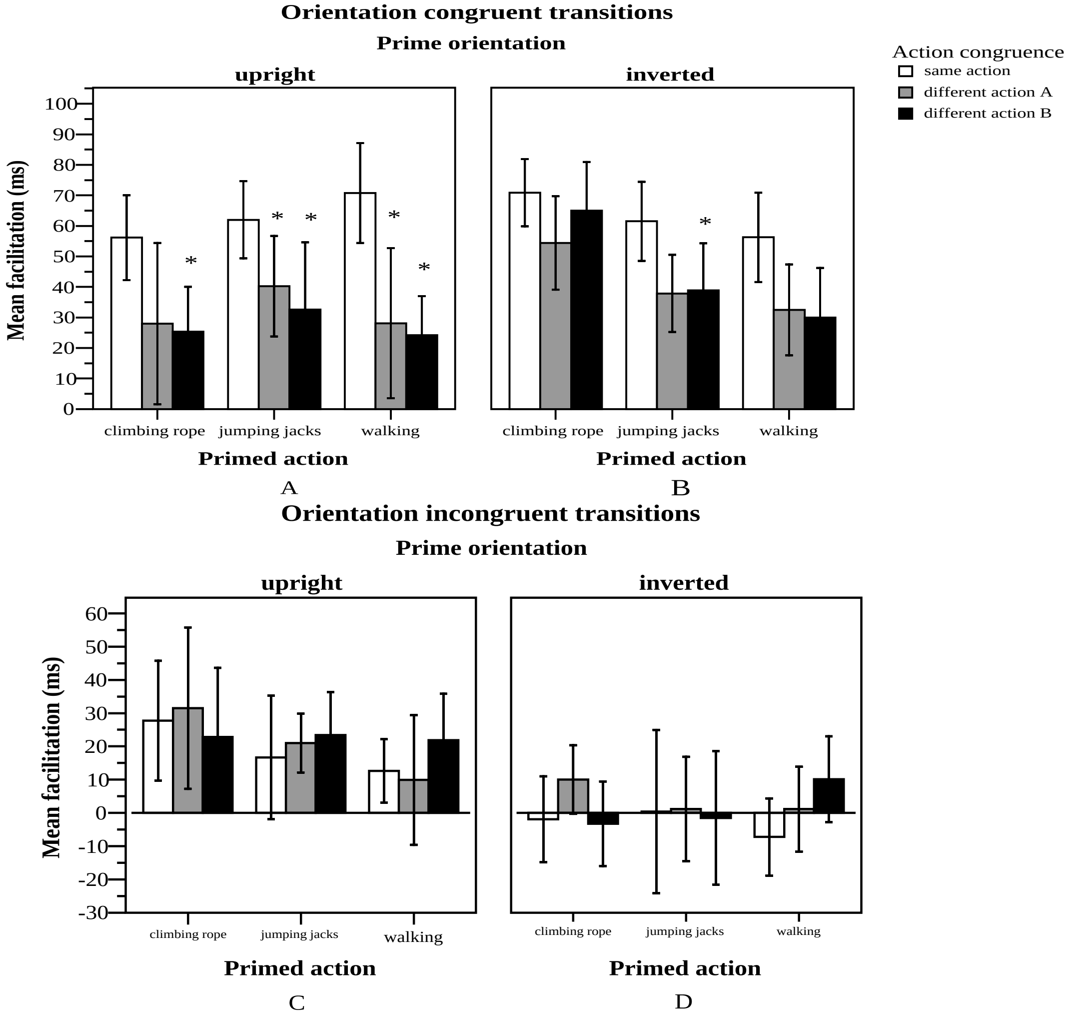

Figure 1: Mean facilitation effect for the orientation decision as a function of orientation congruence, prime orientation, primed action and action congruence. Error bars represent 95% confidence intervals. A star (*) represents a significant Bonferroni adjusted priming level difference between an action priming itself and when it was primed by a different action.

information, then the results for this transition in exp. 1 should be replicated.

An additional prediction concerns the priming effects for inverted-inverted transitions. If the priming effects for these transitions resulted from extensive exposure to inverted actions, then those effects should be significantly reduced if participants are only given one exposure to inverted primes.

A long-term priming paradigm was used in these experiments to evaluate the potential long-term effects of priming in contrast to the repetition priming used in exp. 1.

Method

Participants: Forty adults (20 f, 20 m) (mean age = 27yrs) participated in exp. 2 and 40 adults (21 f, 19 m) (mean age = 32 yrs) participated in exp. 3.

Materials and procedure: Forty different point-light actions were used in this experiment, e.g., walking, jumping, running, kicking, throwing, sawing, etc. The stimuli were presented on a television monitor.

