

Multiscale kinematics of action intention

S.F. Corbin¹, C.H. Moore¹, T. Davis¹, K. Shockley^{1, *}, and T. Lorenz^{1, 2, 3}

¹Department of Psychology, ²Department of Electrical and Computer Engineering, ³Department of Mechanical and Materials Engineering, Center for Cognition, Action, & Perception, University of Cincinnati, Cincinnati OH 45219-0376, USA

*Correspondence: shocklkn@ucmail.uc.edu

DOI: <https://doi.org/10.56280/1565382896>

This is an open access article under the CC BY 4.0 license (<https://creativecommons.org/licenses/by/4.0/>).

Submitted: 25 November 2022 Revised: 6 January 2023 Accepted: 10 February 2023 Published: 28 April 2023

Abstract

Human motion contains rich contextual information about not only action, but action intention. In two experiments, we investigated whether the multiscale kinematic information that differentiates intentional actions is the same information to which observers attend when asked to observe an actor's intended movement. To do so, we first recorded an actor's movement kinematics while performing four different intentional sit-to-stand actions. Analyzing the differences in movement kinematics, we then identified the joints that contributed to differentiating the actions using principal components analysis and multinomial regression. Observers were then shown point-light displays of these movements and given a forced-choice task to select which action the actor intended to complete and were highly accurate at this task. We hypothesized that if perceptual information used to perceive action intention corresponds to the kinematic information that differentiates among the four possible actions, then observers' gaze should center more on the joints identified in the movement analysis. This hypothesis was supported, suggesting that joint kinematics that differentiate possible actions are the same joint kinematics to which observers attend in order to successfully differentiate movement intentions in others.

Keywords: Kinematic information, embodiment, point-light displays, action intention, intention recognition, biological motion

1. Introduction

Intention does not live exclusively in the brain. Indeed, intentions translate into a person's movement patterns and human observers are well attuned to such biological motion (e.g. [Bosbach et al., 2005](#); [Manera et al., 2011](#); [Sartori et al., 2011](#); [Troschianko et al., 2004](#)). Perceptual attunement to biological motion was initially demonstrated by Johansson (1973) who used point-light displays (i.e., only the motion at each of the major joints [i.e., 10-12 locations] depicted as points of light in video displays) to demonstrate the robust ability of humans to perceive what actions an actor is engaged in. Via these displays, observers not only identified what the actor was doing but even more subtle aspects of the action such as how much effort the actor was putting into their actions. Moreover, subsequent studies have demonstrated perceptual sensitivity to not just actions but a variety of actor characteristics including biological sex ([Kozlowski & Cutting, 1977](#)) and even their identity ([Cutting & Kozlowski, 1977](#)).

[Runeson \(1983\)](#) proposed the kinematic specification of dynamics (KSD) principle as the basis for the human ability to perceive the dynamics (e.g., the action) from observation of kinematics (e.g., motion of point-light displays) ([Runeson, 1983](#); [Runeson & Frykholm, 1983](#)). He argued that because the motion of an object is lawfully related to the forces that produced that motion, the kinematics, accordingly, bear a 1:1 relationship to the original dynamics and thus provide an unambiguous basis to perceive the action for what it is. The KSD principle was initially investigated in non-biological motion (specifically, event perception) by having observers report on the relative masses of colliding balls ([Runeson, 1983](#); [Runeson & Frykholm, 1983](#)) and was extended to perception of biological motion via studying the perception of lifted objects using point-light displays ([Runeson & Frykholm, 1981](#)). Actors were asked to lift boxes of varying weight while their joint motion was recorded using reflective tape on the major joints. The corresponding point-light displays were presented to observers who were asked to report on the weight of the

boxes being lifted. In other words, they were asked to “see” the mass of the object lifted while provided only the motion of the major joints (i.e., kinematics). Observers were quite good at this task, supporting Runeson’s KSD principle as the perceptual basis for this ability. Subsequently, this perceptual sensitivity has been demonstrated across a wide range of dispositions and intentions (e.g., Cutting & Kozlowski, 1977; Dittrich et al., 1996; Groner & Schollerer, 2005; Jokisch et al., 2006; Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; Montepare & Zebrowitz-McArthur, 1988; Steel et al., 2006; Troje et al., 2005).

In addition to the ability to identify current actions, there is also evidence that the intention of an actor can be ascertained via observation of their movement. For example, Bosbach et al. (2005) recorded videos of actors lifting boxes of varying weight (cf. Runeson & Frykholm, 1981) in which the actors were told how much the box-to-be-lifted weighed. However, in some trials the weight given to lift was greater than the actor was told, thus surprising the actor who intended to lift less weight. Observers were then asked to report when the actor received incorrect information about the weight (i.e., when the actor intended to lift less weight than they were actually given to lift). Bosbach et al. (2005) found that patients with neurogenerative peripheral nerve disease (who were, accordingly, limited in recent lifting experience) were worse than controls at differentiating the intention of the actors’ when the weight of the object was greater than the actor expected. Not only does this further evidence that actor movement embodies the intention of the actor, but also suggests that an observer’s ability to ascertain intention is exquisitely linked to one’s own *action capabilities*.

The ability to perceive action capabilities from kinematics is further evidenced in a study by Weast et al. (2014) who investigated whether an observer could perceive action capabilities based on kinematics. Specifically, they investigated whether observers could perceive the maximum reach-with-jump height and long-jump distance of actors based on point-light displays. They found that jumping-related kinematics (walking and squatting) yielded more accurate perception than non-jumping-related kinematics (standing and twisting). Importantly, the kinematics displayed to the observers were not directly related to the action to be judged (jump height or jump distance) but still informed the observer about the action capabilities of the actor and hence allowed for an extrapolation of the required information.

