

Virtues and Economics

Peter Róna
László Zsolnai *Editors*

Agency and Causal Explanation in Economics

OPEN

 Springer

www.dbooks.org

Virtues and Economics

Volume 5

Series Editors

Peter Róna, University of Oxford

László Zsolnai, Corvinus University of Budapest

Editorial Advisory Board

Helen Alford, Pontifical University of St. Thomas Aquinas (“Angelicum”), Rome, Italy

Luk Bouckaert, Catholic University of Leuven, Belgium

Luigino Bruni, LUMSA University, Rome and Sophia University Institute, Loppiano

Georges Enderle, University of Notre Dame, USA

Carlos Hoevel, Catholic University of Argentina, Buenos Aires, Argentina

John Loughlin, Cardiff University, Emeritus Professor, Wales, UK

David W. Miller, Princeton University, USA

Sanjoy Mukherjee, Rajiv Gandhi Indian Institute of Management, Shillong, India

Mike Thompson, GoodBrand, London, CEIBS Shanghai, and University of Victoria, Vancouver, Canada

Johan Verstraeten, Catholic University of Leuven, Belgium

Stefano Zamagni, University of Bologna, and Johns Hopkins University – SAIS Europe and Pontifical Academy of Social Sciences, Italy

The series is dedicated to virtue ethics and economics. Its purpose is to relocate economic theory to a domain where the connection between the virtues and economic decisions, as that connection is actually experienced in everyday life, is an organic component of theory rather than some sort of an optionally added ingredient. The goal is to help develop a virtue-based economic theory which connects virtues with the contents of economic activities of individuals, unincorporated and incorporated economic agents. The primary context is Catholic Social Teaching but other faith traditions (especially Judaism, Islam, Hinduism, Buddhism, and Confucianism) will also be explored for their construction of virtues in economic action. Special attention will be made to regulatory and policy issues in promoting economic justice.

The series connects virtue ethics with the core of economic theory and practice. It examines the basic and irreducible intentionality of human activities concerned with the production, distribution and consumption of goods and services. It considers the incommensurability of values as the central problem of economic decision making and examines whether that problem can be overcome by any means other than practical reason.

This series will cover high quality edited volumes and monographs.

More information about this series at <http://www.springer.com/series/15627>

Peter Róna • László Zsolnai
Editors

Agency and Causal Explanation in Economics

 Springer

OPEN

Editors

Peter Róna
Blackfriars Hall
University of Oxford
Oxford, UK

László Zsolnai
Business Ethics Center
Corvinus University of Budapest
Budapest, Hungary



ISSN 2520-1794

ISSN 2520-1808 (electronic)

Virtues and Economics

ISBN 978-3-030-26113-9

ISBN 978-3-030-26114-6 (eBook)

<https://doi.org/10.1007/978-3-030-26114-6>

© The Editor(s) (if applicable) and The Author(s) 2020. This book is an open access publication.

Open Access This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The editors dedicate this volume to the memory of Prof. Paul Clough who tragically died a few days after submitting the final draft his manuscript for this volume.

Preface and Acknowledgement

The papers contained in this volume were delivered at Blackfriars Hall on 6 and 7 July 2018 in a symposium organized under the auspices of the Las Casas Institute. This was the fourth of a series of annual symposia held at Blackfriars with Economics as a Moral Science as the general title of the series, organized to explore the philosophical foundations of mainstream economics. The papers of the previous symposia and from workshops organized with the European SPES Institute in Cambridge, Oxford and Leuven have been edited by Peter Rona and Laszlo Zsolnai and published by Springer in two previous volumes, the first one, titled *Economics as a Moral Science*, published in 2017, while the second, titled *Economic Objects and the Objects of Economics*, appeared in 2018. The symposia at Blackfriars Hall grew out of a series of public lectures delivered in 2012–2014 by Peter Rona.

Blackfriars Hall, the Las Casas Institute and the editors of this volume would like to express their gratitude to the Mallinckrodt Foundation for the generous financial support of the symposia.

Oxford, UK

Peter Rona

Introduction

It is quite difficult to make sense of an event without having a notion as to why and how it happened. Indeed, we often have an anxious sense of doubt and uncertainty about something that we know has happened if we have no or only an inadequate idea of the circumstances bringing it about. As Elizabeth Anscombe recollected in the first two sentences of her introduction to Volume II of her collected papers,¹ ‘My first strenuous interest in philosophy was in the topic of causality. I didn’t know that what I was interested in belonged to philosophy’. Causality and – as some of the papers in this volume argue – agency are with us even when we are not aware of it, so much so that the questions of the ‘why’ and the ‘how’ not only affect what we know but also are quite fundamental to judgements; no system of morality, no ethical norm can do without them, and even aesthetics cannot lack some conception of the agent. Causation and agency, therefore, affect and permeate all of philosophy ranging from metaphysics through epistemology and ethics all the way to aesthetics.

Causal and agency questions are fundamental to all branches of the social sciences as well, and the failure to thoroughly explore them, to specify their role in the theory or model being defended, lies behind many of the disappointments the social sciences, particularly economics, have suffered. The unfulfilled aspiration of the latter to keep pace with the successes of the natural sciences has been regularly noted, at least since the birth of rationalist thought. Kant, for example, in a footnote to the introductory chapter to his *Critique of Pure Reason*² objects to the complaints about the ‘shallowness of the present age, and the decay of profound science’ but acknowledges that there is a problem with the social sciences:

...I do not think that those which rest upon solid foundation, such as Mathematics, Physical Science, etc. in the least deserve this reproach, but that they rather retain their ancient fame, and in the latter case, indeed, far surpass it. The same would be the case with the other kinds of cognition, if their principles were but firmly established.

