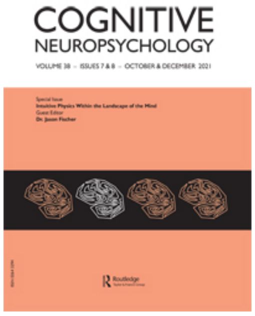

Cognitive Neuropsychology

ISSN: 0264-3294 (Print) 1464-0627 (Online) Journal homepage: www.tandfonline.com/journals/pcgn20

Naturalistic embodied interactions elicit intuitive physical behaviour in accordance with Newtonian physics

Nils Neupärtl, Fabian Tatai & Constantin A. Rothkopf

To cite this article: Nils Neupärtl, Fabian Tatai & Constantin A. Rothkopf (2021) Naturalistic embodied interactions elicit intuitive physical behaviour in accordance with Newtonian physics, Cognitive Neuropsychology, 38:7-8, 440-454, DOI: 10.1080/02643294.2021.2008890

To link to this article: https://doi.org/10.1080/02643294.2021.2008890

Published online: 08 Dec 2021.

COGNITIVE NEUROPSYCHOLOGY
2021, VOL. 38, NOS. 7–8, 440–454
https://doi.org/10.1080/02643294.2021.2008890

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Naturalistic embodied interactions elicit intuitive physical behaviour in accordance with Newtonian physics

Nils Neupärtl [a,b], Fabian Tatai [a,b] and Constantin A. Rothkopf [a,b,c]

[a]Institute of Psychology, TU Darmstadt, Darmstadt, Germany; [b]Centre for Cognitive Science, TU Darmstadt, Darmstadt, Germany; [c]Frankfurt Institute for Advanced Studies, Goethe University, Frankfurt, Germany

ABSTRACT
The success of visuomotor interactions in everyday activities such as grasping or sliding a cup is inescapably governed by the laws of physics. Research on intuitive physics has predominantly investigated reasoning about objects' behaviour involving binary forced choice responses. We investigated how the type of visuomotor response influences participants' beliefs about physical quantities and their lawful relationship implicit in their active behaviour. Participants propelled pucks towards targets positioned at different distances. Analysis with a probabilistic model of interactions showed that subjects adopted the non-linear control prescribed by Newtonian physics when sliding real pucks in a virtual environment even in the absence of visual feedback. However, they used a linear heuristic when viewing the scene on a monitor and interactions were implemented through key presses. These results support the notion of probabilistic internal physics models but additionally suggest that humans can take advantage of embodied, sensorimotor, multimodal representations in physical scenarios.

ARTICLE HISTORY
Received 10 April 2021
Revised 26 August 2021
Accepted 13 November 2021

KEYWORDS
perception and action; embodied cognition; intuitive physics; probabilistic modelling; continuous action control task

Introduction

In everyday situations, observations of and visuomotor interactions with our environment are inescapably governed by the laws of physics. Whether sliding a cup of coffee on a counter-top in your kitchen or observing a puck sliding on a sheet of ice in sports, objects' dynamics and interactions are well described by Newton's laws. Despite our sense of understanding such scenes, numerous psychological experiments have revealed pervasive biases and errors in human judgement of physical relationships, from the walker paradigm (McCloskey et al., 1983) to projectile motion after cutting a pendulum (Caramazza et al., 1981), and contended that these stem from fundamental physical misconceptions. Similarly, systematic errors in judgements of relative masses when observing object collisions have been attributed to participants relying on approximate rules of thumb, so called heuristics (Cohen, 2006; Gilden & Proffitt, 1994; Todd & Warren, 1982).

By contrast, a stream of recent works have suggested that pervasive errors and biases in human intuitive physical judgements do not stem from fundamental misconceptions about physics but can be explained on the basis of probabilistic inference involving generative models of physical scenarios: Sensory measurements are inherently uncertain, ambiguous and probabilistically interact with our internal physical models and prior knowledge. By reverse engineering participants' thought processes (Griffiths et al., 2010), human behaviour in experiments on intuitive physics (Kubricht et al., 2017) has been successfully explained with probabilistic models based on exact or approximate physical laws while also taking human limitations such as lack of precision in perception and action execution into account: From object perception (Kersten et al., 2004; Knill & Richards, 1996), reasoning about object masses (Bramley et al., 2018; Sanborn et al., 2013), efficient tool use (Allen et al., 2020), and fluid motion (Bates et al., 2015; Kubricht et al., 2016) to

CONTACT Nils Neupärtl neupaertl@psychologie.tu-darmstadt.de Institute of Psychology, TU Darmstadt, Darmstadt 64283, Germany; Centre for Cognitive Science, TU Darmstadt, Darmstadt64283, Germany

hierarchical probabilistic reasoning about physical properties (Ullman et al., 2018) and the transfer of physical reasoning to a visuomotor task (Neupärtl et al., 2020).

Common to probabilistic models of intuitive physical reasoning is a rather disembodied approach to cognition: Inference about physical scenarios is based on probabilistic representations of physical quantities (Sanborn et al., 2013), symbolic representations of objects and relationships (Ullman et al., 2018), or geometric descriptions with physical properties (Battaglia et al., 2013), which are thought to be extracted from two-dimensional computer rendered images of scenes. Reasoning unfolds in mental models akin to physics game engines (Ullman et al., 2017) or by generating programmes through probabilistic programming (Ullman et al., 2018). In line with this view, participants' responses mostly consist of binary judgements (Kubricht et al., 2016; Sanborn et al., 2013) or estimation of a single parameter (Battaglia et al., 2013). Very rarely subjects can interact with scenes at all, but then by simulating different manual interactions and tool use with a computer mouse (Smith et al., 2013) or by simulating a touch by clicking with a computer mouse on simulated two-dimensional objects rendered on a computer screen (Bramley et al., 2018; Smith et al., 2018) instead of through visuomotor actions, as in everyday situations. Given evidence that cognition is at least partly grounded in mechanisms for interaction with the environment, that is, mechanisms of sensory processing and motor control in specific situations (Anderson, 2003; Foglia & Wilson, 2013; Wilson, 2002), this raises the question, whether intuitive physical reasoning may take advantage of such embodied representations. Accordingly, we hypothesized that the mode of visuomotor interaction in an intuitive physical reasoning scenario may affect the responses, whether it is a button press simulating an interaction or a multimodal visuomotor interaction with physical objects.

Here, we consider two variants of a visuomotor control task to investigate whether naturalistic, multimodal, embodied interactions elicit the same physical behaviour as less representative, more abstract task designs. Subjects were asked to propel pucks into a target's bulls-eye positioned at different distances across trials. In a first condition, the scene was rendered on a monitor and the interaction was achieved through a button press on a keyboard, as in Neupärtl et al. (2020). In a second condition, subjects were immersed in a virtual environment viewed through a head-mounted display (HMD) and interacted with a real hockey puck, sensed its weight, and slid it on a real table. The visual displays were adjusted in exploratory experiments to result in comparable uncertainties about the target's distance. The physical simulations were identical. In both conditions, subjects were not given any feedback about the puck's movement and final position to ensure that they could only rely on their a priori internal model and their beliefs about physical factors, such as the table's friction coefficient and the laws of motion.

Method

Participants

Sixteen subjects had performed the experiment using a keyboard and computer screen, as described in Neupärtl et al. (2020). Sixteen additional subjects were recorded in the virtual reality (VR)-based experiment. All participants were undergraduate or graduate students recruited at the Technical University of Darmstadt, who were paid 10 € or received course credit for participation. All experimental procedures were carried out in accordance with the guidelines of the German Psychological Society and approved by the ethics committee of the Technical University of Darmstadt. Informed consent was obtained from all participants prior to carrying out the experiment. All subjects had normal or corrected to normal vision. One participant from the keyboard condition was excluded from further analysis, because variability of estimated mass beliefs was more than two standard deviations larger than those of the other participants.