Not only do kinematics contain information about current actions and action possibilities, but there is also evidence that differences in kinematics unfolding *early* in a movement process can allow for a differentiation of the intentions of the actor prior to the action being completed (i.e., what type of action they intend prior to the action actually happening). There is evidence both in the analysis of the kinematics as well as of the perceptual sensitivity to the kinematics. Becchio et al. (2010) asked participants to reach toward an object and grasp it to move it from one place to another or to grasp it to hand it to another person. Analysis of the kinematics of the finger movement during the reach-to-grasp movements showed changes in maximal finger aperture and peak grip closing velocity for these two actions, see also Ferri et al. (2010, 2011). Likewise, in a reaching task in which the outcome of movement was occluded, Sartori et al. (2011) demonstrated that observers could discriminate among cooperative, competitive, and individual reaching types of movement. Manera et al. (2011) showed the same perceptual capability of observers to differentiate action intention¹ using point-light displays of the end-occluded grasping kinematics (though less accurately than under full-light conditions).

Although the outcomes of actions (such as whether a thrown dart will hit its target or whether a penalty shot will reach the net) have been rather extensively studied in the action prediction literature, there have been few studies investigating the perceptual identification of the action itself that is intended (i.e., when the full action is not visible to the observer). For example, Troscianko et al. (2004) showed video displays from closed-circuit television cameras to observers who were quite effective at perceiving whether the actor’s intent was violent or non-violent, but this was based on video (cut short) not on point-light displays. Likewise, Dittrich (1993) demonstrated that observers can identify if a social interaction was intended to be threatening based on point-light displays, but this did not involve identifying the particular action that was intended without seeing the kinematics of the full act. A variety of different classes of action were studied (locomotory, instrumental, and social) but the closest to intention was if the interaction was threatening. Fewer still have investigated the kinematics that support this ability.

¹ In Sartori et al. (2011) and Manera et al. (2011) intention is conceptualized as social task intention, i.e., cooperation vs. competition.

Our present focus is on the informational basis for our ability to perceive the array of action-related aspects of biological motion. Specifically, given that there is evidence that intention is embodied in movement kinematics, the present study investigated whether the structure that allows for differentiating intention in movement is the same structure to which the brain is sensitive. If so, this would suggest that intention operates at the scales of brain and body simultaneously in ways that are principled, reliable, and quantifiable. In order to evaluate this possibility, we completed two experiments involving four intentional sit-to-stand (STS) transitions: a baseline STS transition (*ST-Stand*), an STS transition that includes a forward reach for a button (resembling reaching for an object while standing up, *ST-Press-Stand*), an STS transition that is interrupted after a forward reach for a button (resembling reaching for an object and sitting back down, *ST-Press-Sit*), and an STS transition that includes an upwards reach for a pull string (resembling reaching for a light or fan pull string, *ST-Reach-Stand*; see **Figure 1**). A single actor produced the four different actions (for which the position of all the major joints were recorded) to provide the kinematic data to analyze and serve as stimuli.

Our central hypothesis was that we are perceptually sensitive to the same kinematic information that differentiates action intention during motion execution. This hypothesis is rooted in the previously established claim that movement intention is embodied in the kinematic information of the actor (e.g., in point-light displays) and that we are perceptually sensitive to the kinematic information that is specific to the intention of the actor. Our strategy for testing this hypothesis was to determine the joint kinematics that serve to differentiate the actions and determine if these correspond to the joints that observers attend to when they correctly differentiate among the action intentions. We executed this strategy by conducting two experiments. The first experiment aimed to identify the joint kinematics that differentiate the four STS transitions described above. The second experiment aimed to determine if observers could perceptually discriminate the intention of the actor (that is, could participants differentiate which action was being initiated prior to it being executed [e.g., prior to the actor pressing a button, standing, stretching, or sitting back down]) and if their gaze focus corresponded to the joint kinematics identified in the first experiment. We predicted that 1) we could statistically identify the joint kinematics that differentiate among the four possible actions (Experiment I); 2) that observers would be able

to differentiate among the four possible action intentions based on early kinematic information (Experiment II); and 3) that observers' gaze would focus more on the joints that differentiate among the four possible actions than on the joints that do not differentiate among the four possible actions (Experiment II). These findings would suggest that the same kinematic information is involved at the levels of perception and motor control for the actions under study.

2. Experiment I – Actor movement

The purpose of Experiment I was to determine which joint kinematics differentiated the four possible STS actions. Our strategy was inspired by a study by [Weast-Knapp et al. \(2019\)](#) who used Principal Components Analysis (PCA) to identify the kinematic information to predict maximum reach-with-jump height of actors. PCA is a statistical technique that is used to reduce the dimensionality of a multivariate data set while preserving as much of its original variability as possible ([Jolliffe & Cadima, 2016](#)). This is achieved by transforming the original variables into new variables which are uncorrelated and capture the maximum amount of variation in the data. PCA works by finding the directions in which the data vary the most, and then representing the data along these directions forming new variables, or principal components. The first principal component represents the data projected along the direction that most maximizes the variance in the data; the second principal component represents the direction that maximizes the variance while remaining orthogonal to the first principal component; and so on where the number of principal components is equal to the number of variables in the original dataset—e.g., if the data has 7 dimensions, then there will be 7 principal components (however, to use PCA for dimensionality reduction, we can elect to keep only the first few principal components and ignore the rest). One especially useful measure that may be derived from this procedure are the PCA *loadings*. The loadings are the coefficients of the linear combination that produces the principal component scores—factors by which each original variable is multiplied to obtain the principal component. The magnitude of the loadings may be squared to calculate the relative contribution of each variable to the principal component.

