What to do if I can do what you can do?

Sander Looise

January 3, 2011

Bachelor thesis
Written by Sander Looise

Supervisors:
Janny Stapel, Iris van Rooij and Sabine Hunnius

Department of Artificial Intelligence,
Radboud University
Donders Institute for Brain, Cognition and Behaviour

Abstract

Action prediction is a complex and interesting quality from human beings. It is necessary to live in a normal way with other living creatures and even robots in the future. It is commonly believed that motor resonance (i.e., the overlap in neural activity during action execution and action observation) plays a role in action prediction. This idea motivates the hypothesis that action prediction may be more effective when the action one observes resembles actions oneself could perform. This hypothesis is tested by comparing the action prediction of infants when they look at someone walking the way they do to (i.e., other infants) or when they look at someone who walks differently (i.e., adults). This experiment is to find out how the human action prediction system works, so we know more about it and can use it for creating robots. To take all cues besides movement away, there were point-light displays created of walking adults and infants. In this thesis is a new technique described how and why to use point-light displays. Eye tracking was used to measure the action prediction during the experiment, the results were to discuss action prediction theories.

1 Introduction

1.1 Action prediction in humans

When a thirty-year old man is walking down the street with normal speed, step size and direction the most adults are capable of predicting where the man is going. When adults see an infant of 2 years walking over the street, this action prediction seems harder because infants of that age are sometimes walking in no specific direction. The infants are able to walk for just a few months and they are walking to whatever their attention is drawn. An interesting question for the looking behaviour of infants is how infants themselves see this difference in walking between adults and infants of their own age. How do we learn to predict each others actions? A hypothesis about action prediction is that it may be more accurate when the action one observes resembles actions one oneself could perform. Adults are worse in predicting the action of an infant than predicting the same action an adult is performing. This could be because infants are hard to predict, but maybe infants are better the other way around, better in predicting the walking behavior of someone who walks like their own motor behavior, namely an infant of the same age. Builders of robots may learn from the human action prediction system for creating robots who are able to predict other agents' actions.

1.2 The role of motor resonance in action prediction

A commonly proposed mechanism for predicting actions is activation of someone's own motor system during action observation. A widely used term in this activation is motor resonance, which is the overlap in neural activity during action execution and action observation. Research has shown that there is more neural activity visible in those brain regions that are associated with executing actions when someone is looking at a person performing an action when he or she is experienced in executing this action than when he or she is not well able to execute the action. (Aglioti, Cesari, Romani Urgesi 2008) Neuroimaging findings indicate an overlap in neural activation during action observation, execution and simulation (Grezes Decety, 2001) Motor resonance seemed stronger when one has a high level of expertise at performing the observed actions and, or perceives one's own previously performed actions than when someone doesn't have a high level of expertise. (Kilner, Friston Frith, 2007) Fecteau et al. (2004) showed that motor resonance arises already early in development in the brain, in children from 36 months there was some overlap visible. One thought about the relation between motor resonance and action prediction is, when motor resonance is stronger for a certain action, action predicting might be better (see the first perspective in figure 1). This hypothesis is supported by research that showed that actions are not purely coded in terms of visual properties of the observed movement but rather in terms of action goals (Sebanz, Bekkering Knoblich, 2006). The second perspective is that motor resonance is a reflection of predictions made about other people's behavior.

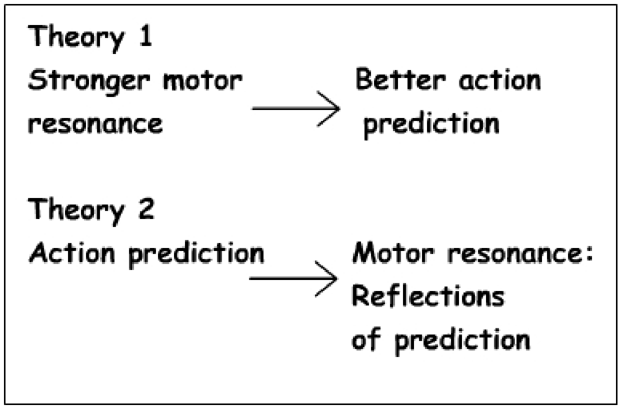

Fig. 1: Motor resonance theories

In this study motor resonance itself is not examined, but it may be an underlying mechanism to explain action prediction. This study focuses more on the action prediction in relation to the motor system.

1.3 Deeper look into action prediction

The ability to predict actions of others is very important for living beings to understand and live with each other. Humans can use prediction to improve a certain movement action or understand others behavior (Knoblich Flach 2003). That is why action prediction is a central issue in experimental and computational studies of motor control. According to these studies the parietal cortex and the cerebellum play a role in action prediction. These parts become active when performing an action prediction experiment. Which role these parts exactly play is not clear yet (Blakemore Sirigu, 2003). Research has already shown that when one is able to perform an action one is also better in predicting how that action will unfold over time (van Elk, van Schie, Hunnius, Vesper Bekkering, 2008). For example when individuals observed a person stacking blocks, their gaze preceded the action and predicted a forthcoming grip, just like when they performed the block-stacking task themselves. Flanagan et al. performed and experiment to test whether this predicting is worse when someone performs a task differently than the way the observer would perform the action himself (Flanagan Johansson, 2003). This research builds further on the finding of Flanagan and Johansson whether someone predicts a walking action better when the movement is like the way he or she would perform it.

1.4 Contribution to other research fields

This study focuses on action prediction by children; the question is if children are better in predicting an action when the action they see lies closely to their own motor system. Because this would provide evidence that being able to perform an action causes better action prediction and possibly stronger motor resonance. Another research goal of this study is to help improve the social interaction of robots. Humans are perfect examples of self-operating systems, for building robots a perfect inspiration. Robots are becoming a bigger part of our contemporary society. Robots can help us by certain actions, like driving, cooking or buying groceries, but it is important that they are able to predict actions else we would have to tell them all of the time what we are about to do. To work with them in the future the robots have to understand somehow the human actions and may be able to predict these actions. Saegusa et al. (2007) showed that a robot could learn to improve the interaction between the robot and its environment, when using an external source of information. But the robot just learns to better predict another agents action when someone else performs this action, perhaps the robot learns faster when he himself is able to perform the action perfectly. Gaussier et al. (1998) tried to let robots imitate each other, but did not really learned the robots to predict each other actions. Maybe our human action prediction system is a good inspiration for building this mechanism into robots. This is another reason for the choice to research the action prediction from children, their action prediction is in full development and so it is easier to see how action prediction develops. Sommerville et al. (2005) showed that infants are better in predicting an action when they just performed the action themselves as when they observed an action without performing it before. This mechanism could be the base for building action prediction in robots, by practicing and seeing other agent's performing certain actions.

