# Modeling human intention inference in continuous 3D domains by inverse planning and body kinematics

Yingdong Qian[1] and Marta Kryven[2] and Tao Gao[3] and Hanbyul Joo[4] and Josh Tenenbaum[5]

***Abstract*— How to build AI that understands human intentions, and uses this knowledge to collaborate with people? We describe a computational framework for evaluating models of goal inference in the domain of 3D motor actions. The framework receives as input the 3D coordinates of an agent's body, and of possible targets, to produce a continuously updated inference of the intended target. We develop three models of intention inference, based on the principle of least effort, formalized using different assumptions about the actor's constraints: Linear Extrapolation Heuristic, Parametric Curve Extrapolation Heuristic, and Generative Body Kinematics. We evaluate the models in three behavioural experiments using a novel Target Reaching Task, in which human observers infer intentions of actors reaching for targets among distracts. We show that the Generative Body Kinematics model outperforms the two heuristics in environments with obstacles, as well as during early stages of reaching actions, while the two heuristics make increasingly accurate predictions during later stages of reaching actions. Our results suggest that human efficiency in real-life intention inference is, at least in part, due to using information about body kinematics, and show that modeling body kinematic can improve performance of inference algorithms.**

## I. Introduction

Intention inference is one of the central questions in artificial intelligence (AI), critical to efficient human-AI collaboration, autonomous self-driving vehicles and automation. In humans, intention inference has been perfected by evolutionary pressures, and emerges early – during the first six months of life [27]. Six to 12 month old infants are already sensitive to the costs of other agents' actions, showing surprise when someone takes a longer route to a goal when a shorter route was available [18, 5]. The efficiency of such inferences relies on two core cognitive abilities. First, humans understand the actions of others in terms of optimized utilities – the *rationality principle* [3, 11] or principle of least effort [30]. Second, we understand physical interactions in terms of *intuitive physics* – an ability to approximately simulate physics in daily interactions [23].

Assuming rationality has proved useful to quantitative modeling of human inference of goals and intentions in 2D domains [1, 11, 17, 7]. Notably, the rationality principle states that agents minimize effort to achieve their goals, but does not specify that the agent's utilities are computed exactly [11]. For example, humans use a variety of approximations when computing utilities in simple planning tasks [16, 10, 9].

Understanding which utility functions do people use to plan their actions, and to understand the actions of others, is an active area of research, with a potential impact on building explainable AI that aligns its values with those of the humans [6, 4]. Previous work on inferring human intentions applied inverse inference and inverse reinforcement learning to observations of human behaviour [29, 6, 2, 21]. Deep-learning has also proved useful for recovering intentions of highly stereotyped behaviours, such as inferring a driver's lane-change intention, from video [28]. For general-case scenarios in continuous motor domains, reasoning about motion trajectories [8] and kinematics [24] plays an important role in how humans attribute intentions. Thus, we hypothesized that people could simulate the movements of an actor's body to anticipate the future end-point of a continuous motor-action. However, the high-dimensional nature of the continuous motor domains makes online inference in them computationally challenging [25].

Here we propose three models of human intention inference, which assume that humans minimize effort in continuous motor trajectories, formalized with various amounts of computations: Linear Extrapolation Heuristic (LinH), Parametric Curve Extrapolation Heuristic (ParamH), and Generative Body Kinematics (BodyGen). We show that the BodyGen model outperforms the two heuristics during the early stages of reaching actions and in environments with obstacles, suggesting that the human intention inference can be improved by simulating body kinematics. The two heuristics predict human inferences increasingly well at later stages of reaching actions, in contexts were simple extrapolations of trajectory can achieve accurate results. To our knowledge, our model is the first to combine a 3D physics engine and inverse planning in a computational framework of intention inference in a real-life physical environment.

## II. Target Reaching Task

*a) Hypothesis.:* To perform online intention inference humans could simulate the agents' physical constraints and affordances to predict how they may move to minimize physical effort while achieving their goals. To test this hypothesis we define our task domain as follows. A human actor is tasked with grasping an object among multiple distractors. The actor knows the location of the target, and may need to walk, turn, or reach over an obstacle to retrieve the target.

[1]Department of Computer Science, UCLA yd@ucla.edu
[2]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology mkryven@mit.edu
[3]Department of Statistics and Communication, UCLA end tao.gao@stat.ucla.edu
[4]The Robotics Institute, Carnegie Mellon University hanbyulj@cs.cmu.edu
[5]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology jbt@mit.edu

