Agency and Causation:
Formal and Conceptual Issues
Book of abstracts

Date: 27-29 October 2016

Venue: Royal Academy of Dutch Language and Literature (KANTL)

Koningstraat 18, Ghent, Belgium

Organization:

Centre for Logic and Philosophy of Science (Ghent University)

Centre for Philosophical Psychology (University of Antwerp)

Chairs	**Program Committee**	**Local Organizing Committee**
Bert Leuridan	Mathieu Beirlaen	Stef Frijters
Frederik Van De Putte	Inge De Bal	Thomas Lodewyckx
	Leen De Vreese	Julie Mennes
	Raoul Gervais	Annelies Monseré
	Joke Meheus	Pawel Pawlowski
	Dingmar Van Eck	Dietlinde Wouters

ABSTRACTS

Invited speakers

JAN BROERSEN (UTRECHT UNIVERSITY)
Getting a Formal Grip on Responsibility

Our goal is to develop a formal language in which we can talk about responsibilities of individual agents. Through agency, responsibility is connected to causality. The literature has several suggestions on how to formally define this connection. We will critically review some of these definitions and put forward some suggestions for improvement.

AGNES MOORS (CATHOLIC UNIVERSITY OF LEUVEN)
The power of goal-directed processes in emotional and other actions

Standard dual process models in the action domain hold that stimulus-driven processes are responsible for suboptimal behavior because they take them to be rigid but automatic and therefore the default. We propose an alternative dual process model in which goal-directed processes are the default instead. This amounts to shrinking the explanatory territory of stimulus-driven processes and expand that of goal-directed ones. We then transfer the dual process logic from the action domain to the emotion domain. This reveals that emotional action tendencies are often attributed to stimulus-driven processes. Our alternative model submits that emotional action tendencies can also be caused by goal-directed processes. We evaluate the type of empirical evidence required for validating our alternative model and we consider the implications of our model for behavior change, encouraging strategies focused on the expectancies and values of action outcomes.

ELIZABETH PACHERIE (INSTITUT JEAN NICOD)
Intentions and motor representations: the interface challenge

A full account of purposive action must appeal not only to propositional attitude states like beliefs, desires, and intentions, but also to motor representations, i.e., non-propositional states that are thought to represent, among other things, action outcomes as well as detailed kinematic features of bodily movements. This raises the puzzle of how it is that these two distinct types of state successfully coordinate. I examine this so-called "Interface Problem". First, I clarify and expand on the respective nature and functions of motor representations and intentions in explaining intentional action. Next, I characterize the interface challenge that their differences in representational format and content imply. I then evaluate Butterfill and Sinigaglia's (2014) recent answer to this interface challenge, according to which intentions refer to action outcomes by way of demonstrative deference to motor representations. I present some worries for this proposal, arguing that, among other things, it implicitly presupposes a solution to the problem, and so cannot help to resolve it. Finally, I suggest that we may make some progress on this puzzle by positing a "content-preserving causal process" taking place between intentions and motor representations, and offer a proposal for how this might work.

Contributed speakers

Lieke Asma (Free University of Amsterdam)

Dual action control, conscious guidance, and first-order consciousness

In response to threats to the causal efficacy of conscious agency, Pacherie (2014) suggests that actions that are produced by unconscious motor programs can still be under conscious guidance. For this, conscious processes do not actually have to affect the intentional movements, as long as the conscious agent 'was prepared to intervene if necessary and that he was in a position to do so more or less effectively' (Frankfurt 1978, p. 160). Pacherie defends this cooperation between conscious and unconscious processes by referring to experiments by Fourneret and colleagues (1998, 2001, 2002) that show that when discrepancies between the desired state and actual state become too large while performing a simple task, subjects will become conscious of their bodily movements. This seems to show that only when needed conscious processes intervene. However if we accept, as Pacherie (2006, p. 160) does in earlier work, that conscious causation is about first-order consciousness according to which an intention is conscious if the creature is conscious of the state of affairs the intention is about, then the subjects' conscious intention to perform the task plays a causal role throughout the experiment. The only difference is that when the discrepancy becomes too large, they also become conscious of their specific bodily movements.

