Inferring epistemic intention in simulated physical microworlds

Stephanie Droop (stephanie.droop@ed.ac.uk)
Institute for Language, Cognition and Computation
University of Edinburgh

Neil Bramley (neil.bramley@ed.ac.uk
Department of Psychology
University of Edinburgh

Abstract

We explore whether people can recognise the *epistemic goal* or intention of active learners interacting with simulated physical objects. In a novel online experiment, 110 adults watched screen recordings of other adults ("players") manipulating objects in a 2D simulated physical microworld. Players had either the goal of identifying the nature of a hidden magnet-like force connecting two of the objects, or the relative masses of those two objects. Observers were then asked to identify the learning goal of the player they were observing. By drawing from a previously collected dataset of active physical learning interactions and an ideal observer analysis, we systematically manipulated how informative the actions of the player were about the target property, while also manipulating observers' level of access to the behaviour such that some participants could see the players' micro-control actions as well as their impact on the physical objects. We found observers were better at identifying the goals of successful players and of players trying to identify the force than the mass property, while the micro-dynamic evidence trace seemed to improve accuracy on identifying the mass goal. We use mixed methods to explore what cues our observers used to make these judgments, and discuss implications for social cognition in the wild.

Keywords: active learning; epistemic goal; social inference; intention; intuitive physics

Introduction

When making sense of others' behaviour, we are said to take an *intentional stance* (Dennett, 1987), assuming others pursue their best interests rationally given their beliefs. This means behaviour provides a window on what others' beliefs and interests are likely to be. However, we know people are not perfectly rational, and even if they were, reverse engineering their goals and beliefs is computationally challenging even in the simplest of toy scenarios. In more naturalistic contexts, the problem seems to only get harder, since moment-to-moment behaviour may reflect interim or epistemic goals—i.e. solving subproblems or learning things about the environment only indirectly related to one's ultimate intentions. Nevertheless, a number of recent computational accounts have modelled social cognition as Bayesian inverse planning (Baker et al., 2017; Blokpoel et al., 2013; Jara-Ettinger, 2019). For example, studies have had participants reason about the preferences and beliefs of artificial agents based on observing their trajectories through simple "gridworld" environments containing obstacles and potentially rewarding destinations (Baker et al., 2009, 2017), reverse engineering what the agent must believe or like in order to behave that way. In one scenario, participants see the agent walk past a nearby food truck to reach another, and might reasonably conclude that the agent prefers the food sold at the more distant vendor (else, why make the extra effort?). If they instead travel past the first food truck around an occlusion, coming into view of a second food truck, then turn around and return to the first, we might conclude that they prefer the closer food option than the occluded one, but that they had anticipated the existence of some third option, preferable to both others. Adults and even children have been shown to be capable of such "naive utility calculation", at least in settings where behaviours and options are salient and unambiguous (Jara-Ettinger, 2019). Related research has shown that people can also make use of finer-grained behaviour traces such as body kinematics (Cavallo et al., 2016; McEllin et al., 2018). But goals come in many kinds: *epistemic* goals—the desire to learn about something, or resolve some form of uncertainty about the world (Sandoval, 2015)—may be rather more inscrutable than a desire for some worldly reward. The challenge here is that, in order to test their hypotheses and resolve their uncertainty, people necessarily take actions whose outcomes are uncertain even to them, meaning that what they want to learn and what they achieve can diverge in complex ways. It is an open question, which we begin to explore here, to what extent people can recognise the epistemic intentions behind others' epistemic actions.

We see this ability or task as important to focus on because it combines our physical and social expertise, two areas of skill sorely lacking in AI systems. If we could find a way of modelling this intersection then we could give AI systems a key to understanding humans and our way of being in the world.

