Visual Action Recognition Study: Orientation Specificity in Mental Representations of Upright and Inverted Biological Motion

(HS-IDA-EA-01-515)

Sigríður Hrefna Pálsdóttir (a97sigpa@student.his.se)

Department of Computer Science
University of Skövde, Box 408
S-54128 Skövde, SWEDEN

Dissertation for the degree of BSc in Cognitive Science, written during the spring term of 2001.

Advisor: Paul Hemeren

Visual Action Recognition Study: Orientation Specificity in Mental Representations of Upright and Inverted Biological Motion

Submitted by Sigríður Hrefna Pálsdóttir to Högskolan Skövde as a dissertation for the degree of B.Sc., in the Department of Computer Science.

June 11, 2001

I certify that all material in this dissertation which is not my own work has been identified and that no material is included for which a degree has previously been conferred on me.

Signed: _____

Visual Action Recognition Study: Orientation Specificity in Mental Representations of Upright and Inverted Biological Motion

Sigríður Hrefna Pálsdóttir (a97sigpa@student.his.se)

Abstract

Research on biological motion, using point-light displays to present the motions, have been unravelling what information factors are still embedded in those impoverished stimuli and which of these factors are essential in visual processing of biological motion. Earlier studies suggest that orientation is a crucial factor in biological motion processing. The short-term priming experiment presented in this paper will further investigate the legitimacy of the primacy of orientation and suggest different solutions based on contradicting findings in previously published studies.

In a serial two-choice reaction-time task, participants were presented with a patch-light display of a human engaged in one of three possible actions: climbing up a rope, jumping jacks, and walking. Participants had to identify the in-plane orientation of the human figure emerging from the moving patch-lights. Reliable facilitation effect was established for transitions containing same-oriented upright trails and same-oriented inverted trials. Interestingly, transitions of same-oriented upright trials produced significantly greater facilitation effect than transitions of same-oriented inverted trials.

Keywords: Visual action recognition, orientation specificity, biological motion, point-light displays, short-term priming.

Acknowledgements

Index

1 Introduction

Is it too far fetched to say that human beings live their lives through recognizing the world they live in? That they are dependent on recognition to survive? People are equipped with at least five senses; we are receptive to all kinds of different elementary particles of information (or stimuli) existing in our environment; like chemicals, sound waves, light rays, temperature changes and pressures. Receiving these stimuli, though, only can help us survive when we know how to sort out the meanings of the stimuli and recognize the identity of the source of the stimuli. Most of the time we recognize what we perceive, e.g. when we fumble for the light switch in the dark, eventually we will find it because it feels differently than what it is attached to, we recognize our date in a place crowded with people, we can differentiate between Pepsi and Coca Cola (some can anyway), know when a smell of something burning is dangerous and when it is just a cigarette or an incense and it is not unusual to recognize friends and relatives on the telephone just from hearing "Hi". We do all this through perception and recognition, which makes us able to react appropriately to any given situation.

Research interest in visual perception has been great for a long time. Much of research is going on in different scientific fields like neuropsychology, AI and cognitive psychology, etc. Perhaps the area that has received most attention is the area of visual recognition. Studies in visual recognition are further divided into research areas such as face recognition, object recognition and action recognition. An action recognition study is presented in this paper where the focus is set on factors involved in the recognition of human biological motion. The rest of this section discusses the phenomena of human recognition. In later sections I will discuss visual recognition in general and object recognition and action recognition in particular. Current theories of how we use mental representations in both object and action recognition will be reviewed and the notion of orientation specificity will get special attention.

> [Recognition is] The action or fact of perceiving that some thing, person, etc., is the same as one previously known; the mental process of identifying what has been known before; the fact of being thus known or identified (The Oxford English Dictionary, 1933, p. 252).

The above definition of recognition gives a simplified everyday explanation of what recognition is in terms of perception and cognition. Of course it is limited in its explanatory power but it gives a hint to the fact that recognition is a specific part of perception that involves cognition. The cognition, or the mental process of identifying what has been previously encountered, is strongly dependent on what we perceive.

What is then perception, and how is it different from recognition and cognition? There is no clear-cut answer to this question, and it is not within the scope of this paper to try to answer it either. Perception can be defined as the whole process e.g. from the moment light rays reach the retina in the back of our eyes, through millions and billions of neuronal responses until it reaches our visual consciousness as being a red square, or something else more or less complex. What we take in through our senses is dependent on the abilities and limitations of our senses, how they are constructed, but what we make of what we take in is dependent on other structures, cognitive structures or representations we make about the world and which can be different from one person to another (Lundh, Montgomery & Waern, 1992). Recognition is about using these representations and mental structures to identify what we take in and is probably an important link in making new representations and structures for

novel phenomena we have not encountered before. For example, say I would go on a blind date with someone and am supposed to meet that person in a place crowded with strangers. The description I have to go on is a medium tall man, in a black shirt, with brown hair and a pink hat. I spot the pink hat immediately, and if we decide to meet somewhere else on another occasion, he will not have to wear that pink hat again because now I have a good enough representation of him in my head to be able to recognize him.

Recognition of information perceived through any of our senses generally happens quickly, effortlessly and accurately. Because recognition is so fast and easy (see discussion in the section about visual recognition), the process seems simple and straightforward, we perceive what is out there and we can rely on that as a source of absolute truth, or that was what the philosophers, Hume, Berkeley and Locke, thought (Gregory & Colman, 1995). But we do not always correctly interpret what we perceive. We sometimes e.g. think we hear or see things that are not there, which is impossible to explain if perception is just a reflection of the external world. Since scientists have realized that there might be something more to perception, experiments have been done and evidence has been accumulated to support a theory of perception being more of a creative activity, requiring assumptions from knowledge stored from the past to complement sensory data (Gregory & Colman, 1995). Perception may still be straightforward in a way, though constrained by the structural limitations of each of the senses, but it is also a complex system of structures or internal representations where experience and familiarity play an important role (Lundh et al., 1992). We become familiar with things we encounter or experience often, and we become better at identifying things we are familiar with than things we seldom experience.

So what do these internal structures or mental representations in recognition look like and how do they work? In neuropsychology, observations and studies with patients with different kinds of agnosia have revealed some of the factors that are involved in recognition. "Agnosias are generally confined to a single perceptual modality, such as auditory, tactile, or visual, suggesting that for each perceptual modality there is a stage of processing beyond elementary perceptual processes that is nevertheless modality-specific, and that represents learned information about objects' sounds, tactile qualities, and visual appearances." (Wilson & Keil, 1999, p. 615). In visual recognition alone, different kinds of agnosia (face agnosia, object agnosia, alexia, etc.) have been identified. Face recognition is a part of the study of visual recognition that has got much attention from scientists in recent history. To give an example of what is happening in the field of recognition, a short "overview" of results in face recognition studies and related theories will be given.

Human beings have an incredible memory for faces, we not only know our closest relatives and friends that we see everyday but can easily recognize people we have not seen in years and even if they have changed hair colour, gained weight, are grown up or had a nose job. Face agnosia (prosopagnosia) is a selective deficit in recognizing faces. The fact that people suffering from brain damage, causing a deficit only in the recognition of faces, suggests that there are specialized cognitive processes involved in face recognition that are not involved in other types of visual recognition. Prosopagnosic patients are aware of faces being faces but they cannot recognize even their best friends and relatives unless they get a hint from their clothes, the sound of their voice, hairstyle or other non-facial factors. Neurophysiological studies of recordings of neurons in both humans' and monkeys' brains support the theory of special processing of facial stimuli. Neurons that fire selectively to face stimuli have been located in the temporal neocortex. Different explanations have been given for

the apparent "specialness" of faces, such as 1) that different strategies for processing faces are necessary because of the structural uniqueness of faces compared to most other object classes or 2) that faces and other objects are processed and stored in a similar manner, but that faces are simply harder to tell apart than other kinds of objects (Wilson & Keil, 1999).

We human beings recognize things in our environment all the time and our functioning in the world depends on it. The effortlessness of recognizing thousands and thousands of different things from different stimuli has fooled some people in the past to believe recognition and perception to be simple and straightforward when they really are complex. The aim of this paper is to further explore how biological motion is represented in the mind of human beings through visual recognition, and that is the reason for pointing out the importance of mental representations in recognition. Review of the development that has occurred in the research area and what main theories are available are presented and compared to relevant issues in visual object recognition. The following sections address general facts about visual recognition, what factors are known to influence visual recognition (object and action recognition especially) and what these factors can tell us about the internal representations we use in visual recognition. Two special areas of visual recognition, object and action recognition will be presented and the latest findings and theories discussed, in particularly findings supporting orientation specificity in the recognition of biological motion. Finally I will propose a study that will further explore the notion of orientation specificity.

1.1 Visual recognition

Of all our different senses, we human beings seem to trust vision the best. That might not come as a surprise as it has been estimated that 80% of the knowledge we obtain through perception is accounted for by vision (Gregory & Colman, 1995), and almost 50% of the human brain's cortex is involved in vision (Hoffman, 1998). These facts also partly explain why vision, of all the senses, has received the most attention from scientists so far.

Vision provides us with detailed information about the world beyond our body surface. Vision enables us to track events as they take place in front of our eyes. It enables us to read, to detect objects, actions, animals, people and other living things, static or moving. Vision also plays an important role in allowing us to interact with other people, objects (both natural objects and artefacts), and other structures in our environment. To be able to interact, we first must identify an action or an object, before we can decide on an appropriate action. We see the ball (or at least we see the pitcher throw), then we hit the ball.

Researchers have suggested two separate, but interacting, visual systems; one evolved for the perception of the external world and the other evolved for controlling of actions directed at the external world and objects within it (for more detailed account see: Milner & Goodale, 1996, Tarr & Bülthoff, 1999 (chapter 7, Goodale & Humphrey)). Most research has been done on how we visually perceive the external world, where the main task of the visual system is considered to be to construct a representation of the three-dimensional layout of the world and the objects and events within it. These kinds of representations or internal models can be used in the recognition of objects and understanding their interrelations (Tarr & Bülthoff, 1999). The research is focused on what kind of information the visual system provides us

with, how this information is used to construct representations, on how the representations are constructed, what their advantages and disadvantages are, and on how we use these representations in recognition and other kinds of cognising of the visual world.

Various theories have surfaced through the years in attempts to explain what kind of information is important for us to perceive in order to create representations or activate existing representations to be able to recognize. Some have believed the overall shape to be the most important factor, others the texture, the colour, the context in which we see the phenomena, etc. At least we know much about what we can see, shape, shadows, textures, colour, spatial information, temporal information, etc., and we are able to see these features under a wide range of conditions, more or less optimal for visual perception, like in various degrees of lightning, shaking ourselves or watching a shaky world, proportionally distorted, presented in different degrees of detail and abstraction etc. A central issue is what features are preserved in the mental representation and if any of the different features preserved are more essential for recognition than the other.

Visual recognition is about recognizing phenomena in the world through information perceived by the eyes using internal representations of previously perceived phenomena. The research area of visual recognition is further divided into independent research areas, where face recognition, object recognition and action recognition are the most prominent. Face recognition, our ability to recognize faces and facial expressions, has already been mentioned as an example of recognition (see previous section). The study of object recognition revolves, most of the time, around what factors influence the way we perceive and represent objects we see in the world whereas research in action recognition studies the factors that influence the way we perceive and represent actions (motion).

We know what we see, and we do it effortlessly, we can differentiate between immense amount of objects, people and actions in a fraction of a second and tell if we have previously seen what we see despite the time elapsed in between and changes that have occurred in that interval. Visual recognition is a very important function in our everyday lives but we are far away from having solved the problem of how we go about doing it. One of the most researched questions of visual recognition is how we manage to recognize an object viewed from different perspectives as the same object, this ability is often referred to as visual object constancy and for which Bechtel and Graham give a good example (1998, p. 266):

> An elephant is an elephant whether one is looking at the front, back, or side. For this reason, perceptual representations are often assumed to have a viewpoint-independent frame of reference.

The next two sections will deal with this question within the areas of visual object recognition and action recognition. Different theories accounting for the ability of visual object constancy, viewpoint-independent recognition being one of those, are discussed in both areas, and factors that have been found to influence (or not to influence) visual object constancy will be highlighted. This discussion will ultimately lead to the notion of orientation specificity in both object and action recognition. Theories in the different domains of action and object recognition are compared to a certain extent and both domain specific and overlapping studies will be discussed.

1.2 Object recognition

The area of visual recognition is divided into smaller research areas. The area that has received much attention is the area of visual object recognition, and still, no artificial system exists today with the capabilities of even recognizing common objects (Tarr & Bülthoff, 1999). Visual object recognition in AI has primarily focused on techniques that use models of specific objects, based on properties such as shape and appearance. These techniques require prior knowledge of the objects and are referred to as model-based recognition. The human visual recognition system, on the other hand, can recognize novel objects as well as previously encountered objects (Wilson & Keil, 1999). How we go about identifying and categorizing objects as well as understanding new objects is certainly a problem worth the pursuit of solving it. The focus of this paper is not specifically aimed at examining how we recognize new objects and actions but to examine how we manage to recognize objects and actions viewed from different viewpoints (see figure 1 below) as being the same object (visual object constancy). The answer to how we manage to recognize novel objects and actions might partly be found within the explanation of how we manage visual object constancy.
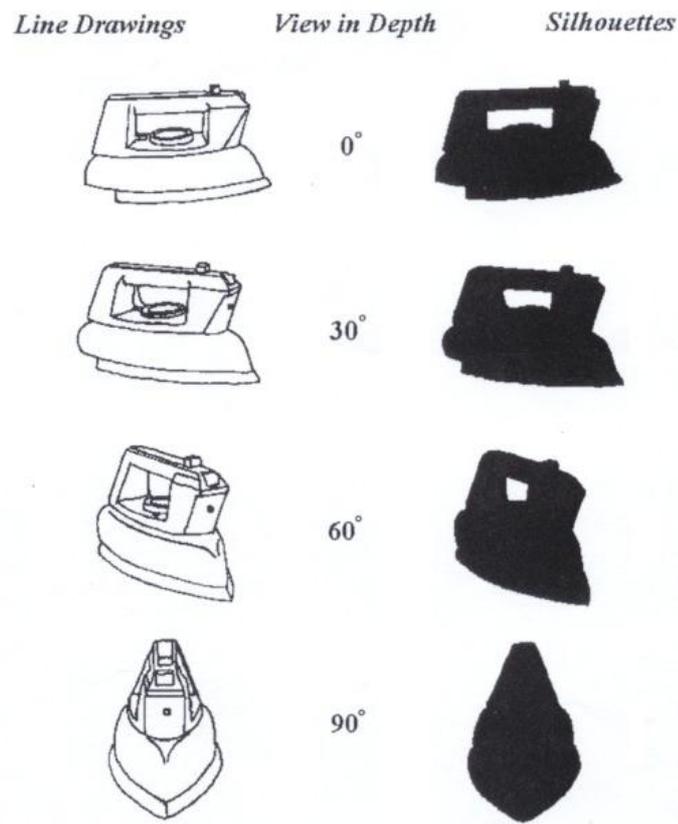
Figure 1. An iron and its silhouette seen from four different angles (Lawson, 1999, p. 223). The numbers indicate the degree of rotation in depth from the first angle.

Images of the same object viewed from different angles can vary in size, outline shape, and some features may be partly or totally invisible from certain viewpoints but we still manage visual object constancy. This ability suggests that the structure of our representations, or how they are interconnected somehow enables us to recognize objects from different viewpoints.