¹ Anscombe, G. E. M. (1981) *Metaphysics and the Philosophy of Mind*, Basil Blackwell, Oxford.

² Kant, I. (1781) Preface to the First Edition, included in the Dover edition, Kant, I. (2003) *Critique of Pure Reason* translated by J. M. D. Meiklejohn and published by the Colonial Press in 1900. p. ix.

The distinction between mathematics and the physical sciences (sometimes lumped together under the ‘exact sciences’ label) and ‘the other kinds of cognition’ on the other runs deep and features throughout much of the history of philosophy, and most thinkers analysed the problem, like Kant, in the context of cognition.

The implication of the last clause of the quote above – namely, that the principles of these other types of cognition *could* be placed on just as firm a footing as mathematics and physical science – is a striking feature of the Enlightenment. It has stayed with us ever since and has conditioned the development of the social sciences, obscuring the great Augustinian tradition which saw the will as a distinct part of cognition, a faculty of the mind with its own properties and propensities. In a previous volume in this series,³ we defended the view that the question is not a matter of cognition but a matter of ontology. Economic events and processes are ontologically different from the objects studied by the natural sciences because they are the product of human will: their better understanding is not obtained by the increasingly rigorous application of mathematical techniques but, rather, by a deeper understanding of their ontology.

The modern commitment to the unicity of reality has necessarily conflated causation and agency and has given the social sciences a determinist colouration that is at odds with both experience and traditional notions of free will. In its extreme modern forms, such as rational choice theory and game theory, the modelled construction of reason has been posited as the sole source of social action notwithstanding the obvious lack of either predictive or explanatory powers of these theories. The papers included in this volume challenge in various ways the exclusive sovereignty of mathematically ordered reason in the dynamics of social life and also question the scientism implicit in the conflation of causation with agency.

Of course, the difficulties in economic theory described in these papers have been known at least since the beginning of the twentieth century.⁴ What is relatively new is the realization that Walrasian economic theory is not only mistaken but also harmful. In separating what it deems to be endogenous from the exogenous, in specifying the terms of ‘rational’ choice and conduct, in constructing its protagonist, the *homo economicus* economic theory also legislates the rules for authorized human conduct, and it also generates its system of rule-based roles. Unlike the hypotheses and laws of the natural sciences, economic ‘laws’ condition human conduct that, but for such laws, would be different. The pursuit of utility maximization replaces the pursuit of virtue, and the inherent normativity of this displacement takes place under the guise of the claim to scientific objectivity. These papers examine, from various perspectives, the foundations of this claim.

Blackfriars Hall
University of Oxford, Oxford, UK

Peter Rona

³Róna, P. and Zsolnai, L. (eds.) (2018) *Economic Objects and the Objects of Economics*, Springer.

⁴See Stephen Pratten’s paper below.

Contents

Part I Theory

1	Free Will & Empirical Arguments for Epiphenomenalism.	3
	Nadine Elzein	
2	Causality, Agency and Change.	21
	Stephen Pratten	
3	How Economics Becomes Ideology: The Uses and Abuses of Rational Choice Theory	37
	Jason Blakely	
4	Economics, Agency, and Causal Explanation.	53
	William Child	
5	Causation and Agency	69
	Peter Róna	

Part II Praxis

6	Why Aquinas Would Agree That Human Economic Behaviour Is Largely Predictable	93
	Richard Conrad and Peter Hunter	
7	Agency, Time and Morality: An Argument from Social and Economic Anthropology	115
	Paul Clough	
8	The Switch from Agency to Causation in Marx.	125
	Scott Meikle	

9	The Morphogenetic Approach; Critical Realism's Explanatory Framework Approach	137
	Margaret Scotford Archer	
10	'God Created Man αὐτεξούσιον': Grotius's Theological Anthropology and Modern Contract Doctrine	151
	Jonathan Price	
	Index.....	169

Editors and Contributors

About the Editors

Peter Róna is fellow of Blackfriars Hall, University of Oxford, where he teaches courses in economics and lectures on the philosophical foundations of the social sciences. He obtained his BA degree in Economic History from the University of Pennsylvania in 1964 and his law degree from the University of Oxford in 1966. He was an associate of the Washington, DC, law firm, Arnold & Porter, and legal counsel to the US Department of Commerce before becoming the personal assistant of Lord Richardson, the governor of the Bank of England. He joined Schrodgers PLC in 1969 as the general counsel of its operations in the United States and became the president and chief executive officer of the IBJ Schroder Bank & Trust Co. in 1985. In 2003, he joined the faculty of Eötvös Loránd University in Budapest and in 2006 was made honorary professor.

László Zsolnai is professor and director of the Business Ethics Center at Corvinus University of Budapest. He is president of the European SPES Institute in Leuven, Belgium, and co-chair of the Future Earth Finance and Economics Knowledge Action Network in Montreal. His recent books include *The Palgrave Handbook of Spirituality and Business* (2011), *Beyond Self: Ethical and Spiritual Dimensions of Economics* (2014), *The Spiritual Dimension of Business Ethics and Sustainability Management* (2015), *Post-Materialist Business: Spiritual Value-Oriented in Renewing Management* (2016) and, most recently, *The Routledge International Handbook of Spirituality in Society and the Professions* (2019).