The 40 actions were randomly divided into two sets, 20 actions in each set. The two sets were used in the study and test phases of the experiment. The study phase consisted of 20 actions. The test phase, on the other hand, contained all 40 actions, including the 20 actions shown during the study phase and 20 new actions not previously viewed. The purpose of the 20 new actions was to establish a baseline from which to evaluate potential priming effects. Study and test action sets were rotated across participants. Each set of actions was further divided into two subsets where one subset was presented in an upright orientation, and the other subset was presented inverted. Subsets were counterbalanced across study sets. All actions appeared an equal number of times in the study and test sets as well as in upright and inverted orientations.

Participants were instructed to attend to the monitor and view the actions as they appeared. Participants did not respond verbally to the actions. Each action was presented for 5 seconds. The actions were masked with a static checker board pattern to reduce ceiling effects. Half of the actions were presented upright, and the other half was inverted. Prior to the study phase, participants received practice trials to acquaint them to the masked point-light actions. The practice actions were not included in any further phases of the experiment.

After the study phase (between 5 and 10 minutes), participants were individually tested on an identification task (test phase) where they viewed a total of 40 actions. Twenty actions were previously presented in the study phase, and 20 actions were new. In the test phase, all actions had the same orientation. This provided the test of orientation congruence. So, 10 'old' actions presented in the study and test phases had the same orientation (orientation congruence) and 10 'old' actions were orientation incongruent between study and test. The 20 new actions served as a baseline for possible priming effects. For half of the participants, the 20 new actions were upright and for the other half, the new actions were inverted. Participants were randomly assigned to these conditions. Participants verbally identified the actions as soon as they could. The experimenter wrote down the responses.

The materials and procedure from exp. 2 were also used in exp. 3 with the exception of viewing time. Instead of viewing the displays for 5 seconds in the study and test phases in exp. 2, display duration was reduced to 1 second because of an apparent ceiling effect for upright-upright transitions in exp. 2.

Results and discussion

Results for exp. 2 and 3 for correct percent identification for the different priming transitions are presented in Table 1.

The means show differences between identification performance for upright and inverted actions not previously viewed in the study phase, i.e., new actions, 85% vs. 60% for exp. 2, $t(38) = 4.91$, $p < .001$, and 44% vs. 19% for exp. 3, $t(38) = 4.68$, $p < .001$. These results confirm the general finding of orientation specificity in the literature.

For the priming results, a comparison of upright-upright transitions between study and test (88% vs. 85%) did not reach significance in exp. 2. This was largely due to a ceiling effect as indicated by the significant difference for the same transition in exp. 3, 63% vs. 44 %, $t(19) = 3.28$, $p < .01$. This shows that there is significant facilitation for upright displays when the task requires access to categorical level information, i.e., identification of specific actions.

Table 1: Mean proportions of correctly identified actions for exp. 2 & 3. Transitions indicate display orientation for study-test phases. UPR=upright and INV=inverted. 'Old actions' indicate actions presented in study and test. 'New actions' indicate actions presented only in the test phase.

	Old Actions				New Actions	
Transition	UPR-UPR	INV-INV	UPR-INV	INV-UPR	UPR	INV
Exp. 2	.88	.62	.73	.84	.85	.60
Exp. 3	.63	.26	.27	.49	.44	.19

Results for the inverted-inverted transitions show no significant priming, exp. 2 62% vs. 60%, $t(19) = .42$, $p > .60$, exp. 3, 26% vs. 19%, $t(19) = 1.37$, $p > .18$. The results from exp. 1 did not replicate in these experiments. This suggests that inverted actions, in contrast to upright actions, do not sufficiently activate categorical level information such that later categorical processing is facilitated.

Results for the upright-inverted priming transitions in exp. 2 indicate that upright actions facilitate the later identification of inverted displays, 73% vs. 60%, $t(19) = 2.90$, $p < .01$. This result is similar to the short-term priming effect in exp. 1 (Panel C). When participants are given 5 seconds to view the upright actions in the study phase, identification of inverted actions significantly improves. However, this effect did not hold for the same transition in exp. 3. With a display duration of 1 second, there was no significant difference, 27% vs. 19%, $t(19) = 1.53$, $p > .10$.

Finally, there was no reliable long-term priming for the inverted-upright transitions in either experiment, exp. 2, 85% vs. 84%, exp. 3, 49% vs. 44%, $t(19) = 1.21$, $p > .20$. This result is consistent with the results from exp. 1 showing no significant priming for this transition. Viewing an inverted action has no effect on later identification of the same action presented in an upright orientation.

General Discussion

Previous findings suggest that high-level access to categorical level information for inverted actions is impaired relative to upright actions. The further claim here is that this access in turn is a result of limited global level processing for inverted displays. Boucart & Humphreys (1992) showed this to be the case for static objects. Two major predictions were formulated to evaluate the role of categorical level information in biological motion perception. First, if access to categorical information is automatic then we should see categorical effects in priming for a task that does not require access to categorical level information. Secondly, if this access is greater for upright than for inverted actions, we should see greater effects of action congruence for upright than for inverted actions. These predictions were confirmed in exp. 1.

