

## Summer School 2018

### THINKING ABOUT THE POSSIBLE 9 JULY – 14 JULY, 2018, BUDAPEST, HUNGARY

*Patricia Ganea*, University of Toronto & *Ágnes Kovács*, Central European University

#### **Course Aims:**

The proposed course will bring together perspectives on thinking about imagination and counterfactual reasoning from a range of disciplines, with seminars offered by faculty from the fields of Developmental and Cognitive Psychology, Philosophy, Linguistics, Cognitive Science, Neuroscience, and Informatics. Course participants will learn about empirical techniques and findings from studies in cognitive development, cognitive psychology, and neuroscience. They will engage with theoretical perspectives on the nature of counterfactual, causal, conditional, and temporal reasoning. A presentation session for course participants will give early-career researchers the opportunity to share their research and receive feedback from one another and from established scholars on the course faculty.

The foremost goal of the proposed course is student training and networking. Additional aims of the proposed course are to 1) stimulate discussion about theory, methods, and empirical evidence on imagination and counterfactual reasoning by bringing together perspectives from different disciplines, 2) promote interdisciplinary collaborations and conversations, and 3) increase conversations about the applications of research on imagination and counterfactual reasoning.

#### **Abstracts and Bibliography**

##### **Ruth Byrne –Thinking about counterfactual alternatives**

I consider how people create counterfactual alternatives to reality when they imagine how things would have turned out differently 'if only...'. The cognitive processes that compute counterfactuals mutate aspects of the mental representation of reality to create an imagined alternative. I discuss alternative views of the sorts of cognitive processes that may underlie this ability, and review evidence that suggests that people understand counterfactuals by constructing mental models that compare the conjecture and the presupposed facts. This dual meaning account of counterfactuals has led to the discovery of a range of phenomena, for example, people make more inferences from counterfactuals compared to ordinary conditionals. I also consider the many reasons that lead people to create counterfactual alternatives to reality. Counterfactuals help to explain the past, in some cases by helping people to identify causal or intentional links between events; they also help people to prepare for the future, by ensuring that they learn from their mistakes and can prevent bad things from happening again, and so they can help people to make better decisions. I review evidence to illustrate these functions, focusing particularly on the role of counterfactuals in moral reasoning. Finally, people create counterfactuals by changing an aspect of reality, and an intriguing observation is that they seem to do so by focusing on a 'fault-line' in their representation of reality, a juncture or pivotal joint in a sequence of events, such as an exceptional event or a controllable one. I consider evidence that these fault-lines reflect aspects of reality that have been represented explicitly in models. Evidence that the 'fault-lines' can be manipulated is discussed in the context of the modulation of what is represented in models by knowledge.

**Readings:**

Byrne, R.M.J. (2016). Counterfactual thoughts. *Annual Review of Psychology*. 67.135- 157.

Kahneman D, & Miller D.T. (1986). Norm theory: comparing reality to its alternatives. *Psychological Review*, 93.2, 136–53.

**Discussion Questions:**

What new experiments could distinguish between different theories about how people create counterfactual alternatives? And what new experiments could distinguish between different theories about how people reason from counterfactual conditionals?

What additional functions or impacts could counterfactual thinking have, that have not yet been studied?

**Chris Lucas – Rationality and utility in counterfactual reasoning**

Counterfactual reasoning is at the heart of our attributions of credit and blame in legal contexts, as well as in everyday life. In order to make better decisions in these settings and to understand how people think about counterfactuals, it is useful to reflect on what makes a counterfactual inference better or worse. Numerous studies show that when people make counterfactual inferences and judgments about the actual causes of past events, they attend to factors that signal how an individual or community can make useful decisions in the future. For instance, people tend to focus on events that are under human control, and are sensitive to probabilities in a way that licenses sensible distinctions between bad luck and bad judgment. Despite these hints that human counterfactual reasoning is well-adapted to support efficient learning and sound decisions, it is still unclear what counts as a good counterfactual judgment in many situations. Pearl has offered one theory of how an agent should answer counterfactual questions, but it disagrees with human judgments in ways that could be taken as evidence for human irrationality, or as evidence that the model's assumptions should be relaxed. How can we assess the correctness or utility of a counterfactual judgment? When should we regret our actions and learn not to repeat them? Can we describe the counterfactual inferences that a rational or optimal agent should make, and use these to understand and guide our own behavior? In an effort to answer these questions, I will discuss (1) models of counterfactual reasoning; (2) their relationships with accounts of optimal or rational decision-making; (3) some of the theoretical and empirical challenges in assessing the correctness or value of these models; and (4) previously-offered and new answers to these challenges.