Intuitive physics

A key domain in the cognitive science of learning is *intuitive physics*: the understanding shared by embodied beings of how physical objects move in space (Kubricht et al., 2017; Ludwin-Peery et al., 2021; McCloskey, 1983). One popular experimental setup borrowed from computer vision involves reasoning about *billiard worlds* (Fragkiadaki et al., 2015): realistic dynamic simulations of 2D, billiard ball-like objects interacting within a bounded space in ways controlled by a physics simulator (see Figure 1). Because the objects move independently but also respond realistically to contact, such billiard worlds are a useful testbed for studying how people infer the latent causally relevant properties of objects, both passively (Ullman et al., 2018) and actively (Bramley et al., 2018), as well as how they reason hypothetically and counterfactually about what will or did happen and why (Gerstenberg et al., 2021). Thus, we see billiard worlds as a useful setting for exploring how people infer the learning goals of others from relatively fine-grained behaviour; a grown-up version of toddlers at nursery learning how a toy works by watching one another play.

Active learning

Numerous experiments have demonstrated the value of active learning—in which the learner chooses what to do or look at next—over passive learning—in which the learner merely observes the situation of interest (Markant & Gureckis, 2014; McCormack et al., 2016; Sobel & Kushnir, 2006). Active learning is critical in domains where experiments can reveal latent properties and relationships that are confounded in passive observation (Bramley et al., 2015, 2018; Pearl, 2000). A typical comparison group in active learning experiments is a *yoked* condition, where participants observe the actions of an active learner but cannot influence them, so matching the information available but removing the control and the privileged access to the learners' epistemic goals. Often yoked participants are less accurate than their active counterparts (Markant & Gureckis, 2014) but occasionally they do as well or better. In active learning about physical properties, yoked learners have been found to approach the accuracy of their active counterparts *provided* they share a learning goal. For example, Bramley et al. (2018) found yoked learners were able to identify the force relation or relative mass of two objects in a billiard world as accurately as the active learner they observed, but were substantially worse when the learner they observed was focused on a different property. An information gain analysis revealed that active learners generally generated more evidence about their learning goal than the alternative goal. Secondary analyses also identified a number of micro-experimental strategies used by participants, with some associated with the mass goal—shaking the objects back and forth, launching, knocking and throwing them at each other—and some associated with the force goal—holding the target objects close together, bringing them to a stop, moving the distractor objects out of the way.

Experiment

In our experiment, we ask whether observers can recognise what active learners are trying to learn on the basis of observing their interactions with the objects through the relatively fine-grained medium of real-time touchscreen control. We use a similar setup to Bramley et al. (2018), using videos of adult active learners collected as controls for a developmental project Bramley & Ruggeri (in revision). The videos show participants (hereafter, "players") using touchscreen control to investigate one of two latent properties of objects in a billiard world. We selected a subset of the videos so as to systematically manipulate the information content of the recordings under an Ideal Observer model of physical property induction (described below) and also manipulate whether participants can see players' touchscreen control or simply the results of this control in the movements of the objects.

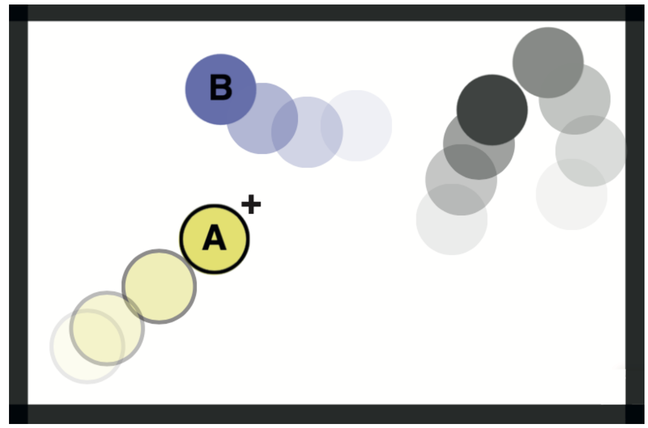

Figure 1: Conceptual depiction of dynamic motion of objects. Faint shading indicates the trajectory of the objects over a couple of seconds. Here object A is controlled by the player as shown by its black border. Their dragging action, pulling A toward the top right, is further indicated by inclusion of the cross showing their finger position.

