Jointly Perceiving Physics and Mind: Motion, Force and Intention

Ning Tang[2,3] ningtangcog@gmail.com
Siyi Gong[4] alicegong@g.ucla.edu
Ziqian Liao[2,4] liaoziqian101@outlook.com
Haokui Xu[2,3] haokuixu.psy@gmail.com
Jifan Zhou[3] Jifanzhou@zju.edu.cn
Mowei Shen[3] mwshen@zju.edu.cn
Tao Gao[1,2] tao.gao@stat.ucla.edu

[1] Department of Communication, UCLA
[2] Department of Statistics, UCLA
[3] Department of Psychology, Zhejiang University
[4] Department of Psychology, UCLA

Abstract

Physics and mind are two major causes of motion. In a leash-chasing display, a disc ("sheep") is being chased by another disc ("wolf"), which is physically constrained by a leash attached to a third disc ("master"). A number of interesting motions can emerge from this simple system, such as a wolf being dragged away from its target. Therefore, it is important for vision to jointly infer a physics-mind combination that can best explain the motions. Here we reported two discoveries from studying this display to support this theory. First, an intuitive physical system like a leash can greatly lessen the detrimental effects of spatial deviation and the diminishing objecthood on perceived chasing, strengthening its robustness. Second, a mutual dependency exists between physics and mind, where disrupting one will inevitably result in an impaired perception on the other. These results collectively support a joint perception of physics and mind.

Keywords: Vision; Perceived Animacy; Chasing; Intention; Theory of Mind; Intuitive Physics

Introduction

"English policy is to float lazily downstream, occasionally putting out a diplomatic boat-hook to avoid collisions." —Robert Gascoyne-Cecil (1830-1903)

Despite all the underlying political implications, this quote from a 3-time prime minister of the U.K. reveals an essential fact: agents live in a world rich in physics, and they interact with physics in a wide variety of ways, sometimes taking advantage of the physical system by doing nothing (e.g., "float downstream"), and sometimes even fighting against the physics (e.g., "put out a boat-hook to avoid collision"). Apart from experiencing the world and producing actions from the first-person perspective, humans are also constantly exposed to other agents' body movements. The question remains – as third-party observers, how do humans explain others' movements in terms of the underlying physics and mind?

Heider and Simmel's classic animation display (1944) has demonstrated that people are remarkably successful at perceiving intention from the geometric shapes in a physical environment. Yet more specifically, we are interested in finding out the interaction between physics and mind. Anecdotal video demonstration from our lab (https://www.youtube.com/playlist?list=PLe7BbCETnjAnrZ_Ma296z2XiPMjB2N9m, Fig. 1) suggests that humans seem to be exceptionally good at perceiving physics-mind interaction from motions. This demonstration contains a circle accelerating and decelerating in two opposite directions, a static square, and a mass of dots flowing in one direction with constant speed. The motion pattern of each geometric shape is rather repetitive and tedious. However, when all components are put into a single physical system (an object moving in a field of dots), many observers reported perceiving animacy out of the scene, such that the circle is "fighting against" a flow of obstacles to "reach" the square, gradually "depleting its energy" and eventually being "carried away" by the flow. The dynamics even evoked strong psychological reactions from one viewer, who vividly reported "The circle is basically me when I was trying to catch up with all my emails."

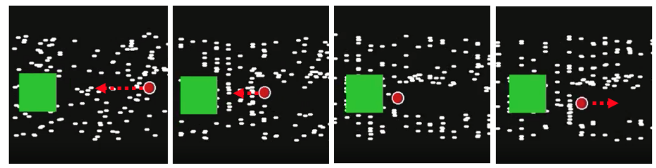


Fig. 1. Illustration of pursuing a goal against a flow. The dashed arrows represent velocity.

The richness in this display largely comes from the constraint imposed on the agent, and the same phenomenon can be often observed in real life. For example, you may spot a stationary cat, a dog being constrained by a leash while moving back-and-forth in various directions, and an owner holding the other end of the leash. In this case, the dog is being dragged by the leash, making its motion deviate significantly from its intended direction toward the cat. Despite the wavering direction of motion, we nevertheless infer the dog's persistent intention, perceiving the scene as the dog is "chasing" after its prey while "fighting against" the leash that is holding it back.