The projected image of a 3D object can vary considerably when viewed from different positions. As mentioned before, we still manage to recognize these projection as being the same object despite the fact that we only have viewer-centred 2D patterns of light to work with (Tarr & Bülthoff, 1999). Marr and Nishihara suggested a solution based on the "belief" that the goal of the visual system is to reconstruct the 3D scene. According to Marr and Nishihara's theory (as presented in Gregory and Colman, 1995), representations used in recognition must be more general than the 2D (2½D in their account of visual information processing, for further information see Gregory & Colman, 1995) patterns of light projected on our retina, describing the basic 3D structure of an object so that its appearance can easily be obtained from any viewpoint. Objects are therefore represented as made up of a hierarchy of cones, where the most general representation is one or a few cones described 3-dimensionally according to an axis running through the elongated part of the cone. More specific representations of the same object can be made through breaking the cone up into smaller cones (see figure 2 below). This type of object recognition is object-centred, as the visual system has to derive a 3D model description of the 2D retinal image before a comparison of the model description and the pre-stored 3D representations can occur. Recognition is then achieved if there is a match between the derived 3D model and the stored 3D representation corresponding to known objects (Bruce & Green, 1990).
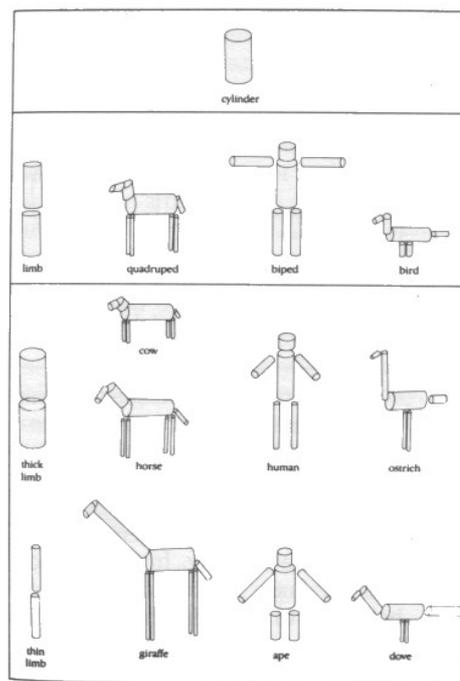
Figure 2. Illustration of Marr and Nishihara's object-centred representation theory (Bruce & Green, 1990, p. 193).

The theory developed by Marr and Nishihara was a so-called viewpoint-independent recognition theory. Their theory had a great impact on the field of visual object recognition but since then different theories have come forth. In fact research results have shown that the speed of object recognition is influenced by the familiarity of a given view, its similarity to views of other objects from which it must be discriminated, and how well it depicts the object (Lawson, 1999). These facts suggest that the object representations in our mind are not viewpoint-independent or at least not entirely.

Lawson (1999) divides all theoretical accounts of visual object constancy into three classes; invariant features, multiple view and transformation. She argues that the visual system probably employs all accounts to some degree depending on the stimuli presented, the context in which the stimuli must be recognized and the task required, which is good to have in mind when designing an experiment and when drawing conclusions. Recognizing objects using invariant features suggests that we extract a specific feature, unique or special to an object, and compare it with stored representation of objects in order to recognize and discriminate. Lawson argues that such accounts of visual object constancy might well hold when the task is to discriminate between highly distinctive objects or when the objects are very well known (familiar). Managing visual object constancy according to the multiple view account involves storing several view-specific representations of each object (from different visual angles) where recognition is achieved through matching the perceived image to the nearest view-specific, stored representation. Lawson (1999) refers to studies demonstrating strong view-specific priming effects (section 1.2.1 explains priming), which means that images of an object are recognized faster when followed by an image seen from a similar point of view than when followed by an image seen from a totally different viewpoint. These results suggest that we use view-specific, stored representations in recognizing familiar objects. Transformation theories propose that the visual system transforms (e.g. by mental rotation) the retinal image to reduce the differences between the image and a view-specific, stored representation. Tasks that require subjects to determine the direction that an object, rotated in the frontoparallel plane (explained later in this section), would have if it were upright, are assumed to employ mental rotation (for more detailed discussion see Lawson, 1999).

Not only is recognition of an object effected by how detailed our representation of it is, recognizing an animal to be a dog but not which breed, but also by the viewpoint by which the object is pictured from, recognizing a house pictured from the front yard is much easier than recognizing a house pictured from above. We usually manage to recognize the object, nonetheless, but it takes more effort (measured in time) than to recognize the object from a different viewpoint. The viewpoint from which we are fastest to recognize an object is often called the canonical or preferred view of an object. Identifying the factors that determine the canonical view is less certain. One factor could be familiarity, the fact that we normally see a house from the ground in front of the house and not from above. Another factor could be the object's attributes. Palmer (1981, cited by Blanz, Tarr & Bülthoff, 1999), for example, suggested that canonical views of objects are determined by the number of surfaces visible from a given viewpoint where the viewpoint showing largest number of surfaces is therefore the preferred view.

In studies trying to determine what factors influence our recognition of objects and actions, and especially in studies trying to determine how view-invariant or viewpoint-dependent our mental representations are, the viewpoint from which the object is pictured is manipulated. The object can either be rotated in the frontoparallel plane or in depth. Plane rotations are a special case of image transformations about the horizontal axis in the line of sight of the viewer and are rarely observed in everyday situations. Depth rotations on the other hand are transformations about the vertical axis, running through the middle of the object and are more common in everyday situations. Human beings' performance in visual object constancy across an object's depth and plane rotation is not perfect. Studies have invariably shown that objects in their canonical, upright view are easiest to detect, this shown in response time (RT) in different kinds of recognition studies. An object shown from other views

rotated in the plane take significantly longer time to recognize and the RT increases linearly with the degrees of rotation from 0° to 180° (180° being an exception though) (Lawson, 1999).

How does a view-invariant, object-centred theory explain these findings? Marr suggested that it is more difficult to assign the axis of elongation to some objects than others and that it could also be more difficult to assign the axis of elongation to some views when compared with other views of the same object. This explanation might be acceptable if there were an RT difference for objects rotated in depth, because that kind of rotation can extensively change the shape, illumination and features visible from view to view, but it does not explain the difference in RT noted for objects rotated in the plane. Verfaillie (1992) has a theory of less viewpoint-invariant nature, suggesting that stored object representations are orientation sensitive.

In this section I have introduced some theories and concepts relevant to the discussion of visual object constancy. In recent years more and more evidence, especially from neurophysiological studies, has accumulated to support the fact that visual object recognition is somehow orientation-dependent, but there are conflicting results in behavioural studies and room for different hypothesizing. In the following section I will shortly explain how priming studies work, as they will be referred to later on. After that I will discuss the notion of orientation specificity, comparing results from behavioural, and neurophysiological studies.

1.2.1 Priming

Priming studies are designed to test if a previous exposure to a stimulus affects the recognition of an identical stimulus viewed later. Two different kinds of priming studies exist, studies testing short-term priming and studies testing long-term priming. The difference between short-term priming and long-term priming studies is that a long-term priming study consists of separate study phase and testing phase with time intervals in between, the intervals varying in length from minutes to months (Olofsson, Nyberg & Nilsson, 1997).

1.2.2 Orientation specificity

Simply put, orientation specificity in object recognition means that manipulation of the orientation of a viewed object has an effect on visual processing of an object. Both studies manipulating the in-depth orientation and in-plane orientation of an object have created systematic effects, leading scientists to theorise about the role of orientation in visual processing. Some scientists interpret these findings as evidence for orientation specificity of the mental representations we retain of a processed object, where the representations are somehow organized by orientation and recognition of an object is therefore dependent on the extraction of the orientation of the viewed object. Some studies manipulating other factors of the presentation of an object than orientation have, however, produced systematic effects. Long-term priming studies have shown that naming latency reduces for drawings viewed on a second occasion, regardless of whether the same drawing is shown, or the same object is presented in a different position in the frontoparallel plane, a different size, or a left-right reflection (Verfaillie, 1992). If the mental representations were strictly organized by orientation then naming latency should only be reduced for drawings

viewed on a second occasion that shared the same absolute orientation as the drawings viewed in the study phase, irrespective of size and in the position in frontoparallel plane.

Jolicoeur and colleagues (Jolicoeur, 1985, 1988; Jolicoeur & Milliken, 1989; cited by Verfaillie, 1992) have found that rotation of mono-oriented drawings (drawings usually seen in upright orientation) in the frontoparallel plane increases naming latency for drawings departing from the upright. On the other hand, rotation of mono-oriented drawings around their top-bottom axis does not affect recognition. These findings can be related to view-dependent coding and such theories more than view-invariant theories, like those of Marr. The impact of the object's orientation seems to grow as the object becomes more complex (the human body is taken as an example of a complex object as well as faces and confining results from perception of human biological motion). Humans are, however, affected by in-depth rotation of an object when doing a matching task when the objects are shown sequentially. The response latency is greater for two images of an object shown from two different in-depth rotation viewpoint than when they are identical. Short-term priming effects have also been noticed due to orientation. Subjects asked to classify images in two predefined categories, one image showed at a time. RTs are shorter for images of objects preceded by an image of the same object from the same viewpoint than when preceded by an image of the same object viewed from a different perspective (Verfaillie 1992). Despite this, experiments have been reported where no effect of orientation (rotation) was obtained, especially in identification task studies and matching tasks. There is also invariance in results done with short-term priming and long-term priming where long-term priming tasks show no orientation-dependency.

Verfaillie (1992) thinks that the object's orientation is the most promising candidate for being the primary factor that makes the visual system of recognition less view-invariant and more view-dependent, i.e. that orientation is the view-dependent factor, other factors do not affect recognition. Or as Verfaillie puts it: the solution to the problem of object recognition employed by the animate perceptual system would involve the computation of an orientation-dependent object description, yet independent of other variations (Humphreys and Quinlan, 1987, cited by Verfaillie, 1992).

Neurophysiological studies have examined the effect of other factors on object recognition and so far human beings seem to be unaffected by changes in illumination, colour, changes in position in the picture plane (then we are not talking about rotation), and changes in position in depth (again not talking about rotation) as size and viewing distance have only small effects (Verfaillie, 1992).

1.3 Action recognition

So far, only theories and scientific findings regarding orientation specificity in object recognition have been presented. Action recognition is different from object recognition in fundamental ways. Action recognition involves temporal and spatial factors, where the continuous changes of the target create the action. Objects on the other hand are not normally affected by the elapsing of time. It is possible to look at a vase of flowers for hours without anything happening, while a person looking at a snail for hours will experience changes in the shape of the snail as it moves, and as times goes, the snail will change position in the space, moving with time from place to place.

Action recognition can probably apply to other kinds of visual recognition than the recognition of moving things, but in this paper action recognition will refer to the recognition of motion, and of biological motion in particular. Biological motion is simply put, a product of a moving biological creature. Recognizing movements is just as important for humans as recognizing objects. We see people around us produce motion in different ways, walking, running, dancing, lifting an arm, tapping a foot etc., and identifying what action people are performing does not cost us any more concentration than identifying objects or people.

One would though think that action recognition is different from object recognition, if not for other reasons than different information is required to identify objects than actions. Seeing a snapshot of a familiar person gives us no problem in identifying the person, but we would have trouble identifying what action he was engaged in at the moment the picture was taken, for example recognizing if he was sitting down or standing up from a table.

Of course some kind of information might be equally important to object recognition and action recognition, but if there were two separate mechanisms working behind the two different kinds of recognition, using different kinds of information, then two mechanisms would possible respond differently to manipulations of different factors. Orientation seems to be an important factor in processing of objects but that does not imply that it has to be primary in biological motion processing. Mental representations of biological motion might be organized by totally different factors than object recognition. The following section will introduce a method to study biological motion, stripping it of as much object related information as possible. After that, studies testing the effect of manipulation of orientation of biological motion will be presented.

1.3.1 Point-light displays

In 1973 Johansson published an article about visual perception of biological motion. With his article he was testing a model, which he and his colleagues had been working on in relation to research done on perception of mechanical motion, to see if it also could be applied to the perception of biological motion. The relevance of the article to the work presented here has mostly to do with how he presented the stimuli for the biological motion. Johansson (1973, 1975) used point-light displays (or patch-light displays), which he produced by filming his assistant doing different kinds of natural movements, for example walking, running, jumping, etc. To prevent the motion from being recognized by simply recognizing the contours of the human figure, the contours were made invisible. The assistant was dressed in tight-fitting

dark clothes with small patches of retroreflective tape attached to his head and main joints (see figure 3 below). Filmed against a dark background, only the reflections of the patches on the moving head and joints were visible. Later, when the recordings were shown to subjects, the structure of a human body was invariably extracted from the motion of the cluster of lights, but if the lights were stationary a human was not identified.
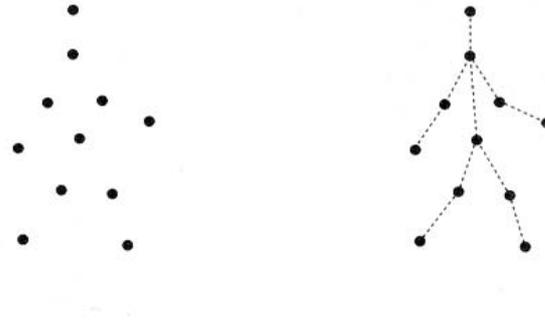
Figure 3. On the left, a static point-light figure like presented in some point-light studies, and on the right, the same figure with lines connecting the spots to show the right structure of the figure. The human figure and the action performed are easily recognized as soon as a sequence of a motion is played (Hoffman, 1998, p. 161).

Since 1973 much research on perception of biological motion has been done using point-light displays as visual stimuli. The point-light display has been manipulated in different ways in attempt to further identify the factors that affect the visual perception of motion and biological motion in particular.

1.3.2 Orientation specificity

Several studies have been done manipulating both the in-depth and the in-plane orientation of biological motion stimuli, presented in point-light displays. Not all experiments involve priming studies, but studies employing naming and identification tasks have reported differential results for point-light biological motion in upright and inverted orientation. Upright biological motion stimuli have been reported to be identified correctly more often than inverted stimuli, both when viewed without a mask of moving point-lights (Sumi, 1984; Dittrich, 1993), and even when viewed with a mask (Bertenthal & Pinto, 1994).

Subsequent sections will present priming studies manipulating both in-depth and in-plane orientation of biological motion and introduce the theories that have been put forth as to explain the findings.

1.3.3 Priming effects of in-depth rotations

Karl Verfaillie (1992, 1993, 1999, 2000) has adopted a notion of orientation-dependent mental representations of both object and action recognition. To support his theory, he obtained results from short-term priming experiments (1993) done with biological motion point-light displays, other behavioural studies and findings from neurophysiological single-cell recordings.

As can be seen in the figure below (see figure 4), same orientation for both priming and primed display had a greater facilitating effect (the two first columns) as

compared to the trials where the motion had different orientation in the priming and the primed displays (columns 3 and 4). Verfaillie therefore suggests that view-specific representations are employed in visual recognition of biological motion and that the representations are view-specific in terms of orientation.
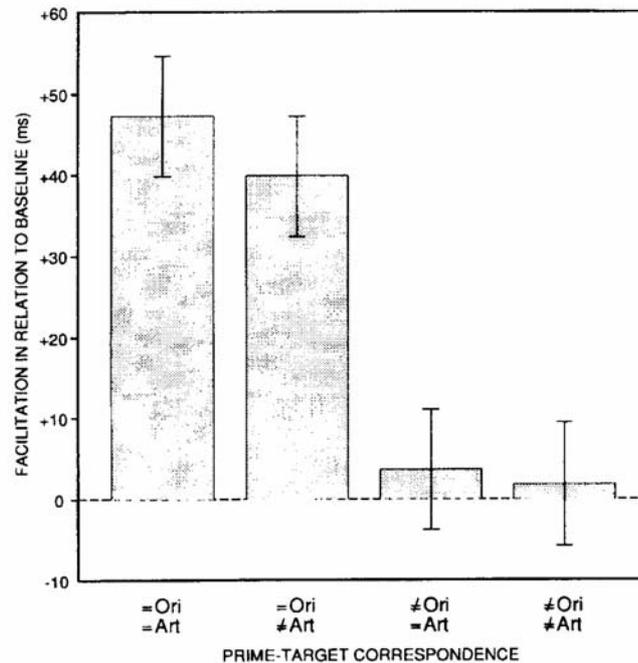

Figure 4. This diagram shows orientation specific priming effects for human trials in Verfaillie's 1993 study (Experiment 1), where the task was an object-decision task, discriminating between a human and a nonhuman walker. The movements' direction of articulation had no significant priming effect (Verfaillie, 2000, p. 199). The bars indicate 95% confidence intervals. =Ori, same orientation; ≠Ori, different orientation; =Art, same articulation; ≠Art, different articulation.