When this study shows that action prediction is better when the action is close to the own motor system a recommendation for the robotic part of AI can be made: if we want robots to be able to predict the actions of humans, e.g., by a mechanism similar to how humans do the same, then we would do well to ensure we design the robots to also be able to perform the action we want them to be able to predict. When the results of this study are significant it may also contribute to models about action prediction.

1.5 Research plan

The main research question is whether infants are better in predicting a walking action when the person they're seeing walks just like the way they would walk. The experiment in this study is performed on infants (16 to 18 months), who are able to walk for a few weeks. The infants will watch movies in which infants of the same age walk from one side to the other side of the screen, and of adults walking the same trajectory. The variable tested in this experiment is whether infants are better in predicting the walking actions of someone who walks like themselves (i.e. infants) than predicting the walking actions of adults. Next to this is analyzed if the infants look more at the movies of infants of their own age.

Because it is easy to distinguish infants from adults by seeing their length, clothes and distances between body parts and it is important that infants distinguish between the way of walking and not all this other cues it is needed to take some of these recognizable cues away. So cues such as color, shading and contours are taken away from the movies by creating point-light displays of the original movies (Saygin, Wilson, Hagler, Bates Sereno 2004). In this method just the joints of a body, like ankles, knees, wrists and elbows, are shown, together with the head, shoulders and middle, by some simple dots. Research showed that adults are capable of recognizing a human person and what he or she is doing in the moving dots (Cutting Kozlowski, 1977). Also the gender of the human walker was recognizable for the viewers of the point-light displays. Even emotions from the point-light figures, where only the motion was visible, were quite clear for the viewers (Dittrich, Troscianko, Lea Morgan, 1996). Also newborns seem able to distinguish between a biological movement and some random moving dots (Simion, Regolin Bulf, 2008). This research may contribute to these findings when infants around 1.5 years appear to be also capable of distinguishing between adults and infants, when all cues beside movement are taken away. In the early days, in the experiment from Cutting and Kozlowski, they wrapped reflecting tape around the joints, or put lights on those places, and filmed this to create point-light movies. Nowadays it is easier to just create animated movies, made from normal motion films, by the computer (Shipley Brumberg, 2003). In the study from Shipley and Brumberg study a newly created technique on the computer is used to create point-light displays from movies where adults and infants are walking from left-to-right or right-to-left towards a goal object. When the experiment shows some good result, the point-light displays and the mechanism to create point-light displays may be used for other experiments in different research fields.

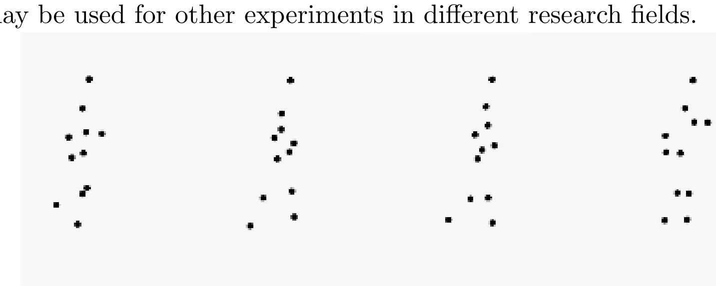

Fig. 2: Point-light movie

The point-light displays of walking adults and walking infants will be shown to infants to gain some information about their interest. In the point-light displays a goal object is created where the walking figures are walking to. The action prediction manifests itself in different ways, one of which is anticipatory eye movements (Falck-Ytter, 2008). By using eye tracking, data will be obtained to see if the infants are better in action prediction when the walking is more similar to how they would walk so when they are watching to the created point-light displays of walking infants. Eye tracking is a widely used method for studying babies or other non-speaking and non-writing humans (Johnson, Amso Slemmer, 2003). Eye tracking already showed that infants are able to predict actions in an early stadium of their development (von Hofsten, Vishton, Spelke, Feng Rosander, 1998). By eye tracking the attention of the infants is visible by following their gaze. If infants look earlier to the goal object when they are watching to the point-light displays of walking infants than when they are watching the point-light displays of walking adults action prediction is better when they see someone close to their own motor system than when they see someone with a different walking pattern. If infants look longer to point-light displays of walking infants than to the point-light displays of walking adults they are more interested in the point-light displays of walking infants. If this hypothesis holds, the theory that action prediction relates to the motor system, the human mechanism of action prediction can be used to inspire us building robots.

2 Method

2.1 Participants

In this study seventeen infants were tested, eight boys and nine girls with no known physical problems. The data from one girl and two boys was excluded from analysis because they were too restless and did not watch enough of the stimuli or sometimes the calibration did not work very well. When there was watched for less than 10 % of the presented stimuli the data was excluded. The data of six boys and eight girls was used for analysis, they were on average 1 month and 30 days old (range 16 months and 18 days to 17 months and 10 days). To recruit infants of this age we used a database from infants to contact parents, who earlier agreed that they wanted to contribute to baby research. It was important that the children could walk for themselves independently, at least know how to walk. The remaining infants followed most of the time the walking figures on the screen. It was not important that they watched all of the stimulus movies because the movies were showed for a long time. So if they watched a quarter of the presented stimuli it was already useful.

2.2 Design

To test the hypothesis a within-subject test with actor type (infants versus adults) as independent variable and percentage of anticipatory eye movements and timing of the eye movements as dependent variables was performed. The infants watched all movies in semi-randomized order. The twelve created point-light displays were randomized, and then four of these randomizations were put after each other so the stimuli consisted of forty-eight point-light displays. There was chosen for not fully randomization because the chance existed that most movies of walking infants were at the end of the stimuli and it could be that the watching infants only pay attention in the beginning of the showed point-light displays. The choice for percentage of anticipatory eye movements as dependent variable was to see if the infants anticipated the goal of the walking infant more frequently than the goal of the walking adult. The timing of the eye movements is useful to see if the infants are better in predicting actions of someone who walks the way they do.

2.3 Stimuli

The experiment consisted of a sequence of forty-eight stimuli of walking point-light figures. The procedure to create this sequence will be described in this section.