Apparatus

In the keyboard condition, participants saw the scene containing the puck and the target displayed on a computer monitor and responded through button presses on a keyboard. All trials were rendered using unity. Participants were seated so that their eyes were approximately 40 cm away from the display and the monitor subtended approximately 66° of visual angle horizontally and 41° vertically. For more details, see Neupärtl et al. (2020). The haptic condition was also implemented in unity but using the SteamVR plugin and the HTC-Vive Pro Eye HMD with a resolution of 1440 × 1600 pixels per eye and a field of view of 110° horizontally and 110° vertically. For motion tracking purposes, the Qualisys motion tracking system with six 6+ cameras was used. To easily change the weight of the object to be propelled, we custom built a puck by drilling multiple holes into the puck and filling them with different metal weights, resulting in a mass of 0.25 kg or 0.35 kg. Drill holes were covered with a 3D printed plastic covering staffed with four passive markers. Elastic fabric was fixed on the table with bench vice to protect the motion tracking cameras and to facilitate trial resets. Thus, the puck was restricted to a smaller area, allowing the subject to grab it by themselves at the beginning of a trial.

Experimental design

In both experiments, participants were instructed to slide pucks into the bulls-eye of a target, see Figure 1. The two experiments however differed in the way subjects were able to make the puck slide. In the keyboard condition, the puck was a two-dimensional rendition on a computer screen. Subjects carried out 50 trials in which the target was placed at distances drawn uniformly at random between 1 m and 5 m, where the entire scene displayed on the monitor was 7.5 m in the vertical dimension. The duration of the key press determined the duration of the impact of a constant force, i.e. the change in momentum. Participants were told that they were able to adjust this force, which initially was going to accelerate the puck and thus the initial velocity of the puck, by the duration of their press. However, they were not explicitly told about the linear relationship between the press time and the initial velocity. In contrast, participants in the haptic condition were able to pick up the custom built hockey puck and pushed it in 100 trials on a table with their own hands. The objects in the VR scene were carefully designed to match their actual dimensions, the table on which the real pucks were slid and the puck itself. The target was placed at distances drawn uniformly at random between 0.3 m and 2.5 m. Subjects were randomly assigned a puck with a mass of either 0.25 kg or 0.35 kg. Here, the puck and the bulls-eye were shown in VR using an HTC-Vive HMD.

In both setups, subjects did not obtain any feedback about their action by blanking out the resulting movement as well as the final position in which the puck came to a stop. For this purpose, the screen turned dark for half a second in the keyboard condition after participants released the key and thus the puck. In the haptic condition, the field of view was not completely darkened for safety reasons. Instead, both the puck and the target were blanked out. Thus, the table was still visible to avoid dizziness. Explorative trials were carried out prior to obtaining the experimental data to ensure that the perceptual uncertainty about the targets' distances were comparable across the two experimental setups.

Physical description of the sliding task

To be able to compare the actions propelling pucks between both experiments, it is necessary to find a

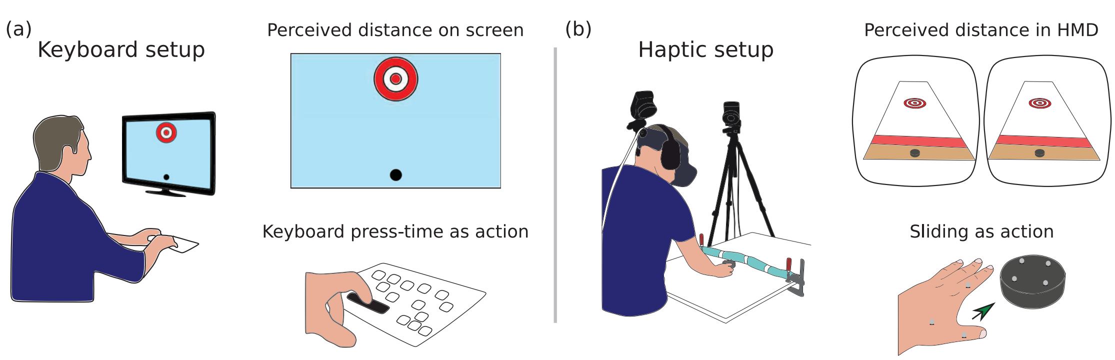

Figure 1. Comparison of both experimental setups. (a) In the keyboard condition, participants saw the target and the puck on a computer screen and adjusted the momentum acting upon the puck via press time of a keyboard button (Neupärtl et al., 2020). (b) In the VR setup, participants viewed the scene including the distance to the target in a HMD and were able to grasp the actual puck and slide it naturally on a table.

single description of the physics governing the puck's motion. Besides the distance to the target, the impulse that subjects intend to use p_{int} to propel the puck to the target is the decisive physical quantity. When adjusting the necessary impulse to reach the target's bulls-eye, both participants' belief about the necessary speed v_{int} and their belief about the puck's mass m_B play a crucial role:

$$p_{int} = m_B \cdot v_{int}.$$

Now, both experiments differ in the way how participants can propel the puck and control the resulting impulse. In the case of the haptic condition, participants could directly interact with the puck and accelerate it to the intended velocity by controlling the release impulse with respect to the perceived mass, see Equation (2a). In the keyboard condition, subjects could control the impulse via the duration of a key press. Here, the magnitude of the force F and the puck's mass m_B were abstract and hypothetical quantities, unknown to the participant, and can be summarized as a constant variable C_1, see Equation (2b):

$$v_{int} = \frac{p_{int}}{m_B}, \tag{2a}$$

$$v_{int} = \frac{F}{m_B} \cdot t_{key} = C_1 \cdot t_{key}. \tag{2b}$$

This means that the use of haptic interaction in the haptic condition enabled a direct naturalistic control of the intended velocity compared to the indirect abstract control in the keyboard condition. The magnitude of the intended velocity v_{int} however needs to be chosen in both scenarios depending on the final distance to be covered x and on the influence of the decelerating friction, i.e. the product of the friction coefficient μ_{fr} and the gravitational acceleration g:

$$v_{int} = \sqrt{2 \cdot \mu_{fr} \cdot g \cdot x} = C_2 \cdot \sqrt{x}.$$

The friction coefficients are unknown to participants in both conditions and their product with the gravitational accelerations can be summarized as constant variable C_2, since both variables do not change during the experiment. Because subjects never obtain feedback about the movement of the puck, they had no possibility to infer the values of these constants. This has the consequence that even if subjects consistently acted under the belief of a specific gravitational acceleration, mass, and friction coefficient, these values are interdependent. To allow recovering subjects' beliefs, we therefore can set two of these values to constants. In the haptic condition, we set the friction coefficient to the true value measured for the table and the true gravitational constant. Note that this does neither affect the physical relationship described by either Newtonian physics or the linear heuristic nor the variability observed in participants' actions, upon which our conclusions rest.

In the keyboard condition, subjects had no opportunity to infer either the initial force or the coefficient of friction or the puck's mass. Note, however, that these constants all enter the computation of the intended velocity linearly. Thus, if subjects acted under a consistent belief for these quantities, we can set two of these values to constants to investigate the variability of the third quantity. To be able to compare the inferred values across both conditions, we set the friction coefficient and initial force so as to result in puck masses with the same mean as the masses in the haptic condition. Note that this does not alter the conclusions that can be drawn from the inferred variables, as this constitutes only a linear scaling in the masses. Subjects' uncertainty about these environmental variables can still be captured by the spread of the posterior over the mass beliefs, which is then compared across subjects. The important point here is that these constraints do not affect the assumptions about the basic relationship between intended velocity and distance to the target and thus still allow the computational analysis of the two experiments based on probabilistic models.

Taken together, two experiments were set up to study human behaviour in the physics-based task of sliding an object, in this case a puck, as accurately as possible across a given distance into a target under the effect of friction, without feedback. The experiments differed primarily in the form of the available action: Via a button press in the keyboard condition and by pushing a real puck by hand in the haptic condition. In the following, the effects of this difference on the behaviour of subjects will be investigated based on participants' beliefs as estimated by a Bayesian model of the interaction task.