Vanessa Carr (University College London)

Davidson's Challenge for Irreducible Agent-Causation

There is a distinct lack of clarity over the notion of reduction at issue in the debate over the reducibility of causation by agents (agent-causation). This lack of clarity obstructs critical evaluation of the various positions that are taken up. Clarification of the notion of reduction here should allow us to draw a line between two coherent mutually exclusive views, placing at least most self-proclaimed reductionists and non-reductionists on their chosen side of the line. I argue here that we should take the reduction at issue to be that which Davidson can be understood as defending in this debate: inter-theoretic explanation. But I also highlight that while Davidson specifically defends the reduction of agent-causation to event-causation, we should not articulate reductionism here specifically in terms of event-causation, contrary to its usual articulation. I then propose that once we have so understood the debate, we can appreciate the content of an important challenge that Davidson put to those defending irreducible agent-causation, and see that it remains unmet.

Inge De Bal & Erik Weber (Ghent University)

Causation and technical problem solving

Everyone is familiar with technical problem solving instructions like "If the light is out, replace the batteries." One element of these instructions describes the remedy, viz. to replace the batteries. In this talk we analyse the causal underpinnings of such remedy claims found in technical problem solving instructions. For these claims to be successful, they need to be based on causal relations that hold in the world and have certain properties. These required properties are the focus of our talk. We first introduce examples from car- and bike repair manuals to demarcate our topic and function as case studies. We then build a framework for analysing these cases. We start from theories of causation by Ronald Giere, Ellery Eels and John Mackie and develop a series of definitions to capture the properties of the aforementioned causal relations. Throughout the talk, we use the case studies to illustrate our point. We conclude that remedy claims need to be based on causal relations with properties (1) positive causal factorhood with weak context-unanimity and (2) sufficiency in maximally normal contexts. We finally show that our analysis is not limited to means of conveyance.

Anna Drozdzewska (Université Catholique de Louvain)
Intentions and causality

Why genuine mental causation is crucial for free will. The causal power of intentions and decisions is often implicitly assumed in various discussions on the problem of free will, where an action to be considered free, should originate from our conscious decision rather than physical brain activation. In my presentation, I will discuss the place of conscious decision in the debates in both neuroscience (starting with Libet's experiments, moving to the more recent versions) and philosophy (Mele, 2009, Searle, 2010). I will hypothesize that efficacious mental causation is one of the conditions for the possibility of free will. I will further focus on interventionism, with its basic assumptions applied also in neuroscience where the cause is identified through its potential influence on the effect. The most detailed formal accounts of this intuition were developed by Menzies and Price (1983), Pearl (2000) and Woodward (2003). Here I focus on List and Menzie's (forthcoming) attempt to solve the mental causation puzzle and show how free will is possible under interventionism. They argue that, thanks to multiple realization of the mental, various brain activations can be supervenient bases of the same mental event, thus only manipulating the mental will likely change the outcoming action. I will then discuss some of the criticisms against interventionism (Baumgartner, 2009, 2010) focused on the impossibility of independent manipulation of the two potential (mental and physical) causes. I will show how, when applied to the case of free will, Baumgartner's arguments provide new insights into the debate relevant both for interventionists as well as neuroscientists.

Hein Duijf (Utrecht University)
Conflicting intentions: rectifying the consistency requirements

Many philosophers are convinced that rationality dictates that one's overall set of intentions be consistent. Our main objective is to argue against this conviction and demonstrate that one can be rational in having mutually inconsistent intentions. Our formal analysis of consistency requirements is based on Bratman's planning theory of intentions and scenarios of choice under uncertainty, as studied in decision theory. Various consistency requirements and an acceptance criterion for such requirements are carefully recollected from Bratman's theory. This acceptance criterion emphasizes that the required unity of one's various intentions is an ability to act rational in pursuit of realizing them. We specify what this means using insights from the study of choice under uncertainty, and introduce a new consistency requirement, coined practical consistency. Our findings are three-fold: First, we prove that having mutually consistent intentions is insufficient for supporting the coordinating role akin to our intentions. The demand for mutually consistent intentions is thus to be rejected on the basis of the acceptance criterion. Second, we show that the consistency requirements are also not necessary for providing this coordinating role. This shows that it can be rational for an agent to have mutually inconsistent intentions. Third, the norm of practical consistency is to be accepted. Finally, when restricting to intentions conforming to the 'own-action condition', we observe that the practical consistency norm is equivalent to overall consistency. In particular, this highlights that the acceptance criterion challenges the consistency requirements only in scenarios of choice under uncertainty.