When applied to kinematic data, PCA can be particularly useful for understanding how different body parts are related and how they interact with each other. Previous work has employed PCA to uncover the most important

factors contributing to variation in movement kinematics in gait (Vallery & Buss, 2006), juggling (Post et al., 2000), and even the movements of cooperating actors (Ramenzoni et al., 2012). Weast-Knapp et al. (2019) used PCA to examine the kinematics of an actor's walking gait. The actors' movements were motion-captured via markers at 14 major joints about their bodies. PCA was used to analyze the patterns of marker movements for the purposes of revealing candidate body segments whose patterns of movements provide perceptually rich information about the actors' reach-with-jumping ability. In order to test whether PCA indeed correctly identified the relevant information in the movement kinematics, they first showed point-light displays of a low-jumper walking and a high-jumper walking to observers who could correctly differentiate jumping ability based on the walking kinematics. In a second experiment they then swapped the PCA loadings to the actors' joints (loadings for a low-jumper were applied to the kinematics of a high-jumper and vice versa) and presented these modified point-light displays to observers. They found that observers' reports of maximum reach-with-jump height reversed, meaning observers reported low-jumpers as being able to jump higher than high-jumpers. Hence, the principal components of human movement can be used to indicate the relevant kinematic information of an action.

Therefore, our general strategy for the analysis of intentional motion was to use PCA to determine the markers most relevant (cf. Weast-Knapp et al., 2019) to differentiating the four possible STS actions. For example, using point-light displays, Abernethy & Zawi (2007) asked participants to determine the direction of badminton strokes under varying occlusion conditions. They found that only a relevant subset of joints was needed to correctly predict the outcome of the action. They termed the kinematics of these relevant joints *essential* kinematic information. Accordingly, our strategy was to use the PCA loadings of joint kinematics with stepwise, multinomial regression to identify the essential kinematic information needed to differentiate the four possible STS actions (i.e., the subset of joints needed to statistically differentiate them). One of our central hypotheses was that the relative contributions of joint kinematics to the first principal component (PC1) would systematically vary as a function of the sit-to-stand intention. This would be reflected in changes in the percent-contribution that each segment makes to the major principal components. This change in arrangement would be indicative of changes in

the higher-order relationships among joint kinematics in action. A second goal of this experiment was to identify the essential kinematic information by using the resulting PCA loadings as a tool for feature selection—thereby reducing our original set of seven joints down to a more parsimonious number that still allowed for reliable differentiation among our four sit-to-stand intentions.

2.1 Method

2.1.1 Participants

We collected intentional sit-to-stand data from one male participant (28 years). The study was approved by the University of Cincinnati Institutional review board (IRB # 2012-2827).

2.1.2 Setup and procedure

A schematic of the experimental setup is shown in **Figure 1b**. A button was placed in front of the participant at shoulder height while sitting and 1.6-times the arm length from the shoulder. A pull switch was positioned above the participant at a height of 0.8 times the arm length and at a distance of 0.5 times the arm length from the shoulder while standing. Motion data was recorded using a 20-camera 3D motion capture system (Motion Analysis Corporation, Santa Rosa, CA). A 29-marker set based on the Helen Hayes body marker placement protocol (Kadaba et al., 1990) was used to track the motion as shown in **Figure 1a**. A screen was positioned at eye level in front of the participant to provide instructions for the specific trial. Every trial started with the participant sitting on a stool (height 45.72 cm) without any hand or back rests. The participant was shown a "Ready" signal on the screen and after 3 seconds the instruction to perform any of the four intentional movements marked the go-signal. Instructions to perform any of the four intentional movements were fully randomized. The participant was allowed to take breaks whenever he felt fatigued.

Overall, the participant performed 100 trials of four intentional STS transition tasks that differed in the intention of the subsequent action only (25 trials per intention, see **Figure 1c-f**) (Corbin et al., 2019; Moore et al., 2019; Patil et al., 2019). For *ST-Stand*, the participant was asked to stand up at a comfortable speed without any intention towards added activity. During *ST-Press-Stand*, the participant was asked to stand up from the chair while pushing a button in front of them before

finishing the standing up movement. During *ST-Press-Sit*, participant was asked to press the same button as in *ST-Press-Stand* but interrupting the standing up movement afterwards and instead sitting back down. Finally, during *ST-Reach-Stand*, the participant was

asked to stand up to pull a fan string above their head before finishing standing up. For all trials, the participant was instructed not to use his hands to push down on the chair or thighs during the STS transition and not to lift his feet from the heel or toes during the trial.