From the above examples, two major characteristics of real-life agents can be derived. First, social agents are not omnipotent – they cannot always get what they want. Agents sometimes even get pushed in undesired directions, resulting in large deviations between intended paths toward the goal and their actual motion track (e.g., circle moving back-and-forth; dog swinging around). Second, agents do not simply exist as individual objects. Instead, they exist as a part of a larger physical system and thus are bound by various restrictions, often in the forms of connections, bonds, and contact (e.g., the circle is in contact with the particles; the dog is chained to a leash).

From the anecdotal video and introspective daily experience, humans seem to understand agency — its intention and how it is constrained by the physical system — rather effectively despite interference from deviations and restrictions in the world. However, there has been a lack of theories on vision that focus on the joint inference of physics-mind combination as a way to explain the observed motion. Understanding of physics and mind has been typically treated as two separate core knowledge systems, each with a unique set of properties. Furthermore, the lack of theoretical guidance is also reflected in empirical research. In fact, existing empirical psychophysical studies seemingly suggest contradictory results, further compromising the validity of a joint physics-mind inference mechanism. For the rest of the introduction, we will first review existing psychophysical evidence as well as theoretical literature on the understanding of physics and mind, and then we will propose our framework that intends to address this challenge.

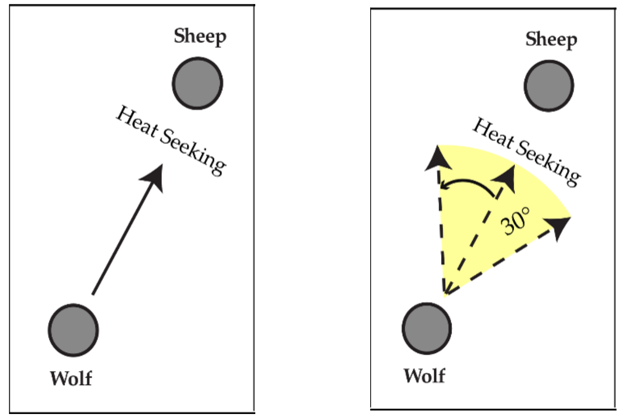
Psychophysical Evidence

In recent years, the psychophysics of chasing has been systematically studied with a search-for-chasing paradigm, in which a display contains one disc (the "wolf") chasing another disc (the "sheep") among several moving distractor discs (Gao, Newman, & Scholl, 2009; Meyerhoff, Huff, & Schwan, 2013; van Buren, Gao, & Scholl, 2017). The task for participants was to identify whether chasing was present in the display.

We focus on two case studies of perceived chasing that are particularly relevant to physics-mind combination. One study manipulated chasing subtlety (Fig. 2) through a maximum range of spatial deviation from the heat-seeking direction (Gao et al., 2009). The result revealed that a subtlety value greater than 30° led to a significant decline in performance on chasing detection, suggesting that perceived chasing is highly sensitive to spatial deviations. Another study revealed the "object-based" nature of perceived chasing, showing that it is severely disrupted when the wolf and sheep are connected to other objects by visible lines, therefore threatening their objecthood (van Buren et al., 2017).

Separately, each of the manipulations above effectively disrupt perceived chasing. When combined, they unsurprisingly make a strong prediction that leash-chasing cannot be readily perceived given the dual presence of spatial deviation and line connection. According to these results, when the dog is connected to a leash and moving in various directions, humans should experience a hard time identifying the dog's intention of chasing, an idea inconceivable to pet owners. The contradiction between psychophysical evidence and real-world observation leaves a question mark on whether humans can genuinely understand agency in a physical system.

Here we approach this question by attributing the fragility in the perception of physics-mind to the arbitrary nature of the physical disruptions introduced in the previous studies (e.g. spatial deviation with no reason, connection between objects that do not belong to a physical system). To take physics seriously in understanding social minds, it is necessary to review the existing theoretical accounts of the perception of physics and mind, after which we shall propose our theory of joint perception of physics and mind.