Contributors

Margaret Scotford Archer is professor of Sociology at the University of Warwick and at the College of Humanities, Ecole Polytechnique Federale de Lausanne. Her theoretical work for much of the past 25 years has been devoted to the problem of structure and agency. She is the author of *Culture and Agency* (1988), *Realist Social Theory* (1995), *Being Human: the Problem of Agency* (2000), *Explanation and Understanding* (2010) and *Reflexivity* (2010). She is currently serving as president of the Pontifical Academy of Social Sciences.

Jason Blakely is assistant professor of Political Science at Pepperdine University. He is the author of *Alasdair MacIntyre, Charles Taylor and the Demise of Naturalism: Reunifying Political Theory and Social Science* (University of Notre Dame Press, 2016) and *Interpretive Social Science: An Anti-naturalist Approach* (with Mark Bevir, Oxford University Press, 2018). He has written on politics in academic journals like *The Review of Politics*, *Theory & Event*, *Political Studies*, *Polity*, and *Interpretation* as well as in popular venues like *The Atlantic*, *America* and *The Washington Post*.

William Child is professor of Philosophy in the University of Oxford and fellow and tutor in Philosophy at the University College, University of Oxford. He has specialized in the field of Philosophy of Mind and has written numerous studies on various aspects of the philosophy of Ludwig Wittgenstein. His books include *Causation and Interpretation: Some Questions in the Philosophy of Mind* (1989), *Causality, Interpretation, and the Mind* (1994), *Wittgenstein Themes: Essays in Honour of David Pears* (ed.) (2001) and *Wittgenstein* (2002).

Paul Clough was professor of Anthropology at the University of Malta. He has conducted anthropological field studies for over 30 years among the Hausa people of West Africa. He has been especially interested in Economic Anthropology and Political Economy. He edited (with Dr. Jon Mitchell) *Powers of Good and Evil: Moralities, Commodities, and Popular Belief* (2001). He was the chief editor of the *Journal of Mediterranean Studies*. His most recent book is *Morality and Economic Growth in West Africa: Indigenous Accumulation in Hausaland* (2019, Berghahn Books).

Richard Conrad fellow of Blackfriars Hall, obtained his Ph.D. in Chemistry from Cambridge University and his license in dogmatic theology from the Angelicum, Rome. He is the author of *The Catholic Faith: A Dominican's Vision* (1994), *The Seven Gifts of the Holy Spirit: Every Spiritual Warrior's Guide to God's Invincible Gifts* (2009), 'St. Thomas on Christ's Passion and Resurrection as Efficient and Exemplar Causes' *Angelicum* (2014) and 'Human Practice and God's Making-Good in Aquinas' *Virtue Ethics* in Carr, Arthur and Kristjanssons (Eds.) *Varieties of Virtue Ethics* (2017).

Nadine Elzein is faculty member at Lady Margaret Hall, University of Oxford. She specializes in the Philosophy of Mind focusing on free will, determinism and moral responsibility having published studies such as ‘Supervenient Freedom and the Free Will Deadlock’, *Disputatio*, 9(45); Freedom of Action, *Encyclopedia of the Mind*, 2013; Review of Vihvelin, Causes, Laws and Free Will, *Mind*, 2015; and Conflicting Reasons and Freedom of the Will, *Proceedings of the Aristotelian Society*, Jan. 1, 2010 as well as numerous articles on ethics.

Peter Hunter is lecturer and tutor in Philosophy at Blackfriars Hall, University of Oxford. He studied mathematics in Cambridge University and philosophy and theology at the University of Oxford before obtaining his doctorate in philosophy from King’s College, University of London. He is the coauthor (with Scott Steinkercner) of *Pictures of the World: Three Views of Life, the Universe, and Everything*.

Scott Meikle is honorary research fellow at the University of Glasgow. He is the author of *Aristotle’s Economic Thought* (1995), *Essentialism in the Thought of Karl Marx* (1985), *Marx: International Library of Critical Essays on the History of Philosophy* Vol 13 (ed.) and numerous articles and studies on Karl Marx.

Stephen Pratten is senior lecturer in Economics and Philosophy at King’s College, London. His primary research area is the methodology of economics, focusing particularly on issues related to social ontology. He is coeditor of the *Cambridge Journal of Economics*. His published work includes *Dewey on Habit, Character, Order and Reform* (2015), *Cambridge Social Ontology: Clarification, Development and Deployment* (2017), *Social Ontology and Modern Economics* (2015).

Jonathan Price is junior research fellow at the Aquinas Institute, Blackfriars Hall, University of Oxford; assistant professor, Institute of Philosophy, University of Warsaw; and fellow of the Center for the Thought of John Paul II, Warsaw. He is the editor of the peer-reviewed journal *Politics & Poetics*.

Part I

Theory

Chapter 1

Free Will & Empirical Arguments for Epiphenomenalism



Nadine Elzein

Abstract While philosophers have worried about mental causation for centuries, worries about the causal relevance of conscious phenomena are also increasingly featuring in neuroscientific literature. Neuroscientists have regarded the threat of epiphenomenalism as interesting primarily because they have supposed that it entails free will scepticism. However, the steps that get us from a premise about the causal irrelevance of conscious phenomena to a conclusion about free will are not entirely clear. In fact, if we examine popular philosophical accounts of free will, we find, for the most part, nothing to suggest that free will is inconsistent with the presence of unconscious neural precursors to choices. It is only if we adopt highly non-naturalistic assumptions about the mind (e.g. if we embrace Cartesian dualism and locate free choice in the non-physical realm) that it seems plausible to suppose that the neuroscientific data generates a threat to free will.