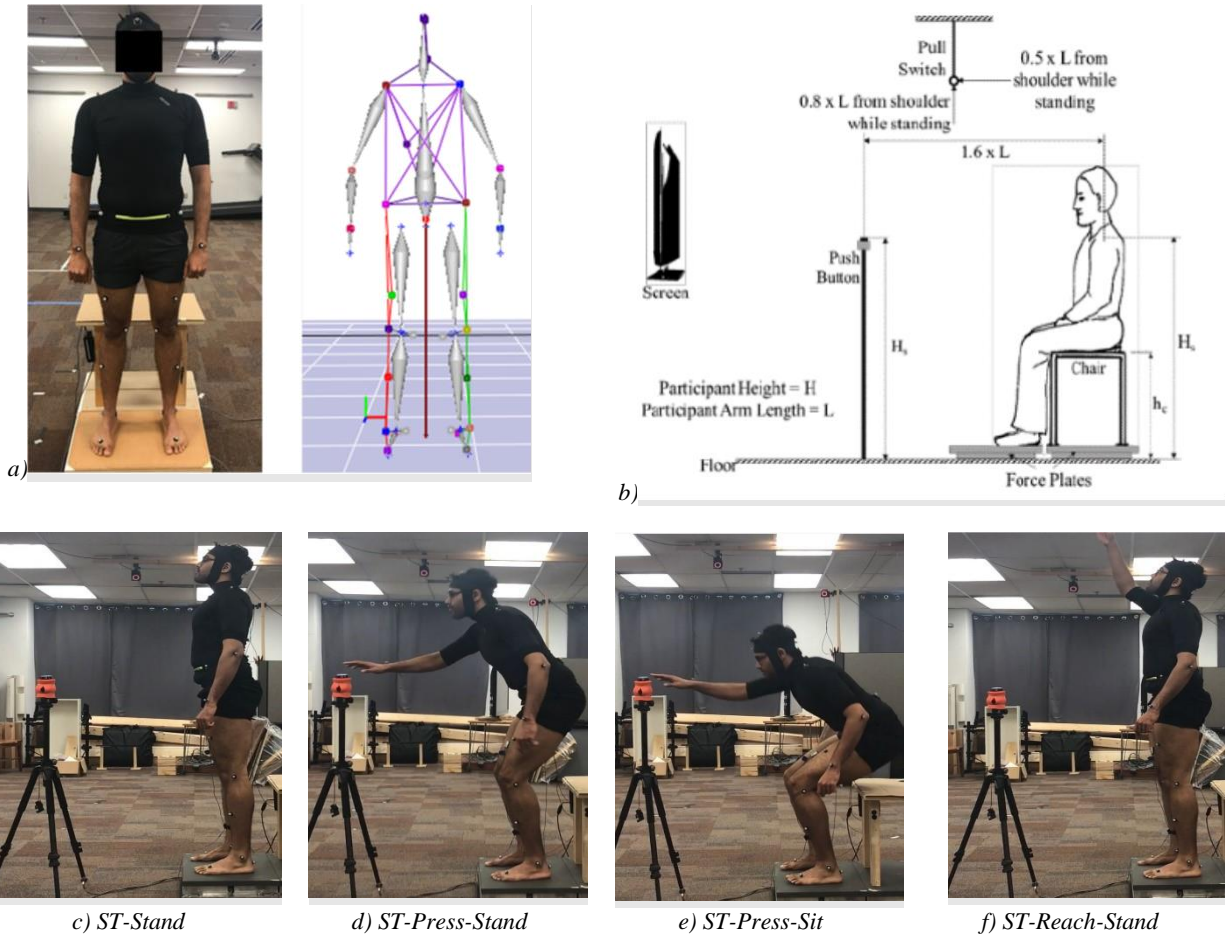


Figure 1 a): 29-marker set based on the Helen Hayes body marker placement protocol (Kadaba et al., 1990), see also (Patil et al., 2018), figure reprinted with permission; b) experimental setup; c-f) Four intentional sit-to-stand transitions: c) *ST-stand*: baseline sit-to-stand transition without any additional activity; d) *ST-Press-Stand*: press a button in front of them while standing up; e) *ST-Press-Sit*: press a button in front of them before sitting back down; f) *ST-Reach-Stand*: sit-to-stand transition with an upward reaching movement to pull a ceiling fan string.

2.1.3. Data processing

Motion-capture of our actor included a set of 29 markers positioned at various joints and segments on his body. Given our interests and goals for this study we only focused our analysis on marker positions along the superior-inferior and anterior-posterior axes of the actor's movement. Therefore, seven markers on the right side of the actor's body were extracted for further analysis: ankle, elbow, wrist (hand), crown of the head, hip, knee, and shoulder. Time series of the two-dimensional (superior-inferior and anterior-posterior)

coordinates for each marker were then transformed into a one-dimensional time series of their Euclidean distances from an origin point in the motion-capture space. These Euclidean distance time series were then scaled to unit variance and submitted to principal components analysis via R (stats::prcomp).

2.2 Results

We applied PCA to each STS movement data series from our model. Provided our seven original marker (Euclidean distance) time series, this resulted in seven

components per trial. For each trial, we calculated the amount of total variance accounted for by the first principal component (PC1) which captures the majority of the variance in the movements (cf. Weast et al., 2019). As seen in **Table 1**, the amount of the variance accounted for by PC1 varied as a function of condition.

Next, we analyzed the loadings of each marker to PC1. Our strategy for identifying the markers relevant to differentiating the four possible actions was to use stepwise, multinomial regression to model the relationship among the four recorded actions and the corresponding body segment contributions. For each movement series, the percentage contribution of each marker was first calculated by squaring the PCA loadings. As seen in **Figure 2** the percent contribution of each marker to PC1 varied as a function of actor intention. The vertical dashed line in each plot represents the expected average contribution assuming the contributions of the markers were uniform.

Next, the relative contributions of each of the remaining markers were used to build multinomial regression models with intention (ST-Press-Sit, ST-Press-Stand, ST-R Reach-Stand, ST-Stand) as the outcome variable predicted by each of the marker contributions. Starting with a null model (intention ~ 1), we conducted a forward

stepwise procedure adding markers in the order of increasing contribution (e.g., knee was the first marker loaded into the model given its overall contribution to PC1 was least). We used the Akaike information criterion (Akaike, 1998) to arrive at the most parsimonious model that might still allow for effective discrimination among intentions. The effectiveness of this model was tested by comparing its performance against the full model using the area under the receiver operating characteristic curve (AUROC) as a measure, where lower values indicate poorer performance. The resulting optimal model included only shoulder, head, and elbow. The model using these features was as effective in distinguishing among sit-to-stand intentions as the full model that included all seven of the markers (AUROC values for both models were nearly identical at 0.98). Comparing the fitted probabilities of the two models also showed no appreciable differences (**Table 1**). Taken together, these results suggest that the shoulder, head, and elbow serve to effectively differentiate among the possible actions.

