Book of Abstracts

- Maria Alvarez (King's College London): Is action explanation factive?

> The explanation of human actions can take various different forms. Discussions of the issue tend to focus on 'belief-desire' explanations, i.e. explanations of the form 'He phi-ed because he wanted to psy and believed that p', for example, 'He bought an electric car because he wanted to reduce his contributions to environmental degradation and believed that driving an electric car would contribute to that aim'. I shall provide a picture of range of types of action explanation, locating these B-D explanations within that picture, and examine whether all these explanations are factive.

- Joe Cunningham (University of Nottingham): Responding to Reasons: From Deviance to Disjunctivism

> An agent who successfully responds to a normative reason is an agent who displays a certain kind of normative achievement: There really is a reason for them to j and they j in a way that manifests sensitivity to this reason and its normative status, so that they end up j-ing as they (pro-tanto) ought where this is a success attributable to them. But what is it to j in a way that manifests the sort of sensitivity at issue? My aim in this paper is twofold. First, I aim to attack an approach to answering that question codified by what I call the Neutrality Assumption: The claim, roughly, that successfully responding to a normative reason factors into j-ing in response to what appears to one to be a normative reason for one to j plus a set of further independent conditions. I argue that there is no plausible account of the success case which meets the constraints laid down by the Neutrality Assumption. My second aim is to develop a positive account of responding to normative reasons premised on a rejection of the Assumption: An account according to which the success case involves the exercise of essentially successful normative capacities and according to which we should offer a disjunctive account of j-ing in response to what appears to one to be normative reason for one to do so.

- Anna Krämer (University of Salzburg): Teleology – good reasons for seeking evidence from developmental psychology

> When it comes to understanding intentional actions by means of objective, publicly available facts, teleology theory (Perner & Roessler, 2010) provides a good explanation with several advantages over other theories like theory theory, simulation, or rationality theory (see Perner et al., 2018; Perner & Esken, 2015). But how do we come to this teleological understanding of intentional actions? Perner and Esken (2015) proposed a developmental trajectory: Around 6 to 12 months, infants understand goal-directed actions as behavior that effects changes that regularly end in the same state (goal). The end point is not yet seen as particularly desirable. This happens around 18 months and enables children to understand that actions are done for good reasons, i.e., effecting a change to the better, explaining the emergence of helping behavior at this age. In this talk, I will present one study carried out and further studies planned to empirically support the teleology theory by investigating these proposed developmental steps.

- Arturs Logins (University of Zürich): What Else Can Knowledge Explain?

> This paper revisits the question of the causal relevance of knowledge. On the one hand, appeals to knowledge are insightful in some instances of causal explanation of action or emotions (e.g., in cases of persistent action or factive emotions). On the other hand, however, constraints of generality and simplicity suggest that knowledge is not indispensable in causal explanations. After all, given that knowledge entails belief and belief is not factive, appeals to belief might provide a higher degree of generality and simplicity in causal explanations of action and emotion. This talk aims to move the debate forward by considering a potentially indispensable role for knowledge in explaining certain forms of inquiry (e.g., knowledge aimed curiosity) and certain forms of assessment (e.g., blame and praise).

- Jennifer Nagel (University of Toronto): The Safety Condition on Knowledge

> There is a difference between knowing and just happening to get it right. Epistemologists sometimes explain this difference in terms of a "safety condition" on knowledge, a condition which is violated in Gettier cases. I examine the connection between safety conditions and the possibility of gaining second-hand knowledge through factive mental state attribution.

- Jonathan Phillips (Dartmouth College): Actual knowledge

> In this talk, I'll argue that when you represent others as knowing something (or represent others' factive mental states more generally), you represent their mind as being related to the actual world. This feature of knowledge explains the limits of knowledge attribution, how knowledge differs from belief, and why knowledge underwrites learning from others. This picture contrasts with accounts of knowledge according to which knowledge, like belief, requires representing the contents of someone's mind independently from the actual world.

- Johannes Roessler (University of Warwick): "How?" and "why?"