Victor Gijsbers & Leon de Bruin (Leiden University)
Folk psychological interventions

In everyday life we attribute a broad range of mental states to others in order to make sense of their behavior. For example, we readily understand that Sarah takes the exit because she wants to avoid the traffic jam ahead, that Michael wears a tie because he thinks it is important to make a good impression, and that John carries an umbrella because he believes that it is going to rain and has the intention to stay dry. The capacity to understand

the behavior of others in terms of their mental states is commonly referred to as 'folk psychology'. A central assumption of folk psychology is that mental states cause behavior. But how should this be understood?

An approach to causation that has achieved widespread popularity recently is James Woodward's interventionism (Woodward 2003). The main idea of Woodward's theory is that causation should be analysed in terms of intervention: roughly, X causes Y if and only if there is a possible intervention I on X that changes the value of Y. Much of the theory's work is done by conditions which spell out exactly which interactions count as interventions.

The aim of our paper is to investigate whether the interventionist approach can be successfully applied to the practice of folk psychology. First, we will show that one of the main obstacles for such an approach is that everyday mental state ascriptions do not seem to meet Woodward's conditions for an intervention (in particular condition I2, which states that an intervention I on X should act as a switch for all the other variables that influence the value of X). We argue that this does not so much show that folk psychological interventions are 'fat-handed' instead of 'surgical', but rather that the interventionist approach to folk psychology is lacking in some significant respect.

Second, we propose to solve this problem by supplementing interventionism with an agency theory of causation. Such a solution is attractive because an agency theory of causation: a) fits naturally with our folk psychological practice, b) imposes less stringent conditions compared to interventionism, yet does not suffer from problems related to common cause structures, and c) solves the circularity problem for interventionism (Gijsbers & De Bruin (2014)).

Johannes Himmelreich (Humboldt-Universität zu Berlin)
What Agency Does: A Challenge for Production Theories of Agency

There is an important link between agency and responsibility. In this paper I argue that many theories of agency fail to account for this link. Nevertheless, there are other theories of agency that hold the promise of doing so. While agency theories of the former kind have been widely explored in philosophy, theories of the latter kind have often been overlooked. Theories of the former kind, which I call production theories of agency, include proposals such as the theory of Donald Davidson. Theories of the latter kind, which I call counterfactual theories of agency, include proposals such as stit-logics. This paper raises a challenge for production accounts of agency.

Carolyn Jennings (University of California, Merced)
The Puzzle of Skilled Behavior

The puzzle of skilled behavior is that of determining whether and how it counts as a form of intentional action. Although skilled behavior seems more deeply agent-involving than other behaviors, making it seem the perfect candidate for intentional behavior, it also seems sometimes to proceed without and despite the agent, giving it qualities of unintentional and automatic behavior. In this paper I review the puzzle of skilled behavior and some attempted solutions to the puzzle, before presenting my own solution: the puzzle of skilled behavior calls for us to distinguish two types of control, rational control and expert control, both of which allow for intentional action. The natural tension between these forms of control, at its most extreme for highly skilled activities, explains the tension felt in the case of skilled behavior. Along the way I will discuss the case of choking and the empirical research on this phenomenon to draw out the puzzle.

Naftali Weinberger & Alessandra Marra (Tilburg University)
Equilibrium Systems and Intentions: A Causal Analogy

In our talk, we explore an analogy between the role of intentions – in particular future-directed intentions (Bratman, 1987, 1992) – and a certain type of non-agential causal systems, namely equilibrium systems. Our reason for doing so is to resolve several difficulties involved in understanding the causal role of future-directed

intentions. Intentions are: 1) extended in time; 2) do not provide reasons for action; 3) do not determine a complete plan of action; and 4) involve reasoning backwards from a goal to the actions needed to accomplish it. After explaining why these features are potentially problematic, we show that these concerns may be addressed using an analogy between intentional agents and equilibrium systems. By "equilibrium systems", we mean systems whose variables tend towards stable states. When modeling these systems at a time scale that is sufficiently long for the system's variables to reach their equilibrium states in response to perturbations, one can abstract away from the shorter-term dynamic interactions among the variables (Iwasaki & Simon 1994). The analogy we draw is that between future-directed intentions and the conditions that are required for a system to reach equilibrium. Just as these conditions influence a system's causal relationships without actually being variables in the system, intentions influence the roles of desires and beliefs in action without providing independent reasons for action. The analogy helps us resolve our difficulties by showing how intentions could explain how agents reliably bring about their goals without the intentions themselves counting as causes.