(a) When the chasing subtlety is 0°, the wolf always heads directly toward the (moving) sheep, in a "heat-seeking" manner.

(b) When the chasing subtlety is 30°, the wolf is not perfectly heat-seeking: Instead, it can move in any direction within a 60° window.

Fig. 2. An illustration of the chasing subtlety.

Intuitive Physics and Intuitive Psychology

Previous research has emphasized two distinct core systems embedded in humans: intuitive physics and intuitive psychology. On one hand, studies on the intuitive physics system suggest that humans are endowed with naive understandings of physical objects and physical rules, allowing them to effectively simulate how the physical world operates (Baillargeon, Spelke, & Wasserman, 1985; Spelke, 1990; Carey, 2000; Battaglia, Hamrick, & Tenenbaum, 2013; Ullman, Spelke, Battaglia, & Tenenbaum, 2017). Based on physics, a Theory of Body (ToB) was developed that explains agency from a physical perspective (Stewart, 1982; Premack, 1990; Leslie, 1995; Luo, Kaufman, & Baillargeon, 2009). ToB makes a prediction that inanimate objects should move in observation of physical laws such as the conservation of energy. From this prediction, it naturally follows that any object that violates this predication should be categorized as animate with an internal energy source. Since the theory is essentially based on the prediction of inanimate objects, we refer to ToB as the "weak" definition of agency. On the other hand, extensive studies on Theory of Mind (ToM) suggest that humans are able to interpret observed motions in terms of their underlying mental states (Wellman, 1992; Gelman, Durgin, & Kaufman, 1995; Leslie, 1995; Gergely, Nádasdy, Csibra, & Bíró, 1995; Saxe, Carey, & Kanwisher, 2004). ToM fundamentally differs from ToB in the way it concerns the unique properties of agency. Specifically, it explicitly predicts that agents should act rationally in pursuit of desires and intentions, given its beliefs. Here we refer to ToM as the "strong" criterion of agency. Understanding physically constrained agents like a leashed dog falls between the weak and strong criteria of agency. On one hand, it goes beyond merely violating prediction of inanimate objects, and must be explained in terms of desires and intention. On the other hand, it is not entirely social in a way that its mind exceeds the scope of a visual physical scene. Instead, understanding the composition of physical forces driving its motion is still the key in revealing its intention. Building upon the two theories, our work aims to fathom how ToM works on top of ToB as a result of interactions between the social mind and the physical environment. But first, we will discuss a few limitations that lie within how ToB describes physical events.

One important assumption of ToB is that physics is certain, in which case observers should form clear, certain predictions of objects' motions and corresponding energy sources based on physics principles. Any violation of those predictions can then be conclusively attributed to agency with internal energy. Yet for humans, physics in the real world is less certain, with many latent physical properties such as mass, friction, and unlimited possibilities of field of force (Ullman et al., 2017). Having sole access to the observed motions on the surface, humans may not be immediately aware of objects' energy sources. For example, when we observe an object moving downward at an accelerating speed, we may hold doubts on whether the object is free-falling by gravity or intentionally propelling itself downward. In this scenario, the uncertainty of physics leads to an uncertainty of the agent's mind.

Another limitation resides in the fact that ToB heavily relies on the rigid distinction between internal and external forces. In fact, forces in the world are often more complicated. When an agent acts on a physical system, the system "re-acts" back. In the leash-chasing example, when the dog pulls the leash forward, the leash "pulls" it back. The force of "pulling back" neither comes from internal energy (the leash is certainly inanimate without internal energy for self-propulsion) nor external energy (it is from the leash, not gravity or wind). According to the physics of rigid body dynamics, it is a "constraint" force that balances internal and external forces so that the motion will not break the physical system. Without accounting for this force, it would be impossible to make a realistic description of the physical world. Therefore, our analysis here establishes a solid theoretical foundation for understanding physically constrained agents.