1.1 Introduction

In philosophy, while concerns about mental causation span back centuries, the question of whether epiphenomenalism undermines free will is surprisingly underexplored. In contemporary literature, worries about mental causation tend to derive from the concern that if there is an adequate physical explanation for every event, this renders the mental causally superfluous (Malcolm 1968, Kim 1989, 1993, 1998, 2005; O'Connor and Churchill 2010). Philosophers have puzzled over the implications of this problem for the viability of non-reductive, emergentist, or dualist theories of mind, but there has been little connection between this dispute and the traditional free will problem.

In contrast, in the neurosciences, researchers typically suppose that any threat to the causal efficacy of the conscious mind is also a threat to free will. Research suggesting that our consciousness of choices occurs too late to causally influence them

N. Elzein (✉)
Lady Margaret Hall, Oxford, UK
e-mail: nadine.elzein@lmh.ox.ac.uk

(Libet et al. 1979; Libet 1982, 1985; Soon et al. 2008; Wegner 2002, 2004) has been taken to be interesting primarily *because* it's taken to have drastic implications for free will.

There are three possible explanations for this curious disparity:

1. The empirical case for epiphenomenalism reveals a distinct epiphenomenal threat; one that has more serious implications for freedom than any threat identified by philosophical arguments for epiphenomenalism.
2. Philosophers are oblivious to the threat that epiphenomenalism poses to free will.
3. Empirical researchers are mistaken in supposing that the epiphenomenalist threat they identify really has serious implications for free will.

I will argue that explanation (3) is right. It is not obvious that the empirical arguments for epiphenomenalism really *do* have serious implications for freedom and moral responsibility.

1.2 The Philosophical Worries

Puzzles about how mind and body causally interact have been discussed at least since Elisabeth of Bohemia's famous correspondence with Descartes in 1643 (Atherton 1994, pp. 11–21). Such puzzles led theorists, even close to Descartes' time, to seek alternatives to Cartesian interactionism, such as occasionalism (e.g. Malebranche 1674, 1997), parallelism (e.g. Leibniz 1695, 1989), and monism, both idealist (e.g. Berkeley 1710, 1982) and materialist (e.g. Hobbes 1651, 1994; Cavendish 1664, 2017). And at least since Hodgson (1880) some philosophers have embraced epiphenomenalism.

While contemporary philosophers have predominantly hoped (much like Hobbes and Cavendish in Descartes' own time) that we can make sense of mental causation by rejecting dualism in favour of physicalism, many have been pessimistic about the prospects of completely *reducing* the mental to the physical (Putnam 1967; Fodor 1974; Block and Fodor 1972; Pylyshyn 1984). And "causal exclusion" arguments purport to show that non-reductive physicalists (no less than dualists) may be stuck with problems of mental causation.

As Kim argues, if physics is "causally complete", every physical event has a sufficient causal explanation that appeals only to other physical events. This poses a problem: if we have a sufficient causal explanation of an event in physical terms, this seems to render mental phenomena causally redundant. Unless we are willing (implausibly) to posit constant overdetermination, the mental realm will turn out to be epiphenomenal (Kim 1989, 1993, 1998).

Non-reductive physicalists have supposed that every mental event is realised by a physical event, and hence that mental events may be causally effective on account of their physical underpinnings (e.g. Davidson 1970). But even if this does make

mental *events* causally effective, it does not make them causally effective *qua* their mental features (Stoutland 1980; Honderich 1982; Sosa 1984).

Suppose that event *A* is Zina's spitefully dropping a statue into a pond and event *B* is the statue sinking. Event *A* has certain characteristics, among them its spitefulness. But this looks irrelevant to any explanation of why *A* caused *B*. The spite with which Zina dropped the statue is not part of the reason why it sank. Similarly, if an event's physical realisers sufficiently explain why the event caused another event, the fact that these events also realise mental properties will seem irrelevant. While mental events seem to be realised by causally effective physical events, these events do not seem to be causally effective in virtue of their mental qualities.

These issues are hotly debated, and won't be further explored. Epiphenomenalism remains a highly controversial philosophical thesis.

1.3 The Neuroscientific Worries

Concerns about the irrelevance of the conscious mind in decision-making emerged with Libet's famous study, purporting to show that choices are initiated by unconscious neural events prior to agents becoming consciously aware of forming an intention to act (Libet et al. 1979; Libet 1982, 1985). Recent variations on Libet-style experiments (Soon et al. 2008) provide even stronger evidence of unconscious neural precursors to decisions, and it's often supposed that such research raises serious doubts about free will (e.g. Wegner 2002, 2004).

The alleged threat arises from the suggestion that our sense of conscious will is epiphenomenal. There are notably some serious problems, however, with reaching this conclusion on the basis of the data. Firstly, the neural precursors do not correlate perfectly with the decisions that follow them. Libet himself speculated that agents had some conscious power of veto right up until the last moment, and that this gives us reason to reject the epiphenomenalist conclusion (Libet 1985, 2003). But Libet's reasoning on this matter has garnered serious criticism (e.g. Gallagher 2006).

Studies have also focused on getting participants to make an arbitrary choice about when to press a button or which of two buttons to press, with no strong considerations present in favour of either action. This experimental situation is quite far removed from real life decisions. Perhaps we can delegate this sort of decision-making to unconscious processes more easily than decisions with real consequences, which may require more conscious engagement (Waller 2012).

Waller suggested the studies could be redesigned to test whether morally significant decisions would be similarly initiated unconsciously, by asking participants to make choices with significant outcomes. E.g. gambling with funds that are going to go to a charitable cause. This has since been tested experimentally (Maoz et al. forthcoming) and the results corroborate Waller's contention that greater conscious input may be present in the case of morally significant choices than is present in that of entirely arbitrary ones. This casts doubt on any inference from the traditional

experimental setup to conclusions about the epiphenomenal nature of consciousness more broadly.

