

PHD THESIS SUMMARY:

Agency & Choice: On the Cognitive and Conceptual Foundations of Agency in Economics and Behavioral Decision Research

JAMES GRAYOT

Doctorate in philosophy, February 2019

EIPE, Erasmus University Rotterdam

In this thesis, I offer a philosophical perspective on the different conceptions of agency and choice as they are understood and employed in economics and behavioral decision research—this perspective is two-fold: on the one hand, philosophical analysis can clarify ambiguities in definitions and concepts that arise within interdisciplinary research. This is of particular importance given how philosophical concepts such as *mind*, *cognition*, and *intentionality* feature in economic studies of rational choice. Hence, one project of this thesis is to subject contemporary research on questions about agency and choice to such philosophical scrutiny. On the other hand, the questions and topics discussed in this thesis can be understood as an exercise in philosophy of science: they deal explicitly with questions and topics that pertain to the theoretical and empirical practices of scientists. This includes traditional microeconomic disciplines, such as decision theory and game theory, as well as interdisciplinary research in behavioral economics, neuroeconomics, and experimental psychology.

Chapter 2 illustrates the contentious relationship between agency and choice by focusing on a cluster of debates in the philosophy of economics about the scientific status of *preferences*. These debates revolve around a central question, namely, whether preferences should be construed as mental or behavioral phenomena. The literature indicates that these debates are critical to the disciplinary identity of economics (Hausman 1998; Davis 2006; Bruni and Sugden 2007; Backhouse and Medema 2009; Hands 2009, 2014; Ross 2011, 2014), and further, that disputes about the role of psychological explanations in economics are predicated on how one conceives economics as a science (Dowding 2002; Camerer 2008; Hausman 2008, 2012; Guala 2012, 2017; Gul & Pesendorfer 2008; Hands 2013; Clarke 2016; Okasha 2016;

Dietrich and List 2016). Chapter 2 thus evaluates the relevance of the mentalism-behaviorism dichotomy in economics in light of these debates. In particular, I argue that there are two problems with current conceptions of the dichotomy as it pertains to how economists and decision researchers interpret and gather evidence. First, it is unclear what the dichotomy pertains to or is about exactly—which is to say, economists and decision researchers may have different motivations for endorsing mentalism and/or for opposing behaviorism. Second, and more importantly, it is unclear how the mentalism-behaviorism dichotomy is supposed to improve or advance empirical research in economics—in particular, supporters of mentalism have the difficult task of clarifying what mentalism entails or consists in. In response to the first problem, I consider two common motivations for endorsing mentalism: one motivation appeals to the choice-theoretic foundations of economics; the other appeals to scientific practice in economics. In response to the second problem, I argue that the mentalism-behaviorism dichotomy is not likely to advance or improve scientific practice in contemporary economic settings because neither mentalism nor behaviorism are equipped to analyze and resolve explanatory problems that are specific to the interpretation of psychological and neuroscientific (i.e. non-choice) data. I conclude by discussing the limitations of functionalism, the mainstay of the mentalism defense book, and suggest alternative schemas to the mentalism-behaviorism dichotomy, some of which are employed in neighboring areas of the cognitive and behavioral sciences.

Chapter 3 considers whether, i.e. under what conditions, human persons behave like economic agents. It is now recognized that humans are boundedly rational, which means that persons typically do not think and behave like *homo economicus* agents. Among the methodologies for modeling boundedly rational individuals, multiple-self models have gained considerable popularity as tools for representing the dynamics of intrapersonal choice under various conditions and constraints. Multiple-self models typically work by isolating features endogenous to individuals that motivate them to act in different ways. Generally, these features are taken to correspond to autonomous structures *within* the individual and, as such, are modeled as if they were independent agents (that is, independent agents who can reason together). In contrast to debates discussed in chapter 2, which take an individualistic perspective toward the analysis economic concepts and decision phenomena, this

chapter demonstrates how external forces such as social institutions and informational structures both support and constrain individual behaviors. I argue that individualism is problematic as a basis for investigating social interaction. In so doing I examine the Don Ross' (2005, 2006) account of multiple-selves as a way of reconciling individuals' bounded rationality with their bounded individuality. Ross argues that individual persons are complex aggregations of selves, which arise in response to external pressures to regulate individual behaviors—this enables the tracking of public norms and conventions. I thus investigate the different roles that 'selves' play in Ross' broader philosophy of economics and I identify separate projects that arise therein. I distinguish three different roles for selves, which are *evolutionary*, *narrative*, and *economic*, and I argue that these roles contribute to two distinct, but overlapping, projects. I show that, while it is not problematic to conceive of selves according to their different roles, we should not presume that the functions or properties of selves in one role can serve the same purposes for different projects.