**Readings:**

Lucas, C. G., & Kemp, C. (2015). An improved probabilistic account of counterfactual reasoning. *Psychological Review*, 122(4), 700-734.

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, 12(2), 168-192.

**Discussion Questions:**

\*Describe a few situations where thinking about counterfactuals might change a person's behavior in the future. Try to come up with a case about counterfactuals involving a person's own past actions (e.g., "If only I had..."), a case that involves the actions of other people ("If only they had...") and a case where the focus isn't on human action. For each situation, is the change in behavior likely to lead to a better outcome in the future?

\* We will discuss a couple of ideas of what it means to be "rational" in the context of how we should update our beliefs given new information. Come up with a simple everyday situation where two reasonable people might disagree about what counts as a rational belief or strategy. Explain why they disagree and whether the disagreement is likely to be resolved.

### **Eva Rafetseder – Development of counterfactual reasoning: Conceptual distinctions and relation to theory of mind**

A central question when studying human social interaction is: How do we come to understand each other? How do we figure out what other people believe? And if other people's beliefs are mistaken, how do we predict what they are going to do? A longstanding debate centres on whether we do so by theorizing, simulation or by any other means. Studying children has proven a very useful tool in this debate, particularly to gain deeper insights into the latent cognitive processes. One avenue has focused on counterfactual thinking. Counterfactual thinking is a highly specialised tool of human cognition which allows us to imagine how an event or a situation could have unfolded differently. For example, Max puts his chocolate into the drawer. In his absence his mum moves the chocolate to the cupboard. One can ask counterfactually, where the chocolate would be, if mum had not moved it. Over the past 20 years counterfactual questions have repeatedly been found to correlate with false belief questions (Where will Max look for his chocolate?) in 3- to 4-year-old children. The central idea is that, when our knowledge runs counter to the simulated agent's, the imagined situation becomes counterfactual. Like in the false belief literature, however, there is still a debate about when children are competent in answering counterfactual questions. Several recent studies demonstrate that 4-year-olds (who typically succeed on false belief tasks) can answer some but not all counterfactual questions, raising the question of whether there is actually a relationship between false belief reasoning and counterfactual reasoning, and if so, what the structure of that relationship might be. In my lecture I will outline the current debate on the development of counterfactual reasoning and I will link this debate to the most recent literature of false belief understanding.

#### **Readings:**

- Riggs, J. K., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, 1998, 13, 73-90.
- Rafetseder, E., Cristi-Vargas, R., & Perner, J. (2010). Counterfactual reasoning: developing a sense of "nearest possible world. *Child Development*, 81(1), 376–389. doi:10.1111/j.1467-8624.2009.01401.x

#### **Discussion Questions:**

What are the defining features of counterfactual reasoning? What would you class as counterfactual reasoning and what not?

How can developmental psychology help unravel cognitive underpinnings of counterfactual reasoning? What cognitive factors would you consider?

By what mechanisms might counterfactual reasoning affect children's false belief understanding? How does your answer fit with simulation theory and theory theory?

## **Caren Walker – Imagination and causal cognition**

Conventional wisdom suggests that knowledge and imagination, science and fantasy, are deeply different from one another – even opposites. However, new ideas about children’s causal reasoning suggests that exactly the same abilities that allow children to learn so much about the world, reason so powerfully about it, and act to change it, also allow them to imagine alternative worlds that may never exist at all. A large portion of our psychological lives is spent engaging in counterfactual thought, planning and anticipating future states, and considering near and far alternatives to the actual state of the world. While the imagination has long been assumed to generate counterfactuals, little research has explored how human minds, even the very youngest human minds, manage to produce these counterfactuals, how we know which possibilities will be the most likely to occur, and why imagining new possibilities is important. I will present an argument characterizing the relation between imagination and causal cognition, in light of recent developments in computational theories of cognitive development. According to the Bayesian model of human learning, our ability to imagine possible worlds and engage in counterfactual reasoning is closely tied to our ability to think causally. Indeed, the purpose and distinguishing feature of causal knowledge is that it allows one to generate counterfactual inferences. I will first provide a broad overview of the “probabilistic models” framework of causality, and review empirical work in that framework which shows that adults and children use causal knowledge to generate counterfactuals. I will also outline a theoretical argument that suggests that the imagination is central to the process of causal understanding. Finally, I will offer evidence that Bayesian learning implicates the imaginative process, and conclude with a discussion of how this approach may be applied to the study of the imagination, more classically construed, in the contexts of both pretend play and scientific inquiry.