Bayesian graphical model of physical interaction

The Bayesian model depicted in Figure 2(a) considers the subjects' actions from the experimenter's point of view. This means that variables that were experimentally measurable, such as the distance to the target in the display and the actual magnitude of the sliding action, are included as observed variables in the model. Variables that describe the physical assumptions and perceptual beliefs of the subjects are unknown to the experimenter, and therefore unknown in the model. Thus, this model constitutes a departure from an ideal observer model in that the subject's beliefs about task relevant quantities are explicitly modelled. This in turn requires using Bayesian inference to infer these latent beliefs of the subjects based on the model structure and the observed experimental data. In Figure 2, known observed nodes are shown in grey and unknown latent ones are shown in white.

The perception of the actual distance x is naturally subject to sensory uncertainty. This sensory uncertainty is modelled by the parameter σ_x of the log-normal distribution of the perceived distance x_{per}. Given her perception, the subject internally decides about a velocity v_{int} required to let the puck slide to the target and stop there. However, the subject is now dependent on her internal model, which describes the relationship between target distance and initial velocity. Here, we consider two candidates as possible internal models describing the relationship between the distance x and necessary speed v_{int}: A linear and a square-root relationship:

$$v_{int} \sim x,$$

$$v_{int} \sim \sqrt{x}.$$

The linear relation in Equation (4) corresponds to an approximate heuristic and the square-root relation in Equation (5) is the relationship prescribed by Newtonian physics.

Note that the other physical parameters of the environment such as the coefficient of friction, the gravitational constant, and the force in the keyboard condition were unknown to the participants throughout the experiment. Since subjects never saw the puck gliding, decelerating, and stopping after being propelled, they could not calibrate the impulse with

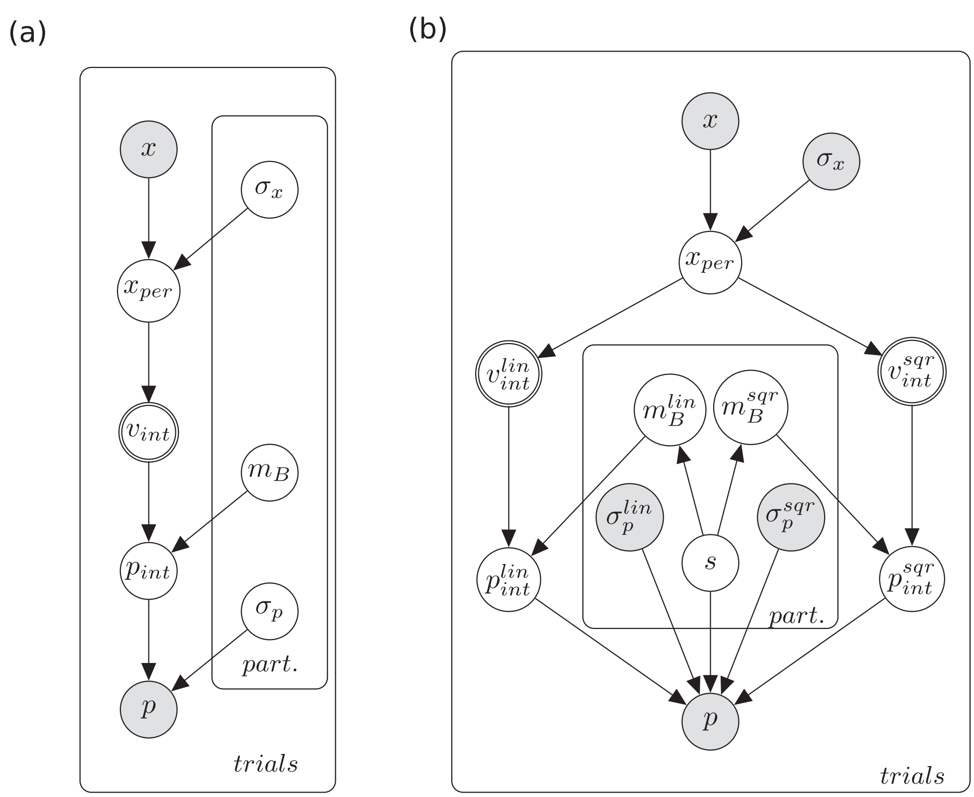

Figure 2. Basic Bayesian interaction model (a) and corresponding nested model (b) for the product space method. Shaded nodes, e.g. the actual distance x and impulse p in the basic impulse model, are observed and known to the experimenter. White nodes are latent and need to be inferred. Participants' observation of the actual distance x is inevitably subject to perceptual uncertainty σ_x and thus leads to a noisy percept x_{per}.

which they push the puck nor the coefficient of friction describing the gliding properties of the surface. Thus, the aforementioned constant variables C_1 and C_2, used to summarize these environmental variables are not included in the figure, as their values are fixed and they no longer functionally affect the inferences, of course apart from their influence on the magnitude of the estimates.

The velocity v_{int} estimated in this manner, together with their belief about the mass m_B of the puck to be accelerated, now guides the subject to produce the necessary impulse p_{int}. Her belief about the mass is relevant since heavier objects need larger impulses to reach the same velocity. The finally measured impulse p is the result of the intended impulse p_{int} and the action variability σ_p of the subject during the execution of the control action.

The basic Bayesian impulse model in Figure 2(a) is based on these assumptions, with the linear and Newtonian models differing only in the calculation of the intended velocity v_{int}. While keeping in mind that variable p still denotes the puck's momentum whereas $p(x)$ denotes the probability of a variable x, the joint posterior probability of the observed data d and all parameters Θ expressed by this model can be written as follows:

$$p(d, \Theta) = p(x)\, p(\sigma_x)\, p(x_{per}|x, \sigma_x)\, p(m_B)$$
$$p(p_{int}|x_{per}, m_B)\, p(\sigma_p)\, p(p|p_{int}, \sigma_p)$$

and accordingly the posterior probability of the parameters $p(\Theta|d)$ given the model and the data d can be computed using Bayes' theorem:

$$p(\Theta|d) = \frac{p(\sigma_x)p(x_{per}|x, \sigma_x)p(m_B)p(p_{int}|x_{per}, m_B)p(\sigma_p)p(p|p_{int}, \sigma_p)}{p(p|x)}.$$

For comparison of the two potential model candidates, we used the *product space method* (Lodewyckx et al., 2011). For this purpose, a hierarchical model is used in which both models are included and an index variable determines which model is selected to explain the data. All inferences were carried out in R via Markov chain Monte Carlo using the JAGS package (Plummer, 2003).

In Figure 2(b), the index variable s selects which model is used to describe the data for each participant. Whenever one model is selected on an iteration, the parameters of this model are updated based on the experimental data. Because the selected model now describes the data better on the basis of the updated parameters, the alternate model would become less likely to be selected on subsequent iterations. To avoid this, we adopt the common technique of sampling the parameters of the unselected model by an already optimized pseudo-prior. This is also why there is not only an arrow from s to p, indicating the model selection process in each iteration but also arrows to m_B^{lin} and m_B^{sqrt} describing the influence of the indicator variable s whether the mass is sampled from the prior or pseudo-prior. The perceptual uncertainties σ_x and action variabilities σ_p as well as the parameter of these pseudo-prior for the mass beliefs m_B for each participant were determined in advance by single model runs for both models. Based on the posterior odds of s, one can calculate the Bayes factors supporting one or the other model. For a more detailed explanation of the method, see Lodewyckx et al. (2011).