Single-cell recordings done with monkeys also show selectivity to in-depth orientation while colour, illumination, and position in plane and in depth have invariant effect. Most of the single-cell recordings have been done in visual object, face, or body recognition (Perrett, Harries, Bevan, Thomas, Benson, Mistlin, Chitty, Hietanen and Ortega, 1989, cited by Verfaillie, 2000), and it is questionable if they can be used to support the notion of orientation specificity observed in biological motion studies with point-light displays. It is not impossible that two different mechanisms are employed in visual recognition, one for object recognition and a different one for action recognition. Interestingly, not all cells show orientation dependency, a minority of cells have been found that are responsive to all views of a particular object.

The only fault with Verfaillie's study (1993) was that the task that the subjects were instructed to do was more of an object recognition task than an action recognition task. The task was a serial two-choice reaction-time task where the participants had to decide as rapidly as possible on each trial if a biological motion configuration depicted a regular human walker or a "nonhuman" version. The task involves categorization where each trial can belong to one of two categories. Categorization in itself is not more of an object recognition task than an action recognition task but the properties focused on in this particular categorization task were object oriented. Concentrating on categorizing a motion as a human or nonhuman walker means that we already know to expect a walker and probably does not require the extraction of action specific features. Being criticised for this, Verfaillie (2000) carried out a new

experiment almost identical to the one he did in 1993 except for the change in the nature of the task. This time participants had to discriminate between forward articulating walkers and backward articulating walkers. The results are shown below (see figure 5).
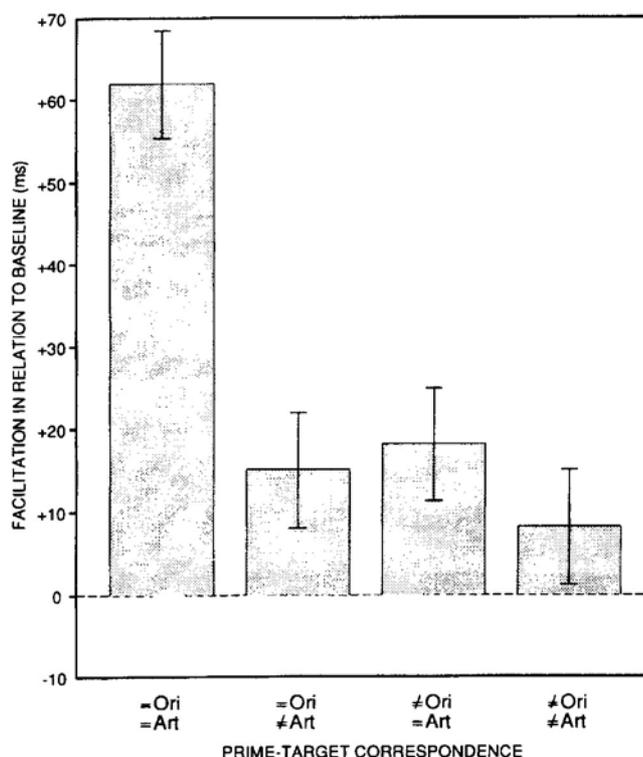

Figure 5. This diagram shows priming effect for both orientation and direction of articulation. The task was an action-oriented task, discriminating between a backward articulating walker and a forward articulating walker (Verfaillie, 2000, p. 206). The bars indicate 95% confidence intervals. =Ori, same orientation; ≠Ori, different orientation; =Art, same articulation; ≠Art, different articulation.

In contrast to the results of the 1993 study, only the combination of direction of articulation and orientation had a pronounced facilitation effect, i.e. the response time to a primed display was shortest if the preceding display had both the same left-right orientation and the same direction of articulation as the primed display. All other conditions had a significant priming effect, at least 10 ms above the baseline, which is also different than in the 1993 experiment.

Verfaillie (2000) explains the difference in the results of these two studies according to the different computational goals of object recognition, on the one hand, and action recognition on the other hand. The purpose of basic-level object recognition is to categorize the stimulus object as a human body and, therefore, to make an abstraction of the specific 3D arrangement of body parts. Irrespective of the 3D organization of the body parts, the object is classified as a human body. Action recognition boils down to categorizing a particular 3D manifestation of a human body as a specific pose or a specific phase of a certain action. Verfaillie goes on to say that recognizing actions of the human being are orientation-dependent because recognizing a human being is orientation specific, meaning that before we can tell which direction a person is articulating in, we have to recognize the figure as a person. This suggests that the same mechanism is used in object and action recognition, or that the mechanism used in action recognition is dependent on the mechanism used in object recognition.

As mentioned before, not only transitions of same-oriented trials produced priming effects, even transitions of different-oriented trials produced significant facilitation effect. These results might be reflecting the orientation-independent neurons at work, but then again, are findings in object recognition studies and action recognition studies measuring the same things?

To explain this small facilitation effect and the orientation independent neurons, Verfaillie adopts and adjusts a hierarchical theory first developed by Perrett et al. (1989, cited by Verfaillie, 2000) in order to explain the orientation insensitive neurons. They suggest that the activity of viewpoint-independent neurons is the result of combining the output of viewpoint-dependent neurons, each tuned to a particular view of the object. This flow of information from view specific to viewpoint insensitive representations proceeds in hierarchical fashion, explaining the slightly shorter response latencies of orientation-sensitive cells than orientation-independent cells. Verfaillie thinks his findings in short-term priming studies with human biological motion point-light displays also supports this orientation-dependent hierarchical notion. In addition, Olofsson et al. (1997) showed that there is also an orientation effect in long-term priming of biological motion, further supporting Verfaillie's theory.

In summary, Verfaillie interprets his findings as supporting the theory of orientation primacy in processing biological motion. The theory is built on another theory explaining orientation specific effects in object recognition as being the results of "… a hierarchical scheme in which orientation-dependent stored representations of objects are accessed (although abstraction is made of other stimulus variations, like image-plane position and size) prior to the activation of representations that are more independent of the observer's vantage point" (Verfaillie, 2000, p. 208). The same kind of hierarchical scheme is supposed to be involved in action recognition because identifying action related information, as for example walking forward, would require the in-depth orientation first to be identified before we can say if the walker is moving in the same direction he is facing.

1.3.4 Priming effects of in-plane rotations

Verfaillie has mainly concentrated on doing experiments in visual recognition of biological motion point-light displays, manipulating the in-depth rotation of the figures, or as a matter of fact, only varying the orientation of a walker between walking to the left or right in comparison to what the viewer sees. Other researchers, Pavlova and Sokolov (2000), have done research in rotating the point-light displays in the frontoparallel plane. They have also produced priming effects depending on the degree in which a walker is rotated in the plane (usually rotated 30° in each step, 0° being at 12 o'clock). Response time was found to vary as a function of the degrees the walker is rotated from its normal upright position, where priming effects are largest for upright walkers preceded by upright walkers. Walkers that are rotated 30° to 45° are also primed by upright walkers and vice versa but the effect is not as great (Pavlova & Sokolov, 2000). Pavlova explains these findings as a result of how our visual system is dependent on other cues from the environment, involving gravity, kinematics and dynamics. These factors are also supposed to help us decide how much weight people are lifting and recognizing our friends just from point-light displays.

Interestingly, Pavlova and Sokolov ruled out knowledge-based constraints to explain obtained priming effects in perception of biological motion. This implies that the power of ecological constraints on perception of biological motion overrides our ability to learn to recognize unnatural motions, so no matter how much training a person would get in identifying inverted walkers she would still produce greater priming effects for upright walkers than inverted walkers. This claim seems quite strong as Pavlova and Sokolov only gave participants 3-10 seconds to familiarize with the motions in the different in-plane orientations. It seems logical that only transitions of upright or close to upright orientations produced priming effects as the participants had had their whole life to get acquainted with upright walking figures. Therefore it also appears a hasty conclusion to draw that humans are not able to learn to recognize motions displayed in other in-plane orientations well enough to also produce priming effects for transitions of motions of those orientations.

1.3.5 Incongruity of priming effects caused by in-depth and in-plane rotations

In previous two sections, priming effects of in-depth rotation and image-plane rotation of a point-light walker on visual recognition (perception) have been reviewed. Summarizing very briefly the findings presented in these sections suggests that varying the orientation of a biological motion display, both by rotation in the plane and in the depth, affects visual recognition. The perception of biological motion is therefore thought to be orientation specific. Verfaillie (1993, 1999, 2000) explains the findings in his studies of in-depth rotated walkers with the theory of orientation specific mental representations, whereas Pavlova and Sokolov (2000) explain the findings of their experiments of image-plane rotated walkers with the theory of how ecological constraints affect human visual processing.

At a first glance these theories might seem to support the same notion of orientation specificity, but a closer look reveals that they contradict each other resulting in an incompatibility (see next paragraph) that may only be solved by either proving one of the theories wrong or by limiting the explanatory power of the theories to the domains within which they have been established. In other words, the theory of orientation specific mental representations is limited to explaining priming effects caused by different in-depth orientations of an upright displayed biological motion and the theory of ecological constraints is limited to explaining priming effects caused by different in-plane orientations of a biological motion. If the latter (domain specific theories) possibility turns out to be true this would suggest that there are two different mechanisms accounting for the different priming effects caused by the two different ways of rotating the point-light displays. One of the mechanisms would account for the priming effects observed when biological motion patterns are rotated in the depth and the other mechanism would account for the priming effects obtained when biological motion patterns are rotated in the plane, both theories could therefore be correct but much more limited in their generalization than what is believed today.

It is obvious that the priming effects obtained in Verfaillie's studies (1993, 2000) cannot be explained by the influence of ecological constraints. A sight of a walking person (in this case a point-light display of a person) facing to the left is no different in ecological terms than a walking person facing to the right. There is no mismatch between perceived kinematics and dynamics to explain why the facilitation effect for recognizing a display of a person is greater when it is preceded by a display of a

person facing in the same direction than when it is preceded by the mirror version of itself.

Maybe not as obvious is the fact that the priming effects found in Pavlova and Sokolov's study (2000) cannot readily be explained by the theory of orientation specific mental representations. Of course it is not natural to see a person walking on the ceiling and it most definitely will give arise to a mismatch between perceived kinematics and dynamics, but that should not matter if the reason for the priming effects seen in priming studies done with rotated point-light displays is caused by orientation specific mental representations. If Verfaillie's theory is right then a priming effect for same orientation should be obtained for perceived biological motion independent of the angle the display is viewed from, i.e. an inverted display should prime an inverted display etc. Pavlova and Sokolov (2000) found only priming effects for upright or near to upright (0-45°) point-light walkers, where an upright prime display had a priming effect on both identical upright walkers and on walkers rotated 45° from upright, suggesting that "… priming effect in biological motion is *partly* independent of the relative orientation of priming and primed displays" (Pavlova & Sokolov, 2000, p. 897). If the angle of view of a biological motion is an essential factor in the way we store mental representations and orientation influences the ease of perceiving and recognizing biological motion then it appears difficult to compromise the notion of orientation specificity for a notion *partly* independent of orientation.

1.3.6 Other factors possibly affecting orientation specificity

The discussion in last section explained how the two theories that have been developed to account for the different effects caused by manipulation of in-depth and in-plane rotations of biological motions, contradict each other, resulting in a need for study to be conducted to clarify the relation between the two theories, that is if they both can be justified. But there might exist other factors interfering with the results of both previously mentioned studies. Both Pavlova and Sokolov (2000), and Verfaillie (2000) only used point-light displays of human engaged in walking. Dittrich (1993) has shown that different actions require different amount of visual processing, resulting in varying recognition rate and reaction time for each action. Dittrich found that locomotory actions (e.g. walking) were recognized better and faster than social actions (e.g. dancing), while instrumental actions (e.g. hammering) were the most difficult to recognize.

Dittrich implicitly showed that properties, embedded in the action a human is engaged in, influence the visual processing of a biological motion. Dittrich used the same two persons to record all 12 different actions he used in his experiment. The point-lights were attached to the same places on the people for the recordings of all the actions (or so I imagine), which should result in the point-light displays for all the actions to show the same figures. It is therefore possible to suggest that it is information inherited in the action that is influencing the visual processing and not the information embedded in the figure itself that emerges through the movements of the point-lights. But how many different kinds of information are embedded in the movement of an action and to what extent they influence processing of biological motion and the production of priming effects is still very much an open question.

Verfaillie (2000) actually indirectly supports the notion of each action being uniquely equipped with a combination of different kinds of information, when he says that in

recognition of point-light displays, not only the figure of the person engaged in the action is identified, but also is the action the figure is performing easily recognized. It would be much more difficult to identify actions if they did not carry unique action properties.

At least one property has been identified and documented by Pinto and Shiffrar (1999). This property is called dynamic symmetry and is a production of the movements of the limbs of certain biological movements. Or as described by Pinto and Shiffrar: "Dynamic symmetry refers to the equal and opposite motions of adjacent limbs (either contralaterally or ipsilaterally)" (Pinto & Shiffrar, 1999, p. 296). This pattern of limb movements is characteristic of human locomotion, when one limb moves forward, its neighbouring limbs move backward, anti-phase to the first limb. Not all actions performed by humans produce dynamic symmetry among their limbs. Walking being one variant of human locomotion produces dynamic symmetry, but not hammering or jumping jacks. Therefore exists a chance that the orientation specific priming effects observed in previously mentioned studies are specific for walking, whereas only point-light displays of walkers were used in those experiments.

1.4 Problem definition

Results of the experiments reviewed here have shown that an entirely viewpoint-independent notion of representing actions is not plausible, or at least not for how the biological motion of a human walker is represented.

It is not impossible that the orientation specificity effect that has been observed in many studies is action dependent. The biological motion of walking has been defined as having dynamic symmetry (Pinto & Shiffrar, 1999) depending on the nature of the movements of the limbs. Not all human biological motions are dynamically symmetric, like jumping a rope or hammering a nail. One of two aims of this experiment is to find out if orientation specific effects found in action recognition might be influenced by the properties of the action under observation.

The other aim is to find some evidence to clarify the apparent incompatibility of priming effects caused by in-plane rotation of biological motion on the one hand and priming effects caused by in-depth rotation of biological motion on the other hand. The results will be compared to Verfaillie's hierarchical theory, and Pavlova and Sokolov's theory of ecological constraints.

1.5 Hypotheses

The above problem definition results in three main hypotheses that will be tested in the following experiment. As mentioned earlier, the backbone of the following research is the conceived mismatch between, on the one hand, Pavlova and Sokolov's (2000) argumentation and, on the other hand, Verfaillie's (1993, 2000) argumentation for orientation specificity. Certainly both Verfaillie, and Pavlova and Sokolov get significant results by manipulating the orientation of a point-light human walker, but the mismatch is mainly found in the way the different authors explain their results in relation to orientation specificity. This mismatch implies that the authors are describing two different phenomena that are affected by manipulation of orientation in two different ways, complicating the notion of orientation specificity in mental representations of biological motion.

The following hypothesis was inspired by the fact that no prior priming study has investigated if the priming effects obtained in previously mentioned studies vary as a function of the action the point-light human is engaged in. Dittrich (1993) experiment shows that some point-light presented actions are harder to identify than others, resulting in more incorrect answers and longer reaction times (RTs) for the hard actions than for the actions easier to identify. Dittrich's work was more concentrated on looking at differences between categories of actions, where each category contains a number of actions, than differences between individual actions. If the three different actions chosen to be investigated in the following experiment turn out to need different amounts of visual processing, resulting in different RTs, it will be interesting to see if correlating differences will also be obtained in the calculations of priming effects or if the priming effects are still the same for the different actions. The following hypothesis regards only RTs for the different actions and does not concern priming effects. The second and the third hypothesis cover priming effects.

Hypothesis one: Point-light displays of a person engaged in different actions will require different visual processing for each action and result in different response latencies for each action.
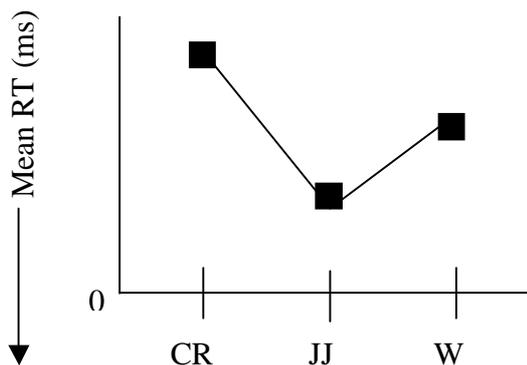
Figure 6. This diagram graphically shows expected results as predicted in the first hypothesis. The squares stand for three different actions, 1) climbing up a rope (CR), 2) jumping jacks (JJ), and 3) walking (W). Reaction time (RT) decreases down the y-axes.