2.3.1 Walking movies

For the experimental stimuli, movies of walking adults and walking infants were used to create point-light displays. For the point light stimuli with infant actor, clips were used that depict goal-directed walking in infants (15-18 months), which had been used previously in an experiment by van Elk and colleagues (2008). They found in their research that infants gained stronger motor resonance when watching a crawling action when they learned to crawl for themselves. For the adult stimuli, movies from eight adults were made in the same form as the original infant movies. Stimuli were matched across conditions on gender of the actor, amount of steps, direction (left-to-right and right-to-left were counterbalanced) and duration of the movies. It was important that they walked the same amount of steps in the same time as the infants, because all other information besides the way of walking should be the same for the adult and infant stimuli. In the end there were six movies of infants and six movies of adults used for creating point-light displays, two male adults and two boys walking to the right end of the screen and one male adult and one boy walking to the left end of the screen. The other movies were one female adult and one girl which walked to the right end of the screen and two female adults and two girls walking to the left end of the screen.

2.3.2 Point-light displays

From the twelve movies, point-light displays were created. The technique to create these movies was based on the technique used by Shipley and Brumberg (2003). However, their technique for creating point-light movies was not completely suitable for this study, because it was too time consuming. By trying different methods based on their system, a method was found with which, in a short amount of time and manpower, point-light displays could be created from the walking movies.

2.3.3 Registering joint positions

To create a point-light display of a movie the positions of the person's most important joints needed to be registered. These joints are the ankles, knees, wrists and elbows. Besides these eight points also the position of the head, bottom of the neck and middle of the body are recorded. The recording of these points is done by positioning the mouse on the positions of the joints on a still frame and use a program which shows the coordinates of the mouse position. Then the coordinates could be written down for every frame (see figure 3a-c). To make it less time consuming, but still accurate, in this study, registration of the coordinates of the joints was done every four frames of each movie. The amount of frames depended on the length of the movie. Most of the time the movies were around seventy to ninety frames, thus 18 to 22 frames joints coordinates were stored. So 200 to 250 coordinates were needed to be registered for every movie of walking persons.

2.3.4 Creating point-light displays

After these coordinates had been registered, there was a program needed to create point-light displays from these positions. To do this, Flash (version: MX 2004) was used, which is a program to make animations, videos and web applications. For every joint/body part, a black dot was drawn on every frame on the stored coordinates, like in figure 3d-e. With Flash, frames were automatically created in between the existing frames so the movements looked fluent. In this way there was a point-light display created from each original video recording of a walking person.

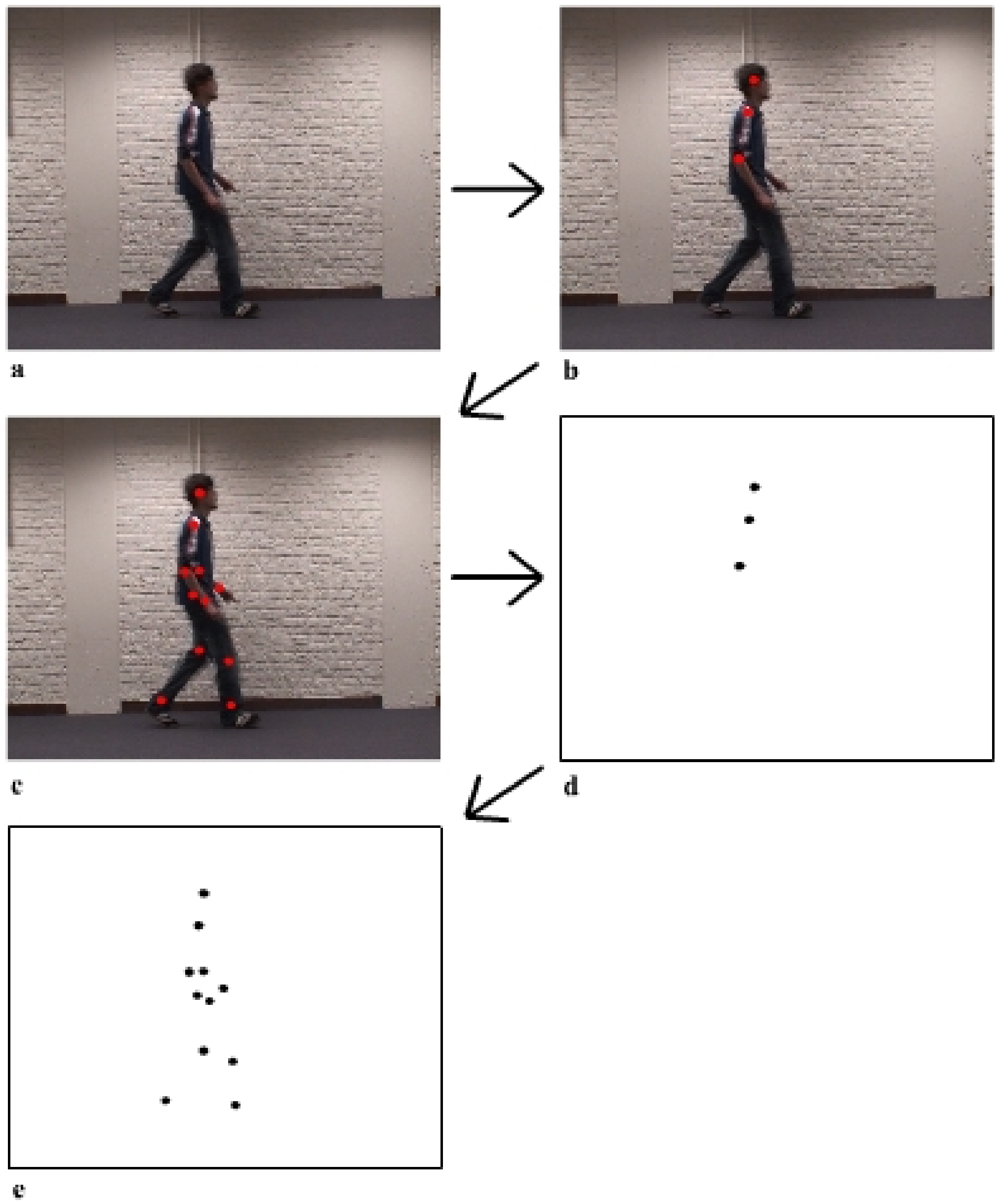

Fig. 3: Creating a point-lights display from one frame.
a: Original frame
b,c: Registering of joints coordinates (red dots are mouse clicks)
d,e: Putting black dots on the registered coordinates to create the point-light displays.

On the side of the screen, where the person was walking to, a goal object was created in the form of a bigger dot on height of the arms. This way it seemed as if the person was walking with the arms towards the dot to do something with it. To take away the difference in length between the adults and infants the length between persons head and feet was measured and the figures were scaled so they all had the same length but kept their own proportions between the body parts. Otherwise it would have been immediately clear whether an animation was an infant or an adult. Finally, the movies were exported into an Audio Video Interleave (AVI) with Flash and compressed by Virtual Dub into a smaller file than the AVI films Flash created.