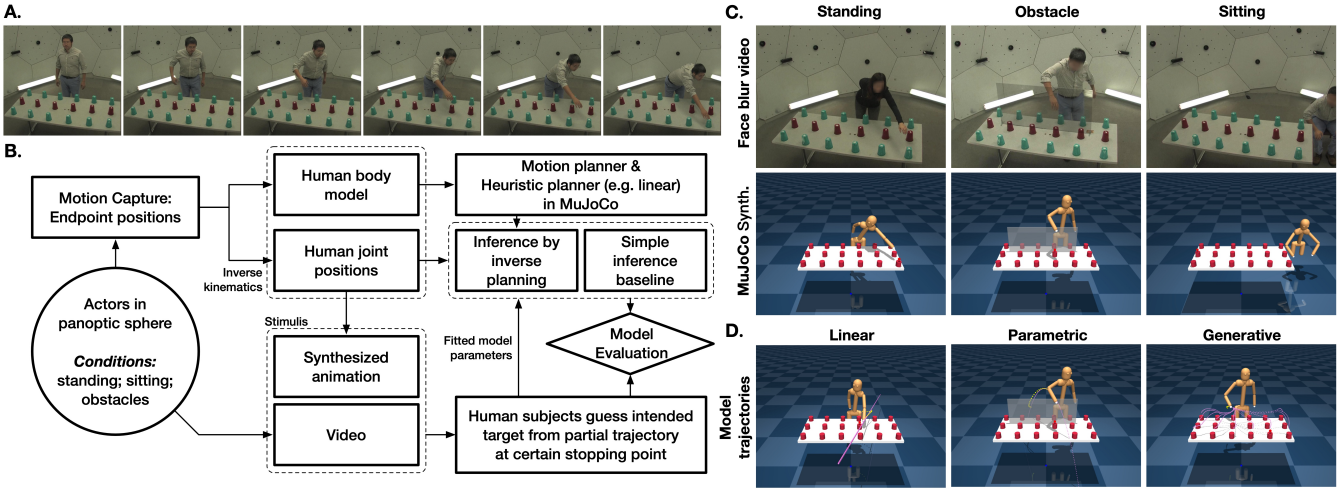


Fig. 1. **A.** A sequence of frames sampled from a stimulus in the Standing video condition. **B.** A diagram explaining our computational framework, stimuli generation procedure, and behavioural experiment. **(C)** Examples of environments shown in the six experimental conditions. In the obstacle condition, a transparent Plexiglas obstacle was placed in the middle of the table. The actor has to reach over the obstacle to reach objects on the outer side of the table. **(D)** Reaching trajectories predicted by the three models. LinH extrapolates a straight line from the history of observed wrist end-points, ParamH extrapolates a parametric curve in 3D space, and BodyGen simulates the most efficient movement of the skeletal figure to reach the 18 targets given the current position, acceleration and joint constraints. BodyGen complies with the laws of physics (e.g. Obstacles cannot be passed through) and follows the least effort principle.

Each trial starts with the actor in a neutral standing position, and ends when the actor's hand touches the target. Fig.1A shows an example of series of frames from such a recording. The Target Reaching Task corresponds to a variety of real-life behaviours, for example, workers manipulating parts at an assembly line; reaching for a car key, as we prepare to go out; picking up an object. Fig.1B illustrates the flow-chart of our experimental framework.

*b) Data recording.:* The actors' actions were recorded on video, as well as in 3D, using Panoptic Sphere – a multi-view camera system designed to re-construct human 3D skeletal structure and motion [15]. We recorded two individuals, a male and a female, who used both left and right hand to grasp the objects. The recordings were segmented into trials to generate individual stimuli. The video recordings were post-processed to blur out the actor's face, to ensure that human observers base their inferences on body motion alone, without relying on gaze cues. The 3D recordings were used as input to the inference models, as well as to generate a set of synthesized skeletal animation stimuli.

*c) Data quality.:* Critically, the 3D coordinate extraction by the Panoptic Sphere is noisy, which introduced ambiguity in the skeletal movement. On major source of noise in the recording occurs because the Panoptic Sphere does not record the positions of hands, or fingers, and so all modeling analysis must based on the position of the wrists. However, as we discovered during post-processing, actors have sometimes interpreted the instructions to reach an object as 'grasping the object with an entire hand' or 'touching the object with a finger-tip', resulting in inter-trial variability (up to $\pm 10\%$ of elbow-wrist length) in how closely the wrist end-point approached the target.

## III. Computational Models

We formalize efficient reaching by a probabilistic planning procedure, which can be inverted for goal inference. The environment is defined by the set of agent states $S$, environment states $W$, the set of goals $G$, and the set of actions $A$. Let $s_t \in S$ be the agent's state at time $t$, let $w \in W$ be the world state, $g \in G$ is the agent's goal (Targets in this case), $a_t \in A$ is the agent's action at time $t$ and $\tau$ is the ending time for a stimuli. By Bayes rule:

$$\Pr(g|s_{1:\tau}, w) \propto \Pr(g|w) \Pr(s_{1:\tau}|g, s_\tau, w) \qquad (1)$$

We approximate $\Pr(s_{1:\tau}|g, w)$ with different heuristic methods and generative models.

### A. Distance Heuristic

The Distance Heuristic sets the probability of each target as inversely proportional to the distance of the actor's wrist to the target, ignoring the history of the actor's movements.

Denote $\mathrm{WP}(s)$ as the wrist position for state $s$, $\mathrm{TP}(g)$ as the target position of goal $g$, and $d$ be the Euclidean distance function. The heuristic probability is defined as follows:

$$\Pr(g|s_{1:\tau}, w) \approx exp(-\theta d(\mathrm{WP}(s_\tau), \mathrm{TP}(g))) \qquad (2)$$

### B. Linear Extrapolation Heuristic (LinH)

The simplest way to extrapolate a reaching action is to assume that the agent intends to move to the target in a straight line. We extrapolate the endpoints of the actor's wrist position to a straight line, compute the shortest distance from each target to this line, and infer the target probability as inversely proportional to the distances. LinH does not model human body constraints. Formally, let $h_1$ be the look-back history length, $\alpha_1$ be the look-back probability length, $p(g)$ be the position of target $g$, let $l(s_{i:j})$ be the linear