MARCUS MISSAL (UNIVERSITÉ CATHOLIQUE DE LOUVAIN)
Experimental evidence in favor of a pluralistic conception of causality

When two objects such as billiard balls collide, human observers perceive that the action of one caused the motion of the other. We have previously shown that this extends to the orientation of gaze: subjects make more predictive eye movements in the expected direction of causal motion than in any other direction. However, we found also that when asked to report whether a simulated causal interaction took place on a computer screen (Michotte launching display), the explicit judgment of human observers and the implicit prediction by the oculomotor system can diverge. These experimental results show that there could be divergent processes underlying oculomotor responses to and judgments of causal stimuli. Furthermore, they support the hypothesis that a single concept of causality is not applicable in cognitive psychology. From the philosophical point of view, our approach strengthens a pluralist conception of causality.

GRIGORY OLKHOVIKOV (RUHR-UNIVERSITÄT BOCHUM, URAL FEDERAL UNIVERSITY)
Inference as doxastic agency: the basics of justification stit logic

In this talk we are going to present a joint work with Prof. H. Wansing on proving as doxastic action. We consider logical inference as an activity that results in proofs and hence produces knowledge. We intend to model the activity component in this picture according to the semantical analysis of deliberatively seeing-to-it-that, which is structured around the crucial distinction between agentive propositions and moment-determinate ones. The proofs themselves and their relations to the proven sentences we propose to model according to S4LP, Artemov logic of proofs augmented with epistemic operator, which exploits the distinction between explicit and implicit modes of knowledge. The two systems are then merged on a basis of a unified semantics. This merger results in a logical system which extends both stit logic and justification logic and allows to apply both crucial distinctions of the component logics within a framework of a relatively rich classification of modes of proving activity in a given community. Intersection of the two distinctions gives us four basic classes of modalities related to proving activity of agents. We look into their semantics and present a number of valid principles that relate the various notions of proving to each other and to notions of justified knowledge, implicit knowledge, and historical possibility. All in all, the talk will focus on conceptual issues rather than technical ones.

AMIT PUNDIK (TEL AVIV UNIVERSITY)
Causation and Generalisation

In a criminal trial, when the prosecution needs to prove that the accused's intentional action caused a certain harm, what could be better than showing how unlikely it is for that type of harm to occur naturally or accidently? When Lucia de Berk was put on trial in the Netherlands for murdering seven babies and attempting to murder another three in the hospital wards where she had worked as a nurse, an expert witness for the prosecution testified that the probability of so many unexplained deaths and resuscitations all occurring naturally during the shifts of a single nurse is 1 in 342 million. This paper argues that evidence on the rarity of natural cause should not be admitted in criminal trials. The argument progresses in two stages. Drawing on the theories of contrastive explanation and contrastive causation from the philosophy of science, the paper first explains why evidence on the rarity of a certain natural cause is only relevant when contrasted, to another piece of evidence: the likelihood of intentional action. To take a simple example, assume that it is disputed whether the deceased's death was caused by strangulation or a rare respiratory disease. That the rare disease occurs only in 1 per million people means one thing if 1 per 100,000 people is a strangler, but entirely different thing if 1 per 100 million people is a strangler. To be relevant to the case at hand, evidence on the rarity of natural cause needs to be contrasted to evidence on the likelihood of the culpable cause, namely to the frequency of the specific intentional action ascribed to the accused among people who are similar to her. While calculating the likelihood of intended action involves technical difficulties of its own, the paper proceeds to explain why using this evidence is objectionable in principle, no matter if and how it can be produced. In previous work I have argued that the kind of intentional action which is relevant to criminal trial, namely free intentional action, cannot be proven with the likelihood of the intended free action among people who are similar to the accused. In this paper, I seek to show that if one piece of evidence (on the rarity of a natural cause) cannot be used without contrasting it to another piece of evidence (on the likelihood of the intended action), and the latter is objectionable in principle, then neither should be admitted in court. Using evidence on the rarity of natural causes is therefore wrong no matter how credible the statistical analysis is.