2.3.5 Final stimuli

Five different sequences were created semi-randomized (not fully randomized because that would give the risk that only point-light displays of adults were visible at the start of the movie) with every point-light display four times so there were five stimuli of forty-eight point-light displays. After every three point-light displays, an attention-grabber (e.g., moving toy, accompanied by a noise) was shown on the screen. Otherwise the infants were likely to get bored by the movies soon.

2.4 Measurements

To measure where the infants looked during the stimulus presentation, an eye tracking system was used. An eye tracker measures the place where the infant looks by using the position of the head in relation with a person's point of gaze (where he is looking), it uses the reflection of an invisible infrared light in the eyes. The eye tracker is built into the screen where the point-light displays were shown.

The eye-tracking system used in this study was a Tobii 1750 from the Swedish company Tobii Technology. The movies were uploaded on the computer and shown to the participants on a 17 inch monitor by using ClearView software.

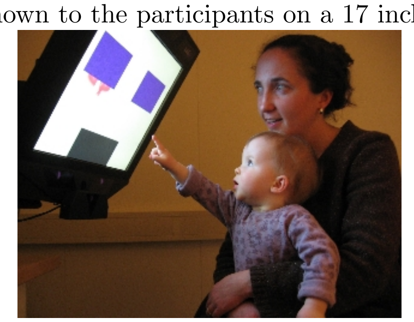

Fig. 4: Eye tracking system

2.5 Procedure

The parents of the children were informed about the procedure of the experiment. During testing, the infants were in an infant car seat on the parent's lap, with minimal distractions from parents or other sources. In the room, there was only the computer with the screen (where the eye tracker was built-in) and some chairs. For the rest there were only curtains and blank walls so the infant was not distracted by other things in the room or outside. Before the point-light displays were shown and the experiment really started there was first a calibration phase. By calibration the eye tracker tries to find the pupils of the infants by showing a 3x3 grid of attention-grabbers on the screen and following the infant's eyes. When on every point on the 3x3 grid the eye gaze of the participant were found, the calibration was approved and the recording could be started. The infant watched the sequence of forty-eight point-light displays. If during the experiment the infant was too distracted by other things and did not look at the screen, the experiment was stopped and sometimes given another try. After the session the infant received a reward in the form of a book or money for the money-box.

2.6 Data analysis

In this experiment, the data analysis took place after the experiments were done. Where the infants fixated on the screen during the stimuli was stored into an AVI file, see Figure 5. This information was also coded into a text file to see exactly where and how long the infants fixated at the screen. Information was retrieved every 20 ms about where the infant looked at the screen. This text file was used to measure whether the infants followed the walking figure and if they anticipated to the goal object before the walking figure arrived at the goal. If this anticipation was the case, there was checked if this anticipation happened early during the movies.

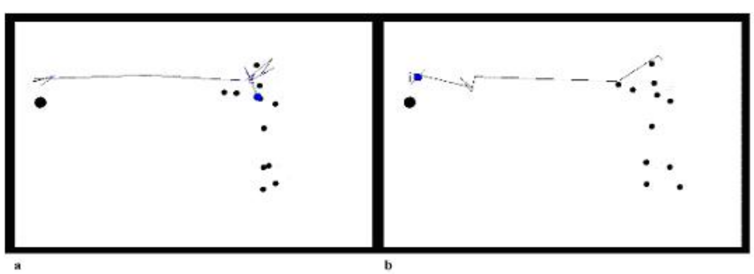

Fig. 5: Fixation on the walking figure, b: Goal fixation

The infants were showed the whole stimulus, if they were distracted or not. The stimulus consisted of forty-eight movies, also named trials in the analysis. For every infant there was counted how many trials they looked at the movie, and how many trials they looked at the goal object. It was important that they really anticipated to the movie, and not just their eyes were flying along the screen. So the infants had to look at least 150 milliseconds and two fixations to be taken into analysis. The counting's were divided into two groups, movies from adult walkers and movies from infant walkers. The percentage of trials looked at the goal with respect to the total amount of attended trials was calculated for every infant that was tested with the eye tracking system. These percentages were used into a t-test as the dependent factor. The adult movies versus the infant movies were the independent within-subject factor.

For every movie (every trial) where the infants fixated to the goal area was measured when the first fixation in the goal area was. This timing of anticipation was averaged by every infant that watched into two categories, the walking adult movies and the walking infant movies. To analyze this effect a t-test was used with point-light displays (made from adults/made from infants) as independent variable and timing of anticipation as dependent variable.

Further it was interesting to see whether the infant were evenly interested in the two types of point-light displays and if they followed the moving dots of the point-light display during the movies. To see this there was information retrieved whether the infants followed the movement of the figures to see if they see were walking figure or some random dots. Also the overall looking time was measured, just as the time they watched at the goal object during the trials.

3 Results

The results of the fourteen participants in this experiment are shown below in Table 1, the number of trials anticipated to the movies and goal objects and in the second table the duration of these anticipations. In table 1, the average number of attended trials and the number of trials in which the infants anticipated the goal is displayed. The anticipations are also divided in anticipations to the adult walking movie and the infant walking movie. The maximum number of trials anticipated was twenty-four for every category. In the second table the mean viewing time in seconds is shown. Here the maximum time the infants watched to the movie was 87 seconds.

	Movies of walking adults				Movies of walking infants		
Participant	Goal object	Movie	Percentage		Goal object	Movie	Percentage
1	3	20	0.15		6	21	0.29
2	1	12	0.08		4	14	0.29
3	11	18	0.61		8	16	0.5
4	3	16	0.19		6	18	0.33
5	7	23	0.30		4	23	0.17
6	0	17	0		1	16	0.06
7	4	21	0.19		4	21	0.19
8	0	8	0		0	7	0
9	6	23	0.26		7	22	0.32
10	1	9	0.11		2	8	0.25
11	2	23	0.09		2	23	0.09
12	8	23	0.35		3	23	0.13
13	2	20	0.1		1	21	0.05
14	6	15	0.4		7	17	0.41
Average	3.86	17.71	0.20		3.93	17.86	0.22

Table 1: Number of trials where was fixated to the screen and anticipated to the goal object

	Movies of walking adults			Movies of walking infants	
Participant	Goal object	Movie		Goal object	Movie
1	0.31	21.45		1.14	19.67
2	0.57	10.69		1.91	15.67
3	5.02	25.19		5.99	20.03
4	1.24	47.08		2.93	50.06
5	1.22	19.26		0.82	17.09
6	0	33.82		0.22	29.89
7	0.87	19.42		1.4	23.20
8	0	4.18		0	3.61
9	4.43	34.74		2.47	34.89
10	0.33	5		0.74	4.00
11	1.57	34.89		0.66	40.15
12	4.76	44.67		1.5	42.52
13	0.63	45.11		0.12	45.89
14	3.58	15.73		1.81	14.62
Average	1.75	25.80		1.55	25.81

Table 2: Length of anticipation towards goal and movie

For every watching infant were two percentages calculated, the number of trials fixated to the goal in relation to the movie. The average of the percentages by watching movies from walking adults was 20,24%, for the walking infant movies this was 21,98%. A t-test for paired samples with type of actor (walking adult, walking infant) as independent variable and percentage anticipations to goal object as dependent variable revealed that the movie category had no significant effect on anticipations to the goal object ($t(13) = 0.57$, $p = 0.58$).