A Theory of Joint Physics-Mind Perception

We further develop the existing theories by proposing a joint inference framework of physics and mind. We employ a "general equation of motion" (Eq. (1)) that has been widely used in modern physics engines (Todorov, Erez, & Tassa, 2012) to describe how an agent in a given physical system accelerates in response to various types of forces, or, in terms of established branches of classical mechanics, how a specific force diagram (kinetics) of a system leads to a set of equations that mathematically describe its motion (kinematics). Specifically, the force diagram of an agent in a given physical system may consist of three types of forces. The bias force (c) refers to forces coming from the global environment that equally apply to every component within a physical system. It can be considered as a major type of external force. In contrast, the control force (τ) is an agent-specific force that sits at the core of our theory. It is willingly exerted by an agent using internal energy. Further, we propose that the control force is endowed with a unique duality nature: it connects the ending portal of mind to the starting point of physical interactions. On one hand, it carries the property of agents in the way it is driven by the mind with beliefs, desires, and intentions (Bratman et al., 1987) following the rationality principle (Gergely et al., 1995). On the other hand, it serves as a mechanical force sending off from the agents that marks the start of physical computations. By applying this force, an agent may elicit various responses from the environment. One type of response the agent may receive is the constraint force (f) imposed by the physical system, as we have discussed in the leash-chasing example.

By placing the control force into a physical system (bias force (c) and constraint matrix (J^T)), the equation of motion takes those components as input and generates two solutions simultaneously: 1) the constraint force f as the physical systems' response to c and τ, and 2) the acceleration ($\dot{v}$) of the system by the composition of all forces. It is the acceleration closest to the unconstrained acceleration from c and τ alone without breaking the system.

$$M\dot{v} + c = J^T f + \tau \tag{1}$$

However, when an agent serves as a third-party observer, it does not have direct access to any underlying force but the observed trajectory of a system, from which the acceleration can be derived. That is, only the kinematics, but not kinetics, of a system are available to an observer. Hence, for humans, an inference takes place from kinematics (observed motions) to kinetics (latent forces), yet it does not end with a force diagram level. The ultimate task of vision is to eventually jointly infer a combination of mental states and physical systems that controls the forces for explaining the observed motions (Fig. 3). This very task can be achieved by Bayesian inference (Eq. (2)). Since motion is a common effect of physics and mind, physics and mind become dependent conditioned on motion (Pearl, 2009), therefore indicating a joint inference process. In fact, the mind node can be further broken down into belief-desire-intention (Bratman et al., 1987), extending the general theory into a joint inference of multiple mental states together with physics. In the present paper, we focus on the perception of chasing as an intention, assuming that there is no uncertainty in belief (the wolf sees the whole display) or desires (the wolf only desires to catch the sheep).

$$P(P, M|\dot{v}) \propto \sum_{\tau} P(P)P(M)P(\tau|M)P(\dot{v}|\tau, P) \tag{2}$$

Related modeling work

Our work is deeply inspired by previous modeling work on intuitive physics (Kubricht, Holyoak, & Lu, 2017; Ullman et

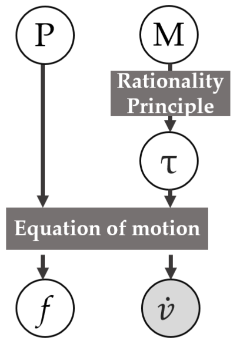


Fig. 3. A casual model of joint perception of physics-mind.

al., 2017) and intuitive psychology (Baker, Saxe, & Tenenbaum, 2009; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Shu, Peng, Fan, Lu, & Zhu, 2018). In a parallel line of work, Shu and his colleagues also investigated the inference of social agency in a physics engine, including animacy from interactions (Shu, Peng, Lu, & Zhu, 2019) and hierarchical social and non-social goals (Shu, Kryven, Ullman, & Tenenbaum, 2020). While our perspective is highly aligned with theirs, our focus differs in the following ways. First, our theoretical analysis is grounded in the equation of motion by explicitly connecting it to the classic ToB and ToM (Gelman et al., 1995; Leslie, 1995). Second, we emphasize the importance of the constraint force beyond the classic internal/external energy distinction. This guides us to study perceived intention where an agent's motion is simultaneously driven by multiple forces. Third, we emphasize a mutual, parallel interaction between physics and mind: not only should the mind be perceived in the context of physics, but the physics should also be perceived in the context of mind. Fourth, we focus on automatic, spontaneous visual perception instead of social inference in general. Therefore, we do not adopt subjective ratings of agency, which could potentially reflect higher-level cognitive inference. Instead, our study is rooted in the psychophysics of chasing using search performance as an objective measurement of perceptual capacity (Gao et al., 2009; Meyerhoff et al., 2013; van Buren et al., 2017).

