Capturing Social Motor Coordination: A comparison of the Microsoft Kinect, Video-Motion Analysis and the Polhemus Latus Motion Tracking System

Veronica Romero (romerovc@mail.uc.edu)
Center for Cognition, Action and Perception, University of Cincinnati, 4150 Edwards Cl., Cincinnati, OH 45221-0376 USA

Joseph Amaral (arumenator@gmail.com)
Department of Psychology, University of Cincinnati, 4150 Edwards Cl., Cincinnati, OH 45221-0376 USA

Paula Fitzpatrick (pfitzpat@assumption.edu)
Department of Psychology, Assumption College, 500 Salisbury St., Worcester, MA 01609-1265 USA

R. C. Schmidt (rschmidt@holycross.edu)
Department of Psychology, College of the Holy Cross, 1 College St., Worcester, MA 01609-2331 USA

Michael J. Richardson (richamo@ucmail.uc.edu)
Center for Cognition, Action and Perception, University of Cincinnati, 4150 Edwards Cl., Cincinnati, OH 45221-0376 USA

Abstract

Social motor coordination remains a relatively overlooked dimension of social behavior in children with ASD. One reason for the lack of research is that the motion tracking equipment historically used for recording body movements of children during social interaction has been very costly, as well as cumbersome and impractical. Here we examined whether two low-cost motion-tracking options can be employed to investigate social motor coordination in children with ASD. Of particular interest was the degree to which these low-cost methods of motion tracking could be used to capture and index the coordination dynamics that occurred between a child and an experimenter in comparison to a much more expensive, laboratory grade, motion tracking system. Overall, the results found the expensive system to be better than the low-cost methods, but that the latter two are still able to index differences in social motor coordination between typically developing and ASD children.

Keywords: Cognitive science, Psychology, action, motor control.

Introduction

Given the importance of social motor coordination for effective social interaction, several researchers have hypothesized that deficits in social movement coordination may play an important role in the interpersonal and social cognitive deficits that characterize autism spectrum disorder (ASD; Fitzpatrick et al., 2013; Marsh et al., 2009). Social motor coordination, however, remains a relatively overlooked dimension of social behavior in children with ASD (as well as for children with developmental delays in general). One reason for the lack of research is that, historically, the motion tracking equipment required to record and objectively measure the limb and body movements of children (or even adults) during social interaction has been very costly, as well as cumbersome and impractical within a non-clinical or non-laboratory setting.

Thankfully, over the last 5 years an increasing number of low-cost motion-tracking systems (e.g., Microsoft Kinect, Microsoft LTD), or alternative video-based methods (e.g., pixel change analysis) of motion capture have become available to researchers and clinicians interested in investigating the behavioral dynamics of human motor control and social motor coordination. In addition to costing only a fraction of the price of their high-end laboratory standard counterparts, these systems are easy to replace, highly portable and can be used almost anywhere (i.e., in both clinical/laboratory and non-clinical/non-laboratory settings). Furthermore, they typically come with companion open-source software or software development kits that enable researchers to develop software applications, testing protocols, and data analysis systems that meet the specific needs of the researcher or research population in question.

The degree to which these systems are able to replace more expensive laboratory grade motion tracking systems for research on social motor coordination in children and adult populations, including research on social motor coordination in children with ASD, is therefore an important question that needs to be addressed. To explicate the viability of these low-cost systems for investigating social motor coordination in children with ASD, we conducted a study comparing social motor control in typically developing children and children with ASD using three methods of motion capture: (1) a high-end laboratory grade Polhemus Latus magnetic motion tracking system, (2) the Microsoft Kinect motion tracking sensor, which is a low-cost optical tracking system; and (3) a video recording based pixel change method of motion extraction. Below, we provide a brief description of these different methods and a detailed comparison of how these methods of motion capture faired with respect to determining the stability and patterning of the social coordination that occurred across a range of interpersonal motor tasks. Of particular interest was how well the low-cost Microsoft Kinect and video pixel change methods performed in comparison to the more expensive, laboratory grade, Polhemus Latus system.