We note that in addition to the shoulder, head, and elbow, the original PCA analysis showed that hand and hip also contribute meaningfully to the first principal component. However, hand and hip were not included in the final model from our stepwise regression. This outcome is

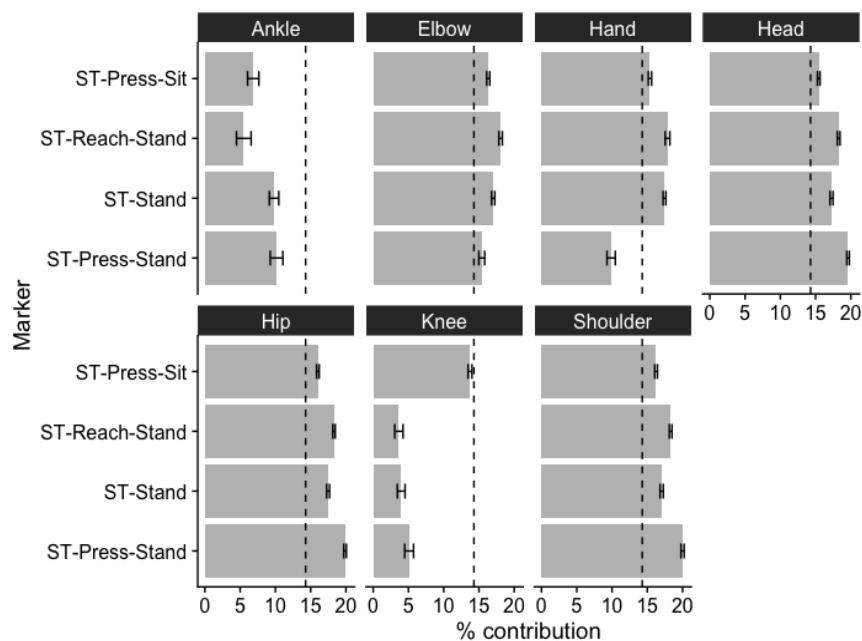


Figure 2: Mean percent contribution of each marker as a function of intention. The vertical dashed line in each plot represents the expected average contribution assuming the contribution of the markers were uniform.

Table 1 Percent variance accounted for by the 1st principal components as a function of the intentional movements.

Intention	PC1 Variance accounted
ST-Press-Sit	85.07 ± 5.09
ST-Press-Stand	70.39 ± 3.89
ST-Reach-Stand	76.99 ± 3.18
ST-Stand	80.50 ± 3.97

Table 2 Predicted probabilities for the parsimonious marker model v. the full marker model

Intention	Reduced model	Full model
ST-Press-Sit	99.8%	99.9%
ST-Press-Stand	99.9%	99.9%
ST-Reach-Stand	88.7%	90.3%
ST-Stand	88.0%	89.4%

likely due to the contributions of both hand and hip strongly correlating with other candidate markers, and thus acting as redundant sources of variation in our model. Given that the correlations between hand and elbow ($r = .84$) as well as head and hip ($r = .98$) were very high, we tested the performance of our current model against alternatives that included either hip in the place of head, hand in the place of elbow, or swapped both. While each of the alternative models performed very well (AUROCs > 0.90) the current model including shoulder, elbow, and head remained the most effective.

3. Experiment II – Perception of action intention

Given that previous research has demonstrated that intention is embodied in our movements and that observers can accurately identify actions based on the essential kinematics in point-light displays, our first aim in Experiment II was to determine if observers can correctly differentiate among the four possible action intentions of the actor from Experiment I. The results of our motion analysis in Experiment I suggest that variations in movement patterns among the shoulder, head, and elbow serve to differentiate the four possible actions. We hypothesized that the variations in contribution (i.e., the change in loadings) among intentional actions (e.g., *ST-Press-Sit* vs. *ST-Stand*, etc.)

may be useful for identifying which body segments are most important for differentiating among intentions. For example, if the relative contribution of a particular segment is consistently low and invariant across conditions, then it may be of low perceptual value-for distinguishing among intentions. Conversely, if the contribution is consistently high and/or varies greatly among intentions, then that body segment may be of high perceptual value. We hypothesized, therefore, that the same joint kinematics that were determined to differentiate the four possible actions in Experiment I, may likewise correspond to the kinematic information to which observers attend in order to differentiate the action intentions.

3.1. Method

3.1.1. Participants

In this experiment, we collected eye-movement data and intention judgements of 14 participants (3 female, 11 male; 18-24 years). Participants were undergraduate students at the University of Cincinnati who participated for research credit. The study was approved by the University of Cincinnati Institutional Review Board (IRB # 2012-2827).

3.1.2 Stimuli and design

To test our hypothesis, we created point-light displays from the processed movement data time series of the seven markers placed on the left ankle, elbow, wrist (hand), crown of the head, hip, knee, and shoulder of the participant in Experiment I (see also Section 2.1.3). Four movement trials were randomly chosen from the 25 performed sequences per intention to create four point-light sequences via Unity 3D. To simplify gaze data processing our “point lights” were represented by QR Codes that allow for automated data processing in D-Lab (Ergoneers Inc.); see **Figure 3**. In addition to the different *intentions*, we also created five different conditions with reduced amounts of available *body information*:

1. full body – all seven markers available
2. full body no arms – without elbow and wrist marker
3. upper body – without knee and ankle marker
4. upper body no arms – only hip, shoulder, and head marker available
5. lower body – only ankle, knee, and hip marker available

3.1.3 Setup and procedure

The point light displays were projected on a screen using a wall-mounted projector. In order to record eye movements and gaze, participants were asked to wear a Dikablis eye-tracking unit (Ergoneers, Inc.). Participants were seated centrally 2.7m back from the screen, and a small table with a keyboard in front of them. After obtaining informed consent, participants were shown the movie clips of a person performing the four intentional movements as recorded in Experiment I,

and were explained the intentions as: Standing with no reaching (*i.e. ST-Stand*), standing and reaching forward to grab something and remain standing (*i.e. ST-Press-Stand*), standing and reaching forward to grab something and then immediately returning to the seated position (*i.e. ST-Press-Sit*), and standing and reaching up as if turning on a ceiling fan (*i.e. ST-Reach-Stand*). No information was given about the different body information levels. Afterwards, the eye-tracker was donned and calibrated. Participants then watched the fully randomized point-light displays of each of the intentional movements and body information levels. For each trial, they were instructed to press a button when they were ready to identify which of the four intentions they were shown. They then verbally responded with their choice which was recorded manually by the experimenter, and the next display was started. Each of the four intentions was shown four times, in five different body information levels, resulting in 80 total trials per participant.