The differences between the mean RTs shown in the diagram are illustrational (see figure 6); no prediction has been made about how long processing time each action needs.

As mentioned before, Pavlova and Sokolov (2000), and Verfaillie (1993, 2000) are getting differential priming effects as a result of manipulating the orientation of point-

light walkers in different ways. Verfaillie (1993, 2000) suggests that humans preserve orientation in their mental representations of biological motion, causing the facilitatory priming effect obtained for the primed trial in a transition of two trials of same orientation. Pavlova and Sokolov (2000), on the other hand, suggest that a mismatch in a motion's kinematic and dynamic information prevents humans from producing facilitation effects for motions shown in other in-plane orientations than upright. They even rule out knowledge-based constraints as being the cause of the results obtained in their long-term priming experiment, suggesting that despite frequent exposure to biological motions in unusual in-plane orientations, people will still be unable to produce priming effects as found for upright motions. These differential findings and theoretical explanations are the basis for the second hypothesis.

Hypothesis two: Transitions between trials of same orientation (the two black squares to the left in figure 7) will have greater priming effect than transitions between trials of different orientations (the two black squares to the right in figure 7). At the same time, no difference in priming effect is expected between the two instances of same orientation transitions (RSU-RSU and INV-INV), nor between the two instances of different orientations (INV-RSU and RSU-INV). Hence no interaction effect is expected as a function of same and different oriented transitions and the two instances of each of those groups.
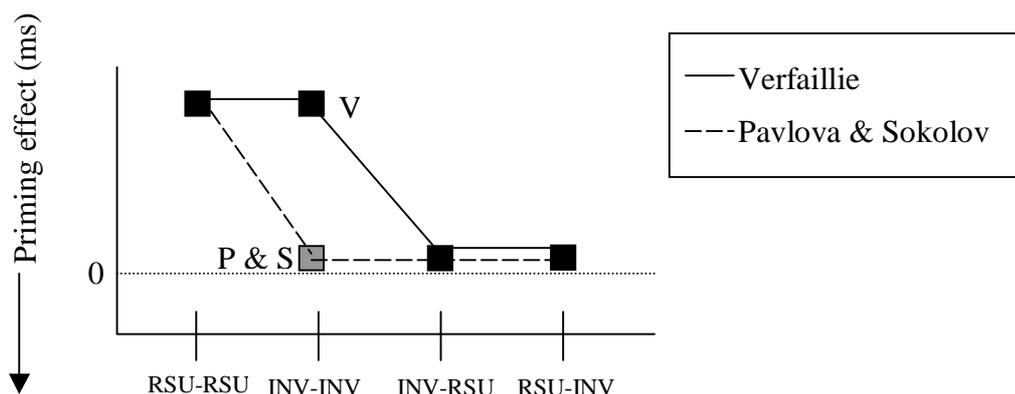

Figure 7. This diagram graphically shows expected results as predicted in the second hypothesis. RSU means right side up and INV inverted. The squares stand for four different kinds of transitions. Priming effect decreases down the y-axes. The zero and the dotted line stand for the neutral baseline(s).

Figure 7 graphically displays the expected results according to the second hypothesis. Hypothesis two is more in concordance with Verfaillie's results and further extending them for inverted biological motions and not only upright. Figure 7 in addition shows possible results of the following experiment according to Pavlova and Sokolov's results and theoretical explanations. If they are correct in excluding knowledge-based constraints, then no facilitatory priming effects should be measured for pairs of inverted human motions even if the same motions produce priming effects in upright orientation. Humans have an incredibly adaptive brain and have great capacities for learning new things. It seems therefore a little bit hasty to rule out knowledge-based constraints as being the source of the results obtained in Pavlova and Sokolov's long-term priming study where the study phase merely lasted 10 seconds. Hypothesis two is more in line with Verfaillie's work as his results stand on a firmer ground (being confirmed in a long-term priming study (Olofsson et al., 1997)), and not in line with Pavlova and Sokolov's work as their reasons for excluding knowledge-based constraints are not convincing enough.

Previously conducted studies investigating priming effects produced by manipulating the orientation of the biological motion display have only used point-light displays of a walker (Pavlova & Sokolov, 2000; Verfaillie, 1993, 2000), or not analysed for differences between different actions when different actions have been used (Olofsson et al., 1997). The results so far, supporting theories about mental representations of biological motion being sensitive to orientation in different ways, can therefore hardly be applied to other actions than walking. The special pattern of point-lights attached to a human walker possibly produces an image in our mind that is more sensitive to orientation than point-light images of other kinds of human motions. This sensitiveness might be produced by features that are special to the motion of walking, like dynamic symmetry. Without trying to pinpoint the source of orientation specific priming effects, it seems reasonable to investigate if different patterns of patch-lights produced by different human motions results in priming effects varying as a function of action, where even some actions do not produce any priming effects. The third hypothesis deals with this matter.

Hypothesis three: Also, due to the different visual processing required for each action, the priming effects will vary as a function of the actions.
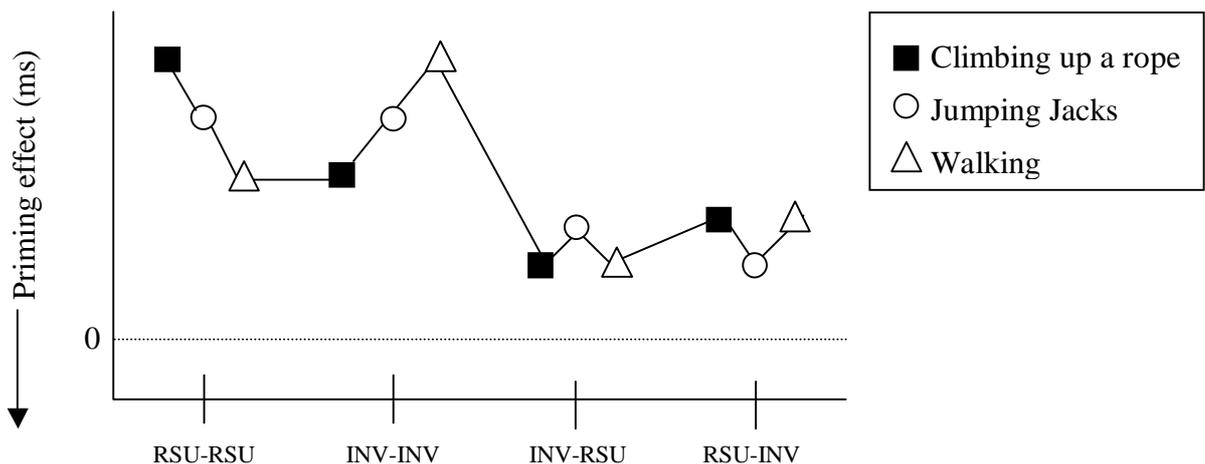
Figure 8. Above diagram graphically shows the expected results as predicted in the third hypothesis. Priming effect decreases down the y-axes. The zero and the dotted line stand for the neutral baseline. RSU, right side up; INV, inverted.

The differences between the actions in each instance of transition as a function of orientation, is illustrational as it has not been predicted how each action effects the time it takes to visually process it.

2 Method

2.1 Choice of method

Verfaillie (1993, 2000), and Pavlova and Sokolov (2000) all describe their findings as being evidence for the theory that mental representations of biological motions are orientation specific. The phenomena they describe are though different form each other. Verfaillie's orientation specificity implies that the most basic property for the organization of mental representations in our heads is orientation. This is illustrated by the fact that orientation is the only factor yet discovered to yield priming effects in both short- (Verfaillie, 1993, 2000) and long-term priming experiments (Olofsson et al., 1997) in visual action recognition studies. These studies manipulated the in-depth orientation of the human figure.

Pavlova and Sokolov (2000) did, amongst other, a long-term priming study, but manipulated the in-plane orientation of the human figure. Also they explain their findings as supporting the notion of orientation specificity. In their case, orientation specificity seems to mean that mental representations are organized and dependent on the matching of dynamic and kinematic information perceived in the biological motion.

These two different phenomena could possibly be coexistent. The grounding belief for the experiment presented in this report is that the above mentioned experiments are describing two DIFFERENT phenomena, justifying the two different theoretical explanations and therefore requiring an empirical study to clarify the notion of orientation specificity.

The method chosen to uncover these differences is a partial replication of Verfaillie's short-term priming studies (1993, 2000). The replication is partial in the sense that the experiment is a short-term priming study and based very much on Verfaillie's design, but at the same time different in number of ways. An important factor in choosing the method was to make the results of the following experiment comparable to both Verfaillie's (1993, 2000), and Pavlova and Sokolov's (2000) results. Partially replicating Verfaillie of course helps compare the different results. The following sections will explain how the experiment is different from Verfaillie, why these changes were made, and how the results of this experiment will also be comparable to Pavlova and Sokolov's results.

2.1.1 Orientation

In a way, the following experiment is also a partial replication of Pavlova and Sokolov's experiments (2000), as the biological motion figure's orientation is manipulated in the frontoparallel plane as in their study. In all of Verfaillie's experiments (1993, 2000), the biological motion figure is rotated in depth, only showing a walker either walking to the right or to the left from the subject's viewpoint.

The reason for changing the way the factor of orientation is manipulated in this experiment is twofold. Firstly to see if the effects observed when manipulating the in-depth orientation of a biological motion displays are also observed when orientation is manipulated in the frontoparallel plane. If inverted point-light displays of biological motion can prime displays of same orientation to the same extent as upright displays

prime displays of same orientation, then more power will be given Verfaillie's theory about orientation being the factor that makes our mental representation less view-invariant. If inverted motions cannot prime motions of same orientation as well as upright do, then it is perhaps time to look for other factors underlying the way we process and store biological motions that are more basic than in-depth orientation. For example, Pavlova and Sokolov's (2000) suggestion about factors involved in ecological constraints, kinematics and dynamics, might be essential in allowing us to identify motions in the first place. Manipulating the in-plane orientation of biological motion displays instead of the in-dept orientation also makes the results more comparable to Pavlova and Sokolov's results as they too used that sort of manipulation in their experiment.

Secondly, Verfaillie has already in his experiments established quite convincing evidence and arguments for a priming effect caused by manipulation of in-depth orientation of biological motion displays. Olofsson et al. (1997) have further extended these findings by obtaining the same effect in a long-term priming study, but the lack of studies investigating the effect of manipulating the in-plane orientation of biological motion displays was the other reason for treating the factor of orientation differently here than Verfaillie did in his experiments.

2.1.2 Action

If the theoretical explanations and empirical findings that Verfaillie (1993, 200) presents regarding his experiments on point-light displays of human walkers also can be accounted for other forms of human point-light motions then the results to the following experiment should show that inverted motions prime inverted motions and upright motions prime upright motions. But, according to Pavlova and Sokolov, an upright point-light motion display should only prime an upright point-light motion display or if inverted (upside-down) primes inverted, that should happen on fewer occasions and with less facilitation effect on reaction time than for upright priming upright.

Predicted results here are that inverted can just as well prime inverted as upright primes upright, at least for the human figure walking. The other factor manipulated in this experiment is the type of action. In previously mentioned studies, the only action investigated was walking. Rotating a walker in the depth might have a unique effect on visual perception and the priming effects observed so far might only be applicable to walking figures. Human beings sometimes engage in actions that produce dynamic symmetry (see section 1.3.6) among the limbs of the body, walking is one of those actions (Pinto & Shiffrar, 1999).

Dynamic symmetry might be the factor contributing to the orientation specific priming effects observed so far. If that is the case, other actions free of dynamic symmetry, should not yield orientation dependent effects. That is why it was decided to try biological motion displays of a person engaged in three different actions. The walker was still used in order to be able to compare more easily the results to previous studies, but a figure doing jumping jacks and a figure climbing up a rope were added to the experiment. Climbing up a rope might also produce some sort of a dynamic symmetry among the limbs, but according to the definition the action of jumping jacks would hardly fit that description. Of course there are other factors that separate these three actions from each other than dynamic symmetry. It cannot therefore be derived that dynamic symmetry caused the orientation specific effects, if it turns out

that the different actions produce different priming effects. It might have been some other factor causing the orientation specificity, but then at least it is clear that the orientation dependent priming effects are somehow action specific.

2.1.3 Task

The task in the following experiment might be categorized as being an object oriented recognition task (see discussion in section 1.3.3), and there have been discussions about if and how that affects the results of studies meaning to investigate action or motion properties and not properties more related to object recognition studies (Olofsson et al., 1997; Verfaillie, 2000). Verfaillie (2000) argued that an action recognition task leads to activation of viewpoint independent representations placed above the basic orientation-dependent representations in the mental representation hierarchy. This notion of hierarchical organization of mental representations is questionable but it was still decided that the argument for using an action oriented recognition task was not strong enough to prevent us from using a less action oriented recognition task. The task was to decide if a figure was shown upright or upside-down on the screen.

2.2 Independent and dependent variables

As mentioned earlier, the following experiment will consider the effect that different orientations of a point-light display of biological motion might have on visual action recognition. Orientation of presented point-light displays will therefore be systematically varied. Independent variable number one, in-plane orientation, will be manipulated resulting in two different levels, biological motion displayed in upright orientation and in inverted orientation. In the same experiment, it will be investigated whether different actions will show different effects in visual action recognition and particularly if there is an interaction effect between different actions and orientations, limiting the generalization of the notion of orientation specificity. The second independent variable therefore consists of the different actions that the point-light figures are engaged in, and has three levels; walking, doing jumping jacks, and climbing up a rope. The two different independent variables and their levels lead to 6 possible conditions: an upright figure walking, doing jumping jacks, and climbing up a rope, and an inverted figure walking, doing jumping jacks, and climbing up a rope.

The experiment is a short-term priming study, which means that a transition of two subsequent trials will be analysed to see if the first trail in the transition has any priming effect on the second trial in the transition. The combination of two trials will result in 36 conditions, though not counting transitions including neutral primes, some of which are used in calculating a neutral baseline (see section 3.1), but the analysis will foremost focus on 12 of those conditions (see table 1).

An example of a transition and how priming effects are measured is a walker shown in upright orientation followed by a walker in an inverted orientation. If seeing a walker in an upright orientation has a priming effect on seeing a walker in inverted orientation then the reaction time for the latter trial should be significantly different from the neutral baseline (RT) for that transition. If the RT is significantly shorter than the RT for the neutral baseline then the first trial in the transition has had a facilitating priming effect on the second trial.

Table 1. The twelve conditions for all combinations of transitions investigated in the experiment. RSU stands for right side up (or upright) and INV stands for inverted (or upside-down).

	Transitions			
Actions	**Orientations**			
Walking	RSU→RSU	RSU→INV	INV→INV	INV→RSU
Jumping jacks	RSU→RSU	RSU→INV	INV→INV	INV→RSU
Climbing up a rope	RSU→RSU	RSU→INV	INV→INV	INV→RSU

Analysing all 36 different transitions would probably lead to interesting results and give more information to work with than what the 12 transitions analysed in this paper will. The main reason for not including analyses on the other 24 transitions is lack of time. Analysis of transitions between trials of same or different orientations will only be carried out within each action (see example in above paragraph), but not between actions. A transition between trials of different actions is for example a pair of trials were the first trial is a walker and the second a jumper, in any combination of orientation. The twelve different transitions that will be analysed in this paper were chosen above the other transition because it is the aim of this study to show that inverted motions can prime inverted motion and eventually to see if different actions require different visual processing, resulting in different reaction times for each action. To find out these things it is not necessary to analyse the rest of the transitions.

For each trial, correct and incorrect responses were registered and the participants' reaction time was measured. The reaction time was the time interval measured between the onset of a motion sequence and the participant pressing a key to respond. The dependent variables were therefore reaction time and correctness of the answers.

2.3 Balancing extraneous variables

Other potential sources for producing systematic variation in the experimental design were identified and balanced. Variables like expectancy, learning, tiredness, and handedness were prevented to have systematic effect by balancing them over subjects or sessions.

To prevent participants from expecting one orientation more than the other, patch-light displays were presented equally often in an upright orientation as in inverted orientation, or 324 times in each orientation in each subsession (843 trials). Stimulus configurations representing a human walking were also displayed as often as configurations of a human jumping or climbing up a rope, or 108 instances of each action in each orientation in a subsession.