For every trial in which anticipation to the goal took place, was the time of the first fixation to the goal area during that trial was noted. A list consisting of all first anticipations was made; a part of this list is shown below.

Participant	Trial number	Child walking	Adult walking
10	3	0.305	
	4	0.056	
	8		2.619
11	1	2.958	
	2	1.699	
	12		0.903
	10		1.917
12	5	1.864	
	6	0.062	
	2	2.925	
	9		1.841
	12		2.789
	8		1.416
	...		...

Table 3: Time of first anticipations per trial

The timing of anticipation to the goal area from walking infants and walking adults was averaged for every participant. If there was no anticipation for one of the categories, the data from the participant was not taken into analysis, and consequently 12 participants remained.

Participant	Child walking	Adult walking
1	2.86	1.19
2	2.34	3.5
3	2.31	2.14
4	1.61	1.04
5	1.82	1.93
6	1.62	1.85
7	1.75	2.20
8	3.23	0.97
9	1.34	2.4
10	1.94	2.29
11	1.41	2.43
12	2.07	2.21
Average	2.03	2.01

Table 4: Averaged time of first anticipations per trial

A t-test for paired samples with type of actor (walking adult, walking infant) as independent variable and timing of anticipations to the goal object as dependent variable revealed that the movie category had no significant effect on the timing of anticipations to the goal object ($t(11) = 0.05$, $p = 0.96$).

Since there were no significant effects found in the results above, two other tests were done to see if there whether there was no effect of sort walking figure (infant or adult) on the degree of interest of the watching infants. In table 5 the total number of fixations of the watching infants is shown, divided into the categories walking infants and walking adults. In table 6 the averaged fixation duration on the two categories of walking figures is shown.

Participant	Child walking	Adult walking
1	233	262
2	56	44
3	150	231
4	153	94
5	181	216
6	110	115
7	208	181
8	22	33
9	169	169
10	35	40
11	214	235
12	265	278
13	139	163
14	97	102
Average	145	155

Table 5: Total number of fixations on the stimuli

Participant	Child walking	Adult walking
1	0,09	0,08
2	0,3	0,25
3	0,13	0,13
4	0,33	0,5
5	0,09	0,01
6	0,27	0,31
7	0,11	0,11
8	0,16	0,13
9	0,21	0,21
10	0,11	0,13
11	0,19	0,15
12	0,16	0,16
13	0,33	0,28
14	0,16	0,16
Average	0,19	0,19

Table 6: Averaged fixation duration

A t-test for paired samples with type of actor (walking adult, walking infant) as independent variable and total number of fixations on the stimuli as dependent variable showed that the movie category had no significant effect on number of fixations on the stimuli ($t(13) = 1.11$, $p = 0.29$).

Another t-test for paired samples with type of actor (walking adult, walking infant) as independent variable and averaged fixation duration as dependent variable revealed that the movie category had no significant effect on the averaged fixation duration ($t(13) = 0.25$, $p = 0.81$).

Four tests with different dependent variables were executed on the type of actor, the walking adult as point-light display and the walking infant as point-light display. And in not one test was a significant result found, which suggests that there is for the watching infants no difference between the walking adults and the walking children. In the discussion section is described what this could mean.

What seemed to happen during the movies was that infants followed the walking figures during the movies when their attention was alert. This could be important to show that the infants were interested in the movies, and maybe recognized humans in the walking figures. They also anticipated towards the goal object sometimes earlier during the movie, but there was no significant difference in timing of anticipation between the movies of walking adults and the movies of walking infants.

4 Discussion / conclusion

The goal of the experiment was to see if action prediction is influenced by the type of actor whose movements are shown on the screen. To see if humans (infants in this experiment) are better in predicting an action when they see someone perform an action the way they would do it. The movement shown in this study was walking from left-to-right or right-to-left towards a goal. To take every cue besides movement away walking infants and walking adults were changed into point-light displays. The eye gaze data of fourteen infants watching these adults and infants walking gave some results for further investigation. Every result points in the direction that type of point-light display (adult walker versus infant walker) has no significant effect on the action prediction from infants. Also the degree of interest in the movies seems not better when the infants see someone walk the way they do. Also the overall looking time showed no difference for the watching infants. This could have different meanings. First it could mean that the infants are not able to distinguish in the point-light displays between an adult and an infant. But informal observation and debriefing with adults suggests that they are capable of seeing this difference in the different point-light displays. Infants seem to walk more uncontrolled movement with the arms and no regular speed of walking. Adults walk more with a continuous speed, this difference is still visible in the point-light displays. Also the relationships between the different joints are different, for example the distance between the knees and ankles is smaller by infants then it is by adults. Maybe adults are more aware of this differences between infants and adults, and are the infants not able to distinguish between these cues. So maybe there is only a conclusion possible for the infants and we need to test adults with the same movies. Another reason for the absence of an effect could be that infants are not able at all to recognize a walking person in the point-light displays and just saw some random moving dots, but this would be not in line with earlier research (Simion et al., 2008). The third, and maybe most acceptable reason for the result is that the hypothesis, that infants are better in predicting a walking action when they see someone walking like they would do it, is false. Infants are not better in predicting an action when the action they're seeing is similar toto their own way of moving.