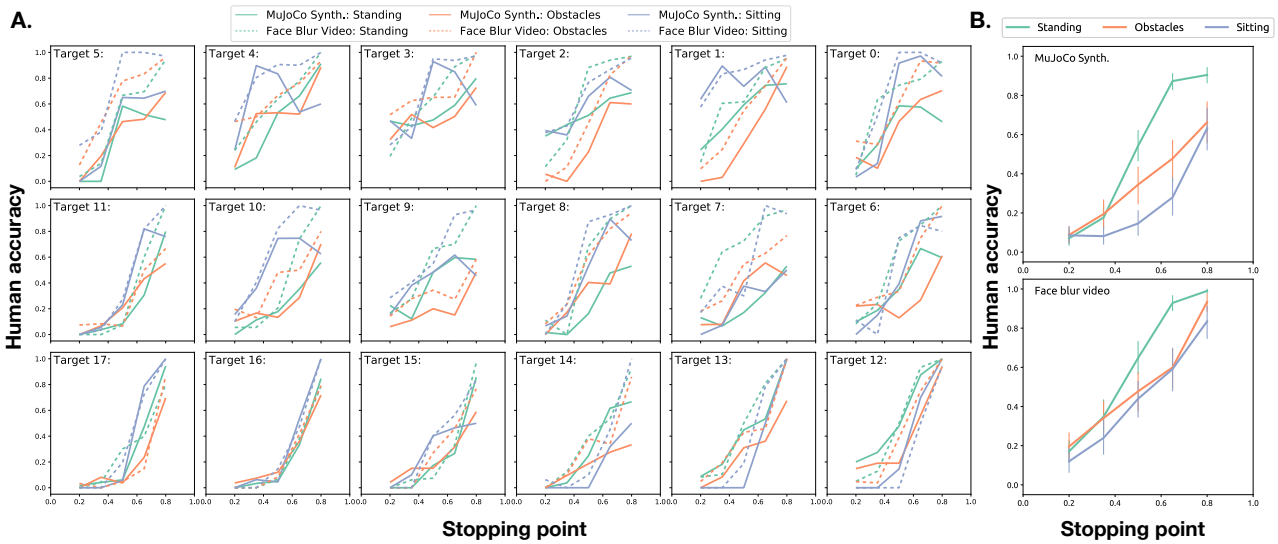


Fig. 2. Human performance across conditions. **(A)** Target detection accuracy in the six conditions. The grid layout corresponds to the object placement on the table in the experimental setup. Human accuracy differ among different targets. **(B)** Human performance in the six conditions aggregated across targets. The Standing conditions are the easiest, humans achieve a near-perfect accuracy after seeing 80% of frames in both Face Blur and Synthetic stimuli. The Sitting conditions are the hardest. Due to 3D noise, accuracy is overall lower in the Synthetic conditions.

fit of endpoints from states $s_i \cdots s_j$ and let $d_{\text{linear}}(l, p)$ be the shortest Euclidean distance from point $p$ to line $l$. The heuristic probability is defined as:

$$\Pr(s_{1:\tau}|g, w) \approx \prod_{k \in 0,1 \cdots \alpha_1} \exp(-\beta_1 d_{\text{linear}}(l(s_{\tau-k-h_1+1:\tau-k}), p(g))) \quad (3)$$

### C. Parametric Curve Extrapolation Heuristic (ParamH)

An important perceptual assumption in studies of biological motion, is that human and animal extremities move along circular trajectories [14]. This assumption is based on an implicit knowledge that limbs are attached to a torso, which limits degrees of freedom. Thus, ParamH assumes a simple constraint on wrist motion, but does not use body kinematics. The ParamH extrapolates the endpoints of the actor's wrists to a parametric parabola curve $(a_1x^2 + b_1x + c_1, a_2x^2 + b_2x + c_2, a_3x^2 + b_3x + c_3)$, and infers the target probability as inversely proportional to the shortest distance from the target to the curve. Let $pc(s_{i:j})$ be the parametric curve fit of endpoints from states $s_i \cdots s_j$ and let $d_{\text{pc}}(l, p)$ be the shortest Euclidean distance from point $p$ to line $l$. The probability is defined as:

$$\Pr(s_{1:\tau}|g, w) \approx \prod_{k \in 0,1 \cdots \alpha_2} \exp(-\beta_2 d_{\text{pc}}(pc(s_{\tau-k-h_2+1:\tau-k}), p(g))) \quad (4)$$

### D. Generative Body Kinematics (BodyGen)

Inference method consists of three parts (see Fig.1B). First two skeletal body models in MuJoCo [22] are fitted with the body proportions of the two actors. Second, a kinematic planning engine generates possible reaching trajectories to the candidate targets. Third, we calculate the probability of a given target based on the proximity of the generated trajectories to the observed human movement. The three steps are described in details below.

**Skeletal model construction.** The skeletal models are built using the sizes, proportions and joint constraints aggregated over the motion captures of each actor. Because of high computational costs of modeling walking and standing up from a chair, we modeled the torso as attached to a moving platform with constraints. We built the skeletal models with joint actuators controlled by parameter set $p_m$ including peak torque of joints, moving speed of torso, etc.

**Forward planning engine.** We implemented a stochastic trajectory optimization algorithm with long-term value approximation similar to [19]. The utility function is defined as $-p_1 d + p_2 h - p_3 \textbf{energy} - p_4 \textbf{contact}$, where $p_{1\cdots4}$ are free parameters, $d$ is the distance between the wrist and target, and $h$ is the average height of body center and head. The model's movements were optimized to minimize energy expenditure, while collision avoidance was enforced by penalizing contact force. We added constraints on the height of the body to encourage upright posture while reaching targets. Since the simulation in MuJoCo is deterministic, neural network based value functions were trained for each actor and condition with body posture and target positions. With value function, we used MPPI[26] algorithm to generate multiple reaching trajectories. We define $\text{Plan}(g, w, cs)$ be the function of generating future states given goal $g$ and starting state $cs$.