Some form of learning was expected and important for the outcome of the experiment. Each participant carried out 5 sessions, each session containing a large number of trials. In doing so many trials, the participants became confident about which key to press for each instance of motion, decreasing the effect hesitation has on measured reaction time. At the same time, learning any sequence of trials was undesirable and therefore the presentation of each instance of any action and orientation was randomised over a whole subsession.

To prevent participants from varying in performance due to their daily pattern of alertness or tiredness, each participant was systematically assigned to carry out a session at different times of the day.

Participants responded to their task by pressing one of two keys on the keyboard. They pressed one of the keys with the index finger of their right hand and the other with the index finger of their left hand. The fact that almost all participants were right-handed could have affected them to automatically want to use the right hand more than the left hand, resulting in more wrong answers for the condition responded to with the left hand. It is difficult if not impossible to prevent this automation or preferentiality. In order to balance this effect, half of the participants used the right hand to respond to upright motions and the other half used the left hand.

2.4 The experiment

2.4.1 Participants

Eight students of the University of Skövde, including the author, participated in the experiment. Four participants were females and 4 were males. All had normal or corrected-to-normal vision. The participants' age ranged from 21 to 24 years, and all but one, were right handed. Participants received a small gift for their participation.

2.4.2 Apparatus

Stimuli were displayed on a Macintosh 17" (33 x 25 cm) monitor set to black and white colour planes and a resolution of 832 x 624 pixels. The monitor operated with a refresh rate of 75Hz. Stimulus generation and presentation was controlled by a Power Macintosh 7100/66Av with a processor speed of 66 MHz. The left- and the right-arrow keys on the computer keyboard were used to mediate the key response to a computer program (DotPlayer[1]) that registered the type of answer and the reaction time (RT) made for each motion. DotPlayer was also used to display the motions on the monitor and is designed to always show all frames. The RT was measured with a temporal resolution of 1 ms and a margin of error of ±4 ms.

2.4.3 Stimuli

There were 8 different stimulus configurations: 3 upright human motions, 3 inverted human motions, and 2 additional configurations that were used as neutral priming stimuli for establishing a neutral baseline (see section 3.1). All patterns consisted of a number of black dots moving on a white background.

The upright human motion stimuli[2] were a patch-light version of the same person (male) engaged in 3 different actions: walking, jumping jacks and climbing a rope. The patch-light displays were created by placing reflective material on the joints and head of a human male wearing dark clothing (Hemeren, unpublished). The person was filmed performing the three actions. A light source was attached to the camera to illuminate the reflective material. The action sequences were then digitised frame-by-frame from the original videotape and re-created in a software animation program. The action sequences were edited so that they would only contain the reflective patches. All other information was removed, including translatory motion.

The number of patch-lights viewed in a motion, and which patch-lights were occluded completely or temporally, varied and depended on the action. All motion stimuli, both human and neutral, were shown in 26 frames. Each frame was viewed for 49.88 ms resulting in a total viewing time of 1296.88 ms for each motion. To prevent participants from recognizing the actions by only perceiving the first frames, each kind of action started randomly at one of three different points in the action cycle. This resulted in 9 files for the upright motions, 3 files for each action. The same files were used for the inverted motions. The orientation of the motions were manipulated in the frontoparallel plane, they were not rotated in the plane though (Pavlova &

---

[1] Thanks to Christian Balkenius the creator.

[2] Originally created by Paul Hemeren.

Sokolov, 2000) but flipped horizontally frame by frame in Macromedia Director™ version 3.13. Flipping a patch-light display only changes the in-plane orientation but rotating it changes the in-depth orientation for some of the actions, like the walker and the climber. There were also three different files for each neutral stimulus (N1 and N2). To create N1, all the bottom halves of each frame of N2 were flipped vertically. The three files of N1 started at the same frames in the motion as the three files of N2.

The actions chosen for the experiment had different properties, such as maximum and minimum extension of height and with, the global direction of the group of patch-lights, and the number of frames needed to complete an action cycle.

As in other studies (Pavlova & Sokolov, 2000; Verfaillie, 1993; Verfaillie, 2000), the patch-light stimulus for a human walking was of a person walking to the right in the frontoparallel plane (see figure 9). The maximum number of patches viewable during the motion was 10, representing the head, the right shoulder, elbow and hip, and both wrists, knees, and ankles. The patches on the left shoulder, elbow were occluded at all times, but the patches on the left wrist, knee and ankle were temporally occluded at the appropriate moments during the step cycle. As mentioned before, all translation of the movement was removed, resulting in the walker appearing as moving on a treadmill. The global direction of the patch-lights involved in the motion was still perceived to be to the right, both for upright and inverted walkers. The walker completed a whole step cycle (two steps) in 26 frames and at the speed of 648 ms per step resulting in a natural looking version of a patch-light walker. The speed would enable the walker to complete 46 cycles per minute, which is within the normal walking speed that ranges from 30 to 70 cycles per minute (Imman, Ralson & Todd, 1981, cited by Pavlova & Sokolov, 2000). Of the three actions, it took the walker the shortest time to complete a cycle, the duration of the other actions was based on that time and all action were shown for 26 frames.
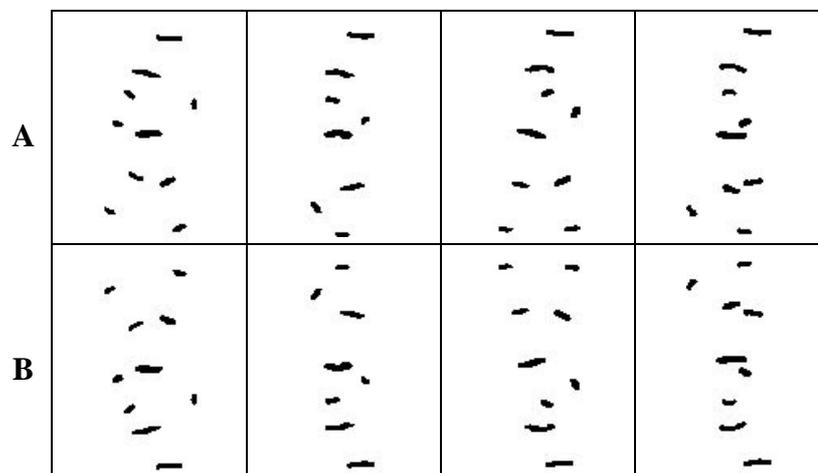
Figure 9. The stimulus configurations for a human engaged in walking. The four different cells in each row contain four frames taken from different moments in the "step cycle". The same moments are shown for upright (A) and inverted (B) walkers.

With participants sitting at a 70 cm distance from the monitor it was decided to try to keep the visual angle constant between 4° and 5° in height. Most motions, though, influence the human posture, affecting the form of the space the human occupies. To somehow control this matter, it was decided that the mean height of each motion should be around 5 cm. The smallest vertical extent measured for the walking figure was 4.93 cm and the greatest vertical extent 5.61 cm. The walker's height ranged

therefore in visual angle from 4.03° to 4.59°. Walking also influences the extent of the width of the perceived figure. The walker's horizontal extent ranged from 1.32 cm to 2.65 cm (1.08°-2.17°).

The patch-light configuration for a human doing jumping jacks was also viewable for 26 frames at the same speed as the walker (see figure 10, A and B). The 26 frames did not suffice to cover the whole action cycle, but was only 2 frames short. The stimulus was composed of 12 patch-lights corresponding to the head, hips (one patch around the hips), shoulders, elbows, wrists, knees and ankles. All patches were viewable at all times. The global direction of these patch-lights was perceived as being slightly upward for the upright jumper but changed for the inverted jumper to downward. The height of the jumper ranged from 4.77 cm to 5.73 cm (3.90°-4.69°) and the width from 1.52 cm to 4.25 cm (1.25°-3.48°).
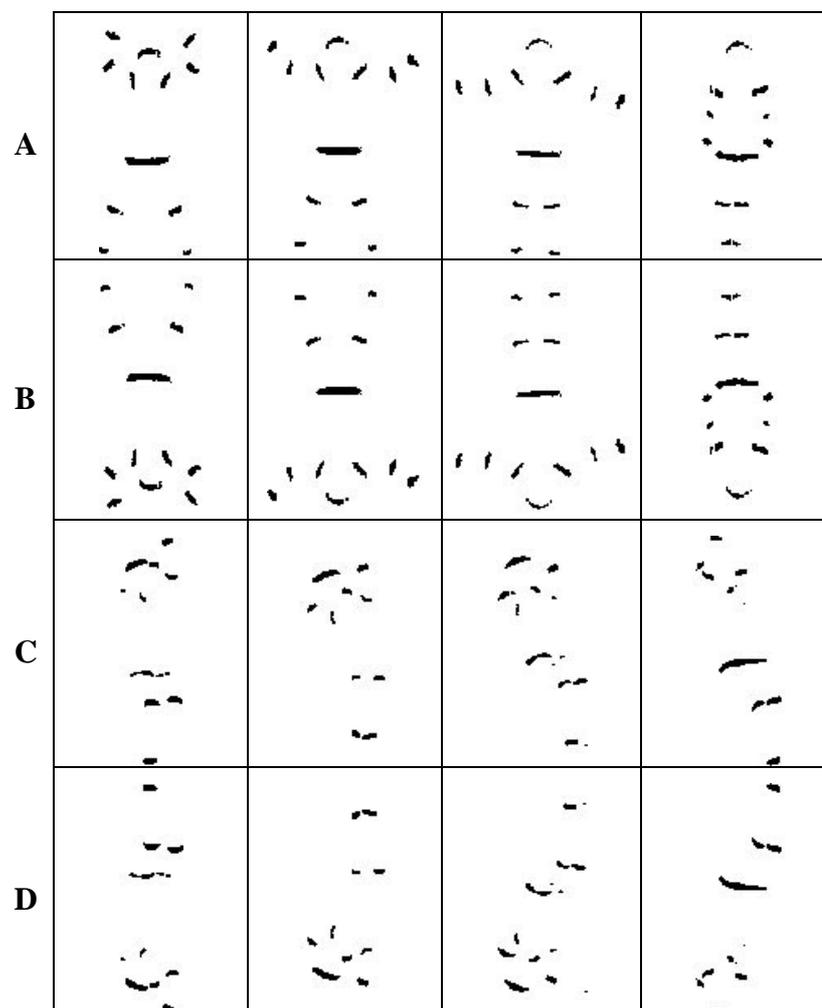
Figure 10. The stimulus configurations of a human doing jumping jacks and climbing up a rope. The four different cells in the two first rows contain four frames taken from different moments in the "jump cycle". The same moments are shown for upright (A) and inverted (B) jumpers. The cells in the two last rows contain four frames taken from the "climbing cycle". The same moments are shown for upright (C) and inverted (D) climbers.

The patch-light configuration for a human climbing up a rope was viewable for 26 frames (see figure 10, C and D). Climbing up a rope is not very easy and therefore the time, which takes to cover one action cycle, can vary from one cycle to the next. The climbing cycle was not completed in 26 frames, but a total viewing time of over a

second gave the participants abundant amount of time to recognize the action given the fact that it only takes a naïve observer a tenth of a second to identify a familiar biological motion presented in a point-light display (Johansson, 1975). The stimulus was composed of 12 patch-lights as in the jumper. The patches on the hips, left shoulder, elbow and ankle, and right shoulder were temporally occluded at the appropriate moments during the climb cycle. As for the jumper, the global direction of the patch-lights changed with orientation and was perceived to be upward for the upright climbers and downward for the inverted climbers. The height of the climber ranged from 4.41 cm to 6.33 cm (3.61°-5.18°) and the width from 1.52 cm to 2.13 cm (1.25°-1.74°).

The abstract stimuli were created with Verfaillie's (1993, 2000) abstract primes in mind. The neutral stimuli consisted of 7 patches where the 3 vertically central points remained stationary. One point was rotating back and forth between 50° and 310° (with 0° at 12 o'clock) above the point in the centre of the 3 stationary points. Another point was rotating below the central point between 130° and 230°. Two additional dots were moving back and forth above and below the previous two rotating points. The two dots above the central point always moved in the same direction and the dots below did too, but the two pairs did not always move in the same direction. The difference between the two abstract configurations was based on the relation between the two pairs moving above and below the central point. In one stimulus (see figure 11, N1), both pairs moved in the same direction (clockwise or counter clockwise). In the other stimulus (see figure 11, N2), the pairs moved in opposite directions.
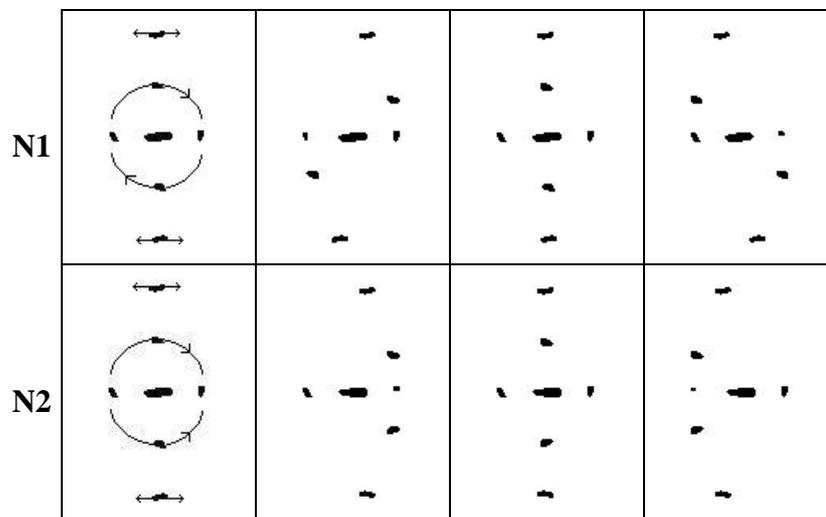
Figure 11. The abstract stimulus configurations that were used as neutral stimuli in the experiment. The four different cells in each row contain four frames taken from different moments in the "action cycle". The same moments are shown for neutral stimuli 1 (N1) and neutral stimuli 2 (N2). The first frame for each neutral stimulus shows the direction of the moving patches at that particular moment.

The height and the width of the global figures of the abstract stimuli were determined according to the mean height and width of the global figures of the other actions. The height remained constant at 5.29 cm (4.33° visual angle), and the width remained constant at 2.41 cm (1.97°). The abstract stimuli cover one and a half action cycle in 26 frames.

2.4.4 Procedure and design

Participants were tested individually, and each of them participated in 5 sessions. The experiment was distributed over 6 consecutive days with each participant only doing one session per day (see appendix 1). Before the first session, each participant was given general information about the reason for the experiment, the procedure of the experiment was explained to them (see appendix 2), general information (age, vision, handedness) about the participants was acquired, and a short practice phase was conducted, lasting approximately 10-15 min, to make sure the participants understood the instructions and especially to give the participants an opportunity to become accustomed to the keys to press for each kind of abstract stimuli. The training session was similar to the subsessions (see next paragraph) in the experiment except that the experimenter was present. It rarely took the participants more than a block of trials to become sufficiently familiar with the task, but all patch-light displays of both human actions and neutral primes were viewed in random order. Before starting each of the subsequent sessions (session 2-5), the participants were able to take a quick look at a sheet explaining what configurations belonged to each key (see appendices 3A-3D) to refresh their memory, but no training was carried out.

Each session was divided into 2 subsessions, where each subsession contained 6 blocks. The first 5 blocks contained 141 trials and the last block contained 138 trials. Each session therefore contained 1686 trials and each participant had to respond to 8430 trials before all 5 sessions were over. Conditions were randomised within a subsession. Before each subsession the participants were instructed to sit approximately 70 cm from the computer screen. After each block, the participants had the opportunity to take a break of a couple of minutes without leaving the room. A ruler of exactly 70 cm was present in the room so that the participants could re-measure and make sure that the distance from their eyes to the screen was correct, for example if they wanted to change position or stand up for a while during the break. After each subsession the participants came out of the room to take a break of about 5-10 min while the author set up the next subsession.