Leash Chasing

To simulate the real-life scenario that integrates physics and intention, we designed a mock-up display of leash-chasing, in which an agent is physically constrained by a leash (Fig. 4). It contains three agents, a sheep, a wolf, master, and a master-wolf leash. An agent's motion is generated by composing the control force from its intention and the constraint force (if any). Here, we first introduce the control force. The wolf is endowed with a control force in a heat-seeking direction toward the sheep (Gao et al., 2009). In contrast, for the sheep, a heat-avoiding heuristic will quickly get it cornered at a border, and thus we used deep reinforcement learning (Mnih et al., 2015) to train an escape policy that returns a control force given all agents' states. In addition, the master simply exerts

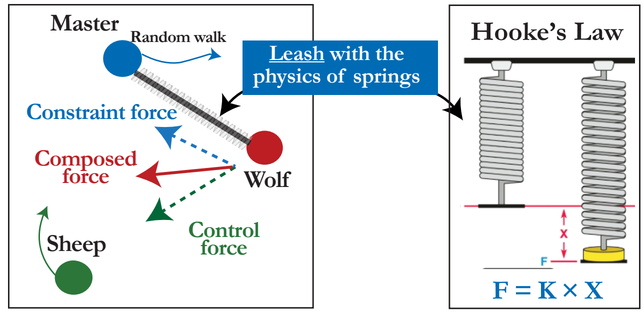


Fig. 4. A leash chasing display.

force randomly. Forwarding to the constraint forces, they are only applied to the wolf and master through the leash, visually represented as a line. Physically, it is modeled by Hooke's law, with the magnitudes of the constraint force a linear function of the leash's length. The constraint forces applied to the wolf and master always have the same magnitudes but opposite directions, summing up to zero. The sheep is unconstrained.

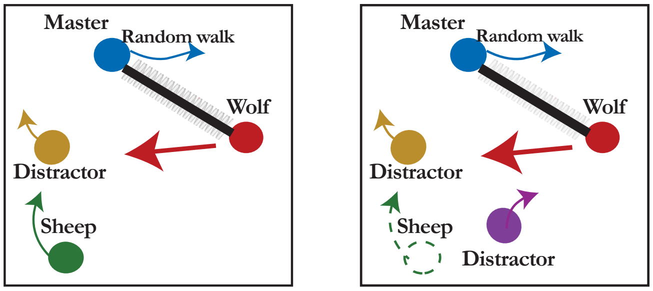
Analyzing trajectories of leash-chasing showed that on average the wolf's motion deviated from the sheep position by 69°, corresponding to a 138° chasing subtlety. Existing results suggested that with such a large subtlety and a line disrupting objecthood, perceived chasing should be severally disrupted. However, the joint perception of physics-mind predicts that vision can use the line as a physical cause to justify the large subtlety, so that perceived chasing can be much more robust.

Psychophysics Experiments

General Method

We adapted the visual search task from Gao et al., (2009). All displays were mutated from the leash-chasing introduced above, with the leash or chasing disrupted in different ways across experiments. To increase the search difficulty, a distractor that moved randomly was added. The dependence between physics and mind given observed motion implies that disrupting one would disrupt the perception of the other. We predicted that performance of identifying chasing would only drop when the physical system was arbitrarily disrupted. In parallel, the performance of identifying a leash system would drop if the wolf's chasing intention was disrupted arbitrarily. Therefore, our strategy here was to manipulate either physics or intention, and then measure its effects on the perception of the other. This also allowed us to study spontaneous visual perception as the manipulations of physics or intention were always task irrelevant. For all experiments, leash was not mentioned at all if the task was to identify chasing, and chasing was not mentioned if the task was to identify leash. There were 12 new participants for each experiment.