Chapter 4 returns to the domain of individual decision-making—it asks: How do interdisciplinary approaches to decision research integrate psychological insights with economic methods? And, what are the conceptual and ontological challenges of such integration? The idea of the 'divided self' has been the source of folk-wisdom for centuries. However, new research into the cognitive and behavioral foundations of decision-making suggests that this idea is more than just a metaphor. Our minds—and brains—appear to be divided in interesting if unexpected ways. Having elaborated the importance of external forces for understanding humans' quasi-economic agency, Chapter 4 critically examines how multiple-self models of intrapersonal and intertemporal choice have been integrated with dual-process and dual-system theories from cognitive psychology. I adopt the term 'multiple-agent model' to denote a special kind of interdisciplinary model which conceives of multiple agents with contrasting psychological abilities. For example, Bénabou and Tirole (2002), Bernheim and Rangel (2004), Benhabib and Bisin (2005), Loewenstein and O'Donoghue (2005), and Fudenberg and Levine (2006) each have sought to characterize the contradictory tendencies of temporally distinct selves by investigating how controlled and automatic processes influence choice behaviors over time. In some instances, the intrapersonal dynamic between sequential selves is taken to establish the limitations on the decision-maker's ability to exhibit

self-control (Bénabou and Tirole 2002; Fudenberg and Levine 2006). In other instances, the conflict between an individual's desire to consume now or later is interpreted as a trade-off between distinct systems, whose aims are regulated by the activation of different cognitive processes (Benhabib and Bisin 2005; Loewenstein and O'Donoghue 2005). Likewise, some neuroeconomic approaches to decision-making have modeled brain processes based on what economists interpret to be optimizing procedures. Research conducted by McClure et al. (2004) and McClure et al. (2007), and further results obtained by Brocas and Carrillo (2008a, 2008b, 2014), suggest that individual decisions are the outcome of strategic interactions between domain-specific systems. This technique presumes that the brain has limited energy resources and that it must allocate those resources efficiently in order to satisfy rewards. In this way, the brain is modeled as an optimizer with budget constraints. These models seem to be growing in popularity given their purported ability to predict and explain reasoning errors and decision anomalies due to internal conflict or lack of self-control. In chapter 4, I analyze how multiple-agent models employ the concepts 'selves' and 'systems' for the purposes of representing intrapersonal and intraneural conflict. Herein I defend three claims. The first and second claims argue that multiple-agent models are conceptually as well as ontologically ambiguous. The third claim argues that such ambiguities can lead to problems in scientific understanding. The examination of multiple-agent models is not only important for understanding how economists and psychologists jointly interpret and model self-control problems, but it further presents an important opportunity to study the effects of cross-disciplinary pollination of concepts and theories.

Chapter 5 builds on the analytical framework developed in Chapter 4 and investigates the role(s) that dual process theory has played as a psychological framework in behavioral economics and neuroeconomics. Cognitive scientists and philosophical psychologists alike have criticized the theoretical foundations of the standard view of dual process theory and have argued against the validity and relevance of evidence used to support it (Gigerenzer and Reiger 1996; Osman 2004; Keren and Schul 2009; Kruglanski and Gigerenzer 2011). Moreover, recent modifications of dual process theory in light of these criticisms have generated additional concerns regarding its applicability and irrefutability (Keren 2013; Mugg 2016; Pennycook 2017; Bonnefon 2018). I argue that this should raise concerns for behavioral economists and neuroeconomists

who see dual process theory as providing psychologically realistic foundations for their models. In particular, it raises the possibility that dualistic models are not as descriptively accurate or reliable as economists presume them to be. In fact, the case can be made that the popularity of dual process theory in economic decision research has less to do with the empirical success of dualistic models, and more to do with the convenience that the dualistic narrative provides economists looking to sort out decision anomalies (cf. Grüne-Yanoff 2017). I argue that the growing number of criticisms of dual process theory leaves economists with something of a dilemma: either they stick to their purported ambitions to give a realistic description of human decision-making, or they stick to dual process theory and revise their scientific ambitions.

