ARTICLE IN PRESS
PLREV:968

Available online at www.sciencedirect.com

ScienceDirect

Physics of Life Reviews ••• (••••) •••–•••

PHYSICS of LIFE reviews

www.elsevier.com/locate/plrev

Comment

Direct perception vs inferential processes in reading an opponent's mind: The case of a goalkeeper facing a soccer penalty kick Comment on "Seeing mental states: An experimental strategy for measuring the observability of other minds" by Cristina Becchio et al.

Giulio Sandini, Pietro Morasso

Istituto Italiano di Tecnologia, Research Unit of Robotics, Brain and Cognitive Sciences, Genoa, Italy

Received 3 January 2018; accepted 4 January 2018

Communicated by J. Fontanari

In engineering cybernetics, observability is a measure of how well internal states of a system can be inferred from knowledge of its external outputs. Moreover, observability and controllability of a system are mathematically inter-related properties in the sense that it does not matter to have access to hidden states if this knowledge is not exploited for achieving a goal. While such issues can be well posed in the engineering field, in cognitive neuroscience it is quite difficult to restrict the analysis in such a way to isolate direct perception from other cognitive processes, named as "inferences" by the authors [1], without losing a great part of the action (unless one trivializes the meaning of "direct" by stating that "all perception is direct": Gallagher and Zahavi [6]). In other words, in spite of the elegance and scientific rigor of the proposed experimental strategy, in our opinion it misses the fact that in real human–human interactions "direct perception" and "inference" are two faces of the same coin and mental states in a social context are, in a general sense, accessible on the basis of directly perceived sensory signals (here and now) tuned by expectations. In the following, we elaborate this opinion with reference to a competitive interaction paradigm, namely the attempt of a goalkeeper to save a soccer penalty kick by "reading the mind" of his opponent.

When a penalty kick is shot from the standard distance of 11 m with a good but not maximum speed of 20–22 m/s, the ball reaches the door frame in a little bit more than half a second. Considering a reaction time of 0.3 s, the goalkeeper has about 0.2 s to organize an effective response and, even if it is therefore impossible for the goalie to plan a complex coordinated action [3], goalies who do not anticipate the kick are twice more successful than anticipators. Thus it is not surprising that in handbooks for goalies (e.g. [3]) it is suggested that a goalie should remain still until the ball is kicked. Therefore, for a goalkeeper preparing and executing a successful dive is a complex task that involves well-tuned perceptual, motor, cognitive, and affective evaluations, including detection of fake actions by the kicker [21] and that is based, using the terminology of the authors [1], on a mixture of "direct perception" and "inference" based on expectation.

DOI of original article: https://doi.org/10.1016/j.plrev.2017.10.002.
E-mail address: giulio.sandini@iit.it (G. Sandini).

https://doi.org/10.1016/j.plrev.2018.01.004

Please cite this article in press as: Sandini G, Morasso P. Direct perception vs inferential processes in reading an opponent's mind: The case of a goalkeeper facing a soccer penalty kick. Phys Life Rev (2018), https://doi.org/10.1016/j.plrev.2018.01.004

Considering that there is a clearly marked t_0, namely the time instant when the player's foot impacts the ball thus producing a sharp sound, the functionally relevant "direct (visual) perceptual information" applies to a short interval on both sides of t_0. What happens in this time interval determines indeed the physics of the shot: direction, height, speed. Even if an experienced goalie facing a professional kicker is exploiting both streams of information, we argue that during the period preceding the kick, starting with the whistle of the referee and the motion of the player towards the kicking spot, (i.e. the period of time when "mind reading" is based only on "direct perception" of the enlarged environment) the role of expectations and inference is fundamental in preparing the dive.

In contrast, the mechanism used by plummeting gannets [9] in order to trigger the closing of the wings at the last moment before impacting with the sea surface is an example of pure "direct perception", based on the estimate of the time-to-collision with the sea surface via the rate of expansion of the optic flow: in this case, we may assume that a Darwinian selection mechanism has been in place to achieve a safe plummeting strategy and a well-tuned algorithm for a direct visual estimate of the time to impact. In the case of humans, there is evidence of pure direct perception in crucial aspects of neonatal social interaction, such as detection of biological motion and human face [19,11]. On the other hand, an experienced goalkeeper has to complement direct perceptual evidence with expectations derived from visual search and visual attention strategies [16].

But is "inference" the right term for describing the mental process of the goalkeeper in his attempt to guess the kicking intentions of his opponent during the few seconds of the approach to the kicking spot? Inference has to do with abstract reasoning, moving from premises to conclusions according to the laws of logic, without any specific consideration of the concrete, time-dependent aspects of actions. A professional goalkeeper is a very unlikely logic agent but certainly his mental processes include prior general knowledge about the properties of "biological motion" [4], the kicking preferences of the opponent player and possibly specific episodic memories of previous kicking experiences. All of this fits quite well in the general framework of Embodied Cognition (EC) which is characterized, according to [23], by four main features: 1) EC is *situated*, in the sense that it is an online process which takes place in the context of task-relevant sensorimotor information; 2) EC is *time pressured*, i.e. it is constrained by the requirements of real-time interaction; 3) the *environment* is an integral part of EC, including both the physical and social environment, 4) EC is intrinsically *action oriented*.

Regarding the four steps strategy proposed by the authors [1] for determining whether and to what extent a given mental state is observable, we agree on the need for empirical evidence and appreciate the rigorousness of the presented approach. In particular, we agree that the analysis of the availability and perceptual efficiency of information to perceive mental states are fundamental steps of this empirical process. However, we think that something is missing in this analysis: in order to assess the observability of mental states and to identify the specific features observers use to detect intention (step 3 of the process proposed), it is not sufficient to analyze the kinematic features of the action in an observer-independent reference frame but it is necessary to assess how such features, possibly reflecting the performer's intentions, are mapped into the perceptual space(s) of the observer. In the example proposed by the authors [1] as well the goalkeeper paradigm, the relative position of the actor and the observer affects the observability of the perceptual feature as well as the perceptual accuracy required to discriminate kinematic differences.

1. From fundamental principles to applications: the case of robotics

In relation to the aspects raised in Box 2 of the article as the potential impacts of the proposed methodology in the field of robotics, we agree with the authors [1] that in order to support the coordination of collaborative action, prospective planning and emphatic interaction, it is crucial for robots to establish a bidirectional communication channel with their human partners that includes some form of access to mental states. This entails both enabling robots to read the behavioral patterns of humans and, crucially, to express their own intentions and internal states with their motion. In the terminology of the authors [1], this means that the robot is able to read its partner's mind and to express its own behavior in a form readable to a human observer. Extensive research has been conducted on both fronts [10,8], with often a particular focus on the role of eyes, proposed as a proxy to convey and read mental states [15,13]. Recently, the attention has been directed toward the intuitive messages associated to body motion and to the design of a computational model able to generate the most legible movements from the perspective of the action partner [5]. The goal is the development of a human-aware movement framework, which synthesizes safe, comfortable, socially acceptable, and readable movements, especially in the context of collaborative manipulation [17, 20]. This is particularly important for humanoid robots since, as far as simple collaborative behaviors are concerned,

Please cite this article in press as: Sandini G, Morasso P. Direct perception vs inferential processes in reading an opponent's mind: The case of a goalkeeper facing a soccer penalty kick. Phys Life Rev (2018), https://doi.org/10.1016/j.plrev.2018.01.004