> Are we ordinarily disposed to make sense of our own and others' intentional actions by reference to reason-giving facts (rather than merely by reference to non-factive mental states)? If so, what is the rationale of that practice? And is the practice defensible, in the light of philosophical (or cognitive-scientific) theorizing about the etiology of our actions? In this talk I approach these questions by examining Elizabeth Anscombe's thesis that practical reasoning and reason-giving explanation are concerned with "the same order" ("an order which is there whenever actions are done with intentions"). My main question will be whether Anscombe's thesis yields an argument for the primacy of "factive action explanation".

- Eva Schmidt (Technical University Dortmund): The Reasons of AI Systems

> Against the backdrop of the program of explainable artificial intelligence (XAI), we explore how to ascribe reasons and reasoning to (non-classical) AI systems to explain their outputs. The outputs of many AI systems are reliably very good and useful, which indicates that these systems robustly respond to reasons, and that their responses can be explained by appeal to reasons. But such systems often have features which make it difficult to get at their exact reasoning processes or reasons - e.g. their

reasons are represented (and their reasoning processes are structured) in sub-symbolic or distributed ways. Here we develop several strategies to overcome these obstacles: approximating reasons and reasoning processes by way of simplifying and idealizing what is actually taking place in a system; and approximating reasons and reasoning processes by providing the most charitable interpretation of a system in terms of reasons and reasoning processes. Finally, we investigate whether the outputs of AI systems can be explained best with the help of psychologistic reasons or rather with the help of factualist reasons.

- Brent Strickland (Institut Jean Nicod): Automatic altercentricy ? New data from two new paradigms

Do we automatically process others' perceptions and resulting beliefs? If so, to what extent does such processing affect our own representations of events and objects? Here I present new work from our group which helps shed light on these questions. For a first set of experiments, we developed a novel object tracking and detection task in which participants and a salient animated agent ("Zoe") watched objects move and become occluded inside one of two domes. Occluders dropped at the end of the trial to reveal object locations that could be congruent or incongruent with the participant's or the agent's perspective. Participants had to indicate as quickly as possible where they saw the object. We found robust and consistent effects of participant perspective across a number of experiments, but no effects of agent perspective. In a second set of studies we manipulated line of sight in a new gaze cueing paradigm in which a central agent shifted their gaze towards (or away from) an object to which there was no clear line of sight (because it was behind a barrier) or to which there was a clear line of sight (because there was a gap in the barrier). We found gaze cueing effects in both line of sight conditions, whereby participants were faster to detect an attentional probe when this appeared in a direction aligned with agent looking direction. However these effects were significantly larger when there was an open line of sight, suggesting that agent perspective automatically influences gaze cueing. On the basis of these findings and others in the literature, I speculate about the likely theoretical scope of "altercentric effects". I suggest is that we are more likely to find automatic effects of others' perspective in cases where that perspective is an ecologically predictive of the behavior the relevant social agent(s). When agent perspective is logically decoupled from event outcomes (as when a participant is tasked to detect the location of an object), we should expect to see fragile altercentric effects or none at all.

- Julia Wolf (Ruhr University Bochum): Before Belief – Knowledge and Pretence

For many years, research on the development of Theory of Mind has focused primarily on the ability to attribute beliefs. Recently, however, arguments have been made that factive mental states like knowledge should play an important role in Theory of Mind. Moreover, it has been argued that knowledge attributions may provide the 'entry points' for mental state attributions (Nagel, 2017), and that knowledge attributions may generally be more basic and developmentally prior to belief attributions (Phillips et al., 2021). A version of this view is also articulated by Phillips and Norby (2019), who argue that the difference between factive ToM and non-factive ToM does not lie in any difference between specific ToM abilities, but rather in the additional demand of being able to construct and maintain a representation which conflicts with how you take the world to be. This domain general ability is required by non-factive ToM, but not factive ToM. In this talk I critically evaluate this proposal, drawing also on evid-

ence from children's abilities to deal with inconsistent perspectives in pretend play. I will argue that the problem lies in being able to access different perspectives, where accessing consistent perspectives is easier than accessing inconsistent perspectives. This can be used to explain the developmental priority of knowledge attributions.