### **Readings:**

Gopnik, A. (2009). *The Philosophical Baby* (Chapters 1). HarperCollins Publishers: NY.

Gopnik, A. & Walker C.M. (2013). Considering counterfactuals: The relationship between causal learning and pretend play. *American Journal of Play* (Special Issue), 6(1): 15-28.

### **Discussion questions:**

At first, the prevalence of pretend play in early childhood seems paradoxical. That is, given that children have so much to learn about the real world, why would they spend so much time and energy engaging with unreality?

What is the relationship between imagination and causal reasoning?

How might causal models and probabilistic inference help to inform the study of counterfactual reasoning?

## **Paul Harris – Children’s ideas about what can and cannot happen**

Young children are often credited with a rich imagination that enables them to enjoy a variety of fantasy worlds. I argue that this portrait is misleading. When children engage in imaginative play and when they make sense of the pretend actions of a play partner, they routinely invoke their understanding of everyday causal regularities. As such, their pretend play tends to reproduce the familiar, pedestrian routines of everyday life even it can also be laced with transgression and disruption. I argue that this restriction on pretend worlds also applies to children’s thinking about counterfactual possibilities.

Although there is evidence that young children are able to think about how reality might have turned out otherwise, I will argue that children invoke only modest departures from what actually happened – departures that typically fall within the range of everyday causality. I will also explore a potentially powerful objection to this insistence on the pedestrian and reality-bound nature of children’s imagination. Children are receptive to religious and fictional narratives in which miraculous or magical events occur. Since these events do depart from ordinary reality, the implication is that children enjoy a fertile imagination. However, I will present evidence that children’s receptivity is quite dependent on input from the surrounding culture. Thus, although children’s imagination can be fed, and arguably enriched, by religious and fictional narratives, they are not prone to generate such narratives autonomously. Finally, I will discuss the extent to which this somewhat withholding portrait of children’s imagination can be extended to that of adults.

### **Readings:**

Lane, J. D., Ronfard, S. L., Francioli, S. P., & Harris, P.L. (2016). Children’s imagination and belief: Prone to flights of fancy or grounded in reality? *Cognition*, 152, 127-140.

Corriveau, K.H., Chen, E.E. & Harris, P.L. (2015). Judgments about fact and fiction by children from religious and non-religious backgrounds. *Cognitive Science*, 39, 353–382.

### **Discussion Questions:**

Do young children have an imagination that plods close to reality?

What conditions enable them to soar?

### **Sarah Beck – Development of counterfactual reasoning and its relation to executive function**

Being able to think about what might have been involves managing thoughts of reality and the unreal. Here I will present evidence that the development of counterfactual thinking is underpinned by improvements in executive functions. Executive functions are the processes that allow us to organise and structure our goal-driven behaviour. I will explore how this evidence fits with a view that the development of counterfactual thinking is protracted and continues through early and middle childhood. Most evidence comes from individual differences studies. Inhibitory control has been implicated in children's ability to respond to counterfactual conditional questions in which they need to resist answering with what they know to be true. Children with better working memory generate more counterfactual possibilities. And, it appears that in order to experience regret, children need to have advanced attentional flexibility. It is less common but some of these claims have been tested with experimental manipulations: if children have to delay a response to a counterfactual question, they are more likely to give the correct answer. This suggests that part of children's difficulty with counterfactual thinking is resisting the impulsive response of answering with how the world is. I will discuss whether executive functions should be thought of as supporting counterfactual competence or counterfactual performance. Furthermore, we will engage with lifespan developmental questions, by reflecting on evidence that suggests that counterfactual thinking becomes automatised in adulthood.

### **Readings:**

Beck, S.R., Carroll, D.J., Brunson, V.E.A., & Gryg, C.K. (2011). Supporting children's counterfactual thinking with alternative modes of responding. *Journal of Experimental Child Psychology*, 108, 190-202.

Drayton, S., Turley-Ames, K.J., & Guajardo, N.R. (2011). Counterfactual thinking and false belief: the role of executive function. *Journal of Experimental Child Psychology*, 108, 532-548.