**Inverse planning as inference.** We used the above system for inverse planning. Let $q$ be the look-back history length. By assuming Markovian property:

$$\Pr(s_{1:\tau}|g, w) \propto \prod_{t \in 2 \cdots \tau} \Pr(s_t|s_{t-1}, g, w) \propto \prod_{t \in 2, 2+q, 2+2q \cdots \tau} \Pr(s_{t:t+q-1}|s_{t-1}, g, w) \quad (5)$$

The key challenge is to align the speed of human movement with our simulated movement, since both are sampled from a continuous sequence. We used Dynamic Time Warping over relatively long period $q$ to reduce the noise involved due to misaligned time step. Notably, human actors may have different rates of acceleration during reaching tasks. Let DTW() be the distance between two temporal sequences. To avoid the curse of dimensionality, and reduce the number of free parameters, we estimated the distance based on the wrist alone. This estimate can be refined by including other state information, such as body-mass centre, wrist velocity and acceleration of the wrist. We also run the planner multiple times to get the mean distances differences.

$$\Pr(s_{t:t+q-1}|s_{t-1}, g, w) \approx exp(-\beta_3 \overline{\text{DTW}(\text{WP}(s_{t:t+q-1}), \text{WP}(\text{PLAN}(g, w, s_{t-1})))}) \tag{6}$$

### E. Model Fitting

We used Maximum Likelihood Estimation to fit the free parameters of each model. We divided all trials in each condition into two sets. Targets with odd indexes (i.e. 1,3,5..) were assigned to the training set, and the remaining targets comprised the test set. We kept this target-wise split to reduce the possibility of over-fitting. We used Nelder-Mead to optimized parameters without gradient and Adam optimizer for the remaining parameters. The reported model performances are based on the aggregation of 100 optimized sets of parameters trained on sampled training set. Denote $H$ as human subjects, $M$ as models, $T$ as trials, $R$ as the ensemble of responses we collected i.e. $\{< h, t, stopping(h,t), target(h,t) >: h \in H, t \in T\}$ where $target(h,t)$ as the target human $h$ select for trial $t$ at stopping point $stopping(h,t)$. $\Pr(M|t,s,g)$ as the probability of model in predicting target $g$ at stopping point $s$ of trial $t$. $p*$ are the parameters of the models (Distance: $p* : \{\theta\}, LinH$ :p*:$\{\theta, h_1, \alpha_1, \beta_1, \}$, ParamH: $p* : \{\theta, h_2, \alpha_2, \beta_2\}$, BodyGen: $p* : \{\theta, q, \beta_3\}$). Our models are optimized as:

$$\underset{p*}{\text{argmax}} \left[\log(\Pr(M|R))\right] = \underset{p*}{\text{argmax}} \left[ \sum_{h \in H, t \in T} \log(\Pr(M|t, stopping(h,t), target(h,t))) \right] \tag{7}$$

## IV. Behavioural experiment

We conducted three behavioural experiments to measure human performance on the Target Reaching Task, and use it to fit our models. If humans use information about body kinematics to infer intentions of humans agents, then the BodyGen model should be better at predicting human inference compared to other models. Since body kinematics is computationally costly, we also hypothesized that people may use kinematic computations only when necessary, and may rely on the heuristics in easy scenarios, such when most of the reaching trajectory is observed.

### A. Experiment 1

We evaluated human performance in 3 types of environments and two presentation styles: $\{$*Standing, Obstacle and Sitting*$\}$ $\times$ $\{$*Face Blur Video, MuJoCo Synthethized*$\}$ – 6 experimental conditions in total. All environments contained 18 targets, placed on a table in a 3x6 layout, as shown in Fig. 1C. In the Standing environment the actor stood behind the table, and could reach all potential targets without walking. In the Sitting environment the actor sat in a chair, placed at the right side of the table, and had to walk to reach objects on the left side of the table. In the Obstacle environment a Plexiglas obstacle was placed in the middle of the table. The actor had to reach over the obstacle, or walk around it, to retrieve certain objects. Each environment was presented as a *Face Blur* – a video recording with actor's face blurred out, and as a *MuJoCo Synthethized* animation – a reconstruction of the 3-D positions of the actor rendered on a stick-man figure. The animated stimuli provide human observers with the same input as the models, to control for the noise in the recorded 3D positions and benchmark the models' performances.

**Subjects** We recruited 80 subjects (30 subjects in the Face Blur condition, and 50 subjects in the Synthesized conditions) on Amazon Mechanical Turk, who were paid for 50 minutes of work. This, and the following experiments, recorded anonymous data, collected no personally identifiable information, and were approved by our institutional IRB board.

**Procedure** The experiment was presented in a web-browser, using an interface developed in our lab. After reading instructions and providing informed consent, subjects completed a session of 6 practice trials, in which they received feedback. On each trial a subject pressed a start button to play a video. The video was then paused at one of the stopping points (20%, 35%, 50%, 65% and 80% of the total video duration), at which point the subject guessed the intended target by clicking on it. During the practice trials the subjects received an immediate feedback (a big green checkmark super-imposed over the video, if they were correct, or a super-imposed big red cross if they guessed incorrectly). After the feedback was presented, only during practice, the video continued playing until the end.

The actual experiment did not include trial-based feedback. During the actual experiment the videos paused for the subject to respond, and after responding, the subject was immediately forwarded to the next trial. People were shown feedback on their overall accuracy after every block (every 20 trials). The experiment lasted on average 45 minutes, and subjects had an opportunity to take a break between blocks. The subjects were told that we will reject submissions with poor performance, however no submissions were rejected. Detailed Experiment instructions and screenshots are given in the Supplement.

**Stimuli** The stimuli comprised 36 videos (18 targets, two actors) for each of the 6 conditions. Each stimulus was shown to each subject once, and paused at a random at one of the