Polhemus Liberty Latus Wireless System. This motion tracking system is a high-end, laboratory-grade wireless motion tracking system developed by Polhemus LTD (Vermont, USA) that uses an electromagnetic field to map the position (Euclidian x, y and z coordinates) and rotation (pitch, yaw, roll) of 1 to 12 small 79.4 gram sensors/markers. The system tracks these 6-Degrees-Of-Freedom sensors within an electromagnetic capture volume that is defined by a map of 1 to 16 receptors. Each receptor has an optimal diametric capture volume of 6 feet and multiple sensors can be aligned by the user (experimenter/clinician) to meet the spatial demands of the behavior(s) performed or recording volume required. The reliability and resolution of this equipment is excellent, with a sampling rate of 188 Hz or 94 Hz (i.e., samples per second) and a positional and rotational resolution of approximately 0.25 cm and 0.5° (if a marker/sensor is no more than 4 feet from a receptor). The system is easy to use with multiple participants and unlike optical tracking systems, the Polhemus Latus is not susceptible to occlusion and can therefore be used for almost any motor task and in almost any environment. The system costs approximately $12,500.00 USD for a 1 marker/1 receptor system and $60,000.00 USD for a 12 marker/16 receptor system.

Microsoft Kinect. The Kinect sensor (version 1)[1] combines a specialized video camera and an infrared depth sensing emitter to optically track the Euclidian x, y and z location (in coordinates relative to sensor placement) of up to 21 skeletal/body joints (i.e., head, left/right shoulders, elbows, wrists, the spine, left/right hips, knees, feet, etc.). The device was originally developed by Microsoft for their Xbox gaming console, but can also be purchased for use on any PC or laptop computer running a Windows 7 operating system or above. The research version costs approximately $225.00 USD and is capable of capturing skeletal/joint data and color BMP/video images at a maximum rate of 30 Hz (i.e., 30 frames per second), with a resolution of 1280x960 pixels. A free C/C++ and C# SDK is available directly from Microsoft and can be used to develop non-commercial applications and recording software. Because it is an optical based motion tracking system, it is completely wireless, and does not require any sensors to be placed on the body of the individual being tracked (which makes it especially useful when collecting data from children with ASD). However, since the skeletal data is based on a combined infrared/video process of depth and a machine learning algorithm trained extensively with the use of synthetic depth images for its inference of motion tracking (Shotton et al., 2011) it requires a constant line of sight of the limbs/bodies being tracked and is especially susceptible to occlusion. It also has a high noise to signal ratio (relative to the Polhemus Latus system for example), such that it is typically unable to reliably capture small or subtle changes in limb or body position, especially when participants are wearing loose clothing or the system is used in a high UV lighted environment.

Video Pixel change Motion Extraction. This method of motion analysis involves calculating the amount of pixel change between adjacent video frames, which can be taken to index the amount of activity of a participant if they are the only source of movement in that part of the frame (Kupper et al. 2010; Paxton & Dale, 2013; Schmidt et al., 2012). This calculation process can be automated using simple video analysis routines written in Matlab (Mathworks, Inc., Natick, MA) or similar data analysis and scripting software, and can even be employed to extract the global movement of two (or more) individuals so long as their movements or activity are within the same recorded frame. That is, video frames can be cropped to include the movements of only one person (i.e., the left half or right half of the screen) and also the absolute difference of pixel change between the adjacent frames of the video when calculated to form an image-change time series for each participant in the interaction.

Materials and Method

Participants

Thirty eight children (7 female) between the ages of 6 and 10 were recruited to participate in the study. Nineteen typically developing children and nineteen children who had previously been diagnosed with ASD took part in the study.