The procedure and data analysis in the study largely followed the search-for-chasing paradigm (Gao et al., 2009). Participants were required to detect whether or not such a chase or leash system was present. If a chase or leash was indicated, they were asked to identify the corresponding agents engaged in the chasing or objects within the leash system. In



(a) chasing-present (b) chasing-absent

Fig. 5. Manipulation of the chasing presence in Expt. 1, 2, 3.

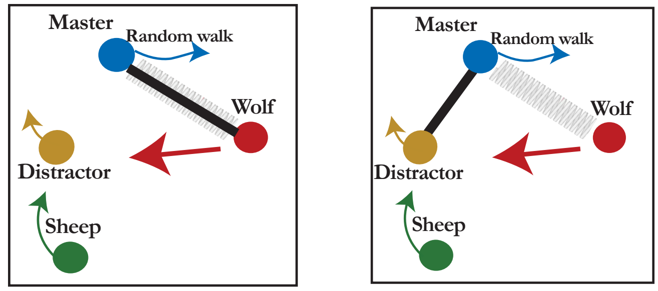
this paradigm, the identification accuracy, unlike a simple hit rate in detection, relied on participants genuinely pointing out the specific roles involved in a relationship. The chance performance in the identification task was 16.7%, much lower than that in the detection task (50%), suggesting a greater difficulty of making a correct guess in the identification task. Due to the advantage in preventing lucky guesses, the identification accuracy is considered as a more reliable measurement of perceived chasing or leash and thus is the main focus of our analysis.

Expt. 1: Perceiving leash chasing

We started by proving perceived chasing could be more robust with a leash, despite large subtleties and line connection. The chasing present condition (Fig. 5a) was simply a leash chasing display with a distractor. The chasing-absent condition (Fig. 5b) was identical except that the sheep turned invisible, and another distractor was added into the display. Since chasing was defined as "catching the sheep", hiding the sheep turned the wolf's motion unintentional, while not changing its motion. Therefore, the motions of the master, wolf, and sheep (if visible) were identical across conditions. To reveal the spontaneous effect from intuitive physics, two line conditions were introduced. In the master-wolf condition (Fig. 6a), a line accurately representing the leash was presented. In the master-distractor condition, intuitive physics was disrupted by hiding the real leash and showing an arbitrary line connecting the master with a distractor (Fig. 6b).

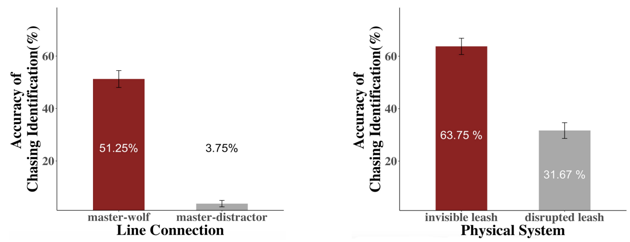
Procedure Each trial started a motion display (18° by 18°) with four discs (0.7°) that continued moving for 10 seconds. After that, participants pressed a button to indicate whether chasing was present (F) or absent (J). When chasing-present was indicated, participants needed to identify the wolf and sheep with mouse clicks. There was no feedback to participants' responses. In total, there were 80 trials, with 20 trials for each combination of the line connection (master-wolf vs. master-distractor) and chasing (present vs. absent).

Results The accuracy of chasing identification was defined as identifying both the "wolf" and "sheep" correctly when chasing was present. It was much higher in the master-wolf condition than the master-distractor condition (Fig. 7a, $t(11)$ = 22.76, $p < 0.001$, Cohen's d = 7.25). This result has two implications. First, perceived chasing is not determined by



(a) master-wolf (b) master-distractor

Fig. 6. Manipulation of the line connection in Expt. 1 & 3.