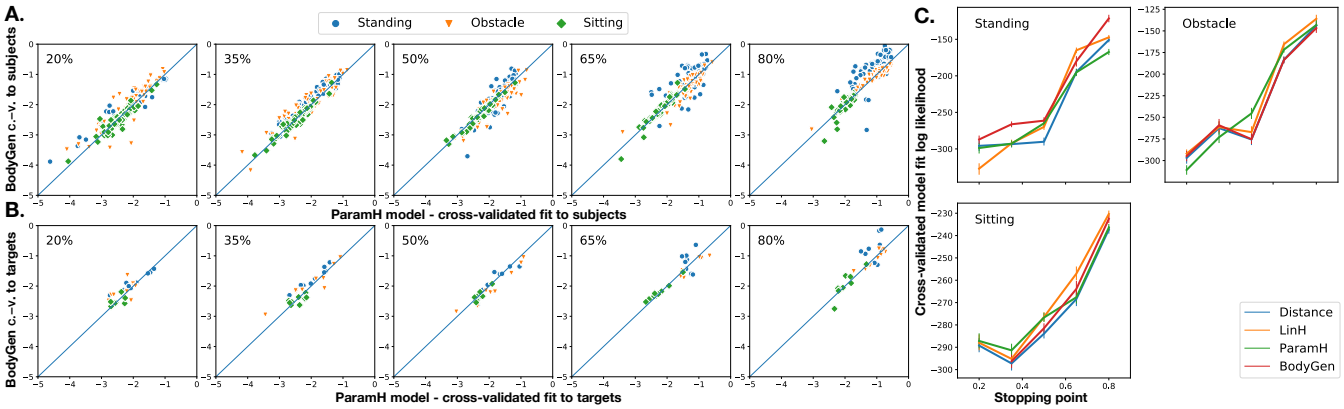


Fig. 3. Model comparison, Experiment 1. **(A)** Comparing fit of the BodyGen and Parametric heuristic to individual subjects across stopping points, with model fit optimized across targets. The Parametric heuristic explains a similar fraction of subjects overall. **(B)** Comparing fit of the BodyGen and Parametric heuristic to targets, with model fit optimized across people. BodyGen was better at predicting human inferences about a number of targets at each stopping point. **(C)** Comparing cross-validated model fits shows that the BodyGen model has a slight advantage over other models in the Standing and Obstacle conditions at early stopping points. Error bars show 95% confidence intervals.

5 stopping points.

**Empirical Results** Fig. 2A summarizes human performance across different targets in six conditions. For simple targets (e.g. target 3 in the center), people achieved $0\%$ accuracy after seeing as little as $50\%$ of the trajectory. In the Obstacle condition, the harder target (e.g. target 14 in middle of the bottom row) was almost never identified correctly. Performance varied between targets within condition. Fig. 2B summarizes human performance aggregated across targets. Overall, humans accuracy increased with the percentage of observed trajectory. Sitting and Obstacles conditions are more challenging to human subjects. Due to noise in the 3D recording the performance was lower in the Synthesized conditions, especially at early stages of Sitting conditions.

**Model-based Results** Fig. 3C summarizes the models' fit to individuals across stopping points and conditions. Overall, all three models made increasingly confident predictions over time. We did not find a significant difference between model at later stopping points, however the BodyGen model had a slight overall advantage over other models in the Standing and Obstacle conditions at early stopping points. Fig. 3A. shows the models' fit to individuals aggregated over targets, indicating that at each stopping point a fraction of individuals relied on BodyGen information for inference, however the majority of individuals was better explained by the heuristics. Comparing the models' performance on individual targets, as shown in Fig. 3B, revealed that for certain targets the BodyGen model was better at predicting human inferences compared to the heuristics(For example, Target 5 and 8).

**Discussion** We evaluated three models of intention inference on the Target Reaching Task in six different conditions, and found that inferences of certain humans, and about certain targets, are better predicted by the BodyGen model, suggesting that humans use body kinematics for inference. The BodyGen also had a slight advantage over other models during early stopping points overall.

Across all stopping points and conditions we found that most subjects were similarly explained by the ParamH and BodyGen. However, the high noise in the recording of the 3-D positions may have reduced the discriminating power of our experiment, given that in the MuJoCo Synthesized conditions, even humans were unable to correctly identify certain targets, even when the entire trajectory was shown. Furthermore, the high number of distractors may have adversely affected the BodyGen performance. In the following experiments we aimed to address these two points.

### *B. Experiment 2*

In the second experiment our goal was to remedy the noise in the 3D recording by shifting the targets to the ending position of the human wrist. We followed the procedure described in Experiment 1 with a different set of stimuli, all of which were synthesized animations made with MuJoCo.

**Subjects** We recruited 30 subjects on Amazon Mechanical Turk, who were paid for 40 minutes of work. None were excluded. Each subject saw stimuli in all three conditions, which were presented in a randomized order, and paused at one of the following stopping points: $35\%, 45\%, 55\%, 65\%$ and $75\%$. The stopping points in the Sitting condition were shifted forward by $10\%$. There were $140$ trials in total.

**Stimuli** The stimuli were created by modifying the three synthesized conditions (Sitting, Standing, Obstacle) of Experiment 1, by shifting the object positions to match the recorded position of the actors wrist at the end of the reaching actions. If the target's adjusted position was outside the table surface, the target was removed, meaning that some stimuli in the second experiment had fewer than 18 targets. Examples of stimuli are shown in Fig. 4A (top).

**Empirical Results** Fig. 4B summarizes human performance across conditions. Expected chance performance was around $0.06$. In the Standing and obstacle condition human performance was above chance at $45\%$, and in the sitting condition at $65\%$. Compared to the first experiment, humans were able to identify targets with higher accuracy and performance difference between Standing and Obstacles is now similar, attesting to the success of our experimental manipulation.