Methods

Participants

We recruited 110 UK-based adults from the Testable Minds crowdsourcing platform (49 female, 2 other, age Mean $\pm$ SD 36.9 $\pm$ 11.4, range 18-71 years). Participants were paid \$4.50 for taking part in the experiment and \$0.10 for each correct answer (mean \$5.46, min \$4.80, max \$5.90). The task lasted 26.9 $\pm$ 9.3 minutes.

Stimuli

Stimuli were 32 video clips, each showing a 45-second interaction between an adult active learner (the "player") and a 2D simulation of four bouncing objects ("balls"). The environment was simulated using Box2D physics engine and the interactions took place through real-time touchscreen control on a tablet screen using a full screen javascript web app. The interactions were chosen from a set of 192 recorded at museums around Berlin (Bramley & Ruggeri, in revision). Players could control the objects by "grabbing" them by holding their finger on the touchscreen over the object, whereupon the balls would become attracted to the finger until touch was released. Players could thus manipulate the objects in real time to produce curated interactions and dynamics but began each trial ignorant about key latent properties of the objects. The physical environment was 6×4 meters and rendered on the tablet screen as 1920×1200 pixels. The physics simulator refreshed 60 times per second resulting in 2,700 frames of evidence making up each clip. Of the four objects in each clip, two were described as "targets" given labels "A" and "B" and coloured pseudo-randomly while the other two were "Distractors" coloured light and dark grey. Players were given one of two possible goals: Mass goal (identifying which of the target objects was heavier) or force goal (identifying whether the target objects attracted or repelled each other). In fact, all objects weighed 1kg except one target which weighed 2kg, and the targets either attracted or repelled one another with a magnet-like force of $\pm\frac{3\,\text{m/s}^2}{\text{distance}^2}$. The other five pairwise combinations of target and distractor had randomly selected force relationships $\in \{\text{attract}, \text{repel}, \text{none}\}$. This resulted in complex and confounded dynamics in the absence of curation by the active learner. All objects began each trial with randomly generated positions and velocities. All stimuli, along with detailed simulator settings, can be viewed in our Repository. Of 192 trials from 24 adults, 16 representative videos were selected based on the following procedure: 8 from each learning goal (force, mass). These comprised 2 from each of the 4 combinations of environment types ({A heavy, B heavy}, {attract, repel}). These were selected such that one was the most, and the other was the least, informative about the target property based on an information gain analysis (below), subject to them also containing 5-25 "actions" (defined as occasions where the player took control of an object for at least 1/6$^{\text{th}}$ of a second). We created two versions of each video. In the *No Cursor* version, the video just showed the objects and indicated which if any was under control with a thick black line. In the *Cursor* version, the video also showed where the player's finger was located whenever touching the screen, with a cross symbol (Figure 1).

Information entropy of stimuli The videos we used are linked with an information entropy analysis in (Bramley & Ruggeri, in revision; Shannon, 1948). This is based on an account of physical parameter inference via simulation. While the details are beyond the scope of the current paper,we provide code for this analysis in our Repository. Roughly, this involved calculating a posterior probability distribution over the unknown physical parameters of each world given the dynamics produced by the players' interactions, under an assumption of a computationally unbounded observer. This involved simulating each environment forward under many combinations of unknown parameter settings, and measuring the instantaneous divergence of these simulations in terms of the direction and velocity of the objects. These discrepancies were run though a simple model of Gaussian perceptual uncertainty so as to assign a likelihood to each frame of observed dynamics under every possible combination of properties in the task, and this was combined with a uniform prior to give a posterior joint distribution. This posterior was then marginalised over to assess the posterior uncertainty of the mass and force properties at the end of each trial. It is important to note that the perceptual uncertainty layer depends on an arbitrarily set precision parameter, meaning these values provide a measure of *relative* informativeness of different dynamic interactions, rather than an absolute value.