In Chapter 6, I offer concluding remarks and consider where one goes from here. Chapters 2-5 project two main approaches to reconciling the tension between agency and choice. One approach views individual persons as the primary objects of study for economics, and as such, indicates that psychology and neuroscience can help improve how economists study rational choice. The second approach views individual persons not as the primary object of study (economic agents are the primary object of study, and they are ontologically distinct from persons). As such, choice should be construed as the outcome of external (market) pressures, which include important socio-cognitive supports. Hence, for each of these approaches, there are new pursuits and new philosophical questions to be considered.

REFERENCES

- Backhouse, Roger E., and Steve G. Medema. 2009. "Defining Economics: The Long Road to Acceptance of the Robbins Definition." *Economica* 76 (S1): 805-820.
- Bénabou, Roland, and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." *The Quarterly Journal of Economics* 117 (3): 871-915.
- Benhabib, Jess, and Alberto Bisin. 2005. "Modeling Internal Commitment Mechanisms and Self-Control: A Neuroeconomics Approach to Consumption-Saving Decisions." *Games and Economic Behavior* 52 (2): 460-492.
- Bernheim, B. Douglas, and Antonio Rangel. 2004. "Addiction and Cue-Triggered Decision Processes." *The American Economic Review* 94 (5): 1558-1590.
- Bonnefon, Jean-François. 2018. "The Pros and Cons of Identifying Critical Thinking with System 2 Processing." *Topoi* 37 (1): 113-119.
- Brocas, Isabelle, and Juan D. Carrillo. 2008a. "The Brain as a Hierarchical Organization." *The American Economic Review* 98 (4): 1312-1346.

- Brocas, Isabelle, and Juan D. Carrillo. 2008b. "Theories of the Mind." *The American Economic Review* 98 (2): 175-180.
- Brocas, Isabelle, and Juan D. Carrillo. 2014. "Dual-Process Theories of Decision-Making: A Selective Survey." *Journal of Economic Psychology* 41: 45-54.
- Bruni, Luigino, and Robert Sugden. 2007. "The Road Not Taken: How Psychology Was Removed from Economics, and How It Might Be Brought Back." *The Economic Journal* 117 (516): 146-173.
- Camerer, Colin. 2008. "The Case for Mindful Economics." In *The Foundations of Positive and Normative Economics: A Handbook*, edited by Andrew Caplin, and Andrew Schotter, 43-69. New York, NY: Oxford University Press.
- Clarke, Christopher. 2016. "Preferences and Positivist Methodology in Economics." *Philosophy of Science* 83: 192-212.
- Davis, John B. 2006. "The Turn in Economics: Neoclassical Dominance to Mainstream Pluralism?" *Journal of Institutional Economics* 2 (1): 1-20.
- Dowding, Keith. 2002. "Revealed Preference and External Reference." *Rationality and Society* 14 (3): 259-284.
- Dietrich, Franz, and Christian List. 2016. "Mentalism Versus Behaviourism in Economics: A Philosophy-of-Science Perspective." *Economics & Philosophy* 32 (2): 249-281.
- Fudenberg, Drew, and David K. Levine. 2006. "A Dual-Self Model of Impulse Control." *The American Economic Review* 96 (5): 1449-1476.
- Gigerenzer, Gerd, and Terry Regier. 1996. "How Do We Tell an Association from a Rule? Comment on Slovic (1996)." *Psychological Bulletin* 119 (1): 23-26.
- Grüne-Yanoff, Till. 2017. "Reflections on the 2017 Nobel Memorial Prize Awarded to Richard Thaler." *Erasmus Journal for Philosophy and Economics* 10 (2): 61-75.
- Guala, Francesco. 2012. "Are Preferences for Real? Choice Theory, Folk Psychology, and the Hard Case for Commonsensible Realism." In *Economics for Real: Uskali Mäki and the Place of Truth in Economics*, edited by Aki Lehtinen, Jaakko Kuorikoski, and Petri Ylikoski, 137-155. Abingdon: Routledge.
- Guala, Francesco. 2017. "Preferences: Neither Behavioural nor Mental." *Economics & Philosophy* 1-19.
- Gul, Faruk, and Pesendorfer, Wolfgang. 2008. "The Case for Mindless Economics." In *The Foundations of Positive and Normative Economics: A Handbook*, edited by Andrew Caplin, and Andrew Schotter, 3-42. New York, NY: Oxford University Press.
- Hands, D. Wade. 2009. "Economics, Psychology and the History of Consumer Choice Theory." *Cambridge Journal of Economics* 34 (4): 633-648.
- Hands, D. Wade. 2013. "Foundations of Contemporary Revealed Preference Theory." *Erkenntnis* 78: 1081-1108.
- Hands, D. Wade. 2014. "Paul Samuelson and Revealed Preference Theory." *History of Political Economy* 46: 85-116.
- Hausman, D. M. (1998). Problems with realism in economics. *Economics & Philosophy*, 14(2), 185-213.
- Hausman, Daniel. M. 2008. "Mindless or Mindful Economics: A Methodological Evaluation." In *The Foundations of Positive and Normative Economics: A Handbook*, edited by Andrew Caplin, and Andrew Schotter, 125-151. New York, NY: Oxford University Press.