**Model-based Results** Fig. 5C summarizes the models' fit to individuals across stopping points and conditions. The

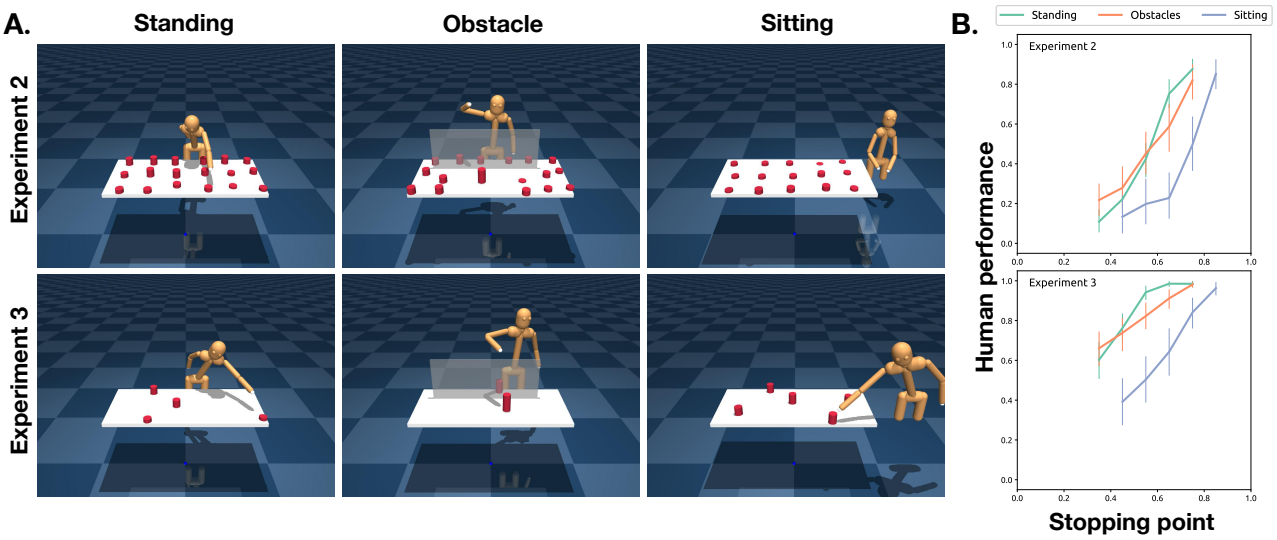


Fig. 4. **A.** Screen-shots of stimuli used in Experiments 2 and 3. **B.** Human performance in Experiments 2 and 3 in the three conditions aggregated across targets. Error bars indicate 95% confidence intervals.

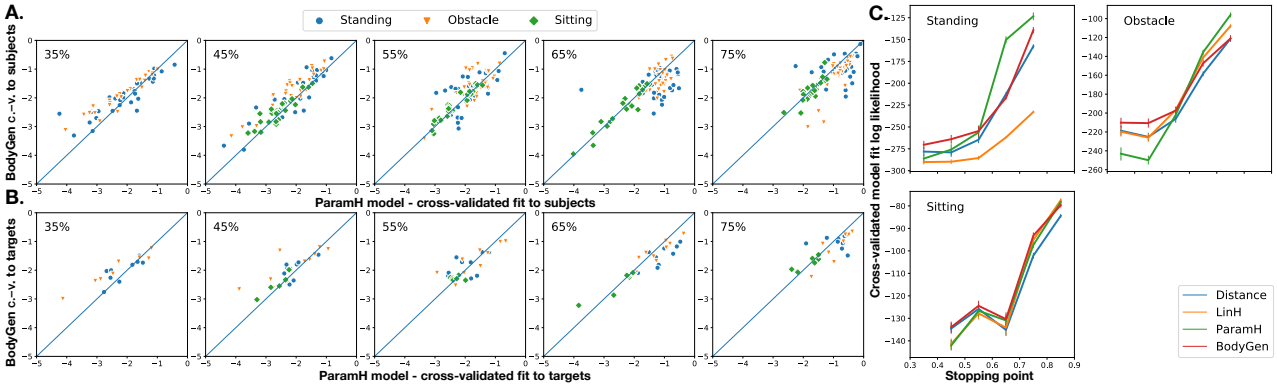


Fig. 5. Model comparison, Experiment 2. **(A)** Comparing fit of the BodyGen and Parametric heuristic to individual subjects across stopping points, with model fit optimized across targets. The BodyGen model explained a higher fraction of individuals during early stopping points. **(B)** Comparing fit of the BodyGen and Parametric heuristic to targets, with model fit optimized across people. BodyGen was better at predicting human inferences about a number of targets at each stopping point. **(C)** Comparing cross-validated model fits shows that the BodyGen model has a slight advantage over other models in the Standing and Obstacle conditions at early stopping points. Error bars show 95% confidence intervals.

three models made increasingly confident predictions over time, and achieved similar performance at later stopping points. The BodyGen model had a slight advantage over other models in the Obstacle conditions at early stopping points, and the Linear heuristic performed relatively worse in the Standing condition. Fig. 5A. shows the models' fit to individuals aggregated over targets, indicating that at each stopping point a fraction of individuals relied on BodyGen information for inference, and this fraction was higher at earlier stopping points. Comparing the models' performance on individual targets, as shown in Fig. 5B, revealed that for certain targets, the BodyGen model was better at predicting human inferences compared to ParamH, such as, for example, target 8 shown in 6D. The relative fit of heuristics improved at later stopping points.

**Discussion** In the second experiment we have replicated the main findings of Experiment 1, while better controlling for the noise in the recorded data. We have shown that the relative reliance on body kinematic information is stronger in the beginning of the trajectory, while in the later stages people tend to increasingly shift toward heuristics, possibly to minimize mental effort and decision time.

### *C. Experiment 3*

In the third experiment we used the target positions computed in Experiment 2, and decreased the number of distractors to simplify inference. For each target, the distractors were chosen by a pseudo-random sampling of 2 to 6 non-adjacent distractors. The experimental procedure was identical to the procedure used in Experiments 1 and 2.

**Subjects** We recruited 30 subjects on Amazon Mechanical Turk, who were paid for 40 minutes of work. None were excluded.

**Stimuli** Each subject saw stimuli in all three conditions, which were presented in a randomized order, and paused at one of the following stopping points: $35\%, 45\%, 55\%, 65\%$ and $75\%$. The stopping points in the Sitting condition were shifted forward by $10\%$. There were 160 trials in total.