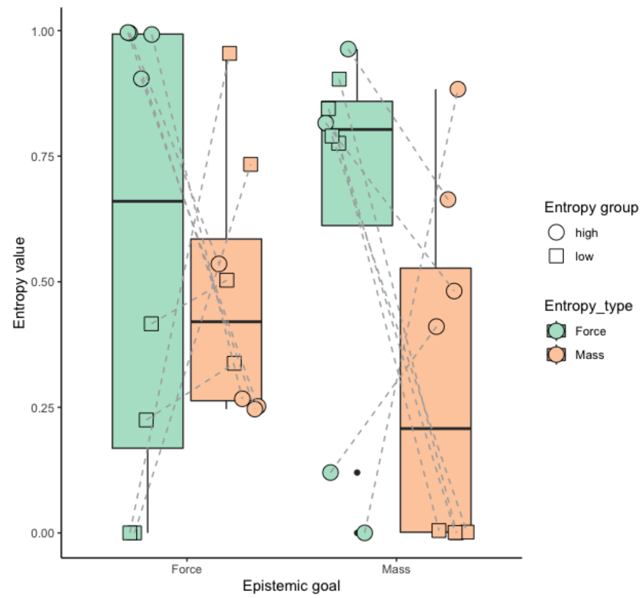

Figure 2: Force and mass entropy of video stimuli. 1 indicates maximal uncertainty (equally likely that A or B is heavier, or equally likely that objects attract vs repel). 0 indicates ideal observer certainty about which object is heavier or about which force relation is correct. Dotted lines connect the force and mass entropies for each video.

Figure 2 shows the resulting mass and force entropy (i.e., an ideal-observer measure of posterior uncertainty) for the clips we used. Mass uncertainty and force uncertainty were negatively correlated (r = -.624), with force entropy lower in trials where the player's goal was force (Mean $\pm$ SD $.57 \pm .45$ vs. $.65 \pm .37$), mass entropy lower on trials where the player's goal was mass ($.31 \pm .35$ vs. $.48 \pm .26$), but mass entropy was lower in general than force entropy ($.39 \pm .30$ vs. $.61 \pm .39$). For selecting trials and predicting guesses, we used the entropy specific to the player's epistemic goal, i.e. mass entropy for mass trials, force entropy for force trials. However, below we also consider the ratio of force entropy to mass entropy in analysing how our participants identified learners' goals.

Design

We ran an online behavioural experiment. We manipulated one factor between subjects (*Trace:*{Cursor, No cursor}), and two factors within-subject (*Goal*: {Force, Mass}; and *Entropy*: {High, Low}. Each participant thus faced all 16 videos described above either with or without the cursor visible, in random order.

Procedure

The experiment was implemented in Testable with participants completing it in the browser on their own devices. Participants first entered their basic demographics then read instructions about the nature of the task explaining the two possible learning goals of the players. Participants then watched a video familiarising them with the virtual physical environment. They all watched the same video, which showed four objects interacting with each other with no active learner intervention, with the properties {Attract; B heavier}. Participants had to correctly complete four True/False comprehension questions before proceeding to the main task. If they answered incorrectly they were looped back to the instructions again. They could attempt the comprehension questions as many times as they needed.

Instructions were the same for all participants with the exception of an additional sentence explaining that the black cross showed the player's touch actions in the Trace: Cursor condition. Apart from this between-subject manipulation of cursor, all participants saw the same 16 stimuli, in random order. At no point did participants interact with the stimuli. Participants received no other information about the player whose goal was the target of their inference.

While each video stimulus was playing, the following prompt appeared: "Was this person testing for force or mass?". When the video finished playing the next screen presented the same prompt but now alongside two response buttons labeled "FORCE" and "MASS". Participants had one minute to respond by clicking one of the buttons. The experiment then proceeded to a screen with a free text response field and the prompt "Why did you answer that? What clues did you base your answer on? How sure were you? (Be as detailed as you like)". Answering this was optional and participants could proceed to the next trial. Participants did not receive feedback during the task but were provided their success rate in the final debrief screen.