**Discussion Questions:**

What are the different types of executive function and how might they relate to counterfactual thinking?

What are the different possible explanations for a developmental relation between executive functions and counterfactual thinking?

What studies would you propose to test whether executive functions underpin counterfactual thinking (in children &/ adults)?

**Christoph Hoerl – Temporal and counterfactual reasoning in regret and relief**

Relief and regret are both emotions often classed as ‘counterfactual emotions’ because they seem to involve an element of counterfactual thinking. Interestingly, regret and relief are also both emotions that seem to bear a special relation to time in that they are both essentially past-directed. Indeed the existence of such past-directed emotions has been taken as the basis of an argument for a certain position in the metaphysics of time – namely that time objectively flows or passes, and that the past is objectively different from the future. In this session, we will look at some puzzles that regret and relief raise, that relate to the role counterfactual and temporal reasoning plays in them. However, I will also use considerations about the nature of regret and relief to draw some more general conclusions about connections between temporal reasoning and modal (or, more specifically counterfactual) reasoning. Specifically, I will argue that there are several features of our everyday thinking about time that seem to require a capacity to engage in counterfactual reasoning. Three such features I will focus on in particular are the fact that we can think of times in an ‘event-independent’ way, the fact that we think of the present as marking the boundary between an ‘open’ future and a ‘fixed’ past, and the fact that we think of time as passing, with moments of time successively becoming present.

**Readings:**

Hoerl, C. (2015): Tense and the psychology of relief. *Topoi* 34, pp. 217-231.

Hoerl, C. & McCormack, T. (2016): Making decisions about the future: Regret and the cognitive function of episodic memory. In: K. Michaelian, S.B. Klein & K. Szpunar (eds.): *Seeing the future: Theoretical perspectives on future-oriented mental time travel*. Oxford University Press, pp. 241-266.

**Discussion Questions:**

To what extent can regret and relief be considered antonyms; and in what respects do they differ from each other?

Is it rational to feel regret or relief about something?

In what ways are relief and regret connected to the idea of the future as 'open' and the past 'fixed'?

## **Teresa McCormack – The development of regret and its impact on children’s decision making**

Experiencing regret is unpleasant, but it is generally assumed to be a functional emotion that helps us make better decisions. If regret does indeed play an important role in decision making, the developmental emergence of regret should have knock-on effects on the quality of children’s decisions. Because regret is a counterfactual emotion that involves thinking about what might have been, it would not be expected to be present in very young children who cannot think counterfactually. Given this, a number of developmental studies conducted over the last five years have examined when regret emerges. Our research suggests that by at least 6-7 years, the majority of children can experience regret, and that this is the case even in the context of more complex decisions involving risk-taking. These initial studies all examined what can be termed intrapersonal regret: regret regarding self-interested choices that did not yield an optimal outcome for the child. However, people also experience regret following choices that have had a negative impact on another person. Our recent studies suggest that the developmental profile of so-called interpersonal regret is similar to that of intrapersonal regret, with children of 6-7 showing this emotion following a failure to make a prosocial choice. We have found that the likelihood that children experience interpersonal regret is related to individual differences in self-reported sympathy, and also to children’s understanding of moral emotions. Having described research establishing the development profile of regret, I will then describe a series of studies suggesting that once children begin to experience regret they make more adaptive decisions. In particular, regret seems to help children switch from making less optimal to making more optimal choices, and may also help them to learn to delay gratification. I will finish by describing our most recent work that indicates that children who regret not acting prosocially will subsequently be more likely to make a prosocial choice, suggesting that the development of regret may also have an impact on prosocial or moral development.

### **Readings:**

- O’Connor, E., McCormack, T., & Feeney, A. (2014). Do children who experience regret make better decisions? A developmental study of the behavioral consequences of regret. *Child Development, 85*(5), 1995-2010.
- Weisberg, D. P., & Beck, S. R. (2012). The development of children's regret and relief. *Cognition & Emotion, 26*(5), 820-835

### **Discussion Questions:**

To what extent can we be confident that the types of tasks used to assess regret in children really do measure regret? How could we make the tasks better?

By what mechanisms might regret affect children’s decision making?

Can you think of ways of looking at the developmental impact of the emergence of regret?