Moreover, even in the case of arbitrary choice, it's not clear whether there may be causally relevant conscious activity correlated with the neural precursors to conscious awareness of the choice. If these neural precursors correlate with *other* conscious activity that typically precedes a choice – contemplating pressing the button, deliberating about pressing the button, etc. – it would be unsurprising that such events tended to both precede choices and influence the chances of the agent making a particular choice. This would not obviously show that the conscious mind was irrelevant to the choices subsequently made (Nahmias 2010; Baumeister et al. 2011).

It would be premature to conclude that consciousness has no causal role to play in initiating choices, given the present state of research. Though it is possible that as such research progresses, stronger neuroscientific evidence may emerge.

1.3.1 What Is “Conscious Will”?

Suppose we take this alleged epiphenomenal threat seriously. There is still some unclarity regarding how the experimental data allows us reach the conclusion that free will is illusory.

Firstly, the notion of “conscious will” is vague. The language of “conscious will” somewhat obscures the fact that willing is usually understood, at least in contemporary thought, to be a sort of mental state, and not a special faculty the agent possesses. The term “consciousness” also needs disambiguation. When we say that “consciousness” or “conscious will” is epiphenomenal, what do we mean by this?

There are numerous different ways in which mental states might be considered “conscious”. At least three categories look directly relevant:

1. *Phenomenal consciousness:*

For present purposes we may group together such things as *qualitative states*, *phenomenal states*, *raw feels*, and *what-it-is-like* states. There seem to be directly sensible qualities to some of our mental states, which provide the character of our subjective experiences. Common examples include sensations of emotion, colour, taste, or pain. As well as ‘qualia’ or ‘raw feels’ this category may include other dimensions of what an experience is like for its subject. It is these phenomenal aspects of consciousness that are sometimes thought to be difficult to capture in objective or physical terms (1958, Nagel 1974; Jackson 1982, 1986, Block 1990).

2. *Access consciousness:*

For present purposes, we may group together *access consciousness*, *informational consciousness*, and *mental states being personally available to awareness*. These categories involve being aware that we are subject to a mental state in such a way that we are able to report being in such a state, and able to take the state

into account in our thoughts and deliberation. It is not obvious that in order to have this sort of access to a state, there must be any particular character to the experience, so access consciousness is typically distinguished from phenomenal consciousness (Block 1995).

3. *Intentional/representational content:*

The philosophical literature, in contrast to neuroscientific literature, often focuses on the intentional or representational contents of mental states; what our beliefs and desires, say, are *about*. Epiphenomenalism is sometimes thought to render these contents causally irrelevant. This seems orthogonal to problems of consciousness understood in either access or phenomenal terms. Even unconscious states can have intentional content. And some philosophers suggest that phenomenal features of consciousness are entirely distinct from intentional ones (Peacocke 1983; Block 1996).

The idea of “conscious will” tends to be invoked frequently, but poorly defined. It is not obvious which sort of consciousness is supposed to (a) have to be causally efficacious in order for free will to be possible and (b) be rendered epiphenomenal by the data.

1.4 Epiphenomenalism and Freedom of the Will

Suppose that we have reason to take the threats to the causal efficacy of consciousness identified in the neuroscientific literature seriously. It is not immediately obvious what the implications are for free will. In order to assess these implications, we need some idea of what free will entails. Let’s take a brief look at some of the leading accounts of free will and at the sorts of requirements typically thought to be necessary preconditions of it.

As the free will dispute has typically focused on determinism rather than epiphenomenalism, the accounts tend to be divided up in terms of their compatibility or incompatibility with determinism; the thesis that the future is fixed by the laws of nature and the way things were in the past.

Popular compatibilist conditions of freedom include the ability to act on the basis of one’s choices (Moore 1903; Ayer 1954; Smart 1961, 1966; Lewis 1981; Berofsky 2002), the ability to respond to reasons (Fischer and Ravizza 1998; Wolf 1990), and the ability to make choices based on one’s deeper values (Frankfurt 1971, Dworkin 1970; Watson 1975).

Incompatibilists typically endorse similar conditions as necessary for free will, but deny that they are sufficient. They tend to also require either that agents are able, in a robust sense, to choose otherwise (Kane 2000, 2002, 2004; Moya 2006, 2007, 2011; Ekstrom 2003; Elzein 2017; Franklin 2018; Kittle 2018) or else that agents are, in some sense, the “ultimate sources” of their own choices (Stump 1999a, b, 2003; Pereboom 2000, 2001, 2003; Zagzebski 2000, 2010; Timpe 2007, 2008;

Shabo 2010; Widerker 2006, 2009). Meeting these conditions is typically thought to require the falsity of determinism.

Note that determinism is unrelated to epiphenomenalism. The question of whether some account of freedom is compatible with determinism has no implications at all for the question of whether it is compatible with epiphenomenalism. While it's surprisingly common to find the two issues conflated (Nahmias (2010) cites a number of instances of this error), the two theses are unconnected.

How do we get from the premise that certain conscious features are epiphenomenal to a conclusion about free will? There must be something required for free will that would plausibly be precluded by the supposed epiphenomenal status of those conscious features. We should therefore look at commonly defended accounts of free will to see whether any of them include requirements that might plausibly be impossible to meet if these conscious features turn out to be epiphenomenal.