The stimuli were presented in the center of the screen, on a white background (see figures 9, 10 and 11) and were viewed binocularly. The room was dimly lit. Stimulus configurations appeared one at a time in a random order. The task for the participants was to decide, each time a figure was presented, whether it was upright or inverted and respond accordingly by pushing one of two keys on a computer keyboard. Each motion was viewable for up to a maximum of 1296.88 ms after which followed a blank screen (the same colour as the background) and stayed blank until the participant responded, but it was suspected that most decisions would be made before the end of the motions. No feedback was given to the participants indicating if they had pushed the correct button or not, but the role of the training phase was to make sure the participants knew when to push each key. Reaction time was measured as the time interval starting at the stimulus onset and ending at the moment a response key was pressed. Immediately after a motion had been responded to, a response-stimulus interval (RSI) of 500 ms was started. During the RSI the screen was white, the same colour as the background was in the point-light figures stimuli.

Half of the participants (2 male and 2 female) pressed the right-arrow key for upright figures and the left-arrow key for inverted figures. For the other half of the participants this mapping was reversed. Moreover, the key to press for the abstract stimuli was varied within each of these two divisions of participants, resulting in a total of 4 different combinations of pressing a key for a movement (see appendices

3A-3D). Half of the participants pressing the right-arrow key for upright figures (1 male and 1 female) pressed the same key for abstract stimuli 1 (see appendix 3A), the other half pressed the right-arrow key for abstract stimuli 2 (see appendix 3B). Half of the participants pressing the left-arrow key for the upright figures (1 male and 1 female) pressed the same key for abstract stimuli 2 (see appendix 3C) and the other half pressed the left-arrow key for abstract stimuli 1 (see appendix 3D). The key responses were balanced, therefore all participants pressed the left-arrow key as many times as the right-arrow key.

The task only stressed the decision to one stimulus at each time and no retention of previous trials was necessary. During instructions, the participants were also informed that accuracy and speed were important when responding to the trials.

3 Results

The purpose of the above experiment was to investigate visual recognition of different human actions, and whether recognition is affected by varying the in-plane orientation of the patch-light pattern displaying each motion. For clarification, the earlier proposed hypotheses are restated below (see section 1.5 for graphical illustration).

> Hypothesis one: Point-light displays of a person engaged in different actions will require different visual processing for each action and result in different response latencies for each action.

> Hypothesis two: Transitions between trials of same orientation will have greater priming effect than transitions between trials of different orientations. At the same time, no difference in priming effect is expected between the two instances of same orientation transitions (RSU-RSU and INV-INV), nor between the two instances of different orientations (INV-RSU and RSU-INV). Hence no interaction effect is expected as a function of same and different oriented transitions and the two instances of each of those groups.

> Hypothesis three: Also, due to the different visual processing required for each action, the priming effects will vary as a function of the actions.

The independent variables manipulated in the experiment were the in-plane orientation of the figure engaged in different actions and the type of action. The patch-light display mediating the biological motion either showed the figure move in an upright orientation or in an inverted orientation. The patch-light figure showed a human engaged in three different actions: climbing up a rope, jumping jacks and walking. In addition to the human motion stimuli, there were two versions of abstract stimuli, also called neutral primes. There were two dependent variables. The first dependent variable was measured RT (ms) for each trial, the trial consisting either of an abstract stimulus or some combination of the three actions and the two orientations (6 combinations). The second dependent variable was the correctness of the answer to each trial; the answer was either correct or incorrect.

The above review renders the analyses in the following sections easier to follow. The analyses are divided into two sections. The first section covers analyses examining the mean RT for each version of a human action, these analyses do not include any analyses concerning priming effects. The second section comprises analyses of priming effects, where the influence of a priming trial on the RT for the immediately following primed trial is examined. In both sections, the results are presented in relation to the working hypotheses and each analysis is justified in accordance to what questions are to be answered. Next section, though, is not directly connected to answering the hypotheses, it explains the purpose of the neutral baselines, how they were calculated and relevant analyses. But before any analyses are presented it should be made clear what responses were included in the analyses.

For each participant, the first session, the first two blocks of the rest of the subsessions and the first five trials in each block served as training and were not included in any analyses. For the remaining trials (34,624 trials), all incorrect responses were eliminated and then all the correct answers, irrespective of action or orientation, were used in calculating the mean RT and the standard deviation for each participant. A cutoff value was then calculated for each participant, set at three

standard deviations above the mean RT. All responses exceeding the cutoff value were also eliminated. The analyses presented in the following three sections only included correct RTs not exceeding the cutoff value (93% of the trials, or 32,111 of 34,624). For the neutral baseline analyses and the priming effect analyses, where transitions of two trials were considered, both responses in each transition had to satisfy those conditions. In other words, only transitions where the response to both priming and the primed figure were correct and did not exceed the cutoff value were included in those analyses.

3.1 Neutral baselines

A neutral baseline is the mean RT for the second trial in a transition where the first trial is an abstract stimulus and the second trial is a human action. Because the first stimulus is not human but abstract, it is supposed to have no priming effect on the RT for the second stimulus, the human action. In other words, the abstract stimulus has a neutralizing effect on the visual processing of the human action, resulting in a "normal" RT that is neither affected by inhibition, nor facilitation of previously processed stimulus. The neutral baseline is therefore used in a comparison to the mean RT calculated for the second trial in a specific transition to see if there was any difference. If there is a significant difference between the neutral baseline and the mean RT for that type of transition the first trial in the transition is believed to have primed the second trial.

Verfaillie (1993, 2000) calculated two different neutral baselines for each participant, one for transitions requiring the same key response (right-right or left-left), and one for transitions requiring different key responses (right-left or left-right). Verfaillie had, however, only one action, walking. As three different human actions were used in the above experiment it was decided to calculate two different neutral baselines for each action, resulting in 6 neutral baselines for each participant. For example, two neutral baselines were calculated for the walkers, one baseline where both trials required the same key response, and another baseline where the two trials required different key responses. The former baseline was used to calculate priming effects for transitions in which participants made the same response to priming and primed walking figures (i.e., when priming and primed walkers had the same orientation), and the latter baseline was used to calculate priming effects for transitions in which participants made a different response to the priming and primed walking figures (i.e., when priming and primed walkers had different orientation).

Table 2. The mean RTs (ms) for 6 different neutral baselines, three neutral baselines for same-response transitions and different-response transitions respectively, one for each action: climbing up a rope, jumping jacks and walking.

Action	Transitions	
	Same-response	Different-response
Climbing up a rope (CR)	460	445
Jumping jacks (JJ)	438	425
Walking (W)	450	443

The table above (table 2) presents the mean RTs for each baseline, calculated across subject. The difference between the neutral baselines is not great; the greatest difference is only 35 ms between CR same-response (460 ms) and JJ different response (425 ms). The neutral baselines were subjected to two analyses of variance, one for the three baselines for same-response transitions and one for the baselines for different-response transitions, to see if there was the difference was significant. The six different neutral baselines would only be used if there would be a significant difference between the neutral baselines for the different actions, otherwise two baselines would be sufficient, one for same-response transitions and one for different-response transitions.

The average RTs for the neutral baselines, computed for each subject, were subjected to an ANOVA as a function of action (CR, JJ, and W). The same-response neutral baselines and the different-response neutral baselines were analysed separately. The difference was statistically reliable for same-response baselines, $F(2,14) = 14.728$, $MS_e = 63.589$, $p < .001$, as well as for different-response baselines, $F(2, 14) = 11.361$, $MS_e = 84.399$, $p < .002$. Analytical comparisons for the same-response baselines showed that all actions were significantly different from each other, while for the different-response baselines only the neutral baseline calculated for jumping jacks was significantly different from the other actions. It was therefore decided to use 6 neutral baselines, two for each action.

3.2 Analyses testing the first hypothesis

The first hypothesis stated that RTs would vary as a function of action due to different visual processing needed for each action. The mean RTs calculated for each combination of action and orientation (see table 3) indicate that a figure doing jumping jacks is somewhat easier to recognize than the other actions. Even though, overall, inverted motions took more time to respond to than upright motions, the jumping jacks motion was responded to fastest compared to the other motions, both for inverted figures (JJ = 424 ms vs CR = 443 ms and W= 442 ms) and for upright figures (JJ = 398 ms vs CR = 418 ms and W = 413 ms).

Table 3. The mean RTs (ms) for each action as a function of orientation.

Action	Orientation	
	Inverted	**Right side up**
Climbing up a rope (CR)	443	418
Jumping jacks (JJ)	424	398
Walking (W)	442	413

To find out if the RTs vary as a function of action and orientation, a 2x3 repeated measures ANOVA was performed on the mean RTs for orientation (right side up vs inverted) and action (climbing up a rope vs jumping jacks vs walking). The mean RT for the six human motion conditions was computed for each subject and used in the analysis. The analysis showed a significant main effect of action, $F(2, 14) = 24.205$, $MS_e = 76.143$, $p < .001$, supporting the first hypothesis. The latencies to figures doing jumping jacks (M = 411 ms) were shorter than to walking figures (M = 427 ms), and to figures climbing up a rope (M = 431 ms), but the difference between a walker and a

climber was vanishingly small (4 ms). Analytical comparisons showed that figures doing jumping jacks were responded to significantly faster than to both walking figures, $F(1, 7) = 17.683$, $MS_e = 507.714$, $p < .005$, and to figures climbing up a rope, $F(1, 7) = 44.181$, $MS_e = 289.714$, $p < .001$. The 4 ms difference between the mean RT for walkers and the mean RT for climbers was not significantly different, $F(1, 7) = 2.907$, $MS_e = 116.286$, $p > .13$.

The above described ANOVA showed also a significant main effect of orientation, $F(1, 7) = 16.670$, $MS_e = 524.762$, $p < .006$, where upright motions were responded to faster (M = 409 ms) than inverted motions (M = 436 ms). For the interaction between action and orientation there was not a significant difference, $F(2, 14) = 1.077$, $MS_e = 21.583S$, $p > .36$, indicating that the significant difference between the RTs for the different actions is not influenced by the orientation of the figure performing the actions. This further supports the first hypothesis stating that patch-light displays of a person engaged in different actions will require different visual processing depending on the action, resulting in different RTs.

3.3 Analyses of priming effects

This section includes all analyses concerning priming effects, or those analyses needed to be computed to find out the validity of the second and the third hypothesis. The second hypothesis predicted that transitions including trials of same orientation would produce greater priming effects than transitions including trials of different orientations. Figure 12 shows the difference between the RT for the primed human action in the 12 experimental transitions and the RT for the primed human action in the appropriate neutral baseline, averaged across subjects. The difference is an estimate of the facilitatory or inhibitory priming measured in milliseconds.

As the figure shows, the bars are presented in groups according to the orientation of both priming and primed trials in the transitions. The first group of bars shows that transitions made of two trials of upright motion stimuli have a facilitation effect between 77 and 88 ms. The facilitation effect is found through calculating the difference between the mean RT of the primed human action in the transition pair of interest (for example inverted walker followed by an inverted walker) and mean RT of the appropriate neutral baselines (in this case a same-response neutral baseline for a walker). There is not a great difference between the different actions in this group (CR = 78 ms, JJ = 77 ms, W = 88 ms), nor for any other group for that matter (see table 4).

Table 4. The priming effects, facilitatory and inhibitory, calculated as the difference between the mean RT of the primed human action belonging to a certain transition, and the mean RT of the appropriate neutral baseline. RSU stands for right side up, INV for inverted and times are shown in ms.

	Transitions			
Actions	**RSU-RSU**	**INV-INV**	**RSU-INV**	**INV-RSU**
Climbing up a rope (CR)	78	46	3	15
Jumping jacks (JJ)	77	47	0	17
Walking (W)	88	42	-3	9

The second group of bars shows facilitation effects for transitions, each made of a pair of inverted stimuli. The facilitation effect is quite smaller than for the first group, contrary to the second hypothesis' prediction, though still amounting to a maximum of 47 ms (jumping jacks). The third group of bars shows hardly any priming effects, the maximum reaching to 3 ms (climbing up a rope), and walking figure even showing inhibitory priming effects of –3 ms. This shows that upright motions have little effect on following motions if they are inverted. The fourth group of bars where priming motions were inverted and primed motions upright, shows a facilitatory priming effect of up to 17 ms (jumping jacks).
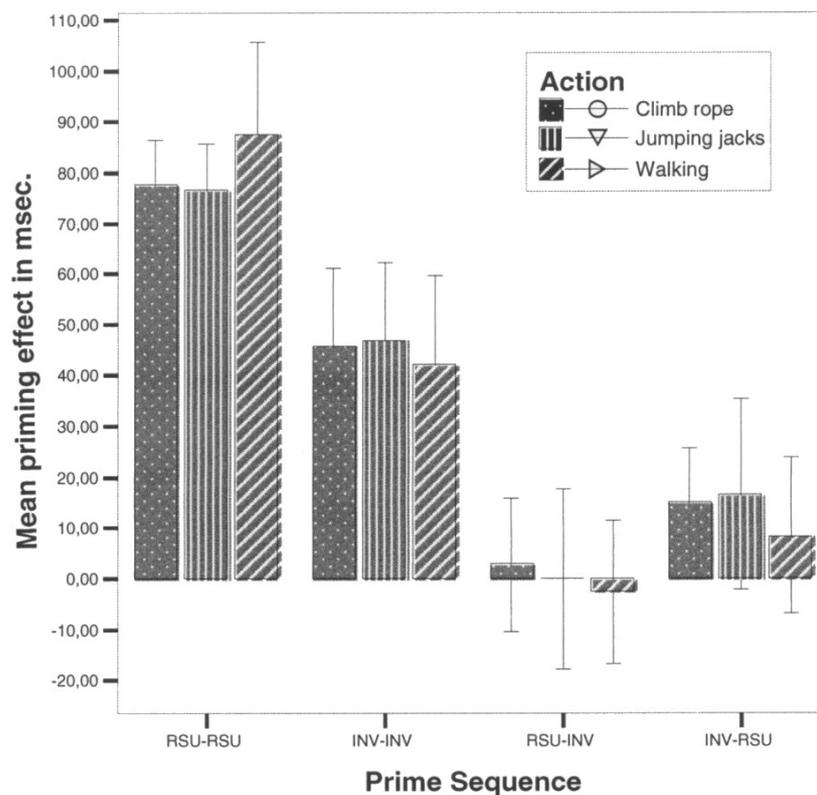

Figure 12. Mean facilitation and inhibition in relation to the appropriate neutral baseline as a function of orientation and action between priming and primed human actions; bars indicate 95% confidence intervals. RSU stands for right side up and INV for inverted.

To see if the unexpected difference between the same orientation upright transitions facilitation effect and the same orientation inverted transitions facilitation effect was significant, and to test the validity of the other predictions made about priming effects, an ANOVA was performed. A 2x2x3 repeated measures ANOVA was performed on the difference between each participant's mean RT for each transition (12) and the mean RT for the appropriate neutral baseline, also calculated separately for each participant. The independent variables were orientation correspondence (RSU-RSU and INV-INV vs RSU-INV and INV-RSU), orientation of the priming trial (RSU-RSU and RSU-INV vs INV-INV and INV-RSU) and action (climbing up a rope vs jumping jacks vs walking).

The analysis yielded a significant main effect of orientation correspondence, $F(1, 7) = 334.089$, $MS_e = 225.784$, $p < .001$, verifying the first part of the second hypothesis. The figure moving in the same orientation as the figure in the preceding trial was responded to faster than when the figure in the preceding trial was of different orientation (M = 63 ms vs 7 ms). The main effect of orientation of the priming trial was also significant, $F(1, 7) = 24.648$, $MS_e = 121.868$, $p < .003$, meaning that there was more facilitatory priming when the primed figure was preceded by a motion of an upright figure than when the primed stimulus was preceded by a motion of an inverted figure (M = 40 ms vs 29 ms). The main effect of action was not significant, contrary to the prediction in the third hypothesis, $F < 1$.

The interaction between orientation correspondence and orientation of the priming trial showed reliable difference, $F(1, 7) = 15.757$, $MS_e = 918.951$, $p < .006$. This significant interaction meant that the last prediction in the second hypothesis, stating that no interaction effect would be found as a function of same and different oriented transitions and the two instances of each of those groups, was wrong. Analytical comparisons were conducted to localize the possible roots of the interaction effect, but figure 12 gives a good hint towards that answer. The analytical comparisons showed that there was significantly greater facilitatory priming effect for transitions of upright priming and primed trials than for transitions of inverted priming and primed trials, $F(1, 7) = 38.343$, $MS_e = 2399.929$, $p < .001$. Also these results contradict the predictions of the second hypothesis, where no difference was expected between the two instances of same oriented transitions. In addition, no difference was predicted between the two instances of different oriented transitions and no significant effect was found, $F(1, 7) = 3.350$, $MS_e = 3844.982$, $p > .11$.