For creating social robotics the results from this experiment are not directly useful, because they do not really show how infants predict actions. It is hard to use the action prediction system from humans to build robots, because were not sure yet how the human action prediction system workds It could still be useful to let the robots be able to perform action before building an action prediction system into the robots. But it is not sure if the way they performing the action should be the same how another agent would perform the same action. Another important part of this research were the point-light displays, to research if they were useful as a new research technique. The research showed that it was hard for infants to see differences between walking adults and walking infants in the point-light displays. The cues besides movement were taken away, but it seemed that was taken away too much. Also a new technique was created, based on the technique from Shipley and Brumberg (2003) and used in this research, to create point-light displays. The technique was after some trying an effective way of creating the displays fast and accurate. Despite the absence of significant results a role of point-light displays in future studies is possible. The infants seemed to follow the walking figure most of the time along the screen, only made no difference between the walking adults and walking infants. It is still a good technique to remove all cues besides movement, but it is possible that too much information cues are taken away.

5 Future research

The best follow-up research is to test adults with the same data to see if they are able to distinguish between walking adults and walking infants in the point-light displays and if they are able to predict the action of the walking adult better than the action of the walking infant. Because this would provide evidence for better action prediction when the showed action is more similar to the watchers own way of moving. Also it is possible to do this experiment with adults and infants without the point-light displays, just showing the original movies. The negative part of this kind of research is that other cues like length, clothes and face are still visible for the watcher, so it is harder to know whether the action prediction is related to the movement. Making use of point-light displays is interesting for future research, because the method can be used to take some information from humans or animals away but still let them possibly be recognized in the movies. The technique described in this study is useful but need to be improved. For instance, registration of the coordinates was time consuming and could maybe be made automatic. Also Flash could be replaced by other programs, but for this study it worked well and would be too time consuming to find other methods.

References

[1] Aglioti, S.M., Cesari, P., Romani, M. and Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11(9), 1109-1116.*

[2] Blakemore S.J. and Sirigu, A. (2003). Action prediction in the cerebellum and in the parietal lobe. *Exp Brain Res. 153:239-245.*

[3] Cutting, J.E. and Kozlowski, L.T. (1977). Recognizing friends by their walk: gait perception without familiarity cues. *Bull. of Psychonomic Soc., 9, 353-356.*

[4] Dittrich, W. H., Troscianko, T., Lea, S. E. G., Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25, 727-738.*

[5] Falck-Ytter, T. (2008). Young children with autism spectrum disorder use predictive eye movements in action observation. *Biology Letters, 6(3), 375-378.*

[6] Falck-Ytter, T., Gredeback, G. and von Hofsten, C. (2006). Infants predict other people's action goals. *Nature neuroscience, 9, 878-879.*

[7] Fecteau, S., Carmant, L., Tremblay, C., Robert, M., Bouthillier, A. and Theoret, H. (2004). A motor resonance mechanism in children? Evidence from subdural electrodes in a 36-month-old child. *Neuroreport, 15, 2625-2627.*

[8] Flanagan, J., and Johansson, R. (2003). Action plans used in action observation. *Nature,424, 769-771.*

[9] Gaussier, P., Moga, S., Banquet, J.P., Quoy, M.: (1998). From perception-action loops to imitation processes. *Applied Artificial Intelligence, 1(7), 701-727.*

[10] Grezes, J., Decety, J. (2001) Functional anatomy of execution, mental simulation, observation, and verb generation of actions: a meta-analysis. *Hum Brain Mapp 12:1-19.*

[11] Johnson, S., Amso, D. and Slemmer, J. (2003). Development of object concepts in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences 100, 10568-10573.*

[12] Kilner, J.M., Friston, K.J. and Frith, C.D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processes, 8, 159-166.*

[13] Knoblich, G. and Flach, R. (2003). Action identity: Evidence from self-recognition, prediction, and coordination. *Consciousness Cognition, 12, 620-63.*

[14] Saegusa, R., Nori, F., Sandini, G., Metta, G. and Sakka, S. (2007). Sensory Prediction for Autonomous Robots. *IEEE-RAS 7th International Conference on Humanoid Robots (Humanoids2007), 29th November - 1st December 2007.*

[15] Saygin, A.P., Wilson, S.M., Hagler D.J., Bates E. and Sereno M.I. (2004). Point-light biological motion perception activates human premotor cortex. *Neuroscience, 24, 6181-8.*

[16] Sebanz, N., Bekkering, H., and Knoblich G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Science, 10(2), 70-76.*

[17] Shipley, T.F. and Brumberg, J.S. (2003). Markerless motion-capture for point-light displays. *Technical Report, Temple University Vision Laboratory, Philadelphia, PA.*

[18] Simion, F., Regolin, L. and Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Science, 105, 809-813.*

[19] Sommerville, J.A., Woodward, A.L., Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96, B1-B11.*

[20] van Elk, M., van Schie, H.T., Hunnius, S., Vesper, C. and Bekkering H.(2008). You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy. *Neuroimage, 43, 808-81.*

[21] von Hofsten, C., Vishton, P., Spelke, E., Feng, Q., Rosander, K. (1998). Predictive action in infancy: tracking and reaching for moving objects. *Cognition, 67, 255-285.*

A Small guide to create Point-light displays

A.1 Movies

To create Point-light displays from humans or animals it is important to have clear videoclips. Also the videoclips need to be short, because this method is too time consuming for a longer movie clip. It is important that the most important joints, for humans the elbows, wrists, knees and ankles, as well as the head, shoulders and middle are clearly visible in the clips. Or at least that you are able to guess where these joints are positioned, because they are important to create a recognizable point-light display. When some joints are not well visible on a frame it may be possible to overtake these positions from previously frames, it is easier to guess when you know how fast the persons walks and what a normal way of moving is.

A.2 Registering coordinates

The first step is to open the movies with a video-editing program, for instance Virtual Dub (version 1.9.10, for free on the internet). Then you can see your movie frame by frame. There is also a program necessary to see the coordinates from your mouse on the screen, for instance the Mouse Position (v1.0) application from the google toolbar (see figure 1). It may be even better to find a program that registers for every mouse click your position on the screen.

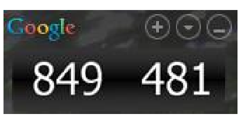

Figure 1: Google toolbar Mouse Position (v1.0) application

For every frame:

- Hold your mouse on every ankle, knee, wrist and elbow and write down the coordinates of these joints
- Hold your mouse on the head, in between the shoulders and on the middle (just above the hips) and write down these coordinates as well. You have now eleven coordinates for this frame

This registering of the coordinates is a lot of work and a movie clip of 4 seconds already consist around the 100 frames, so it is not feasible to do this for every frame. So register this positions for every 5 frames, maybe even every 10 frames. The more frames the better the end result will be, but it is really not necessary.