There are, however, some restrictions on what sorts of requirement would actually be fruitful here. Proponents of the empirical argument would be well advised to search for requirements which are not only plausibly necessary for free will, but which also meets the following criteria:

1. The requirement must be plausibly inconsistent with epiphenomenalism.
2. While the requirement must be a necessary condition of *free and responsible* action, it should not also be a prerequisite for *any* action (even unfree ones).
3. The requirement must be consistent with the thesis that the mental supervenes on the physical.

The reason for the first condition is obvious. Unless some requirement for freedom is also ruled out by epiphenomenalism, we will not be able to use it in conjunction with epiphenomenalism as a basis from which to infer free will scepticism.

The second requirement might look mysterious. If something is required for action, then it must, *a fortiori*, be required for *free* action. Hence if we can provide a necessary condition of action that would be ruled out by epiphenomenalism, we certainly have a good case for free will scepticism. This is true, but it means the argument is likely to prove too much. While scepticism about free will is commonly thought to be a surprising but nonetheless potentially compelling conclusion, scepticism about agency is far less plausible, and is likely to be seen as an unpalatable implication. Some may be willing to bite this bullet, but for others the implication that there are no agents or actions is more likely to be regarded as a *reductio* of any argument that led to it. Hence the argument will lose a good deal of plausibility if it can only proceed to a sceptical conclusion about free will via a much stronger sceptical conclusion about agency.

The final requirement, that it must be consistent with the claim that the mental supervenes on the physical, seems a basic prerequisite of our having any reason to take the neuroscientific literature seriously in the first place. The entire idea of studying the mind via an investigation of the brain would be misguided unless we were willing to grant this thesis.

The question is whether any of the standard accounts of free will posit requirements that meet these criteria. While the categories of "incompatibilist" and

“compatibilist” are useful in relation to the dispute about determinism, they are not so relevant to the threat from epiphenomenalism. For present purposes, I will divide purported preconditions of free will into those that are also preconditions for any action and those that are preconditions specifically for *free* action. The latter will also be divided into naturalistic requirements and non-naturalistic ones.

1.4.1 *Purported Conditions of Action*

Under what conditions does behaviour count as genuine action? At a minimum, it must be purposive or goal-directed. But that is not enough; the behaviour of a simple machine could be goal-directed. Plausibly, it must be *consciously* goal directed. The agent must have some intention in acting, and must be aware of it. This is what seems to be lacking in the case of automatistic action, like sleepwalking, which is not typically classed as genuine agency. While sleepwalkers seem to carry out intentions, they are not aware of what they are doing.

This certainly includes a lack of access consciousness; they are unable to report or further reflect on what they are doing. It may be plausible to suppose that agents who are dreaming enjoy phenomenal consciousness; they are, in some sense, aware of the phenomenal qualities presented to them in their dreams. But such agents certainly lack phenomenal consciousness with respect to what they are *actually* doing. The intentions with which they act are disconnected from the behaviour that’s actually occurring.

It seems plausible that agents need access consciousness in order to count as acting at all, and plausible, though more controversial, that agents *might* need phenomenal consciousness to count as acting at all (Shepherd 2015a). The latter will depend on whether a “phenomenal zombie” could still perform actions; a question which is not east to answer (Smithies 2012).

But it is one thing to suppose that agency requires consciousness and quite another to suppose that it requires our mental states to be causally efficacious in *virtue* of their conscious qualities, or that the conscious awareness must *temporally precede* the initiation of choices.

Consider automatic behaviour; the sort that’s performed on “autopilot”. Often agents lack awareness when performing actions that are highly rehearsed; but it’s typically thought to count as acting nonetheless. This is because they are easily able to *acquire* conscious awareness. If you are driving, you might not be aware of indicating to turn, but if someone asks you to explain yourself, you can easily *become* aware of what you are doing. In this case, it is not obvious that the awareness must *precede* the initiation of behaviour; rather, the behaviour and the intentions that motivate it must be easily *accessible* to conscious awareness (Levy and Bayne 2004; Levy 2011; Levy 2013, 2014a).

The reason why access consciousness seems important is because it seems to characterise the sorts of mental processes that might look especially relevant for our control over our behaviour. We tend to be able to consciously report the sorts of

processing that that directly pertain to our deliberate goals, and which might be amenable to rational reflection and scrutiny. In contrast, much of the activity that occurs below the level of awareness is outside of our *direct* control, but may nonetheless be broadly governed by these higher level aims.

Various authors have tried to elucidate the importance of the sorts of processing that tend to be available to conscious awareness in this way. Shepherd (2015b) argues that the sorts of processing that are typically available to conscious introspection are those involved in “executive functions”. The sorts of processes that are *not* available to conscious awareness are those that occur at a sub-personal level, and are not directly accessible to conscious reflection. These are typically subordinate to executive processing. E.g. our explicit intention to pick something up directs our motor processes. Relatedly, Gallagher (2006) notes that we may be expected to directly cognise features of our situation and environment, while subpersonal processes remain largely inaccessible at that level of cognition, despite being broadly directed by processes at that level. Others have noted that our proximal intentions, even if formed fairly automatically, may be directed by distal intentions that are typically available to conscious awareness (Nahmias 2010; Schlosser 2013).

Genuine action typically involves processes of which we are consciously aware, since these have an “executive” or “directing” role. But there is no obvious reason to suppose that the conscious awareness *itself* must be doing the causing. Conscious awareness is an indicator that the processes in question are the sort that characterise genuine agency. But this doesn’t entail that such processes must be *caused* by conscious awareness. And it certainly does not seem as if the “phenomenal feeling” needs to cause the behaviour (Walter 2014). It seems highly dubious to suppose that our intentions must be causally effective *in virtue* of either access consciousness or phenomenal consciousness, even if they must be available to either sort of conscious awareness.