The unexpected significant interaction between orientation correspondence and orientation of the priming trial will be further explained in the discussion section (see section 4.4). The way of calculating the neutral baselines is seen as a potential source of biasing the results and therefore causing the significant difference observed between priming effects for same oriented inverted transitions and same oriented upright transitions.

The meaning of the significant interaction between orientation of the priming trial and action, $F(2, 14) = 4.398$, $MS_e = 54.760$, $p < .034$, is not known and is unimportant in relation to the hypotheses proposed and will therefore not be discussed further. The interaction of orientation correspondence and action, and the three-way interaction of orientation correspondence, orientation of the priming trial and action were not significant, $F(2, 14) = 1.419$, $MS_e = 143.757$, $p > .27$ and $F(2, 14) = 1.119$, $MS_e = 79.933$, $p > .35$ respectively.

Finally, F tests were done to see if the computed priming effects illustrated in figure 12 were significantly different from zero (i.e., the neutral baselines). Like the confidence interval bars in the bar chart indicate, the priming effects of same-oriented transitions and transitions of inverted priming climbers and upright primed climbers were significantly different from the zero.

4 Discussion

Putting the results in context of the theories presented earlier in this paper, subsequent sections will deal with different perspectives of it. The incompatibility described (section 1.3.5), between Verfaillie's theoretical explanations for orientation specificity on the one hand, and Pavlova and Sokolov's theoretical explanations on the other hand, will be discussed and evaluated as much as possible. Like in any other research, more questions arose under the process than got answered, the last section of this chapter deals with factors complicating the interpretation of the results and other problems with the experiment.

4.1 The impact of the results on Verfaillie's theory

Similarly to Verfaillie (1993, 2000) the above experiment produced orientation specific priming effects. Same-oriented transitions produced much greater priming effects than different-oriented transitions did, and in contrast to Verfaillie (2000), only same-oriented transitions produced priming effects significantly different from zero. This part of the results not only supports Verfaillie's theory, but also enhances it, in the sense that it has been shown that the theory applies for inverted human motions as much as upright human motions. Verfaillie's theory becomes even more generalizable as the different actions do not produce different priming effects even if there is a significant difference between the reaction times for some of the actions. The results are however not entirely according to Verfaillie's recipe. Following discussion will shed a light on how the results presented above can be interpreted as supporting Verfaillie's (1993, 2000) findings and how the results contradict them.

The fact that both same-oriented upright transitions and same-oriented inverted transitions produced significant facilitation effect for the primed trials supports Verfaillie's findings. The notion of orientation specificity that Verfaillie and other scientists have adopted, suggests that of all information available to humans through visual processing of a biological motion, the orientation of the motion is the part of the information retained for the construction of the mental representation of that motion. Constructing a mental representation of a motion with orientation as a crucial factor means that each motion creates more than one orientation-dependent representation in our minds, or as many representations as the number of different orientations we have experienced the motion in.

Verfaillie's experiments (1993, 2000) only investigated facilitation effects caused by different in-depth orientations of an upright human walker. It is normal to see people walk in different directions and we see people walk all the time, with their feet on the ground and their heads above their feet. The results of the study presented in this paper showed too, that right facing upright displayed human walkers prime identically oriented walkers, and in addition showed that even inverted human walkers facing to the right can prime identically oriented walkers. This shows that despite the "unnaturalness" of seeing a human motion presented in other in-plane orientations than the orientation the motion normally is performed in, humans can still construct mental representations of it. The results here have therefore further strengthened the theory Verfaillie is a spokesman of, in showing that the theory holds for inverted human motions as much as for upright human motions.

The results of the above experiment have enhanced the theory of orientation-specific representations in other ways than for different in-plane orientations of a human walker. The experiment contained three different human motions, climbing up a rope, jumping jacks and walking, which all produced the same pattern of priming effects in all four different transitions (RSU-RSU, INV-INV, RSU-INV, INV-RSU) despite the fact that significant difference was obtained between the mean RTs for the three actions. The patch-light displays of a human doing jumping jacks was responded to significantly faster than to the other actions, climbing up a rope and walking, in both inverted and upright views. This shows that the facilitation effects obtained by Verfaillie (1993, 2000) are not only observed for patch-light displays of walkers, but of human doing jumping jacks and climbing up a rope too. This fact increases the probabilities for finding this effect in all human motions, though further research should be conducted with different actions before it can be confirmed.

The fact that jumping jacks was responded to significantly faster than to the other two actions but that it, at the same time, did not produce significantly different priming effects compared to the other actions in any of the four kinds of transitions, suggests that orientation dependent priming effects are produced independent of the amount of visual calculations needed to process a human motion. In other words, if the significant difference found between the mean RTs for the different actions is a result of the actions being more or less complicated to compute visually (implying that jumping jacks is less complicated to compute visually than climbing up a rope and walking), then this difference in computation does not appear to effect priming effects due to orientation. Research on yet other actions, varying in different ways, should reveal if this is the case.

The part of the results that contradicts or perhaps gives a new dimension to Verfaillie's theory is the fact that even though same-oriented inverted transitions produced a significant priming effect it was significantly less than the priming effect produced by same-oriented upright transitions. The impact these findings make on Verfaillie's theory depends on how they are interpreted but in the least they give arise to some interesting questions and, in my opinion, weaken the power of the theory that orientation is the primary factor in organizing mental representations.

These results might be showing that upright motions, or motions seen in their "natural" in-plane orientation, prime better because our mental representations of these motions are stronger in some way, due to the fact that we naturally experience them more often than we experience these same motions in any other in-plane orientations, inverted being one of those other in-plane orientations. If lack of experience is the only reason for this difference in facilitation effect between upright and inverted motions then increased exposure to inverted motions should eventually eliminate this difference, increasing the facilitation effect inverted motions can produce. On the other hand, these results might be caused by some other factor that prevents us from creating mental representations of inverted motions of the same "quality" as for upright motions.

Whatever the reason is behind this difference (assuming it is not caused by faulty calculations) it weakens Verfaillie's theory a bit. There appears to exist another factor involved in creating mental representations more basic than orientation. Orientation is still important in the creation of representations but another factor seems to organize these representations on an even more basic level, in this case grouping representations of upright motions in one group and inverted in another.

4.2 The impact of the results on Pavlova and Sokolov's theory

Even if the results at a first glance seem to favour Verfaillie's theory more than Pavlova and Sokolov's theory, then a closer look changes the view. If the basic factor organizing all our mental representations is orientation then there should not be any difference between the priming effects caused by same-oriented upright transitions and the priming effects caused by same-oriented inverted transitions, but according to the results there was a significant difference between those instances. This difference can be interpreted as to support Pavlova and Sokolov's findings, but at the same time contradict some of the explanations they give for their findings.

In Pavlova and Sokolov's long-term priming experiments (2000), inverted walkers did not have any priming effect on identical trials while upright walkers primed upright walkers and even walkers rotated up to 45° from upright. Walkers rotated 45° from upright also primed identically oriented walkers and upright walkers, though significantly less than upright primed upright. The scientists interpreted these findings as being a strong support in favour of the primacy of ecological constraints (dynamic constraints in this case) in biological motion processing.

Pavlova and Sokolov (2000) concluded that the reason for their findings was involved in a perceived mismatch between the event kinematic and natural dynamic information in a point-light display. Dynamic constraints reflect sensitivity to a match between event kinematics and dynamics. A mismatch arises when a human motion is rotated in the plane to any degree that is different from the original or normal in-plane orientation of the motion. The mismatch is supposed to impair visual processing of biological motion, causing the prevention of producing priming effects and impeding spontaneous recognition of motions in orientations from inverted to 90° (12 o'clock is 0°). These dynamic constraints are though graded in their influences on processing of biological motion, because despite this mismatch being perceived in displays rotated 45° from upright, they still produce priming effect, and displays rotated up to 60° are still spontaneously recognized (Pavlova & Sokolov, 2000). Pavlova and Sokolov further ruled out knowledge-based constraints from influencing the results as they expected that priming effects should be observed with all the prime orientations if knowledge-based constraints were decisive in biological motion processing.

If the second hypothesis, stating that inverted motions can produce as great facilitation effects as upright motions, had been confirmed in the results presented above then that would have contradicted Pavlova and Sokolov's findings. The second hypothesis was not confirmed but the results still impose a need for reconsideration of the explanations given for the findings. The results showed that same-oriented inverted transitions produced only half as great priming effect compared to the same-oriented upright transitions. These results correspond to Pavlova and Sokolov's findings in the way that upright motions produce strongest priming effects compared to motions in other in-plane orientations. The results further suggest that there is something more basic than orientation that interferes with the construction of representations of biological motion. But different from Pavlova and Sokolov, it is impossible to establish only one possible factor causing inverted motions to prime less than upright motions.

Knowledge-base constraints are again in the picture as being a potential factor for causing these results because, different from Pavlova and Sokolov, transitions of inverted motion did produce significant priming effect. This shows that even though a mismatch is perceived between kinematic and dynamic information in biological

motion display a representation can still be created and eventually produces priming effects. The reason for the difference between upright priming effects and inverted might be dynamic constraints as suggested by Pavlova and Sokolov, but further research has to be conducted to confirm that. In the meantime knowledge-based constraints cannot be eliminated from being a potential basic element in creating mental representations of biological motion. Research involving frequent exposure to motions in different in-plane orientations should reveal if training is all that is required to enable inverted motions to produce as great facilitation effect as upright motions.

If future research will reveal that training has a key role in creating representations then it might be suggested that Pavlova and Sokolov's (2000) findings reflect the task's degree of difficulty employed in the experiments. The study phase in the long-term priming experiments Pavlova and Sokolov carried out was 3 seconds in one experiment and 10 seconds in another. The effect they got of upright motion producing the most pronouncing facilitation effect is then probably caused by the fact that we already have representations of those motions since earlier, experiencing humans walking around us all the time. Also implying that the 3-10 seconds study phase was not enough to create representations of motions displayed in an in-plane orientation different from normal.

4.3 Coexistence of the two theories?

The impact of the results on Verfaillie's theory, and Pavlova and Sokolov's theory has been discussed in the last two sections. Both theories explain in different ways orientation specificity in biological motion processing. The discussion showed that the results support both theories in different ways, showing that transitions of same-oriented motions produce facilitation effects, both for motions rotated in depth and in plane, but also showing that there appears to exist a factor, primary to orientation, interfering with the processing of biological motion. The latter is demonstrated by the fact that inverted motions do not produce facilitation effect to the same degree as those of upright motions. The question remains open as to what factor produces this difference between facilitation effects of upright, or normal motions, and motions viewed in other in-plane orientations. Potential factors discussed are dynamic constraints and knowledge-based constraints, but future research will decide if either one of those two factors is involved or if the third factor will be discovered.

Even if the factor causing the above mentioned difference couldn't be determined in this report, a coexistence of the theories is indicated in the results. The results indicate that Pavlova and Sokolov have found a factor that affects processing of biological motion more fundamentally than orientation. Whether it is varying amount of experience of the motions in the different in-plane orientations, mismatch between perceived kinematic and dynamic information of motions in orientations other than upright, or something else that causes inverted motions to prime less than upright motions, it seems clear that mental representations are orientation specific in the sense that only transitions of identical orientation, regarding both in-depth and in-plane orientation, produce priming effects. This kind of orientation specificity is strongly related to the theory described by Verfaillie (1993, 2000), only that it appears to be less fundamental in organizing mental representations of biological motion than described by Verfaillie.

Figure 13 illustrates the coexistence of the theories, or a possible interpretation of the findings in the present study, Verfaillie's (1993, 2000) results, and Pavlova and Sokolov's (2000) results. The illustration indicates that our visual system rates our mental representations by familiarity, or groups them by "normality", but that within each potential group there exist different representations of each motion according to the in-depth orientations that we have experienced the motion in. In this sense all different findings can be accounted for, but by what factor our representations are grouped, how they are grouped, and how many groups there are, only future research can answer.
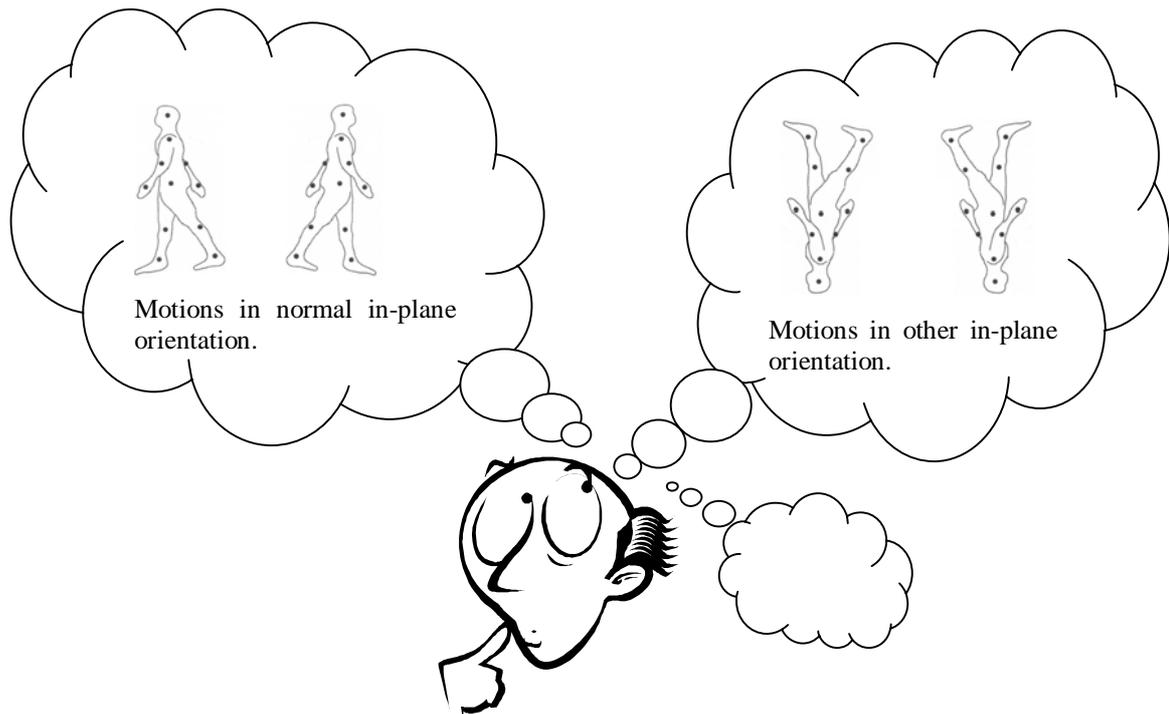

Figure 13. The illustration views a hypothetical explanation of the coexistence of the Verfaillie's and Pavlova and Sokolov's theoretical explanations. Representations for motions are grouped in different groups depending on in-plane orientation, one group for motions of normal orientation and another group for motions of other in-plane orientations. The extra cloud leaves space for alternative explanations.

It remains also an open question as to what extent the motions are represented in our minds. Is every action represented individually or is the orientation of the figure engaged in the action only retained in the representation? The latter alternative should make possible for a range of similar motions to prime each other as long as they share the same absolute orientation, e.g. an inverted display of a human running to the left could prime an inverted display of a human walking to the left. Further analyses of the data collected in the experiment presented above, might reveal that a right facing climber can prime a right facing walker, while the jumper should not succeed in priming anything beside itself, because it is of a different in-depth orientation, facing forward. But then again, climbing up a rope and walking are perhaps not the most similar actions a human can produce, so even here there is a source for new experiments.

4.4 Problems with the experiment

Even in experiments conducted with best of care, where all details are thought of, problems tend to arise complicating the interpretation of the results and weakening the conclusions. This experiment is no exception, though in this case analyses of the raw data were maybe more of a problem than the design or the procedure of the experiment. Following paragraphs will mention practical limitations that rendered the procedure of the partial replication less similar to Verfaillie's design than wished for, and a misunderstanding regarding the role of the neutral baselines, possibly leading to faulty calculations of those baselines, influencing the difference between the facilitation effects of inverted and upright motions.