MANUEL REBUSCHI (UNIVERSITY OF LORRAINE NANCY)
Intentionality inside? Towards an intensional account of action verbs

I can kiss a statue and while doing that, kiss my favourite rock star, my grandmother, or even a non-existent goddess. Actions seem to go far beyond physicalist behaviours. This makes the relationship between action and causality rather complex to analyse. The goal of the talk is to offer a general framework based on first-order modal logic that makes it possible to handle this relationship. Intentionality need not be considered a substance or a relation constitutive of actions. One can be antirealistic regarding intentionality hence regarding action. Behaviours are caused by physical events and should be explained as such. But actions can be analysed following other lines. Actions are ascribed, sometimes self-ascribed, like mental attitudes. The talk will focus onto transitive action verbs (TAVs) like "kiss", "kick", "kill"... These verbs are similar intensional transitive verbs (ITVs) like "love", "seek", "worship"... which raise well-known difficulties. TAVs and ITVs share a common feature, namely intensionality, that put them apart from verbs denoting physicalist behaviours. In order to analyse TAVs (as well as ITVs) I will follow a three-step strategy: (i) analyze TAV-sentences as modal sentences, (ii) extend modal logic with independent quantifiers, and (iii) use cross-world extensions of predicates. The proposal provides a general picture according to which transitive actions are cross-world relations whereas physically caused behaviours are wholly contained in the actual world.

GUNNAR SCHUMANN (FERNUNIVERSITÄT IN HAGEN)
Why causalist theories of action are wrong

Causalists of agency think that actions are effects of causes, be it of mental causes of their own sort, be it of causes of a physical, mainly neurological sort. This picture is fuelled from several ideas: 1) that actions are physical events, 2) that actions are physical events accompanied by volitional mental occurrences, 3) that these volitional mental occurrences are entities of some kind and that these are causing actions just like natural phenomena cause each other. In my talk I want to show that these assumptions are deeply mistaken. In particular, I want to show that actions are not caused by intentions. I will make use of my own renewed version of the "Logical Connection Argument". According to it, intentions and actions are connected logically. More specifically, the connection is a normative one. Our concept of intention is such that the first person use of the expression "to have an intention to do X" implies that the agent is committed to perform the action when he has the opportunity to do so. This will be spelled out for simple and complex actions. From cases of complex, long-term actions or cases of hindrances the expression "to have the intention to..." gets its sense and meaning - by it we do not mean the possession of some entity (be it mental or neuronal) - as it is commonly misunderstood by causalists. The function of a practical syllogism then is to show - given the premises - what action must or should be done by the agent in order to be called "rational".

GEORGE STAMETS (UNIVERSITY OF LEEDS)
A Non-Causal Libertarian Account of Intentional Action

Contemporary philosophy of action is dominated by a set of assumptions that comprise the Causal Theory of Action (CTA) - an approach so widely accepted that it is sometimes called the "standard story" of intentional action. Among these includes the idea that causation is a relation between events, and the thought that any behavior that could count as an intentional action must be caused in an appropriate way by mental items like belief-desire pairs or intentions (or the onsets of these, depending on what one takes an "event" to be). In this talk, I charge that views that adopt the CTA, in addition to being committed to unsatisfactory accounts of causation in general, cannot successfully accommodate the possibility of intentional action: behaviors that are somehow caused by mental items like belief-desire pairs or intentions, however commonly they might be thought to occur in ordinary human life, cannot themselves count as intentional of the persons who perform them. Taking the term "intentional" in this context to be equivalent to "deliberately" or "on purpose", I suggest that any genuinely intentional action must also be a free action. I then argue that free actions must be entirely uncaused - that is, must be neither deterministically nor indeterministically caused, whether by agent-involving events (contrary to the CTA) or by agents themselves. Following E.J. Lowe, I briefly sketch and defend a non-causal libertarian account of free and intentional action, built upon a neo-Aristotelian metaphysics that takes seriously the notion of causal powers.