A.3 Creating point-light displays with Flash (MX 2004)

Create layers (like in figure 2 and 3) for every joint you registered the coordinates from. Every layer stands for a joint or body part. There is in figure 3 also a timeline visible with the frame numbers, here you insert a keyframe (see figure 4) for every frame you registered the coordinates of during step 2. It is important that you create this frames on the same spot where you registered them in the movie. So when you registered the coordinates from frames 1, 6, 11, 16 etc. it is important that you create keyframes on frames 1, 6, 11, 16 etc.

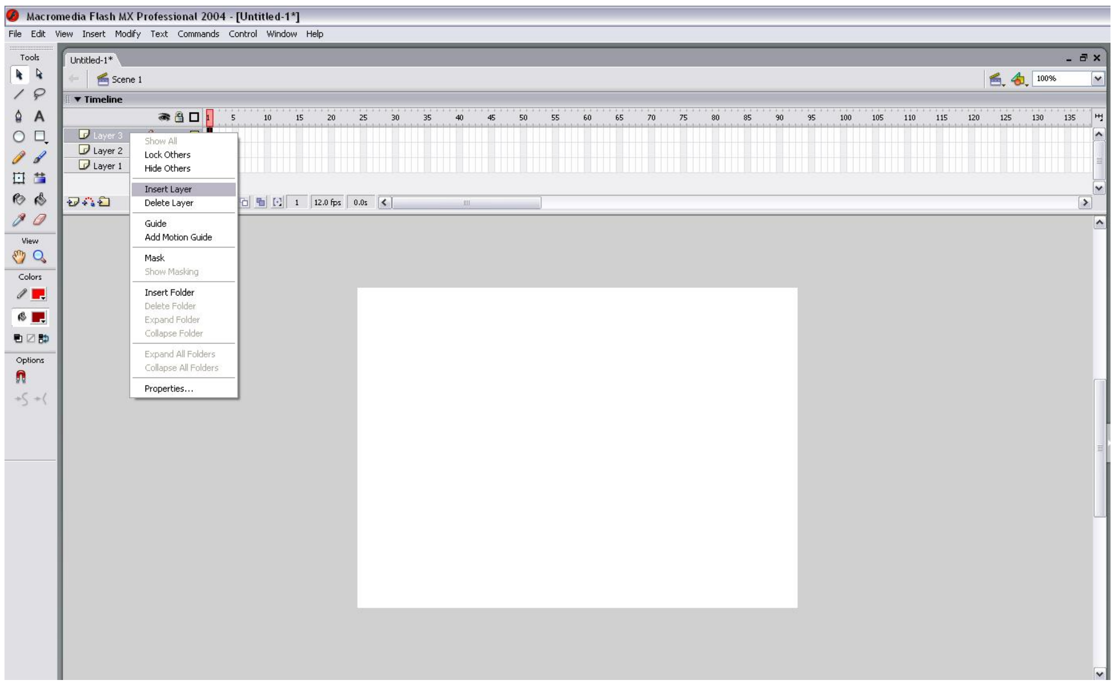

Figure 2: Insert layers for every body part

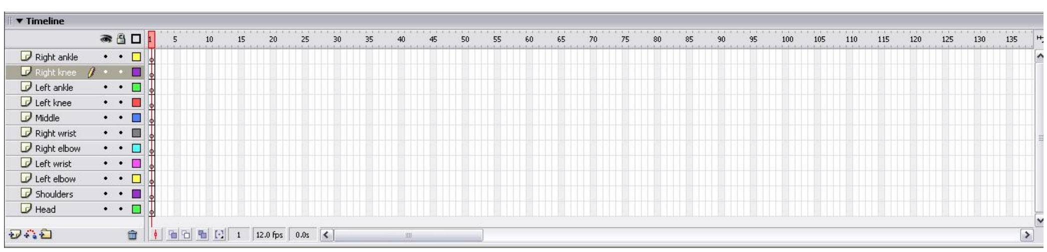

Figure 3: The timeline panel with eleven created layers

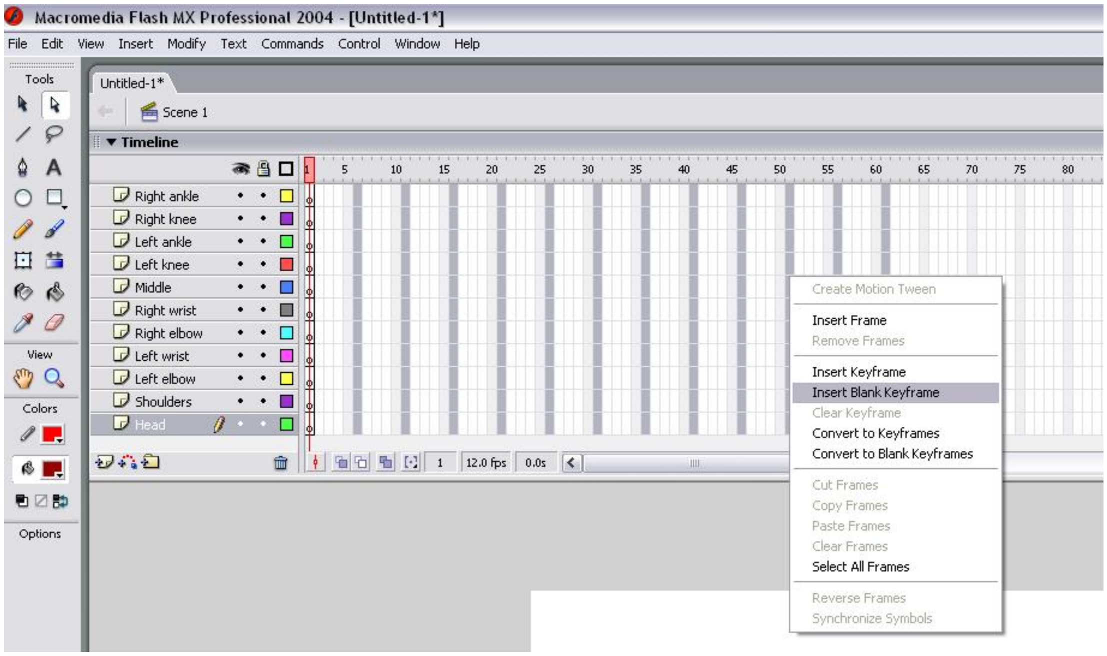

Figure 4: Select the frames using Ctrl, then right mouseclick on the selected frames and choose Insert Blank Keyframe

Select the frame up-left in the corner, and create with the oval tool a black circle (see figure 5). Then select the circle again and you see in the properties box down below you can change the properties of the circle (see figure 6). Put down here the coordinates of the selected joint and frame (here the first frame of the right ankle), and make the circle 15 height and 15 width. Make sure you copy the circle after that, so you can use it easier for the others frame.