## **Karen Lewis – The Semantics and Pragmatics of Counterfactual Conditionals**

The standard semantics for counterfactual conditionals is that a counterfactual conditional like (1) *If I had dropped my coffee mug on the morning of April 4, then it would have fallen to the floor* is true if and only if *all* the closest possible worlds in which I dropped my coffee mug on the morning of April 4 are all worlds in which it falls to the floor. Might-counterfactuals are the duals of woulds; (2) *If I had dropped my coffee mug on the morning of April 4, then it might*

*have fallen to the floor* is true if and only if *at least one* of the closest possible mug-dropping-worlds is a falling-to-the-floor-world. Closeness is standardly taken to be a measure of similarity; that is, roughly speaking, worlds that are more similar to the actual world are closer than worlds that are less similar. I argue that this is not entirely right. Closeness is a measure of both similarity and conversational relevance; sometimes there are less similar worlds among the closest (if they are relevant) and sometimes some of the most similar worlds are not among the closest (if they are irrelevant). I further argue that might-counterfactuals sometimes play a pragmatic role of expanding the possibilities that are conversationally relevant. This semantics and pragmatics together account for some puzzling data. For example, they can account for the fact that (1) is true even though it is also true that (3) *If I had dropped my coffee mug on the morning of April 4, it might have quantum tunneled to China.*

### **Readings:**

Lewis, David (1986). "Counterfactual Dependence and Time's Arrow" in *Philosophical Papers Volume II*. Oxford University Press.

Lewis, Karen (2016). "Elusive Counterfactuals". *Noûs*. 50:2. 286-313.

### **Discussion Questions:**

What worlds (possibilities) do competent speakers consider when evaluating counterfactuals as true or false?

How sensitive are counterfactuals to the conversational context, and in what ways?

How should we think of relevance (as it pertains to counterfactuals)? How might we flesh out the notion so that it is fully predictive as to which worlds are relevant?

### **Felipe De Brigard – Toward a Cognitive Neuroscience of Modal Cognition**

In philosophy, the term "modality" refers to the different ways or modes in which things could be. Some things are necessary, some things are possible, and some things are impossible. There are also different kinds of possibility. Some things are physically possible but biologically impossible, for instance, or logically possible but physically impossible. Understanding the boundaries between what is necessary, possible, and impossible is a difficult task, and a large number of logicians, metaphysicians and epistemologists have endeavored their whole careers to clarify this difficult issue. In this talk, I hope to show that cognitive psychology and neuroscience can contribute to this ongoing and fascinating research by uncovering the neural and cognitive mechanisms underlying certain kinds of modal judgments. In particular, I hope to show that thoughts about alternative ways in which the past could, should, or would have been, draw resources from different memory systems depending on certain features of the counterfactual content, such as the perceived likelihood of the counterfactual event, or whether or not the counterfactual event involves a person one is familiar with. My hope is to show that modal cognition is a promising and highly interdisciplinary area of research in cognitive neuro-philosophy.



**Readings:**

Van Hoeck, N., Watson, P.D., and Barbey, A.K. (2015). Cognitive neuroscience of human counterfactual reasoning. *Frontiers in Human Neuroscience*, 9: 420.

De Brigard, F., Addis, D., Ford, J.H., Schacter, D.L., & Giovanello, K.S. (2013) Remembering what could have happened: Neural correlates of episodic counterfactual thinking. *Neuropsychologia*.51(12): 2401-2414

**Discussion Questions:**

What is the relationship between our capacity to imagine possible future events (episodic future thinking) and our capacity to imagine possible ways in which past personal events could have occurred (episodic counterfactual thinking)? In which ways are they similar and in which ways are there different?

How does the brain manage to distinguish these two kinds of hypothetical thoughts, and what is the relationship between these two kinds of hypothetical simulations and the episodic memories they are derived from? Are there cognitive and neural differences between different kinds of counterfactual simulations?

**Ferenc Huoranszki – commentary****Panel discussion on “Precursors of abstract combinatorial thought in infants”****Patricia Ganea, Gergely Csibra, Erno Teglas, Agnes Kovacs**

In this Panel Discussion we will address specific abilities that may be related to the development of counterfactual reasoning, but emerge earlier in development. The ability to think counterfactually is an important – and likely uniquely human – ability, thus specific precursors for such abilities may be present from very early on. When thinking counterfactually, one engages in representing alternatives of reality, changes in reality and the causal implications of these changes. Here we will target a series of potentially related processes that have an early onset and allow young infants to go beyond the here and now, engage in abstract thinking, interpret and anticipate causal and probabilistic events, and recruit abstract combinatorial thought involving negation or disjunction.