(a) Chasing identification in Expt. 1 (b) Chasing identification in Expt. 2

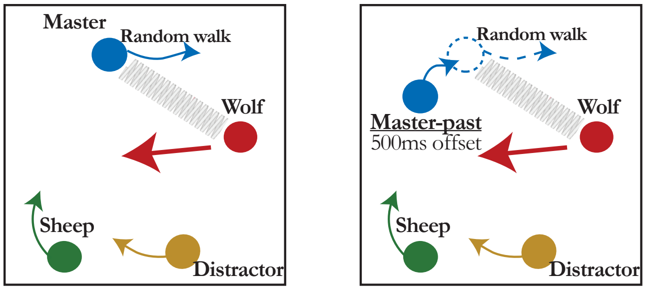
Fig. 7. Results of chasing identification in Expt. 1 & 2.

the absolute value of chasing subtlety, but largely influenced by whether the subtlety could be explained by an intuitive physical constraint. Search performance here (51.3%) also was much higher than those reported in Gao et al., (2009) with similar subtleties (15% $\sim$ 30%), although their displays were not exactly the same. Second, line connection as a disruption of objecthood does not necessarily impair perceived chasing. Here perceived chasing was much more robust when it was the wolf connected to a line. This was inconsistent with the object-based account of perceived chasing, but consistent with the joint physics-mind account, since the line highlighting the physical system largely explained the deviation observed in the wolf. Overall, we found it very intriguing that a line connecting non-chasing objects yielded worse perceived chasing.

Expt. 2: Perceiving chasing with an invisible leash

Our core argument is that vision jointly infers physics and mind, which can best explain the observed motion. It brings the uncertainty of physics into the perception of mind. To realize this physical uncertainty, we turned the leash invisible (Fig. 8a), so that the physical system became latent and must be inferred from motions. Without the line connection, we introduced a new baseline: a disrupted leash, in which the motion of the master was delayed by 500 ms (Fig. 8b). In other words, the display mixed the current sheep and wolf with a master from 500ms ago. Therefore, while the wolf and master's motion may still be correlated, that correlation could not be explained by intuitive physics. Other aspects of the design were identical to Expt. 1.

Results Performance of chasing identification was much higher in the invisible leash condition than the disrupted-leash



(a) invisible leash (b) disrupted leash

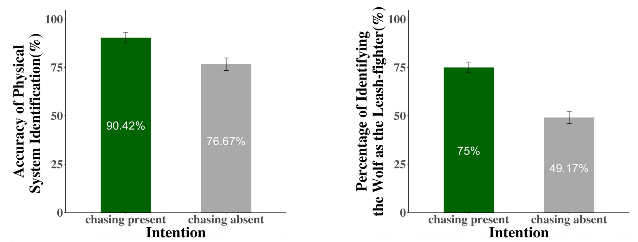
Fig. 8. Manipulation of the invisible leash system in Expt. 2.

condition (Fig. 7b, $t(11) = 5.81$, $p < 0.001$, Cohen's $d = 1.49$). We found it very intriguing that manipulating the motion of the master, a non-chasing object, could significantly impact perceived chasing. This result indicated that even when leash was not mentioned at all and there was no visual cue implying a leash, participants still spontaneously used the master's motion to explain the wolf's large chasing subtlety. With the disrupted-leash, performance was similar to those reported in Gao et al., (2009), suggesting that it was the presence of an intuitive physical system that improved perceived chasing.

Expt. 3: Perceiving leash through chasing

Finally we demonstrated the mutual physics-intention effect by turning the table around: asking participants to identify a physical system (Fig. 6) while intention was manipulated (Fig. 5). As in Expt.1-2, chasing presence was still manipulated by the visibility of the sheep. However, it was no longer the search task. Participants were instructed to identify whether the line connecting two objects was a leash or just an arbitrary line connecting two independent objects. If a leash was reported, they then identified which object was more actively fighting against the leash's constraint. This was to measure the perception of the wolf's control force (or internal energy), which was exerted regardless of whether the sheep was visible or not. According to the joint perception of physics-mind, turning the sheep invisible would disrupt identifications of the leash and the wolf's control force against the leash.