Equipment Setup

The study was conducted in a 10 x 12 foot laboratory room at Cincinnati Children's Hospital Medical Center (University of Cincinnati, Cincinnati, OH). Children came into the laboratory room and were asked to sit at a 2 foot wide x 4 foot long x 2 foot high table next to the seated experimenter. Four Polhemus Latus receptors were attached to the underside of the table top, one in each corner, to create a 10 x 12 x 8 foot capture volume around the table. As soon as the child was seated, the four Polhemus Liberty Latus wireless markers/sensors were placed in wristbands and slipped over the child's and experimenter's wrists (one marker on each wrist of the child and experimenter). The motion of the Polhemus sensors was recorded at 94 Hz on a PC computer using a custom software application written by the authors using the Polhemus Latus C/C++ SDK Library.

The Microsoft Kinect sensor was placed at a height of 1.5 m, 3 m away from corner of the table top closest to the participant and experimenter at approximately a 45 degree angle. A custom software application (www.xkiwilabs.com) using the free Windows Kinect SDK version 1.5 (Microsoft LTD) was used to record video images and the head, spine

[1] Since completing this study, Microsoft released a new version of the Kinect Sensor (i.e., version 2). Although this new version has improved voice and person recognition features, the temporal and spatial resolution of skeletal (motion capture) tracking has remained the same. Thus the current results should generalize to the Kinect Sensor Version 2.

and upper body skeletal data (11 skeletal points in total; no hip, leg or foot data was recorded) of the seated child and experimenter at a sample rate of 30 Hz.

Coordination Tasks

The data presented here was part of a bigger project, in which participants performed a large range of motor, social and cognitive tasks. Here, we selected three social motor coordination tasks that were performed by all of the children. The first coordination task was a sequence of *tapping movements*, which involved children using a finger from one hand to tap/hit three drum-like cylinders from left to right in synchrony with the experimenter. Children repeated this left-to right drumming sequence six times with the experimenter in a continuous manner. The second task involved a sequence of *pointing movements*, in which children were required to point at approximately shoulder height to the right, center and left of their body midline in synchrony with the experimenter. Again, children repeated this pointing sequence six times with the experimenter in a continuous manner. The third task was an interpersonal hand clapping game (pat-a-cake), in which children completed a simple repetitive sequence of clapping their hands together and then with the experimenter. The hand clapping game was completed twice, with each sequence involving six consecutive intrapersonal and interpersonal clapping movements. The data presented here is only the second hand clapping trial. All participants were asked to do these three tasks in synchrony with the seated experimenter.

Motion Data Reduction

All the data extraction and analysis methods presented below were completed using custom MATLAB (Mathworks, Inc., Natick, MA) applications developed by the authors (download from www.xkiwilabs.com).

Polhmeus Latus. The *x*-plane (left-right), *y*-plane (forward-back) and *z*-plane (up-down) positional coordinates of the sensors placed on the wrists of the experimenter and child were recorded for each task. To best determine the stability and pattering of the behavioral coordination that occurred between the child and experimenter we first isolated the primary plane of motion for each task. Since the primary plane of motion for the tapping and pointing tasks was in the left-right plane, the x-*plane* movement time-series were used to assess the behavioral coordination that occurred for these two tasks. For the hand clapping game, the largest amplitude of movement was in the up-down, *z*-plane, with the intrapersonal clapping events occurring at a lower height than the interpersonal clap events. Accordingly, this plane of motion was employed to assess the behavioral coordination that occurred for this task[2].

Microsoft Kinect. The data recorded form the Kinect was extracted for analysis using two different methods. The first method was comparable to the method used for the Polhemus Latus system described in the preceding section. That is, the child's and experimenter's forearm movements in the *x*-, *y*-, and *z*-planes were extracted from the skeletal tracker for the tapping, pointing, and hand clapping tasks, and an additive time-series was created.

The second method involved creating a unified 1-dimensional movement time-series for both the child and experimenter from the *x*-, *y*-, and *z*-plane motion of all of the upper-body joints recorded by the Kinect sensor (i.e., the spine, head, and the left and right shoulder, elbow, hand, and wrist). This was achieved by simply creating a vector based on the sum of the values of each movement/joint dimension at each time-step. This method of normalization was chosen in order to produce a 'collective' whole body motion time-series for the child and experimenter that would be similar to the collective motion time-series obtained from the pixel change method.