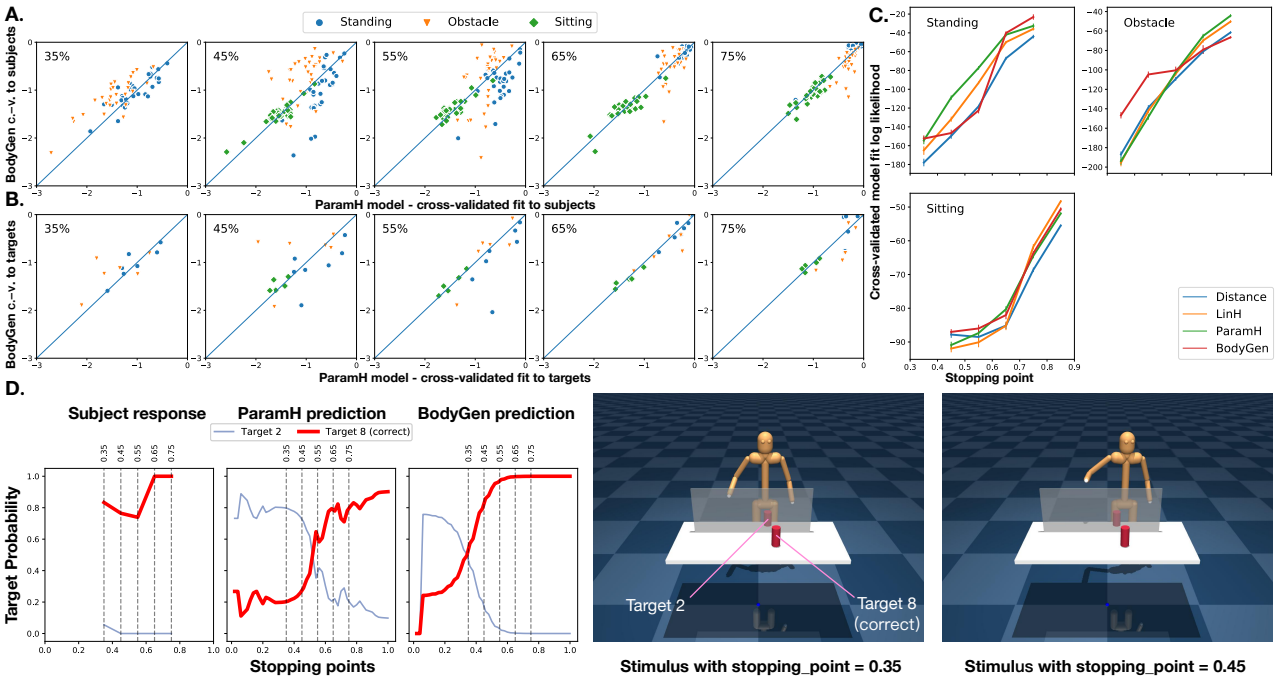


Fig. 6. Model comparison, Experiment 3. **(A)** Comparing fit of the BodyGen and Parametric heuristic to individual subjects across stopping points and conditions, with model fit optimized across targets. In the Obstacle condition during the 35% and the 45% stopping points most individuals were better predicted by the BodyGen model. **(B)** Comparing fit of the BodyGen and Parametric heuristic to targets, with model fit optimized across people. BodyGen was better at predicting human inferences about a number of targets during early stopping points. **(C)** Comparing cross-validated model fits shows that the BodyGen model has an advantage over other models in the Obstacle condition at early stopping points. Error bars show 95% confidence intervals. **(D)** An example of a stimulus on which BodyGen outperforms ParamH. Target 8 is correctly identified by the BodyGen model as soon as the robot starts to lift its arm to reach over the obstacle. ParamH is able to identify this target only after the actor's hand has passed over the obstacle, toward its other side.

**Empirical Results** Fig. 4B summarizes human performance across conditions. Expected chance performance was between 0.2 and 0.5 depending on number of targets presented. Human performance was above chance at all stopping points and in all conditions.

**Model-based Results** Fig. 6C summarizes the models' fit to individuals across stopping points and conditions. The three models made increasingly confident predictions over time, and achieved similar performance at later stopping points. The BodyGen model had a advantage over other models in the Obstacle conditions at early stopping points. Fig. 6A. shows the models' fit to individuals aggregated over targets, indicating that in the Obstacle condition, at each stopping point a fraction of individuals relied on kinematic body simulation approximated by BodyGen. Comparing the models' performance on individual targets, as shown in Fig. 6B, revealed that for certain targets the BodyGen model was better at predicting human inferences compared to the Parametric heuristic during early stopping points. Figure 6D shows an example of a stimulus on which BodyGen outperforms ParamH during the early stage of the reaching action. Like humans, BodyGen is able to infer the correct target from seeing the robot rise its hand to begin reaching over the barrier. In contrast, ParamH is unable to make a correct prediction until later in the trial, after the robot's wrist has passed the high point above the barrier, which allows the model to complete the parabolic extrapolation.

## V. General Discussion

textbfContributions We have developed three models of human intention inference, and evaluated their performance in six interactive physical scenarios, and in three behavioural experiments, using a novel Target Reaching Task. We found that while simple extrapolations of wrist trajectory can reasonably predict human intention inferences, in certain scenarios prediction of human intention inference can be improved by modeling body kinematics. We show that body kinematics is particularly important during the early stages of reaching actions, as well as in environments with obstacles, where the human agent's trajectory can not be extrapolated in a trivial way. Understanding how humans interpret each others' motor trajectories has many multidisciplinary benefits to robotics (better collaboration), AI (modeling planning), and cognitive science (theory of mind).

**Future work** While we found differences in the models' fit to individual observers and to specific targets, future work needs to address these differences in more detail. In future work we intend to investigate the consistency in the use of goal inference strategies between individuals, and between tasks. If people revert to heuristics to minimize cognitive effort, then people should perform more heuristic inferences under conditions of cognitive load, even when the use of heuristic decreases accuracy. We also intend to investigate under which conditions people can identify the best-performing heuristic in a given environment, and when do they rely on the generalizable but costly inference based on body kinematics.