Analysis

Data were analysed using R version 4.1. Package *lme4* (Bates et al., 2014) (glmer, family "binomial") was used for logistic regression mixed effects models following recommendations of Meteyard & Davies (2020). Participants' correct responses were predicted with main effects of trace, goal and entropy group and a random intercept for participant. Our model was contrast coded (-.5,.5) and included all interactions.

Results

Here we first report participants' performance in the task overall and by condition. We then tease out the contribution of the information content of players' actions. Finally we describe an exploratory mixed-methods analysis investigating what cues participants might be guided by in their answers.

Performance

Participants were able to infer the intention of the active learner correctly Mean $\pm$ SD 60.1$\pm$14.2% of the time overall, doing so 62.2$\pm$14.9% of the time in the trace present condition, 58.0$\pm$13.1% in the trace absent condition, 67.2$\pm$17.1% for the force goal and 53.4$\pm$16.2% for the mass goal, and 54.3$\pm$16.9% for high and 66.4$\pm$22.6% for low entropy videos. χ^2 tests show that participants were above chance

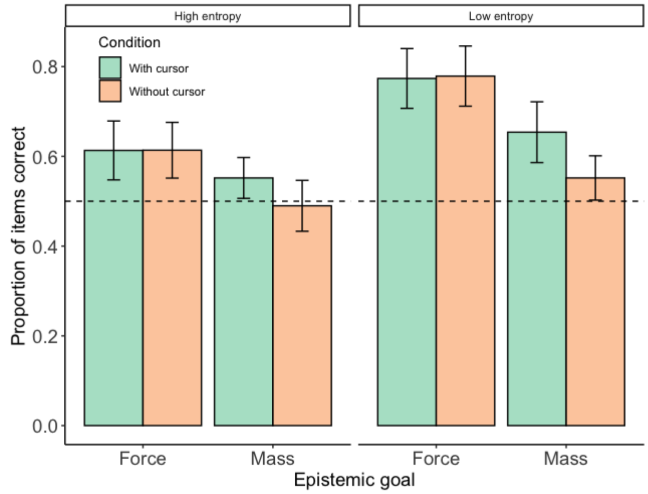

Figure 3: Proportion of correct answers in all conditions. Dashed line shows chance performance. Error bars show standard error for individuals.

50% in 5 of the 8 condition combinations, with p values below a Bonferroni corrected value of .0063. Three exceptions were the cursor, mass goal, high entropy condition: χ^2 = 1.17, p = .28; the no cursor, mass goal, high entropy condition, χ^2 = 3.56, p = .06; and the no cursor, mass goal, low entropy condition: χ^2 = 1.03, p = .31. There were 14 timeouts: 1 participant had 4 and 10 others had 1 each. By trial: 2 stimuli had 3 timeouts each, 2 had 2, and 4 had 1. We simply omitted these trials from analyses. No participants were excluded. The full pattern of results are shown in Figure 3. Logistic mixed-effects regression shows a significant main effect on accuracy of epistemic goal (odds ratios, estimate = 1.39, CI [1.25 1.53], Z = 6.41, $p < .001^{***}$) and of entropy (odds ratios, estimate = 1.30, CI [1.18 1.44], Z = 5.18, $p < .001^{***}$), with an interaction between cursor and goal such that participants were more accurate at identifying when the player had the mass goal when the cursor was present (odds ratios, estimate = 1.11, CI [1.01 1.23], Z = 2.11, $p = .035^{*}$ but no main effect of cursor trace (odds ratios, estimate = .92, CI [.82 1.03], Z = -1.48, p = .14)

Performance by ground truth

We also analysed performance by ground truth, the environmental properties present in every trial (whether A and B attracted or repelled each other, and whether A or B was heavier). Irrespective of goal, participants got more correct answers when the ground truth was repulsion rather than attraction (χ^2 = 7.49, $p = .006^{**}$; Table 1). When split by epistemic goal, the repel trials do not influence correct answers in the force condition (χ^2 = .015, p = .90). However, when the goal is mass, repel trials yield more correct answers than attract (χ^2 = 15.87, $p < .001^{***}$). This was not due to less information being produced in the attract conditions.