Pixel change Motion Time-series. Recall that the amount of pixel change within a video frame can be taken to index the amount of activity of a participant if they are the only source of movement in that part of the frame. To calculate the absolute difference of pixel change between adjacent video frames for both the child and the experimenter, we first split all of the video images recorded using the Kinect sensor down the middle into a child half and an experimenter half and then extracted image change time-series from these separate video frame series.

Data Analyses

Prior to analyzing all of the pre- and post- non-task relevant movement transient periods were cropped from the time-series. These final motion time-series were then low-passed filtered using 10Hz 4^{th} order Butterworth filter.

To determine the stability of the social motor coordination that occurred for each task and condition, two standard measures of interpersonal coordination were employed: cross-spectral coherence and distribution of relative phase (see Schmidt & Richardson, 2008 for a review).

Cross-spectral coherence. This measure, commonly referred to as *coherence*, evaluated the coordination that occurred between the child and experimenter by estimating the correlation between their movements at their peak frequencies. Coherence measures the degree of coordination between two movement time-series on a scale from 0 to 1. A coherence of 1 reflects perfect correlation of the movements (perfect coordination/synchrony) and 0 reflects no correlation (no coordination/synchrony).

Distribution of relative phase angles (DRP). This measure evaluated the concentration of relative phase angles between the movements of the child and experimenter (i.e., the relative space-time angular location of the movements of the child and experimenter) across nine 20° regions of relative phase (0–20°, 21–40°, 41–60°, 61–80°, 81–100°, 101–120°, 121–140°, 141–160°, 161–180°). To determine these distributions we computed the continuous relative phase of the two time-series between -180° and 180° using the Hilbert transform (Pikovsky, Rosenblum, & Kurths, 2001). We then computed the percentage of occurrence of the

[2] An analysis of secondary planes of motion produced results that were consistent with those reported here.

absolute value of the relative phase angles across the nine 20° regions of relative phase from 0° to 180°. Previous research has demonstrated that stable social motor coordination is characterized by a concentration of relative phase angles around 0° and 180° (Schmidt & Richardson, 2008).

Results

Object tapping task

Wrist movement. A one-way ANOVA performed on coherence for the Polhemus Latus data showed that the TD children had significantly higher measures of coherence (M = 0.81, SD = 0.11) than the children with ASD (M = 0.65, SD = 0.25; $F(1, 36) = 5.97$, $p = .02$; see Figure 1a). Additionally, the 9 x 2 mixed ANOVA conducted on the DRP revealed a significant main effect of phase region ($F(8, 288) = 237.36$, $p < .01$) and a significant phase region x diagnosis interaction ($F(8, 288) = 6.67$, $p < .01$). Simple effects revealed that TD children had a significantly higher mean occurrence at 0° (M = 50.6, SD = 11.01) than children with ASD (M = 37.92, SD = 15.38; $t(36) = -2.92$, $p < .01$; see Figure 1b). As expected, both groups of children spent the majority of the trial in the 0° phase region, also referred to as in-phase. On the other hand, the analysis of the Kinect forearm time-series showed no significant differences in coherence between the TD and ASD groups ($F(1, 36) = 0.13$, $p = .72$; see Figure 1a), nor any effects for DRP (see Figure 1c).