**Limitations** The Target Reaching Task takes a step toward measuring intention inference in natural behaviors, however it is impossible to do justice to the complexity of human motor behaviours with just one task. In future work we intend to extend the modeling of intention inference to complex task-and-motion scenarios, that involve sequences of actions. Playback of stimuli at a different speed in a screen using a stick-man may also be unnatural for the subjects and results in different inference pattern. For BodyGen model, we could improve the planning quality by adding more realistic human characters such as muscle-driven models [12, 13] and more advanced algorithms in controlling them (e.g. [20]). We could train neural networks which output belief by comparing predicted trajectories and observed trajectory. We could also sample potential target positions and use particle filtering to increase the robustness of the performance.

## REFERENCES


[1] Chris L Baker, Rebecca Saxe, and Joshua B Tenenbaum. "Action understanding as inverse planning". In: *Cognition* 113.3 (2009), pp. 329–349.

[2] Daniel S Brown and Scott Niekum. "Deep bayesian reward learning from preferences". In: *arXiv preprint arXiv:1912.04472* (2019).

[3] Daniel Clement Dennett. *The intentional stance*. MIT press, 1989.

[4] Owain Evans, Andreas Stuhlmüller, and Noah Goodman. "Learning the preferences of ignorant, inconsistent agents". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 30. 1. 2016.

[5] György Gergely and Gergely Csibra. "Teleological reasoning in infancy: The naıve theory of rational action". In: *Trends in cognitive sciences* 7.7 (2003), pp. 287–292.

[6] Dylan Hadfield-Menell et al. "Cooperative Inverse Reinforcement Learning". In: *Advances in neural information processing systems*. 2016, pp. 3909–3917.

[7] Steven Holtzen et al. "Inferring human intent from video by sampling hierarchical plans". In: *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2016, pp. 1489–1496.

[8] De-An Huang et al. "Motion reasoning for goal-based imitation learning". In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2020, pp. 4878–4884.

[9] Quentin JM Huys et al. "Interplay of approximate planning strategies". In: *Proceedings of the National Academy of Sciences* 112.10 (2015), pp. 3098–3103.

[10] Yash Raj Jain, Frederick Callaway, and Falk Lieder. "Measuring how people learn how to plan." In: *CogSci*. 2019, pp. 1956–1962.

[11] Julian Jara-Ettinger et al. "The naıve utility calculus: Computational principles underlying commonsense psychology". In: *Trends in cognitive sciences* 20.8 (2016), pp. 589–604.

[12] Yifeng Jiang and C Karen Liu. "Data-driven approach to simulating realistic human joint constraints". In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE. 2018, pp. 1098–1103.

[13] Yifeng Jiang et al. "Synthesis of biologically realistic human motion using joint torque actuation". In: *ACM Transactions On Graphics (TOG)* 38.4 (2019), pp. 1–12.

[14] Gunnar Johansson. "Visual perception of biological motion and a model for its analysis". In: *Perception & psychophysics* 14.2 (1973), pp. 201–211.

[15] Hanbyul Joo et al. "Panoptic studio: A massively multiview system for social motion capture". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2015, pp. 3334–3342.

[16] Marta Kryven et al. "Adventures of human planners in Maze Search Task". In: *ICRA 2021 workshop: Social Intelligence in Humans and Robots*. 2021.

[17] Marta Kryven et al. "Plans or Outcomes: How do we attribute intelligence to others?" In: (2021).

[18] Shari Liu and Elizabeth S Spelke. "Six-month-old infants expect agents to minimize the cost of their actions". In: *Cognition* 160 (2017), pp. 35–42.

[19] Kendall Lowrey et al. "Plan online, learn offline: Efficient learning and exploration via model-based control". In: *arXiv preprint arXiv:1811.01848* (2018).

[20] Xue Bin Peng et al. "AMP: Adversarial Motion Priors for Stylized Physics-Based Character Control". In: *arXiv preprint arXiv:2104.02180* (2021).

[21] Deepak Ramachandran and Eyal Amir. "Bayesian Inverse Reinforcement Learning." In: *IJCAI*. Vol. 7. 2007, pp. 2586–2591.

[22] Emanuel Todorov, Tom Erez, and Yuval Tassa. "Mujoco: A physics engine for model-based control". In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2012, pp. 5026–5033.

[23] Tomer D Ullman et al. "Mind games: Game engines as an architecture for intuitive physics". In: *Trends in cognitive sciences* 21.9 (2017), pp. 649–665.

[24] Maryam Vaziri-Pashkam, Sarah Cormiea, and Ken Nakayama. "Predicting actions from subtle preparatory movements". In: *Cognition* 168 (2017), pp. 65–75.

[25] Zhikun Wang et al. "Probabilistic movement modeling for intention inference in human–robot interaction". In: *The International Journal of Robotics Research* 32.7 (2013), pp. 841–858.

[26] Grady Williams et al. "Autonomous racing with autorally vehicles and differential games". In: *arXiv preprint arXiv:1707.04540* (2017).

[27] Amanda L Woodward. "Infants selectively encode the goal object of an actor's reach". In: *Cognition* 69.1 (1998), pp. 1–34.

[28] Yang Xing et al. "An ensemble deep learning approach for driver lane change intention inference". In: *Transportation Research Part C: Emerging Technologies* 115 (2020), p. 102615.

[29] Brian D Ziebart et al. “Maximum entropy inverse reinforcement learning.” In: *Aaai*. Vol. 8. Chicago, IL, USA. 2008, pp. 1433–1438.

[30] George Kingsley Zipf. “Human behavior and the principle of least effort: an introd. to human ecology”. In: (1949).