A more plausible suggestion is that we need to act in virtue of the intentional or representational *contents* of our mental states in order to count as genuinely acting; that *what* I desire and believe must be causally relevant to my behaviour. But this is not the focus of any of the empirical literature that aims to undermine free will. It is only really addressed in the philosophical literature. Moreover, this threat, if taken seriously, is far more extensive than *just* a threat to the possibility of free will. It potentially entails that the entire mental realm is devoid of all influence. Few philosophers embrace this conclusion.

Even if agency is possible, however, *free* agency might not be. Epiphenomenal arguments might threaten our ability to act freely or be held morally accountable. Let’s consider broadly naturalistic conditions of freedom; those that do not posit phenomena beyond the reach of scientific investigation.

1.4.2 *Naturalistic Purported Conditions of Freedom*

1.4.2.1 Acting on the Basis of Choices

It is fairly universally supposed that moral responsibility requires the ability to act in accordance with our choices. If an agent is imprisoned or paralysed, or if there are constraints and impediments that hinder her ability to act as she intends to, this would undermine her freedom and her moral accountability (Moore 1903; Ayer 1954; Smart 1961; Lewis 1981; Berofsky 2002).

But epiphenomenal arguments pose no threat to this ability. This requirement says nothing about the way that our decisions are initiated or about the features in virtue of which they count as causally efficacious. It is usually understood simply in terms of counterfactual dependence; an agent meets this requirement if she acts as she has chosen to and would have acted otherwise if she had chosen to act otherwise. Libet-style studies purport to undermine the causal relevance of conscious awareness to agents' choices; they do not purport to undermine the causal relevance of agents' choices to their subsequent actions.

1.4.2.2 Reasons Responsiveness

It's often supposed that our decisions would need to be responsive to reasons in order for us to count as morally responsible (Wolf 1990; Fischer and Ravizza 1998). Schlosser (2013) and Levy (2011, 2013, 2014a, b) have argued that there is a crucial link between conscious awareness and the ability to respond to reasons. They both suppose that it is (broadly) access consciousness that is required for this ability, as opposed to phenomenal consciousness.

The reason why we might suppose that access consciousness is crucial to reasons-responsiveness is that only this sort of conscious processing seems to be governed by norms of consistency in such a way as to be potentially integrated into a rational outlook. Our unconscious processing tends to work in an associative way, and not to be governed by norms of consistency (Levy 2013, 2014a, b).

This gives us reason to suppose that we can only be morally responsible for processes that are accessible to conscious awareness. But it does not seem to entail that our choices must be *initiated* consciously, or that they must be causally efficacious *in virtue* of conscious features.

Automatic action is typically driven by the sorts of intention that can easily be brought to our conscious awareness, even if we often lack awareness when initiating the action (Levy and Bayne 2004; Levy 2011). There seems no reason to suppose that such behaviour is immune to introspection and rational scrutiny.

Moreover, the crucial point for reasons-responsiveness is not that conscious awareness *itself* must be doing the causal work, but that the sorts of processes *of which* we are consciously aware are the sorts that are governed by standards of rational consistency. Even if such processes were initiated by unconscious events,

this would not stop them from being governed by these norms, so it would not preclude them from being reasons-responsive.

1.4.2.3 Harmony with Deeper Values

Plausibly, if our choices are to count as free, they must be driven by desires we don't mind being moved by (Dworkin 1970). Perhaps they must harmonise with our second order volitions, i.e. we must be moved by the first order desires that we want to be moved by (Frankfurt 1971). Perhaps, ultimately, we need our choices to harmonise with our *deepest values* (Watson 1975). It might be thought that our deepest system of values, those with which we rationally identify, constitute the "real self".

There is a strong case for supposing that only conscious processes typically harmonise with our deeper values. Since a moral outlook needs to be integrated into a coherent system, it also needs to be governed by norms of consistency. Unconscious processes do not obey these norms (Levy 2011, 2013). Moreover, it's well known that agents often unconsciously process information in ways that run directly *counter* to their values; in cases of unconscious bias, agents typically find their behaviour reflects attitudes they consciously repudiate; they are often keen to rid themselves of the bias, or to take steps to prevent it from influencing their behaviour (Levy 2013; Levy 2014a, b).

Once again, it is typically argued that access consciousness rather than phenomenal consciousness is relevant here; the sorts of processes we can report and subject to scrutiny are the sorts we can expect to harmonise with our values. What matters is that such processing is governed by norms of rational consistency and may be subject to scrutiny on the basis of values. This is what enables it to harmonise with our values. This does not require the awareness *itself* to cause our decisions or to precede their initiation.

1.4.2.4 Alternative Possibilities

Alternative possibilities are often thought to be important for moral responsibility, though this has been highly controversial since Frankfurt's famous argument for their irrelevance (Frankfurt 1969).

Compatibilists traditionally understand this requirement in terms of a counterfactual dependence between an agent's choices and actions. We have already noted that this is untouched by epiphenomenalism. Incompatibilists, however, tend to have a different understanding of alternative possibilities. They typically require that agents are able to do otherwise holding that past the laws of nature constant. Determinism seems to preclude alternatives so understood.

As already noted, epiphenomenalism does not entail determinism. Perhaps, however, Libet-style studies should *also* be understood as providing evidence for the thesis that our choices are causally determined. Soon et al. (2008) were able to predict an agent's choice between which of two buttons to press (i.e. a decision with

a 50% chance of going either way) with 60% accuracy, well before the agents themselves became consciously aware of deciding. But this does not constitute compelling evidence for determinism; it does not entail that a single outcome is guaranteed. Proponents of indeterminism typically suppose that causation is probabilistic as opposed to deterministic. This is consistent with earlier events raising the probability of later ones.