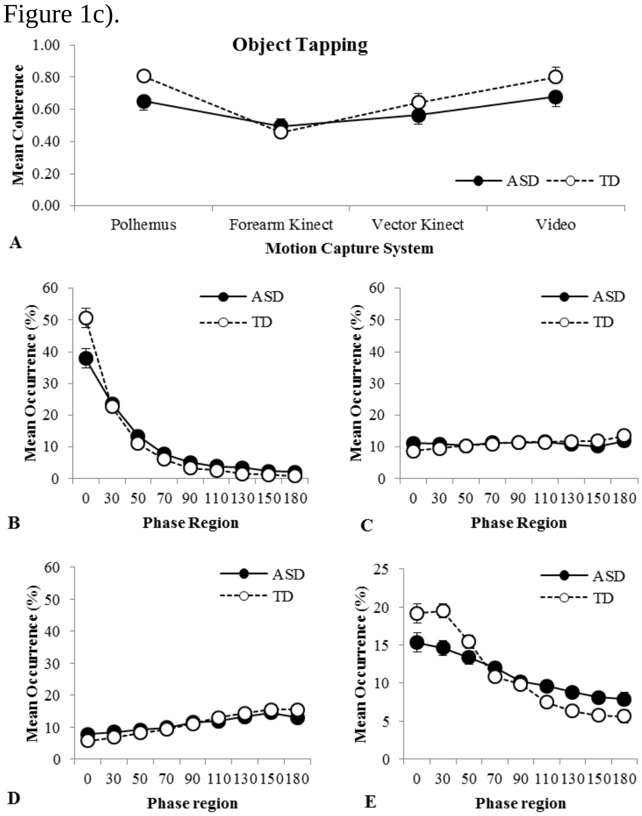

Figure 1. (a) Mean coherence as a function of motion capture system and group for the object tapping task. Distribution of relative phase for the (b) Polhemus Latus, (c) Kinect: forearm movement, (d) Kinect whole body vector, and (e) video methods. Error bars are standard errors.

Whole body movement. The one way ANOVA performed on the Kinect whole body vector movement time-series also revealed no significant differences in mean coherence between the groups ($F(1, 36) = 0.9$, $p = .33$; see Figure 1a). However, the 9 x 2 mixed ANOVA performed on DRP did reveal a significant main effect of phase region ($F(8, 288) = 11.76$, $p < .01$). Planned t-tests showed that participants spent significantly more time in the 180° phase region (M = 14.23, SD = 7.49) than in 0° phase region (M = 6.83, SD = 4.50; $t(37) = -4.17$, p < .01; see Figure 1d). The analysis of the Pixel change motion time-series showed no significant differences in coherence between the TD and ASD groups ($F(1, 36) = 2.11$, $p = .16$; see Figure 1a). The analysis of DRP, however, did reveal a significant main effect of phase region ($F(8, 288) = 51.30$, $p < .01$) and a significant phase region x diagnosis interaction ($F(8, 288) = 5.96$, $p < .01$). Simple effects revealed that TD children had a significantly higher mean occurrence at 0° (M = 19.17, SD = 5.65) than children with ASD (M = 5.22, SD = 15.38; $t(36) = -2.17$, $p = .04$; see Figure 1e). As expected, both groups of children spent the majority of the trial in the 0° phase region, also referred to as in-phase.

Pointing task

Wrist movement. The analysis of the Polhemus Latus data revealed a significant difference in coherence between the ASD and TD groups ($F(1, 36) = 4.18$, $p = .05$), such that children with ASD showed significantly less cross-correlation coherence (M = 0.78, SD = 0.21) than the TD children (M = 0.89, SD = 0.13; see Figure 2a). With regard to the analysis of DRP, there was a significant main effect of phase region distribution ($F(8, 288) = 235.43$, $p < .01$) and a significant phase region by diagnosis interaction ($F(8, 288) = 5.14$, $p < .01$). Simple effect analyses showed that the mean occurrence of a 0° relative phase was significantly higher for the children in the TD group (M = 63.82, SD = 18.88) than the ASD group (M = 49.69, SD = 16.51; $t(36) = -2.46$, $p = .02$; see Figure 2b). The analysis of Kinect forearm data, however, revealed no significant differences in mean coherence between the groups ($F(1, 36) = 0.09$, $p = .77$; see Figure 2a). There was, however, a significant main effect of phase region ($F(8, 288) = 6.68$, $p < .01$). Planned t-tests showed that participants spent significantly more time in the 0° phase region (M = 15.08, SD = 8.22) than the 180° phase region (M = 8.26, SD = 4.79; $t(37) = 3.53$, $p < .01$; see Figure 2c).