Results

First, we can compare subjects' raw responses in the two conditions, i.e. the press times in the screen condition and the puck's release velocity in the haptic condition. These data are shown in Figure 3 as a function of the initial distance of the target aggregated across subjects and trials. In the keyboard condition, subjects were constrained to interact with the keyboard while having only very vague beliefs about the puck's mass and the friction coefficient. Subjects had additional uncertainty about how the duration of a key press translates to an initial velocity, i.e. the strength of the acting force. In the haptic condition, on the other hand, subjects were able to grasp and accelerate the puck with their own hands, giving them a sensory measurement of the puck's mass through haptic feedback and the necessary momentum. In both experiments, participants did not obtain any sensory feedback about the motion trajectories nor about the endpoint of the puck's motion. Therefore, participants could not update their beliefs about the friction coefficient.

To compare subjects' responses, we transformed subjects' press times in the keyboard condition to initial velocities according to equation $v = F \cdot t/m$, see right axis labelling in the first panel of Figure 3. Because subjects were never able to update their beliefs about the initial force and the coefficient of friction and these quantities both enter the target

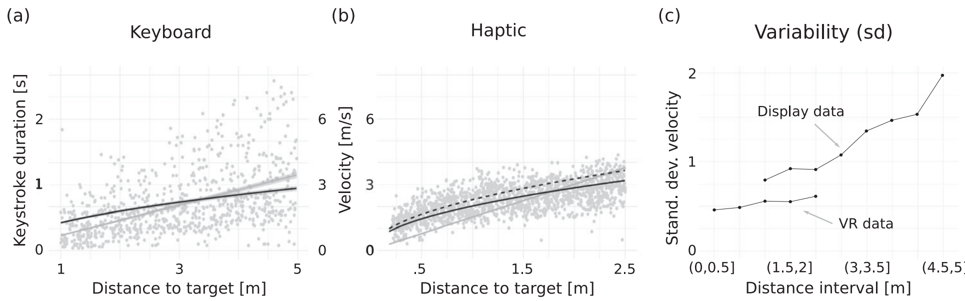

Figure 3. Responses as a function of the initial distance to the target for the keyboard (a) and haptic (b) conditions across all participants. Best generalized additive model fits based on maximum likelihood are shown for a linear and a square-root relationship in light and dark grey, respectively. For the data from the haptic condition, the ideal curve based on the actual weight of the puck and friction coefficients is also drawn as dotted line. (c) Estimated variability of participants' actions as a function of distance to target.

velocity linearly, we can rescale the duration of button presses linearly according to equation $F/m \approx 3$ to be within a comparable range of velocities. Accordingly, the rescaling can be interpreted as setting specific values for force and mass. The values were chosen ($m = 0.3\,\mathrm{kg}$, $F = 0.9\,\mathrm{N}$) so that both graphs are in a comparable range, see first two panels in Figure 3, and to allow a comparison of masses between the keyboard and haptic conditions, see first two panels in Figure 5. The graph clearly illustrates the larger variability in subjects' actions in the keyboard condition compared to the haptic condition. To illustrate differences in variability, the third panel in Figure 3 shows standard deviations for intervals binned over distance for both experiments, ranging from 0.79 to 1.97m/s in the keyboard and 0.45 to 0.61m/s in the haptic condition. The plots also demonstrate the response variability with increased action magnitude, which was captured by the log-normal distribution in the Bayesian interaction model.

Subjects' beliefs

Based on the Bayesian generative model shown in Figure 2(a), we are able to infer the perceptual uncertainty σ_x, the action variability σ_p, and the mass m_B in the Newtonian model or the linear factor in the linear heuristics model for each individual subject. Modes of posterior distributions of perceptual uncertainty σ_x as well as action variability σ_p are plotted in Figure 4 both under the assumption that participants employed a linear heuristic model or a Newtonian physics model. The plots distinguish the putative internal model by the colour of data points and the two conditions by the shape of data points. Which model better accounted for an individual's data was decided based on the resulting Bayes factors obtained through the nested Bayesian model in the product space method, see below.

Figure 4(a) compares the perceptual uncertainties inferred by the linear heuristics model (*x*-axis) and the Newtonian model (*y*-axis) for each subject in both conditions. Through preliminary explorative trials, the experimental setup of the two conditions was adjusted to have comparable perceptual uncertainties about the target's distance. Indeed, inferred perceptual uncertainties do not differ significantly between the two conditions (paired Wilcoxon signed rank test, V=254, p=0.8609). We also tested whether the perceptual uncertainties inferred from data under the two putative internal models differed. No significant difference in inferred perceptual uncertainties was found between the linear heuristics model and the Newtonian physics model (paired Wilcoxon signed rank test, V=253, p=0.8465).

By contrast, inferred action variability was clearly different when comparing the linear heuristics and the Newtonian models in both experimental conditions. Figure 4(b) shows the corresponding plot of inferred action variabilities, demonstrating the separation of the data of the two conditions. Modes of the inferred action variabilities significantly differ between both, the linear and the Newtonian models, when analyzing the keyboard condition's data (pairwise Wilcoxon signed rank test, V=112, p=0.02567) as well as for the haptic condition's data (pairwise Wilcoxon signed rank test, V=0, p<0.001, both p-value adjusted after Benjamini & Hochberg, 1995). In these plots, data points beneath the red

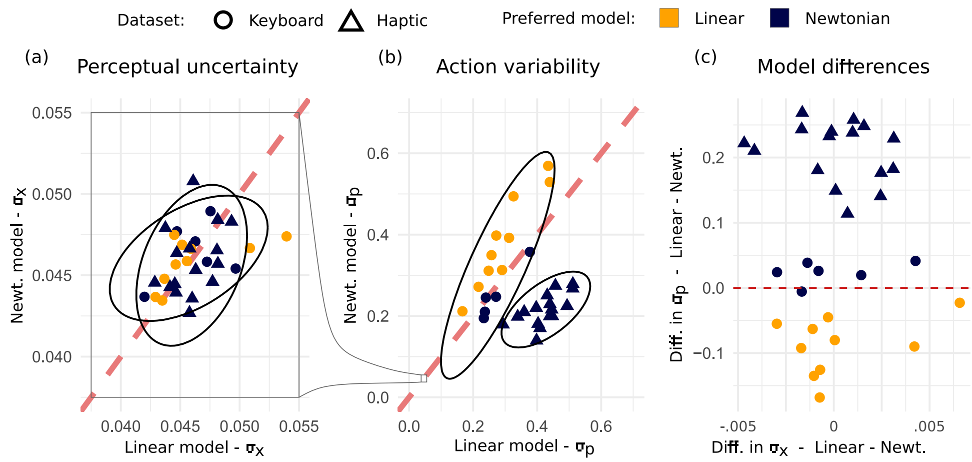

Figure 4. Comparison of inferred perceptual uncertainties and action variabilities for both conditions and models. (a) The *x*-axis shows the perceptual uncertainty σ_x inferred using the linear model and the *y*-axis the one inferred using the Newtonian model. (b) Inferred values for the action variability σ_p, again with values on the *x*-axis for the linear and on the *y*-axis for the Newtonian model. (c) Differences between linear and Newtonian model inferences for σ_x and σ_p.

dotted line, which indicates the equality of variability in both models, require less additive noise to be explained with the Newtonian model compared to the linear heuristics model. The results of both inferences of perceptual uncertainty and action variability are plotted together in Figure 4(c). First, all our participants' behaviour in the haptic condition was better accounted for by the Newtonian model while most of the participants' behaviour was better accounted for by the linear heuristics model in the keyboard condition. Second, while perceptual uncertainties were comparable across the two putative internal models, the linear heuristics model required higher levels of additional noise to account for our subjects' actions in the haptic condition.

The Bayesian generative model also allows inferring individual participants' internal beliefs about the masses of the pucks. The inferred modes of the mass posteriors for each participant are plotted in Figure 5. As subjects in the keyboard condition never had access to any sensory measurement about the puck's physical properties, they had to entirely rely on their prior belief of its mass. Because the puck's intended velocity depends on its mass, the initial force, and the duration of the key press, we can use an arbitrary values for the three factors entering the intended velocity linearly. We adjusted the arbitrary factors in such a way that the masses of the pucks in the keyboard condition had the same mean as the masses in the haptic condition.