3.2. Perception of action intention

We conducted mixed-effects logistic regressions (with random intercepts for each participant) to examine the likelihood of correctly identifying the actor’s intention as a function of the *intention* displayed (*ST-Stand*, *ST-Reach-Stand*, *ST-Press-Sit*, *ST-Press-Stand*), and the *body information level* available in the point light display (*Full body*, *Upper body*, *Full body with no arms*, *Upper body with no arms*, *Lower body*). To address issues of complete separation, we obtained our parameter estimates using Bayesian logistic regression (R package: `bmle::bglmer`) using weakly informative priors (Gelman et al., 2008). Overall, the model indicated a significant interaction between our two predictors, $\chi^2(12) = 59.36, p < .001$.

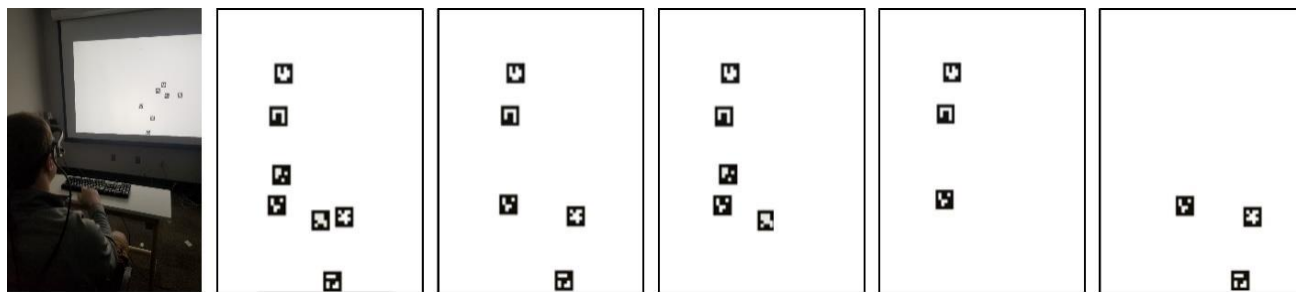


Figure 3: Experiment II Setup and Stimuli (from left to right): full body, full body no arms, upper body, upper body no arms, lower body.

Figure 4 expresses the estimated marginal effects of this model in terms of the participants' correct responses. The light blue bands represent the 95% CI and non-overlapping arrows indicate significant differences (Tukey HSD $p < .05$) between the proportion of correct responses by condition. As can be seen in **Figure 4**, participants were highly accurate at differentiating among intentions when provided full body information and remained accurate in discriminating intentions when only provided upper body information about the actor, with a notable drop-off in correct judgments for *ST-Stand*. All conditions with the notable exception of *ST-Press-Sit* saw a significant decrease in accuracy when arms were removed from the point light model. When only lower body information was present, participants performed at chance levels for *ST-Press-Stand* and *ST-Reach-Stand*.

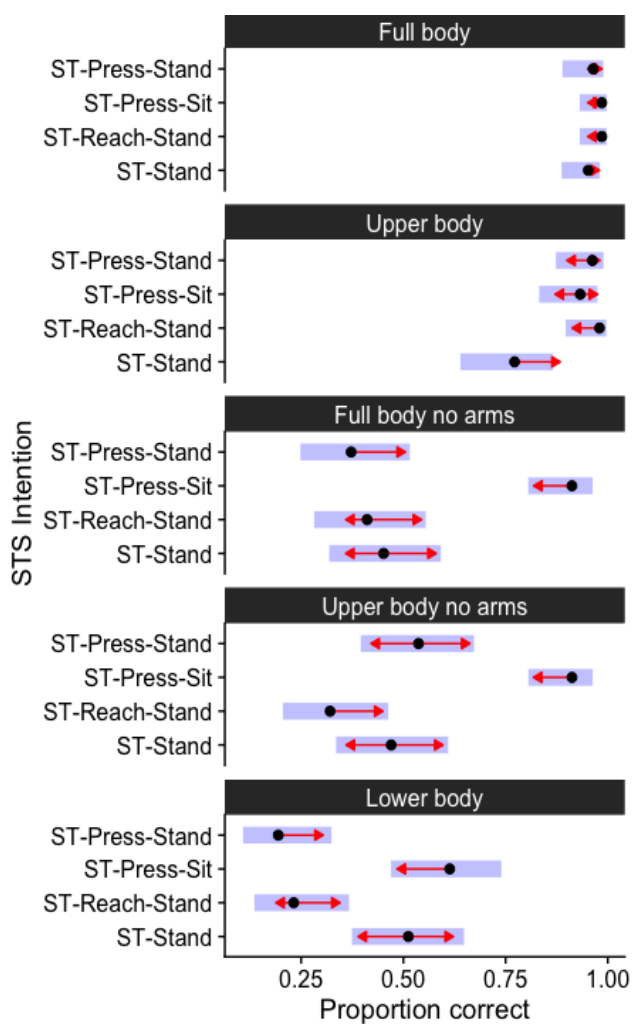


Figure 4: Estimated marginal effects of proportion correct by Intention and Body Information level.