Results Fig. 9a showed that when the leash was presented, the percentage of accurately identifying the leash was significantly higher in the chasing-present condition than the chasing-absent condition ($t(11) = 5.81$, $p < 0.01$, Cohen's $d = 1.49$). In addition, the percentage of identifying the wolf as the leash-fighter was also much higher in the chasing-present condition than the chasing-absent condition (Fig. 9b, $t(11) = 3.35$, $p < 0.001$, Cohen's $d = 0.92$). Note that in all these cases, the wolf, the master, and the leash connecting them were exactly the same, and the only difference came from whether the sheep was visible. We found it intriguing that manipulating the visibility of an object outside of a physical system could nevertheless greatly impact the perception of that system. It demonstrated that participants spontaneously used intention as a cause for inferring control force, which subsequently facilitated the inference of constraint force, from which a physical system could be identified.



(a) Leash identification (b) Leash-fighter identification

Fig. 9. Results of identifying physical system in Expt. 3.

Conclusion

In the present paper, we propose a theory of joint perception of physics-mind by iterating ToB and ToM with the classic equation of motion. The core of this theory is to recognize the duality of the "control force", which is the output of agency and the input to the equation of motion in physics. It explains the motion of an agent when the motion is jointly generated by multiple sources of force. This theory provides a broader perspective for interpreting existing psychophysics of chasing, suggesting that perceived chasing can be more robust than previously expected in the context of an intuitive physical system. In addition, inferring control forces through perceived intention is also critical for identifying a physical system with constraint forces.

References

Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, *20*(3), 191–208.

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

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*, *110*(45), 18327–18332.

Bratman, M., et al. (1987). *Intention, plans, and practical reason* (Vol. 10). Cambridge, MA: Harvard University Press.

Carey, S. (2000). The origin of concepts. *Journal of Cognition and Development*, *1*(1), 37–41.

Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive psychology*, *59*(2), 154–179.

Gelman, R., Durgin, F., & Kaufman, L. (1995). Distinguishing between animates and inanimates: Not by motion alone. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 150–184). Oxford: Clarendon Press.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*(2), 165–193.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American journal of psychology*, *57*(2), 243–259.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in cognitive sciences*, *20*(8), 589–604.

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

Leslie, A. M. (1995). A theory of agency. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (p. 121—141). Oxford: Clarendon Press.

Luo, Y., Kaufman, L., & Baillargeon, R. (2009). Young infants' reasoning about physical events involving inert and self-propelled objects. *Cognitive psychology*, *58*(4), 441–486.

Meyerhoff, H. S., Huff, M., & Schwan, S. (2013). Linking perceptual animacy to attention: Evidence from the chasing detection paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(4), 1003-1015.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... others (2015). Human-level control through deep reinforcement learning. *nature*, *518*(7540), 529–533.

Pearl, J. (2009). *Causality*. Cambridge: Cambridge University Press.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, *36*(1), 1–16.

Saxe, R., Carey, S., & Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.*, *55*, 87–124.

Shu, T., Kryven, M., Ullman, T. D., & Tenenbaum, J. B. (2020). Adventures in flatland: Perceiving social interactions under physical dynamics. In S. Denison., M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the cognitive science society* (pp. 2901–2907). Cognitive Science Society.

Shu, T., Peng, Y., Fan, L., Lu, H., & Zhu, S.-C. (2018). Perception of human interaction based on motion trajectories: From aerial videos to decontextualized animations. *Topics in cognitive science*, *10*(1), 225–241.

Shu, T., Peng, Y., Lu, H., & Zhu, S.-C. (2019). Partitioning the perception of physical and social events within a unified psychological space. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st annual conference of the cognitive science society* (pp. 1048–1054). Montreal, QB: Cognitive Science Society.

Spelke, E. S. (1990). Principles of object perception. *Cognitive science*, *14*(1), 29–56.

Stewart, J. A. (1982). *Perception of animacy*. Phd thesis, University of Pennsylvania, Philadelphia.

Todorov, E., Erez, T., & Tassa, Y. (2012). Mujoco: A physics engine for model-based control. In *2012 ieee/rsj international conference on intelligent robots and systems* (pp. 5026–5033).

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.

van Buren, B., Gao, T., & Scholl, B. J. (2017). What are the underlying units of perceived animacy? chasing detection is intrinsically object-based. *Psychonomic bulletin & review*, *24*(5), 1604–1610.

Wellman, H. M. (1992). *The child's theory of mind.* Cambridge, MA: MIT Press.