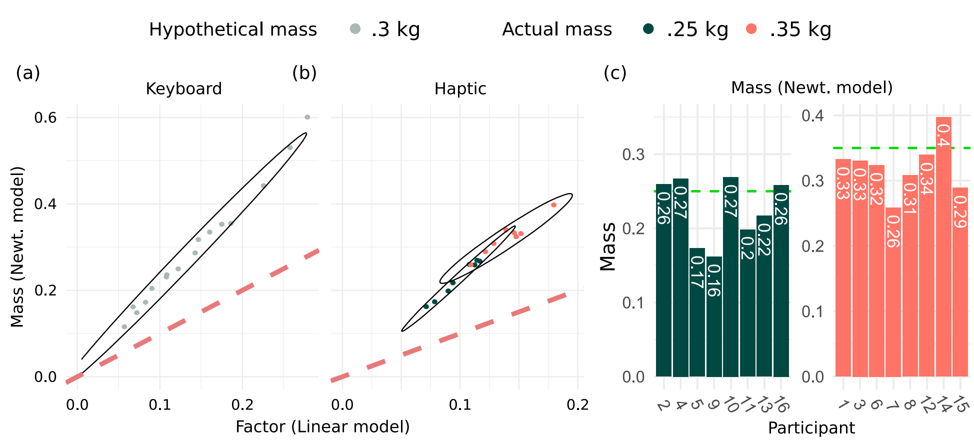

Figure 5. Comparison of inferred mass beliefs (Newtonian model) and linear factors (linear heuristic) for both conditions and models. (a) and (b) The *x*-axis shows the linear factor inferred using the linear model and the *y*-axis the mass inferred using the Newtonian model for the two conditions, respectively. (c) Individually inferred modes of the mass posterior compared to the actual mass of the used puck in the haptic condition.

The resulting inferred masses across participants accordingly have a mean of 0.296 kg and standard deviation of 0.138 kg. By contrast, in the haptic condition, a real puck and its mass was haptically accessible to participants. They were able to grab and lift the puck, and thus to adjust their belief accordingly. In both conditions, inferred linear factors for the heuristic model were smaller than the inferred masses in the Newtonian model (keyboard condition: Wilcoxon signed rank test, V=0, $p < 0.001$; haptic condition: Wilcoxon signed rank test, V=0, $p < 0.001$), accommodating the undershoots that subjects showed with increasing distance of the target, see Figure 3.

In the case of the mass inferences of the Newtonian model in the haptic condition, we can furthermore investigate how close participant's estimates of pucks' masses were based on the available haptic cues. Figure 5(c) shows the posterior modes of the inferred masses for all participants. Remarkably, the mass estimates are quite close to the true values of 0.25 and 0.35 kg, with overall mean values for each puck being slightly smaller at 0.226 and 0.323 kg, respectively. This difference was significant for the heavier puck (one sample t-test, t=−1.9243, p=0.04786) but not for the lighter one (one sample t-test, t=−1.5744, p=0.0797). Subjects' behaviour was well calibrated to the mass of the puck, as modes of the inferred posterior distribution of masses for the lighter puck are significantly smaller than for the heavy one, as expected (Welch two sample t-test, t=−4.5967, p<0.001).

The Bayesian interaction model additionally allows investigating the variability in subjects' mass beliefs and linear factors across trials. The standard deviations of inferred posterior distributions for the masses in the Newtonian model and the linear factors in the linear model are plotted in Figure 6. First, variability of standard deviations in mass beliefs is larger in the keyboard condition compared to the haptic condition (Levene's test, F=29.833, p<0.001), reflecting the larger uncertainty about physical parameters in the keyboard condition. Second, standard deviations inferred with the Newtonian model are significantly higher in the keyboard condition than in haptic (Wilcoxon rank sum test, W=215, p<0.001), which is not true for the linear model (Wilcoxon rank sum test, W=128, p=0.3851). This indicates a more precise computational description of participants' behaviour with the Newtonian model and a higher consistency of subjects' decisions across the experiment in the haptic condition. However, note that in the keyboard condition, variability for the linear factor according to the linear heuristic was on average only 1.08 times larger than in the haptic condition. Remarkably, for some subjects, this variability was even smaller in the keyboard condition compared to some subjects in the haptic condition. This clearly demonstrates that subjects in the keyboard condition were pushing the key to propel the puck under a consistent belief about stable and lawful properties of the puck.

Finally, we evaluated the goodness of fit of the linear and Newtonian models to data from both experimental conditions. First, we obtained posterior predictive distributions of initial impulse for both conditions, as plotted in Figure 7(a,b), respectively. While the linear heuristics model is slightly closer to the observed data in the keyboard condition, the Newtonian model is clearly closer to the data in the haptic condition. Note that the distribution of momentum based on the real weight of the pucks, the friction, and the distance to the target, i.e. the ideal observer's distribution, is additionally shown in green. The two peaks of the ideal distribution are due to the two different masses used in both groups. One particular strength of probabilistic modelling via nested models lies in the possibility of model comparison. The Bayes factor favours the linear model in the keyboard condition, albeit with a value of 1.47, this is only anecdotal evidence. In stark contrast, the Bayes factor of 49.79 in the haptic condition shows very strong evidence for the Newtonian modes.

Deviations from target based on subjects' beliefs

Based on the inferred, best parameters for each subject, we can calculate deviations from the target according to both models. Here, we can compare the performance, i.e. the deviation from the target, as if the best fitting parameters would correspond to the actual environment. This allows investigating participants' consistency with regard to their own, possibly wrong, beliefs. As an example, for the Newtonian model, we can compute how far the puck would have slid using their internal beliefs inferred from the Bayesian interaction model. Figure 8(a,b) shows the mean absolute error for the keyboard and haptic

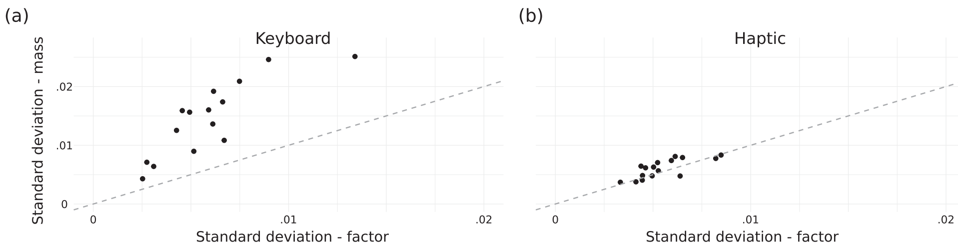

Figure 6. Standard deviation of posterior distributions over mass beliefs in the Newtonian model (y-axis) and linear factor (x-axis) in the linear model for the keyboard (a) and haptic conditions (b).

conditions, again in comparison for the Newtonian model and the linear heuristic. Comparing these errors between both models based on the keyboard condition shows significantly higher error values for the Newtonian model (Wilcoxon signed rank test, V =120, p<0.001; see Figure 8(a)). Based on the haptic condition, however, mean absolute error values are significantly higher for the linear model (Paired t-test, t=−5.8371, p<0.001; see Figure 8(b)). Both results again emphasize the better fit of the linear and Newtonian model to the keyboard and haptic conditions data, respectively. Additionally, we can also compare these errors between both conditions for each model. As expected, the mean absolute errors in the keyboard condition are significantly larger than in the haptic condition, both for the linear model (Wilcoxon rank sum test, W=219, p<0.001) and for the Newtonian model (Wilcoxon rank sum test, W=240, p<0.001). This reflects the higher uncertainty about the pucks mass in the keyboard condition.