Whole body movements. The analysis of the Kinect whole body vector movement time-series revealed a significant difference in coherence between the ASD and TD groups ($F(1, 36) = 4.35$, $p = .04$), such that children with ASD showed significantly less cross-correlation coherence (M = 0.33, SD = 0.19) than the TD children (M = 0.48, SD = 0.25; see Figure 2a). The analysis of DRP revealed a significant main effect of phase region ($F(8, 288) = 18.02$, $p < .01$). Planned t-tests showed that participants spent significantly more time in the 0° phase region ($M = 15.66$, $SD = 6.85$) than the 180° phase region ($M = 7.33$, $SD = 3.43$; $t(37) = 5.63$, $p < .01$; see Figure 2d). On the other hand, the analysis performed on the Pixel change motion time-series showed no significant differences in coherence between the TD and ASD groups ($F(1, 36) = 0.03$, $p = .86$; see Figure 2a). However, there was a significant main effect of phase region ($F(8, 288) = 143.94$, $p < .01$.) Planned t-tests showed that participants spent significantly more time in the 0° phase region ($M = 25.57$, $SD = 7.32$) than the 180° phase region ($M = 3.71$, $SD = 2.18$; $t(37) = 15.48$, $p < .01$; see Figure 2e).

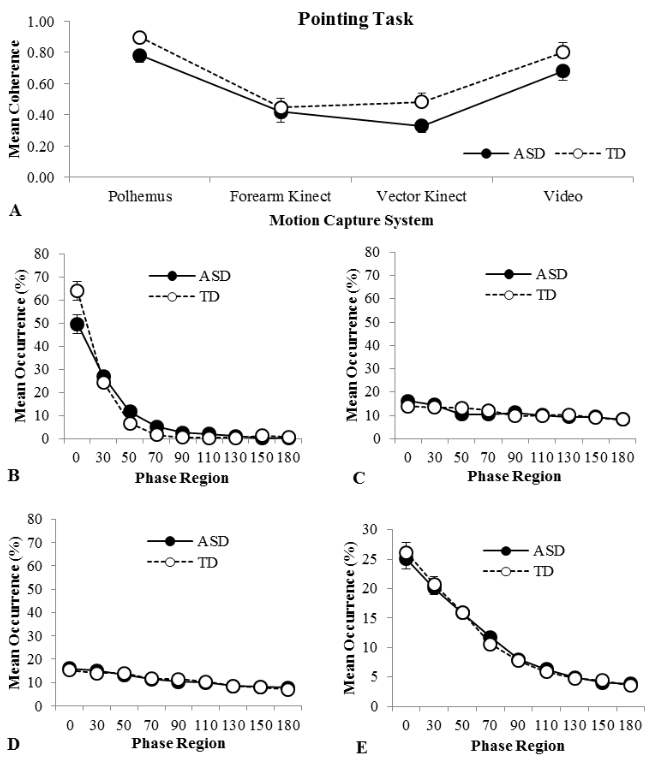

Figure 2. (a) Mean coherence as a function of motion capture system and group for the pointing task. Distribution of relative phase for the (b) Polhemus Latus, (c) Kinect: forearm movement, (d) Kinect whole body vector, and (e) video methods. Error bars are standard errors of the mean.