Experiments 2 & 3 directly assessed the contribution of categorical level information by explicitly asking participants to identify actions. The results from these experiments confirmed the findings from exp. 1. Participants were significantly better at identifying actions when they were previously exposed to upright actions than when they previously viewed inverted actions (with the exception of upright-inverted transitions in exp. 3). There were no significant priming effects with inverted action primes.

As mentioned previously, the significant priming effects for inverted-inverted transitions in exp. 1 may be due to learning to process inverted actions over the course of many trials. Palmeri and Gauthier (2004) suggest that object identity may be automatically activated by expertise. As people become experienced at visually discriminating objects, access to knowledge mediating identification becomes more automatic. The differences in priming effects for inverted primes between exp. 1 and exps. 2 & 3 are consistent with this suggestion. Admittedly, the extensive exposure to inverted actions in exp. 1 and the one-time viewing occasion in exps. 2 & 3 were confounded with priming paradigm. Further analyses and research needs to be done to systematically evaluate the effects of these two factors.

Taken together, the results from the three experiments show a clear interaction between display orientation and access to categorical level information. The role of access to categorical level information is supported by Giese and Poggio's (2003) model of the recognition of biological movement. High-level areas in the form and motion pathways are selective for body shapes and specific human actions like walking and running. Feedforward processing from 'lower' visual areas along the different pathways activates motion pattern neurons that selectively encode motion patterns of human movement. Results from simulations of their model are consistent with the categorical processing of different actions based on psychophysical data.

In addition to the theoretical and modeling framework proposed by Giese and Poggio, the findings from the 3 experiments presented here also suggest that the visual processing of upright displays is indicative of vision at a glance, whereas viewing inverted displays indicates vision with scrutiny. Within the framework of RHT, access to categorical level information for upright displays is fast and automatic. This indicates that participants had early access to high-level stored representations of human motion patterns that depicted specific actions. Findings from exps. 2 & 3 showed that viewing an upright prime led to significantly greater performance in an action identification task. It seems reasonable to suggest that this effect was due to greater access to categorical level information activated by the priming actions.

In contrast to upright actions, orientation decisions for inverted actions took significantly more time and led to relatively less categorical level priming. In terms of RHT, this suggests that the feedforward mechanisms involved in the visual processing of inverted actions do not have the same level of access to stored representations of human motion patterns. The longer processing time for inverted displays could reflect the operation of feedback mechanisms that attempt to rebind local motion components (e.g., local rigidity) into a hierarchical whole for the purpose of identification. This is not to say that inverted displays cannot be reliably detected or recognized. It is rather a relative lack of access to categorical level information that distinguishes the processing of upright actions from inverted actions.

The findings from the work presented here confirm the general findings of orientation specificity for biological motion perception. The contribution from the experiments presented here indicates that access to categorical level information is one important factor that differentiates the visual processing of upright actions from inverted actions.

Acknowledgments

I would like to thank Sigríður H. Pálsdóttir, Maria Nilsson, Patrick Söderberg and Peter Thorvald for running the experiments. This research is supported by the University of Skövde.

References

Ahlström, V., Blake, B. & Ahlström, U. (1997). Perception of biological motion. *Perception*, 26, 1539-1548.

Bertenthal, B. I. & Pinto, J. (1994). Global processing of biological motion. *Psychological Science*, 5, 221-225.

Boucart, M. & Humphreys, G. W. (1992). Global shape cannot be attended without object identification. *Journal of Experimental Psychology: HP&P*, 18(3), 785-806.

Dittrich, W. (1993). Action categories and the perception of biological motion. *Perception*, 22, 15-22.

Giese, M. & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nat. Reviews: Neuroscience*, 4(3), 179-192.

Hochstein, S. & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, 36, 791-804.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14, 201-211.

Jokisch, D., Daum, I., Suchan, B. & Troje, N. F. (2005). Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis. *Behavioural Brain Research*, 157, 195-204.

Palmeri, T. J. & Gauthier, I. (2004). Visual object understanding. *Nat. Reviews: Neuroscience*, 5, 291-303.

Pavlova, M. & Sokolov, A. (2000). Orientation specificity in biological motion perception. *Perception & Psychophysics*, 62(5), 889-899.

Pinto, J. & Shiffrar, M. (1999). Subconfigurations of the human form in the perception of biological motion displays. *Acta Psychologica*, 102, 293-318.

Shiffrar, M., Lichtey, L. & Heptulla Chatterjee, S. (1997). The perception of biological motion across apertures. *Perception & Psychophysics*, 59(1), 51-59.

Thornton, I. M., Rensink, R. A. & Shiffrar, M. (2002). Active and passive processing of biological motion. *Perception*, 31, 837-853.

Verfaillie, K. (1993). Orientation-dependent priming effects in the perception of biological motion. *Journal of Experimental Psychology: HP&P*, 19, 5, 992-1013.