Similar to the mean absolute error, the standard deviation of the errors for the Newtonian model is significantly larger than for the linear model in the keyboard condition (Wilcoxon signed rank test, V=120, $p < 0.001$; see Figure 8(c)). However, for the haptic condition, there is no significant difference between both models (Paired t-test, t=−0.6736, p=0.5108; see Figure 8(d)). This suggests that the Newtonian model is not able to capture the complete range of participants' actions in the keyboard data, potentially caused by larger deviations at higher distances (see Figure 3(a)). The fact that there is no difference for the haptic data set can also be well explained with reference to the data in Figure 3(b). The systematic bias of the linear approximation with its undershots at near and overshots at far distances leads to the higher mean absolute error for the linear model in Figure 8(b) but keeps the standard deviation at these lower values in Figure 8(d). Taken together, these analyses further demonstrate not only the better fit of the linear model for the keyboard condition's data and the Newtonian model for the haptic condition's data but also the increased consistency in subjects' actions within the more naturalistic haptic condition.

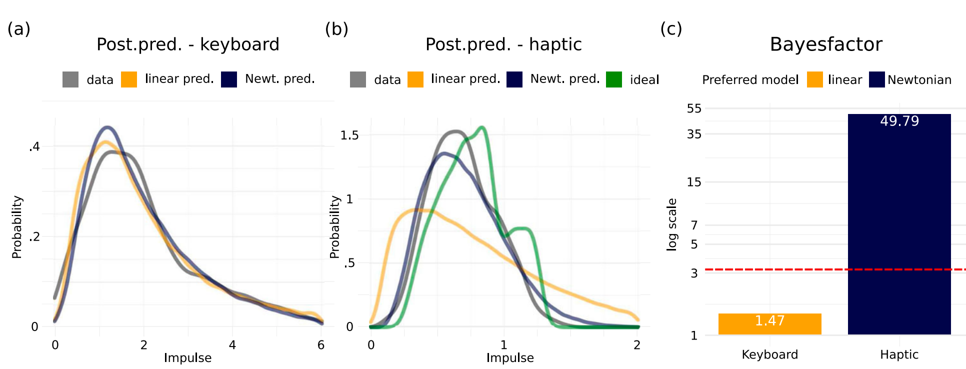

Figure 7. (a) and (b) Posterior predictives for both models and data sets in comparison with the actual data and (c) Bayes factors calculated based on the inferred posterior odds of the nested model. Ideal behaviour shown as additional bi-modal distribution for the haptic condition. For a better overview, Bayes factors in (c) are plotted on a log scale. The dotted line indicates the threshold at 3.2 for substantial evidence that one model is superior to the other.

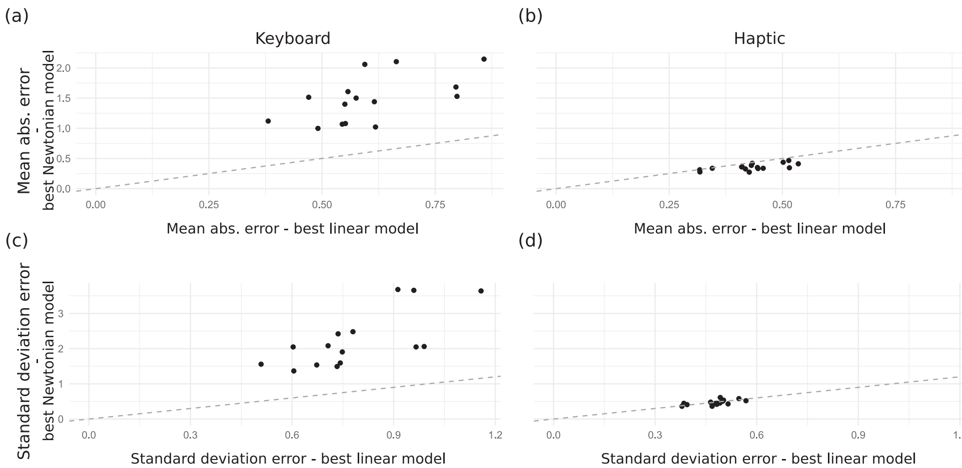

Figure 8. Mean absolute error and standard deviation of errors for each subject based on the best parameters for both models. Values for the best linear model are shown on the *x*- and for the best Newtonian model on the *y*-axis. Grey line marks equal values, above lie data of subjects with higher values for the Newtonian model and below with higher values for the linear model.

Discussion

To investigate whether naturalistic, embodied, multimodal interactions influence visuomtor behaviour involving the judgement of physical relationships between task relevant quantities, we designed an experiment in which subjects needed to propel pucks toward a target's bulls-eye, which was positioned at different distances across trials. While in the keyboard condition, subjects saw the target displayed on a monitor and propelled a virtual puck with the duration of a key press, in the haptic condition the scene was seen through a head-mounted VR display and a real puck with one of two masses could be pushed on a tabletop. Importantly, subjects obtained no visual feedback about their actions in either condition. Therefore, subjects needed to rely on their prior beliefs about physical properties and their lawful relationships to accomplish either task. While in the keyboard condition, the puck's mass, the force with which the duration of the key press was scaled, and the coefficient of friction were unknown, only the coefficient of friction was unknown in the haptic condition.

The task requires participants first to visually estimate the distance that the puck has to travel to reach the target's bulls-eye. Then, subjects need to propel the puck toward the target. For this, subjects need to choose the right impulse, which depends on the puck's mass, the coefficient of friction describing the surface on which the puck is sliding, and the gravitational constant. In the keyboard condition, the initial impulse was achieved through the length of a key press. Importantly, Newtonian physics prescribes a relationship that is linear in the puck's mass and grows with the square-root of the target's distance. Alternatively, subjects may employ a heuristic by which the momentum and therefore the initial velocity of the puck is scaled linearly with the distance to the target.

To be able to compare the behaviour in the two conditions, a Bayesian model of the full interaction task including the perception of distance and the generation of a puck sliding action was devised involving perceptual uncertainty and action variability. The model was fit on an individual-by-individual and trial-by-trial basis under the hypothesis that subjects could either use a linear heuristic or the relationships prescribed by Newtonian physics. Results show very strong evidence that the multimodal naturalistic embodied condition elicited the square-root scaling of initial velocity by the target's distance, which is consistent with Newtonian physics as demonstrated by a Bayes factor of 49.8. By contrast, the keyboard condition, in which subjects interacted with the puck through a key press on a keyboard, resulted in anecdotal evidence that the elicited behaviour was better accounted for by the linear heuristic. Closer evaluation of Bayes factors at the individual subject level showed that 10 of the 16 subjects in the keyboard condition were better accounted for by the linear heuristic while 6 were better accounted for by Newtonian physics. Analyses of the variability of actions and mass beliefs supported this conclusion additionally.

Further analyses of subjects' beliefs required constraining some of the constant parameters in the Bayesian interaction model. Because subjects never obtained feedback about the puck's dynamics, they were not able to infer the scaling of force in the keyboard condition or the coefficient of friction in both conditions. The gravitational constant, which also enters the computation of the initial velocity linearly, can be assumed to be known to subjects based on previous research (Hubbard, 2020; Jörges & López-Moliner, 2017; McIntyre et al., 2001). The coefficient of friction in the haptic condition was set to the actual value of the table used. For the keyboard condition, the initial force and the friction coefficient were adjusted to result in estimated masses with approximately the same mean as in the haptic condition. Note that this does not alter the conclusions about the model used by participants but allows comparing the consistency and variability of participants beliefs. Indeed, subjects acted with remarkable consistency in the keyboard condition. Even though the keyboard condition elicited higher variability of beliefs about mass and linear factors ($\sigma_m = 0.138\,\text{kg}$, $\sigma_f = 0.065\,\text{kg}$) compared to the haptic condition ($\sigma_m = 0.068\,\text{kg}$, $\sigma_f = 0.029\,\text{kg}$), individual subjects showed highly consistent variability within their factor beliefs, being on average only 1.08 times larger than in the haptic condition, see Figures 5 and 6.