- Hausman, Daniel. M. 2012. *Preference, Value, Choice and Welfare*. New York, NY: Cambridge University Press.
- Keren, Gideon. 2013. "A Tale of Two Systems: A Scientific Advance or a Theoretical Stone Soup? Commentary on Evans & Stanovich (2013)." *Perspectives on Psychological Science* 8 (3): 257-262.
- Keren, Gideon, and Yaacov Schul. 2009. "Two is Not Always Better Than One: A Critical Evaluation of Two-System Theories." *Perspectives on Psychological Science* 4 (6): 533-550.
- Kruglanski, Arie W., and Gerd Gigerenzer. 2011. "Intuitive and Deliberative Judgments are Based on Common Principles." *Psychological Review* 118 (1): 97-109.
- Loewenstein, George, and Ted O'Donoghue. 2005. "Animal Spirits: Affective and Deliberative Processes in Economic Behavior." CMU Working Paper.
- McClure, Samuel M., David, I. Laibson, George Loewenstein, and Jonathan D. Cohen. 2004. "Separate Neural Systems Value Immediate and Delayed Monetary Rewards." *Science* 306 (5695): 503-507.
- McClure, Samuel M., Keith M. Ericson, David I. Laibson, George Loewenstein, and Jonathan D. Cohen. 2007. "Time Discounting for Primary Rewards." *Journal of Neuroscience* 27 (21): 5796-5804.
- Mugg, Joshua. 2016. "The Dual-Process Turn: How Recent Defenses of Dual-Process Theories of Reasoning Fail." *Philosophical Psychology* 29 (2): 300-309.
- Okasha, Samir. 2016. "On the Interpretation of Decision Theory." *Economics & Philosophy* 32 (3): 409-433.
- Osman, Magda. 2004. "An Evaluation of Dual-Process Theories of Reasoning." *Psychonomic Bulletin & Review* 11 (6): 988-1010.
- Pennycook, Gordon. 2017. "A Perspective on the Theoretical Foundation of Dual Process Models." In *Dual Process Theory 2.0*, edited by Wim De Neys, 13-35. London: Routledge.
- Ross, Don. 2005. *Economic Theory and Cognitive Science: Microexplanation*. Cambridge, MA: MIT Press.
- Ross, Don. 2006. "The Economic and Evolutionary Basis of Selves." *Cognitive Systems Research* 7 (2-3): 246-258.
- Ross, Don. 2011. "Estranged Parents and a Schizophrenic Child: Choice in Economics, Psychology and Neuroeconomics." *Journal of Economic Methodology* 18 (3): 217-231.
- Ross, Don. 2014. "Psychological Versus Economic Models of Bounded Rationality." *Journal of Economic Methodology* 21 (4): 411-427.

James Grayot completed a PhD and research master in philosophy at the Erasmus Institute for Philosophy & Economics (EIPE) at the Erasmus School of Philosophy. He does research on philosophy of science, working on issues at the intersection of economics, psychology, and cognitive (neuro)science. He also works on philosophy of mind, philosophy of action, and decision theory. He is presently a visiting fellow at the Tilburg Center for Logic, Ethics, and Philosophy of Science

(TiLPS), and a lecturer at the Vrije University Amsterdam, and the University of Groningen.

Contact e-mail: <james.grayot@gmail.com>