Table 1: Accuracy (%) by Ground Truth (Force, A/B Heavy)

Ground Truth:	Attract, A	Attract, B	Repel, A	Repel, B
Force Goal:	63.2	70.5	71.8	63.2
Mass Goal:	33.2	53.2	70.9	54.4

Information

We next explored whether judgments could plausibly be driven by the relative evidence about the two goals. That is, whether participants simply assessed which property was better revealed and decided that was the property the learner wanted to learn about. To assess this, we ran a binomial logistic mixed-effects regression to predict participants' probability of guessing mass, depending on the posterior mass and force entropies, with interaction and a random effect of participant. This regression did not have the predictors from the main performance model of cursor trace, epistemic goal and entropy group. We found a main effect of force information entropy (odds ratios, estimate = 3.69, CI [1.84 7.50], Z = 3.65, $p < .001^{***}$), whereby people were more likely to answer mass when the force entropy is high. Actual frequency of answers was 1001 force; 759 mass. We next ran another binomial logistic mixed-effects regression to check whether the logarithm of the ratio of force entropy to mass entropy could predict whether participants guess mass. We found a main effect of entropy ratio (odds ratios, estimate = 1.03, CI [1.02 1.05], Z = 4.30, $p < .001^{***}$). This regression kept the main predictors of cursor trace and epistemic goal, and their effects were still present and significant, which shows that although the entropy ratio accounts for some variance, it cannot account for all. This suggests people were guided by objective information content to some extent when answering, but not so much that they relied on it as their only signal.

Exploratory analyses

We performed exploratory analysis of participants' free text explanations for why they answered how they did. For this, free text responses were stripped of participant or trial data and manually coded by two independent coders. See our Repository for details of how each response was coded. We observed that participants often mentioned the player's physical actions, that is, they recognised some form of strategic behaviour (e.g. "bumping A into B", "dragging the balls close to each other"). They also sometimes directly mentioned the end goal (e.g. "checking the mass", "seeing if they attract or repel") and sometimes a more specific interim sub-goal which seemed to isolate one property (e.g. "to see which floats higher"). Therefore, we opted to code each response for presence or absence on three dimensions: *strategy*, *end goal*, and *interim goal*. As all the objects had physical properties, we did not accept answers that simply referenced the behaviour of the objects (e.g. "Balls A and B were sticking together throughout"). Inter-rater reliability (Cohen's κ) was calculated for each of the three dimensions: for strategy κ = .75, for end goal κ = .47, for interim goal κ = .56. Out of 1336 text responses received, the raters (Rater1/Rater2) counted that 769/919 mentioned some kind of strategy, 428/741 mentioned an end goal and 455/531 mentioned an interim goal. For further analysis we counted only those responses rated True by both raters. We ran a logistic mixed-effect regression predicting accuracy with fixed effects for strategy, end goal and interim goal and a random effect of participant. We found a main effect of strategy (odds ratios, estimate = 1.27, CI [1.02 1.60], Z = 2.10, $p = .03^{*}$) and of end goal (odds ratios, estimate = 1.43, CI [1.09 1.88], Z = 2.56, $p = .01^{*}$) but not of interim goal (odds ratios, estimate = 1.17, CI [.89 1.55], Z = 1.15, $p = .25$).

Strategies and goals

Participants identified a range of strategies of the players, describing detailed actions that helped them solve the task. For example, many mentioned an action that seems particularly informative as a test for force, characterised by selecting one ball and moving it gently and precisely toward another without touching ("They gently moved the ball towards each other", "holding it close to others"). For some participants a cue indicative of tests for force was more time spent passively observing: some participants mentioned they thought the player had left the balls alone on purpose and merely watched them to allow A and B to reveal their properties. In contrast, participants frequently cited crashing or collisions as a sign the player was testing mass: "they were crashing the balls together". Strategies like this "initiating collisions" were often mentioned alongside what we coded as "end goals", e.g. this was followed by, "to see which ones looked heavier". One other strategy that prompted participants to judge the player as focused on mass was when they intervened on A and B separately: "seeing which bounced faster to compare them", "drag the balls on an individual basis". Interim goals were rarer. People mentioned interim goals after sophisticated and detailed actions which would isolate a property by how the balls moved, e.g. "moved one quickly to see if the other followed", "dragged them to the bottom to see which one floated up more quickly".