A first potential concern might be that interacting with a keyboard could influence participants to use a particular non-linear function. That this is unlikely the case stems from the fact that ample research has employed button presses on keyboards to investigate human time perception and timing of actions, e.g. Buhusi and Meck (2005). The results of these studies show that people seem to be quite unbiased in controlling their button press duration and adhere to Weber's law in that the standard deviation of press times scales linearly with duration. Some studies have reported anchoring effects leading to overshots at smaller durations and undershots at larger durations, but these effects are also described as linear (Jazayeri & Shadlen, 2010). This suggests that the scaling of press times in the keyboard condition were not caused by an idiosyncratic mapping pertaining to the pressing of buttons on a keyboard, particularly, because some participants' press times were better explained by the linear and others' by the square-root relationship.

One may argue that the differences between the two conditions arise because of the difference in the mode of visual presentation of the scenes. Indeed, several previous studies have provided evidence that uncertainty about physical parameters depends on the mode of presentation. Adding motion cues (Kaiser et al., 1992) or auditory cues (Gerstenberg et al., 2018) to stimuli used in probing intuitive physical reasoning has been shown to reduce uncertainty about physical parameters. While in the present experiments subjects saw the puck and the target displayed on a two-dimensional screen in the keyboard condition, they had access to depth cues present in the stereoscopic HMD in the haptic condition. But, importantly, preliminary explorative trials were used to match the perceptual uncertainties across the two conditions. This was furthermore confirmed by the perceptual uncertainties estimated using the Bayesian interaction model. Indeed, no significant difference was found for the perceptual uncertainties between the two conditions. Thus, we do not find evidence that the mode of visual presentation was the cause for the difference in adopted strategies.

A further concern may be that participants' uncertainty about the mass, the coefficient of friction, and the mapping from press times to initial velocities in the keyboard condition led them to use random press times that only in the aggregated data suggest a linear relationship with the initial target distance. Again, the analyses with the Bayesian interaction model suggest otherwise. By estimating subjects' implicit beliefs about the puck's mass or equivalently linear factor on a subject-by-subject basis, one can infer the variability in beliefs across trials. This analysis revealed that in the keyboard condition, standard deviations of subjects mass beliefs were on average only about 2.4 times as large as in the haptic condition and for some subjects even comparable between the two conditions. This is a remarkable result, as it provides evidence, that subjects consistently used a mass belief to propel the puck towards the target in the keyboard condition.

A very much related question is whether subjects may have selected the correct physical relationship only by virtue of touching a real puck and thereby sensing its mass. The ability to grasp and hold the real puck certainly reduced subjects' uncertainty about the pucks' weights in the haptic condition. This was confirmed by participants' inferred mass beliefs. Nevertheless, as state above, the variability in the factor equivalent to mass in the linear relationship in the keyboard condition was only 1.08 times larger than the value in the haptic condition, showing that subjects used a consistent mass belief in their actions in the keyboard condition, in which they could not sense a puck's mass. Thus, it seems rather unlikely that this differences in terms of the uncertainty in belief about the pucks mass could be the sole reason for adopting a different functional relationship. Instead, this suggests that it was primarily the mode of interaction contributing to the difference in adopted physical relationship.

Taken together, the present study is in accordance with previous studies on intuitive physics within the noisy Newton framework (Kubricht et al., 2017), which assumes that internal models based on physical laws interact probabilistically with inherently uncertain and ambiguous sensory measurements. The systematic deviations in our subjects' press times from those prescribed by Newtonian physics under full knowledge of all parameters were explained quantitatively as stemming from perceptual uncertainties interacting with prior beliefs about physical relationships and, additionally, motor variability. The results of the present study furthermore support the notion of structured internal causal models of physical relationships and shows the importance of using structured probabilistic generative models that contain interpretable variables to quantitatively reverse engineer human cognition (Griffiths et al., 2010). Although visual feedback was never given about the pucks' sliding dynamics and final position, subjects showed behaviour that was consistent with the implicit assumption of a stable and lawful world. By employing a full generative model of the interaction task, it was possible to infer subjects' beliefs on an individual-by-individual and trial-by-trial basis.

While ample previous research has investigated the influence of the mode of presentation of physical scenarios on intuitive physical reasoning (Gerstenberg et al., 2018; Kaiser et al., 1992; Smith & Vul, 2013), very little attention has been devoted to the consequences of the mode of interacting with the physical scenario. An exception is the study by Smith et al. (2018), who investigated the degree of accurate physical reasoning across different tasks, which implicitly used different modes of responses. However, the authors used a computer mouse to simulate carrying out different interactions such as cutting a pendulum's rope or moving a bucket to catch a falling pendulum in two-dimensional visual simulations presented on a computer screen. While the authors came to the conclusion that differences in physical reasoning were the result of employing different "systems of knowledge" across tasks, it is not clear, what the underlying reason for having different systems could be, which specific systems could be involved, and by which mechanism a system would be selected. Our results suggest that one reason for differences across tasks is the mode of interaction, specifically, whether physical reasoning involves generating a continuous action of pressing a button on a keyboard in a physical simulation or a sliding action, which is a visuomotor interaction with physical objects practiced and calibrated in everyday activities.

The present study established that the availability of naturalistic multimodal sensorimotor interactions with physical pucks resulted in subjects adopting the functional relationship prescribed by Newtonian physics. Instead, when the same task was presented on a computer monitor lacking depth cues and the interaction was implemented through a button press on a keyboard, subjects' behaviour was more in line with a linear heuristic. This result strongly suggests that our subjects were able to take advantage of the motor planning and motor output generating the physically appropriate responses. This in turn suggests that generating the actions involved in physical reasoning can take advantage of representations that are not independent of the motor control and the body but are thus "embodied". Previous research on physical reasoning has emphasized abstract internal physics models (Battaglia et al., 2013; Bramley et al., 2018; Sanborn et al., 2013) and the few studies allowing subjects to interact with scenes implemented those interactions through abstract mouse clicks in computer simulations (Bramley et al., 2018). Thus, the current results extend our understanding based on previous studies addressing differences in physical reasoning between tasks (Smith et al., 2018) and quantifying different sources of uncertainty in physical reasoning (Smith & Vul, 2013) by adding the mode of physical interaction as an additional factor. This may also reconcile some previous result on intuitive physics, which reported strong deviations from Newtonian physics, but utilized very abstract depictions of scenes and no possibility for interaction (Caramazza et al., 1981; Todd & Warren, 1982).

Overall, the present results contribute to our understanding of how the brain may implement physical reasoning. Indeed, in terms of a computational level account (Marr, 1982) of intuitive physical reasoning, it is not clear why the output of a putative physical simulation engine should depend on the mode of interaction. But at the implementational level of description, there is evidence that different neuronal substrates are involved in physical reasoning, some of which are also implicated in motor planning and visuomotor control of actions involving the body. Previous studies have found evidence for the representation of abstract physical factors in parietal and frontal regions, when physics students thought about verbally presented physics terms (Mason & Just, 2016). Similarly, recent studies involving physical reasoning about objects' dynamics on the basis of short movies also identified frontal and parietal regions representing abstract physical quantities such as mass (Schwettmann et al., 2019) and involved in judging physical interactions (Fischer et al., 2016). These results give credence to the notion of causal generative models of physical objects and their interactions compared to model-free pattern recognition approaches, such as those based on deep neural networks. Nevertheless, the involvement of overlapping parietal regions in the representation of physical quantities such as mass when planning visuomotor interactions (Gallivan et al., 2014) and the additional involvement of motor related regions in such tasks (Chouinard et al., 2005) speak for a crucial role of representations tied to motor planning and motor output with the body (Anderson, 2003; Foglia & Wilson, 2013; Wilson, 2002), i.e. of embodied representations in physical reasoning at the implementational level.