3.3. Gaze focus

In order to determine if participants' gaze focused more on the joints that differentiate the four possible actions (shoulder, head, and elbow, see Experiment I), we submitted the frequencies of gaze fixations to each marker per trial per participant of correct trials (i.e., those trials in which participants identified the displayed intention correctly) to a generalized linear regression Poisson model. As predicted, the results of this model indicated that participants gaze tended to be focused on some body segments more than others, Likelihood Ratio Test, $\chi^2(6) = 3071, p < .001$. Follow-up pairwise contrasts of all segments against one another were significant except for Hand-Hip, Elbow-Head, and Ankle-Knee ($ps > .05$). As can be seen in **Figure 5**, participants overwhelmingly fixated upon the Shoulder, Head, and Elbow compared to the remaining four markers. As anticipated, participants' fixations aligned with the same joints that were identified in our motion analysis in Experiment I. Notably, perceivers fixations on Hand and Hip were relatively low, suggesting that they may have been sensitive to these markers as redundant sources of information and elected to attend to more prominent markers.

4. Discussion

Our goal in this study was to evaluate a strategy for investigating whether the kinematic information that differentiates actions is the same information to which observers attend when asked to perceive an actor's intended movement. An actor performed four different sit-to-stand actions (*ST-Stand*, *ST-Press-Stand*, *ST-Press-Sit*, and *ST-Reach-Stand*). The recorded kinematics were reduced to the kinematics of the major joints and the motion was truncated to end prior to the supra-postural action being completed (e.g., before a press or a reach). Observers were then asked, using a forced-choice paradigm, to select which action the actor intended to complete, and were highly accurate at this task. We then identified the joints that contribute to differentiating the actions using PCA and multinomial regression. The resulting contributing joints included shoulder, head, and elbow. This served to generate our hypothesis regarding participant gaze foci. We hypothesized that if perceptual information used to perceive action intention corresponds to the kinematic information that differentiates among the four possible actions, then observers' gaze should focus more on shoulder, head, and elbow than other joints of the actor.

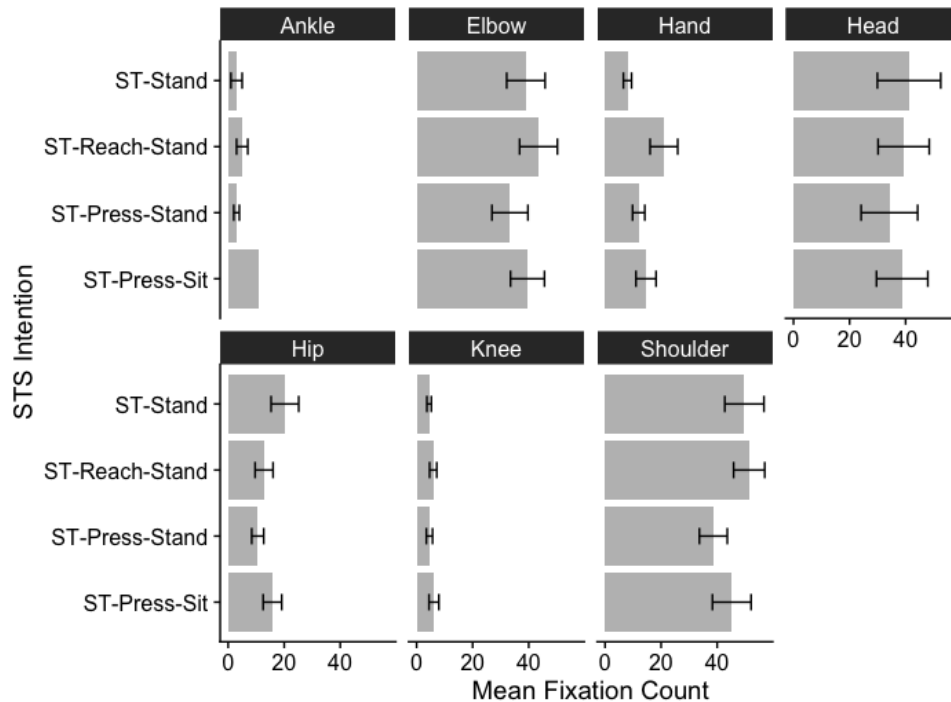


Figure 5: Mean number of fixations to each body marker as a function of the actors STS intention.

This hypothesis was supported by several findings. Perceptual performance was best in the full-body condition and nearly as good when the display included upper-body with arms. Perceptual performance sharply decreased when full- or upper-body was shown with no arms and still further when only lower-body joints were displayed. Further support is provided by the number of fixations observers displayed on certain markers before making a correct judgement on the intentional action displayed. Our findings are, therefore, consistent with central hypothesis that the joint kinematics that differentiate the possible actions under study are the same joint kinematics to which observers attend in order to successfully differentiate the movement intentions of the actor.

Our findings are consistent with previous studies demonstrating that intention is embodied in movements. This is evidenced by the fact that observers could correctly differentiate among the possible intended actions prior to those actions being executed. This is further evidenced by the alignment of our findings of the movement analysis with the findings of the perceptual task. In the movement analysis we found that the joint kinematics that serve to differentiate the

four possible actions included the shoulder, head, and elbow. Not only did participants' gaze focus correspond to those predicted joints, but performance on the perceptual judgments declined in conditions those joints were not visible to participants but did not decline when lower-body information was not visible to participants.

5. Limitations

The present study was limited to the kinematics of a single actor. We, therefore, do not make any claims regarding how generalizable our findings are regarding the particular structure that differentiates the set of actions we investigated. We acknowledge the very distinct possibility that studying different actors may yield different joints that differentiate their actions from each other or that the joints that differentiate their actions collectively may differ from the single actor we studied. Thus, we refrain from making any claims about whether the joints we identified (and/or their loadings with respect to PCs) would serve to differentiate all actors with respect to the actions studied. Rather we can only make claims about the particular actor under study and the structure involved in his particular actions. We would expect, however, that if different joints than

observed in the present study serve to differentiate among the actions of a different actor that observers would focus on those joints in order to successfully differentiate among the actions. The present analyses do not permit us to comment on the particular kinematic structure (meaning the spatio-temporal structure of the joint dynamics that specify the actions in question or which differentiates the actions). Our focus in this study was on the joints most relevant to differentiating the action, not on the dynamics of each joint, nor their relation to one another. The rationale for this focus was to test our strategy for studying the multiscale nature of the kinematics of action intention, rather than to make broad claims about the coordination dynamics themselves.