Discussion

In this paper we explored inference about the epistemic goals of others. We found that adults were often able to judge whether the screen recording they watched was by another adult trying to ascertain the pairwise magnet-like force between two objects, or the relative masses of those two objects. Performance was mixed and responses indistinguishable from chance in some settings, such as when identifying less-successful mass-revealing behaviour (i.e. when mass entropy was still high at the end of the trial according to an ideal observer account). However, they could identify what more successful players were testing for, and seemed to find it easier to identify force-focused behaviour. One limitation is that binary forced-choice design is a maximally lenient way of collecting judgements. It allows pragmatic inferences which we see some evidence for in free responses: several participants wrote e.g., "They didn't do the force test so I knew it was mass". Nevertheless, our experiment shows that people are able to make use of rich behavioural traces in inferences about epistemic goals.

Interactions with environmental properties

People were more accurate when the objects repelled one another than when they attracted. We note the comparison to past work on learning in this setting (Bramley & Ruggeri, in revision; Ullman et al., 2018; Bramley et al., 2018) that showed passive participants struggled to identify repulsion. It seems that being able to act on the scene (bringing repulsive objects closer together than they would normally go) allowed active participants to identify repulsion at least as well as attraction, but here also seemed to make it easier for participants to identify mass-focused behaviour. In future work combining intuitive physics and social inference, we hope to tease out how learners might adapt their goal-directed actions to "piggyback" off helpful properties (like using repulsion to help reveal mass).

Information entropy

There was a relationship between the quality of the evidence the player produced about each property and our participants' judgments. However, this did not fully explain response patterns. If judgments were driven by the amount of evidence generated about the two properties we might expect them to guess mass more often than force, as the ideal observer model considered there to be stronger evidence about mass than force on average. In fact we observed the opposite pattern: participants were more likely to answer force. Participants were better at identifying force-focused than mass-focused behaviour, akin to earlier work which found both passive and active learners to be more accurate at identifying force than mass Ullman et al. (2018); Bramley et al. (2018), and yoked learners as accurate as active participants on force trials but not mass trials.

It could be helpful to tease out what exactly people are sensitive to, and how this differs from what the ideal observer account is sensitive to. The ideal observer learns a lot about mass from collisions (Bramley et al., 2018). Since heavy objects are deflected less than light ones, they exit in different directions depending on their mass. But this evidence depends on seeing exactly how the objects collided and correctly resolving the exchange of momentum. People, however, seem more sensitive to qualitative aspects of the force evidence (objects swerving toward or away from each other).

For our current purposes we should not expect people to benefit from the same evidence as an ideal observer model which has perfect precision and memory and considers all possibilities in parallel with the objective of inferring the latent properties of the objects. In contrast, our participants had the objective of *social* inference of the players' epistemic goal. It seems for this they were especially sensitive to differences in the behaviour they observed. The next section discusses how the free responses we collected suggest people may have recognised behavioural strategies or interim goals that they could map to one or other of the learning goals.

Recognising strategies and goals

We took an innovative approach to gain traction on the question of how people knew what others were trying to do: we *asked* them. This approach may seem naive: given the high dimensional stimuli and impression-based task it might seem unlikely that people would have articulatable insights about the workings of their own mental apparatus. However, we found plenty of hints that people are capable of recognising apparent strategies, sometimes describing them and linking them plausibly to interim and final learning goals. Mention of some strategies recalls the secondary analysis of Bramley et al. (2018), with many participants identifying similar actions, for example "holding the target objects close together" as a test for force. Some strategies may be more easily recognised than others, and we cannot rule out an additional symbolic-interpretability effect linked to participants' ease in *describing* certain actions (a direction for further work). In addition, some people noted that force-focused behaviour had characteristically slower motion and fewer total interventions and some mentioned they could tell the player was "waiting to see" whether the balls moved together or apart. This suggests that social inference is based just as much on absence of any action, or the timing of action, as on the actions themselves.