Interpersonal Hand Clapping Game

Wrist movement. Analysis of the Polhemus Latus time-series data showed that participants in the ASD group had significantly lower cross-spectral coherence ($M = 0.86$, $SD = 0.13$) than those in the TD group ($M = 0.93$, $SD = 0.03$; $F(1, 35) = 6.18$, $p = .02$; see Figure 3a). There was also a significant main effect of phase region ($F(8, 280) = 210.92$, $p < .01$) and a significant phase region x diagnosis interaction ($F(8, 280) = 12.79$, $p < .01$). Simple effects analyses showed a significantly lower occurrence of coordination for TD children in the 0° region ($M = 0.06$, $SD = 0.18$) than children in the ASD group ($M = 0.96$, $SD = 1.61$; $t(35) = 2.36$, $p = .02$). Additionally, the children in the TD group had a higher mean occurrence in the 180° phase region ($M = 59.79$, $SD = 13.03$) than those in the ASD group ($M = 39.08$, $SD = 18.12$; $t(35) = -3.97$, $p < .01$; see Figure 3b). The analysis performed on the Kinect forearm time-series also revealed a significant difference in coherence between the ASD and TD groups ($F(1, 32) = 13.37$, $p < .01$), such that children with ASD showed significantly less cross-correlation coherence ($M = 0.40$, $SD = 0.28$) than the TD children ($M = 0.71$, $SD = 0.20$; see Figure 3a). There was also a significant main effect of phase region ($F(8, 256) = 40.96$, $p < .01$) and a significant phase region by diagnosis interaction ($F(8, 256) = 12.06$, $p < .01$). Simple effects analyses showed a significantly lower occurrence for TD children in the 0° region ($M = 3.00$, $SD = 2.91$) than children in the ASD group ($M = 7.85$, $SD = 5.31$; $t(32) = 3.29$, $p < .01$). Additionally, the children in the TD group had a higher mean occurrence in the 180° phase region ($M = 27.76$, $SD = 12.16$) than those in the ASD group ($M = 15.37$, $SD = 9.32$; $t(32) = -3.34$, $p < .01$; see Figure 3c).

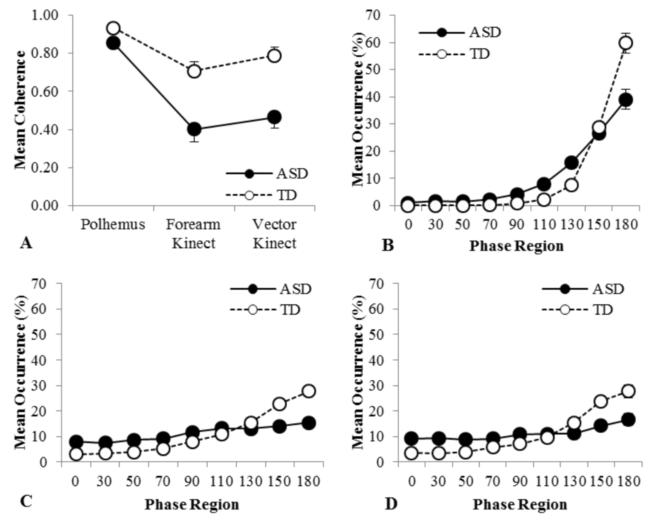

Figure 3. (a) Mean coherence as a function of motion capture system and group for the hand clapping game. Distribution of relative phase for the (b) Polhemus Latus, (c) Kinect: forearm movement, and (d) Kinect whole body vector methods. Error bars represent standard errors of the mean.

Whole body movement. The analysis performed on the Kinect whole body vector time-series for the hand clapping game revealed a significant difference in coherence between the ASD and TD groups ($F(1, 33) = 18.22$, $p < .01$), such that children with ASD showed significantly less cross-correlation coherence ($M = 0.46$, $SD = 0.25$) than the TD children ($M = 0.78$, $SD = 0.19$; see Figure 3a). There was also a significant main effect of phase region ($F(8, 264) = 31.52$, $p < .01$) and a significant phase region by diagnosis interaction ($F(8, 264) = 10.07$, $p < .01$). Simple effects analyses showed a significantly lower occurrence for TD children in the 0° region ($M = 3.51$, $SD = 4.57$) than children in the ASD group ($M = 9.14$, $SD = 7.22$; $t(33) = 2.74$, $p =$

.01). Additionally, the children in the TD group had a higher mean occurrence in the 180° phase region (*M* = 27.72, *SD* = 11.39) than those in the ASD group (*M* = 16.61, *SD* = 27.72; $t(33) = -3.16$, $p < .01$; see Figure 3d). The video pixel change analysis could not be performed in this task because the movements of the participant and experimenter no longer remained in separate sections of the frame throughout the trial.