Acknowledgements

We would like to thank Manfred Haefke for building the puck on which the VR experiment is based and Florian Kadner for his helpful thoughts on the experimental setup.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Nils Neupärtl http://orcid.org/0000-0001-9986-4690
Fabian Tatai http://orcid.org/0000-0002-8957-2683
Constantin A. Rothkopf http://orcid.org/0000-0002-5636-0801

References

Allen, K. R., Smith, K. A., & Tenenbaum, J. B. (2020). Rapid trial-and-error learning with simulation supports flexible tool use and physical reasoning. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(47), 29302–29310. https://doi.org/10.1073/pnas.1912341117

Anderson, M. L. (2003). Embodied cognition: A field guide. *Artificial Intelligence*, *149*(1), 91–130. https://doi.org/10.1016/S0004-3702(03)00054-7

Bates, C., Battaglia, P., Yildirim, I., & Tenenbaum, J. B.. (2015). *Humans predict liquid dynamics using probabilistic simulation*. In Proceedings of the 37th annual conference of the cognitive science society, Pasadena, California, July 22 - 25, 2015.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(45), 18327–18332. https://doi.org/10.1073/pnas.1306572110

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bramley, N. R., Gerstenberg, T., Tenenbaum, J. B., & Gureckis, T. M. (2018). Intuitive experimentation in the physical world. *Cognitive Psychology*, *105*(3), 9–38. https://doi.org/10.1016/j.cogpsych.2018.05.001

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, *6*(10), 755–765. https://doi.org/10.1038/nrn1764

Caramazza, A., McCloskey, M., & Green, B. (1981). Naive beliefs in "sophisticated" subjects: Misconceptions about trajectories of objects. *Cognition*, *9*(2), 117–123. https://doi.org/10.1016/0010-0277(81)90007-X

Chouinard, P. A., Leonard, G., & Paus, T. (2005). Role of the primary motor and dorsal premotor cortices in the anticipation of forces during object lifting. *Journal of Neuroscience*, *25*(9), 2277–2284. https://doi.org/10.1523/JNEUROSCI.4649-04.2005

Cohen, A. L. (2006). Contributions of invariants, heuristics, and exemplars to the visual perception of relative mass. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 574–598. https://psycnet.apa.org/doi/10.1037/0096-1523.32.3.574

Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016). Functional neuroanatomy of intuitive physical inference. *Proceedings of the National Academy of Sciences of the*

United States of America, *113*(34), E5072–E5081. https://doi.org/10.1073/pnas.1610344113

Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(3), 319–325. https://doi.org/10.1002/wcs.1226

Gallivan, J. P., Cant, J. S., Goodale, M. A., & Flanagan, J. R. (2014). Representation of object weight in human ventral visual cortex. *Current Biology*, *24*(16), 1866–1873. https://doi.org/10.1016/j.cub.2014.06.046

Gerstenberg, T., Siegel, M., & Tenenbaum, J. (2018). *What happened? Reconstructing the past through vision and sound*. In Proceedings of the annual meeting of the cognitive science society (Vol. 40), Madison, Wisconsin, July 25 - 28, 2018.

Gilden, D. L., & Proffitt, D. R. (1994). Heuristic judgment of mass ratio in two-body collisions. *Perception & Psychophysics*, *56*(6), 708–720. https://doi.org/10.3758/BF03208364

Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, *14*(8), 357–364. https://doi.org/10.1016/j.tics.2010.05.004

Hubbard, T. L. (2020). Representational gravity: Empirical findings and theoretical implications. *Psychonomic Bulletin & Review*, *27*(1), 36–55. https://doi.org/10.3758/s13423-019-01660-3

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience*, *13*(8), 1020–1026. https://doi.org/10.1038/nn.2590

Jörges, B., & López-Moliner, J. (2017). Gravity as a strong prior: Implications for perception and action. *Frontiers in Human Neuroscience*, *11*, 203. https://doi.org/10.3389/fnhum.2017.00203

Kaiser, M. K., Proffitt, D. R., Whelan, S. M., & Hecht, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 669–689. https://psycnet.apa.org/doi/10.1037/0096-1523.18.3.669

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*(1), 271–304. https://doi.org/10.1146/psych.2004.55.issue-1

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference*. Cambridge University Press.

Kubricht, J. R., Holyoak, K. J., & Lu, H. (2017). Intuitive physics: Current research and controversies. *Trends in Cognitive Sciences*, *21*(10), 749–759. https://doi.org/10.1016/j.tics.2017.06.002

Kubricht, J. R., Jiang, C., Zhu, Y., Zhu, S.-C., Terzopoulos, D., & Lu, H. (2016). *Probabilistic simulation predicts human performance on viscous fluid-pouring problem*. In Proceedings of the 38th annual conference of the cognitive science society, Philadelphia, Aug 10 - 13, 2016.

Lodewyckx, T., Kim, W., Lee, M. D., Tuerlinckx, F., Kuppens, P., & Wagenmakers, E.-J. (2011). A tutorial on Bayes factor estimation with the product space method. *Journal of Mathematical Psychology*, *55*(5), 331–347. https://doi.org/10.1016/j.jmp.2011.06.001

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman.

Mason, R. A., & Just, M. A. (2016). Neural representations of physics concepts. *Psychological Science*, *27*(6), 904–913. https://doi.org/10.1177/0956797616641941

McCloskey, M., Washburn, A., & Felch, L. (1983). Intuitive physics: The straight-down belief and its origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 636–649. https://psycnet.apa.org/doi/10.1037/0278-7393.9.4.636

McIntyre, J., Zago, M., Berthoz, A., & Lacquaniti, F. (2001). Does the brain model Newton's laws?. *Nature Neuroscience*, *4*(7), 693–694. https://doi.org/10.1038/89477

Neupärtl, N., Tatai, F., & Rothkopf, C. A. (2020). Intuitive physical reasoning about objects' masses transfers to a visuomotor decision task consistent with Newtonian physics. *PLoS Computational Biology*, *16*(10), e1007730. https://doi.org/10.1371/journal.pcbi.1007730

Plummer, M. (2003). *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling*. In Proceedings of the 3rd international workshop on distributed statistical computing (Vol. 124, No. 125.10, pp. 1–10), Vienna, March 20 - 22, 2003.

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and newtonian mechanics for colliding objects. *Psychological Review*, *120*(2), 411–437. https://doi.org/10.1037/a0031912

Schwettmann, S., Tenenbaum, J. B., & Kanwisher, N. (2019). Invariant representations of mass in the human brain. *eLife*, *8*, e46619. https://doi.org/10.7554/eLife.46619

Smith, K. A., Battaglia, P. W, & Vul, E. (2013). *Consistent physics underlying ballistic motion prediction*. In Proceedings of the annual meeting of the cognitive science society (Vol. 35), Berlin, July 31 - Aug 3, 2013.

Smith, K. A., Battaglia, P. W., & Vul, E. (2018). Different physical intuitions exist between tasks, not domains. *Computational Brain & Behavior*, *1*(2), 101–118. https://doi.org/10.1007/s42113-018-0007-3

Smith, K. A., & Vul, E. (2013). Sources of uncertainty in intuitive physics. *Topics in Cognitive Science*, *5*(1), 185–199. https://doi.org/10.1111/tops.12009

Todd, J. T., & Warren Jr., W. H. (1982). Visual perception of relative mass in dynamic events. *Perception*, *11*(3), 325–335. https://doi.org/10.1068/p110325

Ullman, T. D., Spelke, E., Battaglia, P., & Tenenbaum, J. B. (2017). Mind games: Game engines as an architecture for intuitive physics. *Trends in Cognitive Sciences*, *21*(9), 649–665. https://doi.org/10.1016/j.tics.2017.05.012

Ullman, T. D., Stuhlmüller, A., Goodman, N. D., & Tenenbaum, J. B. (2018). Learning physical parameters from dynamic scenes. *Cognitive Psychology*, *104*(9), 57–82. https://doi.org/10.1016/j.cogpsych.2017.05.006

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636. https://doi.org/10.3758/BF03196322