Although not a significant predictor of accuracy, the rarer mentions of interim goals give intriguing insight in two ways. Firstly, they show some participants had a sense of which types of motion would effectively reveal certain physical properties (e.g. "to see which bounced slower" as a test for mass). Secondly, these goals are evidence of people imputing fine grained intentional behaviour ("he was trying to drag the balls into a line") which has implications for theory of mind research.

Behavioural trace

Identification of mass-focused behaviour, but not force-focused behaviour, benefited from extra information provided by display of the players' touch control. In one sense this is not surprising: the additional trace increases how much evidence we have, e.g. how quickly the player changed direction or where they paused. It also provides insight into how they produced the actions they did: how quickly they dragged their finger, whether in a straight or curved trajectory. The cursor also reveals actions that were failed or aborted, such as attempting to grab an object but missing it.

Conclusion

In sum, our results, while preliminary, suggest people are capable of recognising active learning strategies, as well as what evidence is produced by the behaviours of others. By combining this information they are able to make sensible guesses about epistemic goals.

References

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1–10.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329–349.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Blokpoel, M., Kwisthout, J., van der Weide, T. P., Wareham, T., & van Rooij, I. (2013). A computational-level explanation of the speed of goal inference. *Journal of Mathematical Psychology*, *57*(3-4), 117–133.

Bramley, N. R., Gerstenberg, T., Tenenbaum, J. B., & Gureckis, T. M. (2018). Intuitive experimentation in the physical world. *Cognitive Psychology*, *105*, 9–38.

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 708.

Bramley, N. R., & Ruggeri, A. (in revision). *Children's active physical learning is as effective and goal-targeted as adults'*.

Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports*, *6*(1), 1–8.

Dennett, D. C. (1987). *The intentional stance*. MIT Press.

Fragkiadaki, K., Agrawal, P., Levine, S., & Malik, J. (2015). Learning visual predictive models of physics for playing billiards. *arXiv preprint arXiv:1511.07404*.

Gerstenberg, T., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological Review*.

Jara-Ettinger, J. (2019). Theory of mind as inverse reinforcement learning. *Current Opinion in Behavioral Sciences*, *29*, 105–110.

Kubricht, J. R., Holyoak, K. J., & Lu, H. (2017). Intuitive physics: Current research and controversies. *Trends in Cognitive Sciences*, *21*(10), 749–759.

Ludwin-Peery, E., Bramley, N. R., Davis, E., & Gureckis, T. M. (2021). Limits on simulation approaches in intuitive physics. *Cognitive Psychology*, *127*, 101396.

Markant, D. B., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94.

McCloskey, M. (1983). Intuitive physics. *Scientific American*, *248*(4), 122–131.

McCormack, T., Bramley, N., Frosch, C., Patrick, F., & Lagnado, D. (2016). Children's use of interventions to learn causal structure. *Journal of Experimental Child Psychology*, *141*, 1–22.

McEllin, L., Sebanz, N., & Knoblich, G. (2018). Identifying others' informative intentions from movement kinematics. *Cognition*, *180*, 246–258.

Meteyard, L., & Davies, R. A. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, *112*, 104092.

Pearl, J. (2000). *Causality*. Cambridge University Press.

Sandoval, W. A. (2015). Epistemic goals. *Encyclopedia of Science Education*, 393–398.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423.

Sobel, D. M., & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition*, *34*(2), 411–419.

Ullman, T. D., Stuhlmüller, A., Goodman, N. D., & Tenenbaum, J. B. (2018). Learning physical parameters from dynamic scenes. *Cognitive Psychology*, *104*, 57–82.