Discussion

The goal of the current paper was to explicate the viability of employing low-cost motion tracking systems for investigating social motor coordination in general and in children with ASD specifically. Of particular interest was how well the low-cost Microsoft Kinect and video pixel change methods performed in comparison to the Polhemus Latus system and the degree to which these differing methods could be employed to differentiate the coordination that occurred for TD and ASD participants.

As expected and consistent with other recent findings (Fitzpatrick et al, 2013; Marsh et al., 2009), ASD participants exhibited a less stable pattern of social motor coordination than TD participants. This difference was apparent in all three social motor tasks, but perhaps most pronounced for the interpersonal hand clapping game. With regard to the questions of whether the different motion capture systems were able to capture data that revealed this difference, the current findings demonstrated that the Polhemus Latus system did in fact provide a finer-grained measure of limb movement than the Kinect and video-based methods and was more robust in differentiating the groups in patterning and stability of the coordination. This suggests that the Polhemus Latus system may be superior for tasks that predominantly involve limb effector movements. The Kinect wrist movement analysis did, however, differentiate the groups in the hand clapping game. One limitation of the Polhemus is that the wireless sensors must be attached to the limbs, which can be problematic for certain participants. In addition, the system's reliance on magnetic signals makes its use incompatible with some other systems (e.g., EEG).

An analysis of the whole body movements using the Kinect and pixel change indicated these methods were able to differentiate the stability of TD and ASD coordination in some instances. For example, the pixel change data did reveal a significant difference in the distribution of relative phase for the tapping and pointing tasks. The whole-body Kinect analysis revealed significant group differences in coherence and the distribution of relative phase for both the pointing and hand clapping tasks. However, due to the reliance on the machine learning algorithm built into the Kinect system, the results presented currently are preliminary. A more rigorous test would be to record participants' movements with the Kinect while recording their movement with Polhemus sensors that correspond to the same skeletal markers in the Kinect in order to measure if the differences observed here are due to errors in the skeletal reconstruction or simple occlusion.

What is apparent, however, is that when employing these low-cost motion-tracking methods, particular care needs to be taken when designing the laboratory environment and the interaction tasks to be employed. In general the current results demonstrate that for both the Kinect sensor and pixel change methods tasks with larger scale movements provide the most accurate and reliable results. Of particular importance when using the Kinect is to choose tasks that have minimal occlusion issues, for example when the arms are not placed in front of the torso and when no props are used. When using the pixel change method the movements of the two people have to be in separate parts of the video frame and may be best-suited to tasks involving less stereotyped movement. More generally, the current study also validates previous research (Fitzpatrick et al., 2013) by demonstrating that children diagnosed with ASD show different social motor coordination patterns when compared to their TD counterparts. The low-cost and completely wireless motion capture systems compared here can therefore provide researchers with new tools to explore social motor coordination and the role it plays not only in ASD, but also in other developmental delays disorders and social functioning pathologies (i.e., schizophrenia).

Acknowledgment. This research was supported by NIH R21MH094659.

References

Fitzpatrick, P., Diorio, R., Richardson, M. J., & Schmidt, R. C. (2013). Dynamical methods for evaluating the time-dependent unfolding of social coordination in children with autism. *Frontiers in Integrative Neuroscience, 7*, 1-13.

Kupper, Z., Ramseyer, F., Hoffmann, H., Kalbermatten, S., & Tschacher, W. (2010). Video-based quantification of body movement during social interaction indicates the severity of negative symptoms in patients with schizophrenia. *Schizophrenia Research, 121*, 90-100.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science, 1*(2), 320-339.

Paxton, A., & Dale, R. (2013). Frame differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods, 45*, 329–334.

Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). Synchronization: A universal concept in nonlinear sciences. Cambridge: Cambridge University Press.

Schmidt, R. C., Morr, S., Fitzpatrick, P. A., & Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *Journal of Nonverbal Behavior, 36*, 263-279.

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of Interpersonal Coordination. In A. Fuchs & V. Jirsa (Eds.). *Coordination: Neural, Behavioral and Social Dynamics*. (pp. 281-308). Heidelberg: Springer-Verlag.