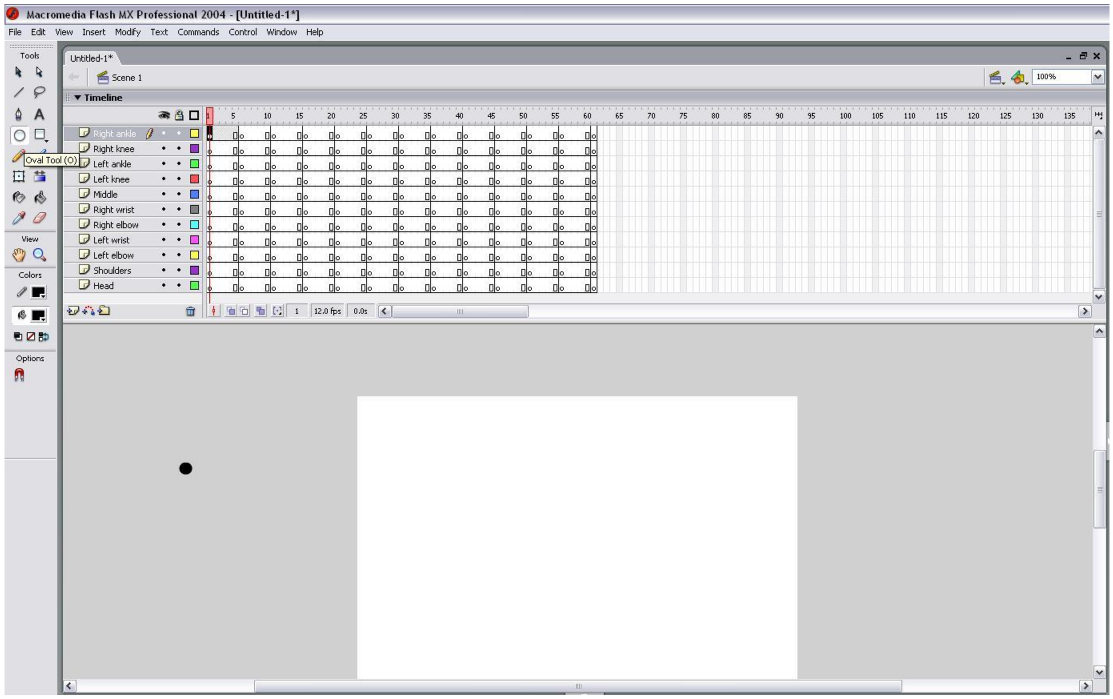

Figure 5: Create a black circle on the first frame of the right ankle

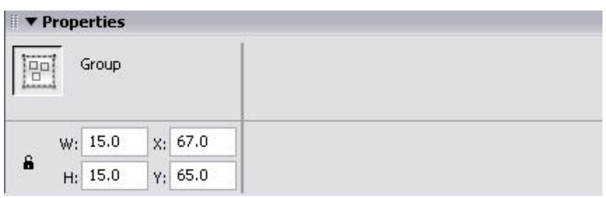

Figure 6: Properties box

You select every frame one by one, paste the copied circle in every frame and changes the coordinates of the circle. This way you will create a figure in every frame consisting of black dots on the body parts you registered the coordinates of. This is very time consuming, but you will get handy in it with some practicing.

In the end it is very easy, select all the frames and select the Create Motion Tween button with right-click (see figure 7). There will be a movie created without hesistations which looks like the figure you created a motion movie of. To check if the movie is good, you can press Enter.

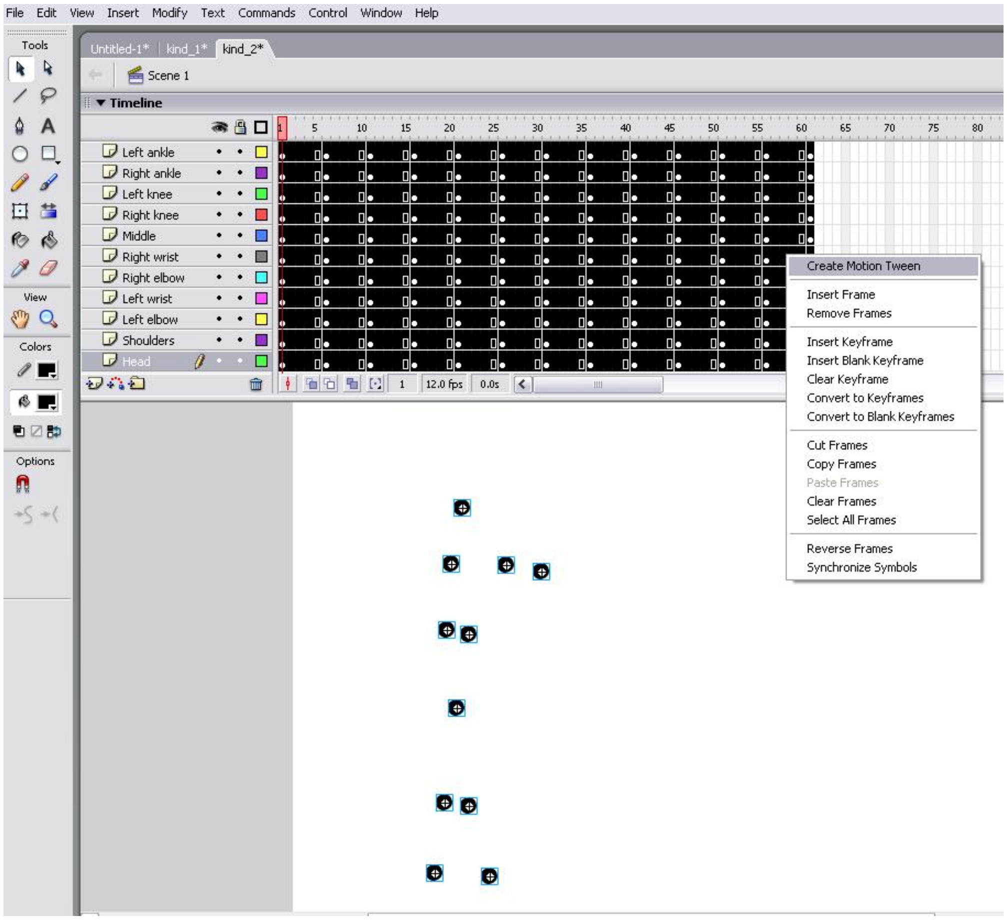

Figure 3: Select the Motion Tween

To export the movie: Select File → *Export* → *ExportMovie*