HORIA TARNOVANU (UNIVERSITY OF ST ANDREWS)
Effect Selection

As formulated by Mill (1843/1947: 213-214), the problem of causal selection is that of explaining the priority given to a determinant as the cause of a target event in contrast to the complex sum of all determinants in its past light cone. While causal selection received considerable attention (Hesslow 1988, Waters 2007, Schaffer 2012, Franklin-Hall forthcoming-a, b, inter alia), there has been little discussion of a similar mechanism operating on the effect side of causal relations. From an analogous Millian angle, the real Effect is the whole set of consequents and no clear explanation is offered for the priority typically given to a consequent or set of consequents (the effect) over other results (by-products, side-effects, after-effects, etc.). The main goal of the paper is to explore effect selection

and assess its relevance to determining the appropriate extent of responsibility for consequences. I argue that the difficulties in the way of a clear answer to the puzzle of effect selection motivate an epistemic argument against consequentialism. Overview: First, I introduce effect selection, review several attempts to formulate selection criteria, and indicate their deficiencies. Second, I examine if two prominent ways of explaining causal selection apply to the effect side. I show that a prospective approach needs to strike a balance between: (a) the claim that effects and by-products are metaphysically distinct, and (b) the claim that effects and by-products are in no sense objectively different and selection is consistently governed by context-dependent pragmatics. I argue that despite the strong sense that effects and by-products are essentially different, the criteria governing their differentiation are neither clear, nor predictable. Third, I consider the relevance of effect selection to determining the appropriate scope of moral responsibility for consequences and develop a novel epistemic objection to consequentialism.

BRANDT VAN DER GAAST (UNIVERSITY OF TWENTE)
Functionalism and Action Individuation

When a believer fails to recognize an individual, he lacks an identity belief that two individuals are one and the same. Gottlob Frege's classic example is that of the Babylonians who used the names "Hesperus" and "Phosphorus" for the same heavenly body without realizing it. A case of mistaken identity, on the other hand, is where a believer misrecognizes one individual as another individual. Here, the believer has a false identity belief. E.g.: Little Red Ridinghood mistaking the wolf for her grandmother. In this paper, I consider how our practice of action explanation interacts with these two types of cognitive failure-ignorance and error. I will claim that in such contexts, believers, even though their cognition is not working as it should, can nevertheless be said to be rational, and that their behavior should therefore be explainable in folk-psychological terms. I will argue that, in order for such explanations to work, their desires and actions have to be described in what I call "proximal terms". I then relate this conclusion to a David Lewis-style functionalist theory of mental states.

PETER VERDÉE (UNIVERSITÉ CATHOLIQUE DE LOUVAIN)
My brain's rules made me do it. Can certain mental properties be seen as causally efficient rules?

In this paper we propose an approach to mental causation in which certain mental properties are essentially modal properties. From a counterfactual difference-making point of view on causation, it is natural to say that modal-structure-giving properties such as rules and laws can cause certain actual physical events. We present evidence for the hypothesis that causally efficient mental properties are better seen as such modal-structure-giving properties than as properties of the actual world.

SIMON WIMMER (UNIVERSITY OF WARWICK)
Infants' understanding of intentional actions and false belief tasks

The belief/desire model of intentional action says that seeing agents as acting intentionally requires thinking of their actions as caused and rationalised by their reasons for acting, which are constituted or provided by belief/desire pairs. (Davidson 1963) Consequently, understanding intentional actions requires understanding the explanatory role of beliefs and desires. There is evidence that 2- and 3-year olds understand intentional actions (e.g. Meltzoff 1995, Tomasello 1999), but whether this evidence is consistent with the belief/desire model has recently been questioned by Perner and Roessler (2010). In contrast to Perner and Roessler, Baillargeon et al. (2016) suggest that it is. In particular, Baillar-geon et al. argue that infants come to understand intentional actions between 0 and 2 years, because they acquire the capacity for "psychological reasoning", including an understanding of the explanato-ry role of beliefs, during that time. In my paper, I argue that Baillargeon et al. are mistaken. I argue that, even if we

grant that 2-year-olds have the capacity for psychological reasoning that Baillargeon et al. ascribe to them, that capacity is not sufficient to understand intentional actions, because intentional actions, on the belief/desire model, are supposed to be both caused by beliefs and desires and rationalised by them. And whilst there is evidence that 2- and 3-year olds understand the former role of beliefs, there is also evidence that they do not understand the latter. Thus, I suggest the belief/desire model is in fact inconsistent with the evidence that 2- and 3-year olds understand intentional action.