6. Conclusions

This study tested the utility of a strategy for studying information about movement intention at different scales of human performance (namely movement coordination and perception). Our findings demonstrate the promise of this approach in investigating how structure in movement embodies intention and how that same structure may inform perception. Our results also serve to expand upon Runeson's KSD principle which provides a theoretical framework for understanding how intention manifests in a consistent form at different scales of human performance.

Acknowledgements

We thank Julie Weast-Knapp for her guidance in the early stages of this project.

References

Abernethy, B., & Zawi, K. (2007). Pickup of Essential Kinematics Underpins Expert Perception of Movement Patterns. *Journal of Motor Behavior* **39**, 353–367.

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer.

Becchio, C., Sartori, L., & Castiello, U. (2010). Toward You: The social side of actions. *Current Directions in Psychological Science* **19**, 183–188.

Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Inferring another's expectation from action: the role of peripheral sensation. *Nature Neuroscience* **8**, 1295–1297.

Corbin, S., Moore, C., Patil, G., Rigoli, L., Davis, T. J., Shockley, K., & Lorenz, T. (2019). Kinematic Specification of Intention in Full-body Motion. *41st Annual Meeting of the Cognitive Science Society*.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society* **9**, 353–356.

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.

Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., & Gentilucci, M. (2010). To me or to you? When the self is advantaged. *Experimental Brain Research* **203**, 637–646.

Ferri, F., Campione, G. C., Dalla Volta, R., Gianelli, C., & Gentilucci, M. (2011). Social Requests and Social Affordances: How They Affect the Kinematics of Motor Sequences during Interactions between Conspecifics. *PLoS One* **6**, e15855.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics* **2**, 1360–1383.

Groner, R., & Schollerer, E. (2005). Perceived velocity of point-light walkers under complex viewing and background conditions 1,2. *Japanese Psychological Research* **47**, 204–215.

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

Jokisch, D., Daum, I., & Troje, N. F. (2006). Self Recognition versus Recognition of others by Biological Motion: Viewpoint-Dependent Effects. *Perception* **35**, 911–920.

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374**, 20150202.

Kadaba, M. P., Ramakrishnan, H. K., & Wootten, M. E. (1990). Measurement of lower extremity kinematics during level walking. *Journal of Orthopaedic Research* **8**, 383–392.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics* **21**, 575–580.

Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale* **211**, 547–556.

Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society B: Biological Sciences* **258**, 273–279.

Montepare, J. M., & Zebrowitz-McArthur, L. (1988). Impressions of People Created by Age-Related Qualities of Their Gaits. *Journal of Personality and Social Psychology* **55**, 547–556.

- Moore, C. H., Corbin, S. F., Lamb, M., Shockley, K., & Lorenz, T. (2019). Visual Extraction of Kinematic Information Specifying Intention in Full-body Motion. In L. van Dijk & R. Withagen (Eds.), *Studies in Perception and Action XV, 20th International Conference on Perception and Action* (pp. 145–148).
- Patil, G., Rigoli, L., Richardson, M. J., Kumar, M., Kiefer, A. W., & Lorenz, T. (2019). Joint Angle Variation in Intentional Sit-to-Stand Transitions. *IFAC-PapersOnLine* **51**, 214–219.
- Patil, G., Rigoli, L., Richardson, M. J., Kumar, M., & Lorenz, T. (2018). Momentum-based trajectory planning for lower-limb exoskeletons supporting sit-to-stand transitions. *International Journal of Intelligent Robotics and Applications* **2**, 180–192.
- Post, A. A., Daffertshofer, A., & Beek, P. J. (2000). Principal components in three-ball cascade juggling. *Biological Cybernetics* **82**, 143–152.
- Ramenzoni, V. C., Riley, M. A., Shockley, K., & Baker, A. A. (2012). Interpersonal and intrapersonal coordinative modes for joint and single task performance. *Human Movement Science* **31**, 1253–1267.
- Runeson, S. (1983). On visual perception of dynamic events. *Acta Universitatis Upsaliensis: Studia Psychologica Upsaliensia*. Academiae Ubsaliensis, Sweden.
- Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance* **7**, 733–740.
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General* **112**, 585–615.
- Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition* **119**, 242–252.
- Steel, K., Adams, R., & Canning, C. (2006). Identifying swimmers as water polo or swim team mates from visual displays of less than one second. *Journal of Science and Medicine in Sport* **9**, 41–42.
- Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics* **67**, 667–675.
- Troscianko, T., Holmes, A., Stillman, J., Mirmehdi, M., Wright, D., & Wilson, A. (2004). What Happens Next? The Predictability of Natural Behaviour Viewed through CCTV Cameras. *Perception* **33**, 87–101.
- Vallery, H., & Buss, M. (2006). Complementary limb motion estimation based on interjoint coordination using principal components analysis. *Proceedings of the IEEE International Conference on Control Applications*, pp. 933–938.
- Weast-Knapp, J. A., Shockley, K., Riley, M. A., Cummins-Sebree, S., Richardson, M. J., Wirth, T. D., & Haibach, P. C. (2019). Perception of another person's maximum reach-with-jump height from walking kinematics. *Quarterly Journal of Experimental Psychology* **72**, 2018–2031.
- Weast, J. A., Walton, A., Chandler, B. C., Shockley, K., & Riley, M. A. (2014). Essential kinematic information, athletic experience, and affordance perception for others. *Psychonomic Bulletin and Review* **21**, 823–829.