The part of Verfaillie's (1993, 2000) design that couldn't be replicated was feedback. Due to limitations in the software (DotPlayer) used for presenting the trials of patch-light motions, no sound files could be used. Verfaillie used two different kinds of tones as feedback that informed the participants of the correctness of the answer they gave for each trial. Immediately after a participant's key response followed a 50 ms tone, a high-frequency tone for correct responses and a low-frequency tone for incorrect responses. Feedback can be important in different ways, first of all it tells the participant without a doubt when he responds correctly and when he responds incorrectly. This gives the participant a chance to realize to which trials he tends to respond incorrectly and concentrate on those trials in order to improve himself and get more correct answers. Secondly, feedback can have boosting effect on participants' motivation. It is quite tedious to respond to 843 trials in a row, despite 5 short breaks, if no encouragement is given along the way. The smallest feedback can be experienced as a kind of encouragement, keeping the participant motivated from the first trial to the last.

The fact that no kind of feedback, as to whether a trial was responded to correctly, was given in the experiment presented in this paper might have lowered the motivation of the participants. Low motivation might then have caused more incorrect answers than would otherwise have been found, as well as uneven distribution of incorrect answers. If a participant has from the start responded incorrectly to a certain motion, then there is nothing to inform him about that, which could result in biased results. On the other hand, the fact of hearing a tone 843 times in a row might drive a person insane. An analysis of incorrect answers for each participant would further reveal if any pattern of lack of motivation exists (probably showing proportionally most incorrect answers towards the end of a subsession compared to the beginning), or if any condition of display (e.g. inverted walker) has suspiciously many incorrect answers compared to the other displays.

The misunderstanding regarding the role of the neutral baselines is related to the effect a neutral baseline has on the facilitation effect calculated for each instance of transition. At first, only two different baselines were calculated for each subject. These baselines were calculated as described in Verfaillie (2000), one baseline for transition of trials requiring the same response and another baseline for different responses. Verfaillie, however, only used one kind of action, walking, while three different actions were used in the present experiment. This fact lead to the thinking that perhaps two neutral baselines would have to be calculated separately for each action, same-response walking transitions, different-response walking transitions etc. This was done and decided that if these neutral baselines turned out to be significantly different than we would use two baselines for each action, otherwise the first two neutral baselines would be used for all actions. Analyses showed significant

difference between neutral baselines as a function of action, meaning that all six different baselines were computed for each subject and used in calculating the priming effects.

At this point no more thought was given towards the calculations of the neutral baselines and all grounds were believed to be covered as different actions had been accounted for. It was not until a significant main effect of orientation on RTs was reported that it was discovered that some grounds might have been missed. Verfaillie (1993, 2000) only studied upright figures, and even though priming effect was reported to be orientation specific, no significant difference was found between left facing primes and right facing primes. That fact is probably the reason for why he did not have separate neutral baselines for left oriented and right oriented walkers.

The results presented in previously in this report showed that inverted trials were responded to significantly slower than upright trials. Using the RTs for both inverted and upright trials in calculating one neutral baseline might bias the facilitation effects that are based on the difference between a specific kind of transition (e.g. a transition of upright walkers) and an appropriate neutral baseline (in this case same-response walker, including RTs to both inverted and upright walkers). Because inverted motions had overall greater RTs than upright motions, combining these RTs in one neutral baseline means that an average of these RTs is calculated, balancing out the RTs characteristic for the different in-plane orientations. Using these neutral baselines in calculating priming effects will then favour transitions including upright motions in the priming trial in the sense that facilitation effect becomes greater. At the same time, facilitation effects for transitions with inverted motions in the priming trial will become smaller. The neutral baselines have therefore biased all priming effects calculated and is a potential factor causing the significantly different priming effects between same-oriented upright transitions and same-oriented inverted transitions.

In other words, if separate baselines would have been calculated for both orientations, the difference between same-oriented upright transitions and same-oriented inverted transitions might not have been found to be significant, completely contradicting Pavlova and Sokolov's findings and strengthening Verfaillie's theory. Different neutral baselines would certainly diminish the difference between previously mentioned transitions, but there might be a difference nonetheless and the difference could still be significant. Further analyses on the data should reveal if any of these speculations are true and these analyses would have been conducted and reported in this paper if time would have allowed.

5 Future research

The whole report has mainly concentrated on showing that there is something more basic interfering with visual processing of biological motion than the absolute orientation of the figure engaged in the motion. However, another issue, lightly touch on in the literature study, is the possible influence of the character of the task on the results. Verfaillie (2000) discusses that the genre of the task can possibly influence the focus participants have on the information embedded in the point-light displays. He implies that some tasks make the participants focus more on the object properties of the figure emerging from the moving dots than on the action properties of the motion creating the figure.

Verfaillie concluded that the task in his 1993 experiments, discriminating between a human and nonhuman walker, was more of an object recognition genre than of an action identification genre. He therefore replicated one of his experiments (Verfaillie, 2000), only changing the task, now participants had to discriminate between forward articulating walkers and backward articulating walkers. Interestingly this change also altered the results compared to the results in the 1993 experiment. Now transitions of trials of both identical in-depth orientation and direction of articulation produced the most pronounced facilitation effect, while transitions of trials of same in-depth orientation and different direction of articulation produced significantly less priming effect. Verfaillie still interpreted these results as supporting his theory of a hierarchical scheme representing biological motion where access to orientation-dependent representations is crucial to any further processing of the motion.

The task in the experiment reported in this paper involved discriminating between inverted and upright displayed biological motions. Criteria for how to categorize tasks into action-oriented and object-oriented tasks have not been established and it is therefore hard to say which kind of a task was conducted in this experiment or if it belongs to a whole new kind of tasks. Future research should focus on clarifying what is meant by object recognition tasks and action identification tasks. When that is clarified then maybe research can be done to investigate if different kinds of tasks really have significantly differential effects and if they have differential effects it would be of utmost interest to know how and why.

Similar experiment to the one described above could be used, but instead of using a whole lot of different actions, different groups of participants could be administered different kinds of tasks with the same stimuli. The results would then show if priming effects were produced as a function of the task type.

Following related questions might get answered through similar experiments. Is it possible to design an action recognition task that is not dependent on the orientation of the agent engaged in the action? Or in other words are there any action parameters that can be recognized that are not dependent on orientation parameters?

References

Bechtel, W. & Graham, G. (eds) (1998) *A companion to cognitive science*. USA: Blackwell Publishers.

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

Blanz, V., Tarr, M. J. & Bülthoff, H. H. (1999) What object attributes determine canonical views? *Perception, 28*, 575-599.

Bruce, V. & Green, P. R. (1990) *Visual perception: Physiology, psychology and ecology* (2nd ed.). London: Lawrence Erlbaum Associates Ltd.

Daems, A. & Verfaillie, K. (1999) Viewpoint-dependent priming effects in the perception of human actions and body postures. *Visual Cognition, 6*, 665-693.

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

Gregory, R. L. & Colman, A. M. (eds) (1995) *Sensation and perception*. New York: Addison Wesley Longman Limited.

Hemeren, P. E. (1997) *Typicality and context effects for action categories using patch-light displays of biological motion.* Unpublished manuscript. The Department of Cognitive Science, Lund University; and the Department of Computer Science, The University of Skövde.

Hoffman, D. D. (1998) *Visual intelligence: How we create what we see*. New York: Norton.

Johansson, G. (1975) Visual motion perception. *Scientific American, 232*, 76-88.

Johansson, G. (1973) Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 15*, 201-211.

Lawson. R. (1999) Achieving visual object constancy across plane rotation and depth rotation. *Acta Psychologica, 102*, 221-245.

Lundh, L.-G., Montgomery, H. & Waern, Y. (1992) *Kognitiv psykologi*. Sweden: Studentlitteratur AB.

Milner, A. D. & Goodale, M. A. (1996) *The visual brain in action*. Oxford: Oxford University Press.

Olofsson, U., Nyberg, L. & Nilsson, L.-G. (1997) Priming and recognition of human motion patterns. *Visual Cognition, 4*, 373-382.

The Oxford English Dictionary, vol. VIII. (1933) Oxford: The University Press.

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

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

Sumi, S. (1984) Upside-down presentation of Johansson moving light-spot pattern. *Perception, 13*, 283-286.

Tarr, M. J. & Bülthoff, H. H. (eds) (1999) *Object recognition in man, monkey, and machine*. Cambridge, Massachusetts: The MIT Press.

Verfaillie, K. (1992) Variant points of view on viewpoint invariance. *Canadian Journal of Psychology, 46*, 215-235.

Verfaillie, K. (1993) Orientation-dependent priming effects in the perception of biological motion. *Journal of Experimental Psychology: Human Perception and Performance, 1*9, 992-1013.

Verfaillie, K. (1999) Viewpoint-dependent priming effects in the perception of human actions and body postures. *Visual Cognition, 6*, 665-693.

Verfaillie, K. (2000) Perceiving human locomotion: Priming effects in direction discrimination. *Brain and Cognition, 44*, 192-213.

Wilson, R. A., & Keil, F. C. (eds) (1999) *The MIT encyclopedia of the cognitive sciences*. Cambridge, Massachusetts: The MIT Press.

Appendices

Participants – Key combinations and Schedule

Participant	Right-arrow key	Schema	Participant	Right-arrow key	Schema
1	RSU R N2 R	8. 13:00 9. 20:00 10. 20:00 11. 08:30 12. 15:30	5	RSU R N1 R	7. 12:30 8. 18:30 9. 09:00 10. 11:00 11. 14:00
2	INV R N1 R	7. 14:00 8. 11:00 9. 15:30 10. 13:00 11. 09:30	6	RSU R N2 R	7. 17:00 9. 19:00 10. 09:00 11. 15:00 12. 11:00
3	INV R N2 R	8. 17:00 9. 14:30 10. 10:00 11. 10:30 12. 16:30	7	RSU R N1 R	7. 15:30 8. 19:30 9. 17:00 11. 13:00 12. 10:00
4	INV R N1 R	7. 10:00 8. 14:30 9. 18:00 10. 19:00 12. 20:00	8	INV R N2 R	8. 15:30 9. 10:00 10. 14:00 11. 16:00 12. 09:00

Instructions for the experiment

Welcome and make yourself comfortable.

Note participant's age, handedness and visual capability.

Thank you for volunteering to participate in my experiment. Before the actual experiment begins I am going to explain to you *1)* what I am studying, *2)* what your task is, *3)* the structure of the experiment, and then we will end this introduction with *4)* a short training session.

1) I am studying how human beings visually perceive biological motion. In this experiment I am focusing on how we perceive normally presented motions (people doing motions in upright position) versus how we perceive unnaturally presented motions (motions originally done in upright positions mechanically inverted).

2) Your task in this experiment is to categorize different kinds of motion into two different categories. You are going to see a total of 8 different motions presented in random order on the computer screen. For 4 of these motions you press the left-arrow key and for the other 4 motions you press the right-arrow key.

   Then show the appropriate illustration sheet to the participant, there are four different according to the different possible combinations.

   You are going to see 3 different kinds of actions performed by a human, a walker, a person jumping and a person climbing a rope. Each of these motions is presented in two different orientations, in an upright orientation and in an upside-down orientation. The figures on the illustration sheet are though different from the ones you are going to see in the way that you will not see any outlines. Instead you will see moving dots that represent patches attached to the head and the10 main joints of the body; the shoulder, elbow, wrist, hip, knee and ankle.

   The two remaining motions are not performed by a human. They are called abstract motions and cannot be naturally categorized as being either upright or inverted. For one of those motions you will press the same key as you do for upright stimuli and for the other abstract motion you will press the same key as you do for the upside-down stimuli. This might seem confusing at the moment but I will explain everything better in the training session.

3) Today is the first time you will do a session but you will need to do 4 more sessions, resulting in a total of 5 sessions. There will only be one session per day per participant but the 5 sessions will be carried out in 5 consecutive days.

   Each session is divided into 2 subsessions. Each subsession is divided into 6 blocks. Each block contains around 140 motions and after each block you will

have the opportunity to have a short break. A screen will pop up saying "Press C to continue" and the experiment will not continue until you press "C". It is good to use these breaks to stretch your legs, blink your eyes, have some candy or change your posture if there is anything annoying you.

After the sixth block, a screen will pop up saying "End of motion –1". That means that the first subsession is over and you may come out of the room. I will be sitting here, just outside and there will be a longer break of 5-10 minutes. I will use some of that time to set up next subsession.

4) Now we will go through a short training session and if there are any questions you are welcome to ask. I will begin to show you and name the different actions we see, say if they are upright, inverted or abstract and what key you should hit for each motion. Right from the beginning you will train to use the index fingers to press the keys. Use the index finger of the left hand to press the left-arrow key and the index finger of the right hand to press the right arrow key. In carrying out the task it is important to respond correctly and as quickly as possible to each motion.

When you feel comfortable identifying the different actions you will take over pressing the keys and please say aloud for each motion which key you press so I can give you feedback for if you are wrong or right. We will train like this until you feel certain about when to press which key for all the motions.

There will not be any more training sessions but before each remaining sessions you will have the opportunity to have a quick look at the sheet with your key-motion combination to refresh you memory. During this first session you are allowed to have this sheet with you so you can refer to it if you forget which key to press.

Now if you are ready then I will show you how to sit in front of the computer, turn off the ceiling lights and leave the room. If you have a mobile telephone, please have it turned off during the session

You will have to sit in a distance of 70 cm from the computer screen. I use this measuring equipment that is exactly 70 cm long to check the distance between your eyes and the screen before each subsession. If you stand up or change the way you sit in the chair during subsession then please use the equipment to make sure the distance is correct. You can move the keyboard forwards or backwards on the table in front of the screen to adjust to your arms but try to keep the arrow-keys as centred compared to the screen as possible.

If you do not have any more questions and you are sitting comfortably, I will turn off the ceiling lights and leave.

Good luck

.

Illustration sheet for RSU R N1 R

Participant presses the right-arrow key for upright motions and neutral stimulus 1.

Right key

Left key

Illustration sheet for RSU R N2 R

Participant presses the right-arrow key for upright motions and neutral stimulus 2.
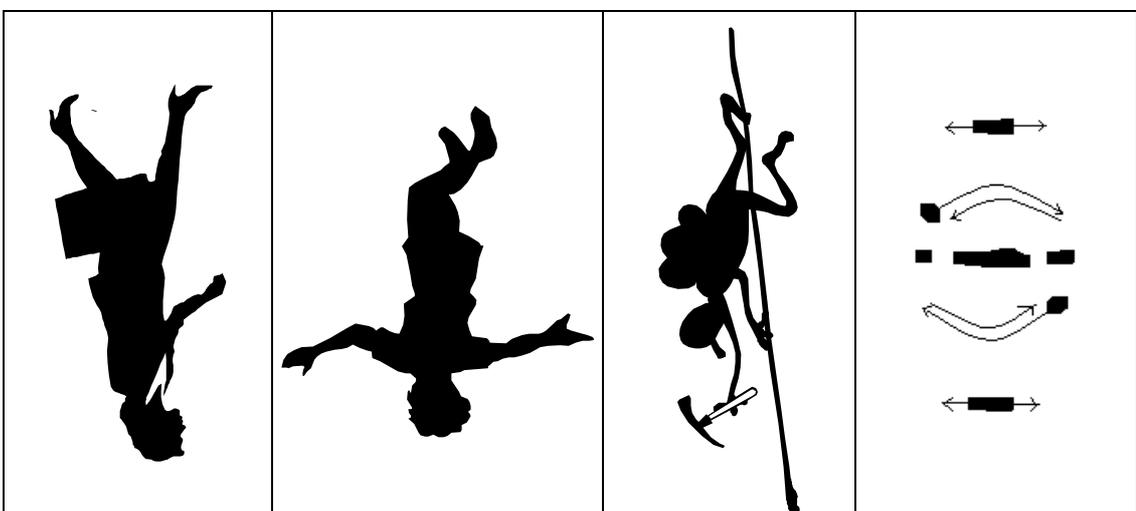
Right key

Left key

Illustration sheet for INV R N1 R

Participant presses the right-arrow key for inverted motions and neutral stimulus 1.

Left key

Right key

Illustration sheet for INV R N2 R

Participant presses the right-arrow key for inverted motions and neutral stimulus 2.

Left key

Right key