1.4.3 *Non-Naturalistic Purported Conditions of Freedom*

The naturalistic conditions of freedom appear to be untouched by arguments for epiphenomenalism, but it is not so obvious that the same will be true of non-naturalistic ones.

Sometimes, it is thought that freedom requires Godlike abilities; that free will involves being an “unmoved mover”, able to originate actions independently of any prior events. There are two features that might seem to be preconditions for this; one involves the conscious mind being the ultimate causal source of choices, and the other involves immunity from prior causal influence.

1.4.3.1 *Conscious Origination*

Source incompatibilists typically suppose that determinism threatens free will because it threatens the agent’s status as the *ultimate source* of her own choices and actions. Do epiphenomenalist arguments challenge this? It depends on how we locate agents. If unconscious brain processes are regarded as part of the agent’s efforts, then it’s not obvious that the agent could only be initiating an action if conscious phenomena are initiating it. An agent’s mental processes could be doing the causal work even if the conscious awareness *itself* were not involved in the initiation of choices.

If consciousness is understood merely as a way of accessing our mental states, then this simply doesn’t look like the sort of thing that could intelligibly figure in a causal relation. It’s not obvious how the mere availability of states to conscious awareness *could* cause our choices. And with respect to phenomenal consciousness, again, it seems strange to suppose that the raw *feeling* of making a choice could cause it (Walter 2014).

If consciousness is going to intelligibly figure as an independent cause, it seems that we will need to understand it as something more than just a way in which mental states present themselves to awareness. Rather, we would need to suppose that conscious states are part of an independent *entity* that might exert its own influence.

Perhaps consciousness is taken to be the crucial element of a Cartesian soul. If we suppose that freedom requires the soul to influence action independently of neural events, then perhaps it *would* make sense to suppose that phenomenal

consciousness might be part of something causally influential in its own right. If this is what free will requires, then the empirical case for epiphenomenalism, by showing that consciousness arrives too late to the game, would indeed be showing that the *agent* (presumably, identical to the Cartesian soul and *not* to any neural precursors) could not be the ultimate source of her choices.

1.4.3.2 Immunity from Prior Influence

This requirement might be supplemented with another; perhaps what is required for freedom is not merely that our choices are caused by an independent conscious entity, but also that these choices are *immune* from prior causal influence.

Some incompatibilists suppose that free choices must be *uncaused* (Ginet 2002, 2007, 2008, 2016; McCann 1998, 2012; Goetz 1988, 2008). Presumably, this rules out even probabilistic causation. Does this entail that there can be no correlation between an agent's choices and neural events that precede those choices? This is not obvious. Non-causalists usually suppose that an agent's choices must be *rationaly* explicable, even if they are not *causally* explicable. But suppose that some pattern of neural activity is typically associated with positively assessing a potential course of action. This pattern would then also be correlated with the presence a positive reason to choose it. On the non-causal view, the agent's choice *also* needs to correlate with the presence of such reasons. We would therefore expect the choice and the neural pattern to be correlated to one another as well.

It is only if we presuppose an explicitly non-naturalistic account of the agent and her mental processes (e.g. we suppose that there are *no* correlates between neural processes and the agent's contemplation of reasons) that a non-causal account rules out *any* correlation between the odds of an agent making a particular decision and neural precursors.

Suppose we regard the *agent* as essentially identical to a Cartesian soul, and we suppose that an agent can only have free will insofar as any rational explanation of her behaviour emanates entirely *from* the soul independently of any brain processes. Libet-style studies certainly seem to cast some doubt on whether agents can meet this requirement.

1.5 Epiphenomenalism and Free Will Scepticism

Recall, I argued that, if we are to reach free will scepticism on the basis of evidence for epiphenomenalism, we would need to identify some requirement for free will that also meets the following criteria:

1. The requirement must be plausibly inconsistent with epiphenomenalism.
2. While the requirement must be a necessary condition of *free and responsible* action, it should not also be a prerequisite for *any* action (even unfree ones).

3. The requirement must be consistent with the thesis that the mental supervenes on the physical.

I examined three categories of purported requirements for freedom: conditions of basic agency, naturalistic conditions of free will, and non-naturalistic conditions of free will.

Obviously, none of those conditions that are alleged to be prerequisites for any action will be able to meet condition 2: An argument from *those* conditions is not only going to establish free will scepticism, but will *also* show that there is no agency of any sort. This would certainly undermine free will, but it comes with the implausible implication that actions do not exist either.

Moreover, we found that the sorts of consciousness that might be potentially rendered epiphenomenal by the empirical arguments were likely to be irrelevant to the conditions of action. While such arguments might show that *phenomenal* consciousness occurs after the initiation of a choice, it is not obvious that this timing issue is especially relevant to the role of access consciousness, and the empirical arguments do not tell us anything about the causal role of intentional or representational content.

It is only if the *latter* is shown to be epiphenomenal that we would plausibly have a serious threat to agency; while the philosophical arguments might be thought to render intentional content epiphenomenal, the empirical arguments do not appear to address it at all. The requirement that we be able to act in virtue of intentional contents, if it is a requirement of free will, does not meet condition 2 in any case, and in relation to the *sorts* of epiphenomenalism plausibly entailed by the empirical research, doesn't meet condition 1 either.

Let's turn to the naturalistic requirements for free will.

The possibility of meeting the purported naturalistic requirements for free will, I suggested is not threatened by epiphenomenalism, so these requirements do not meet criterion 1.

While our choices can only meet the requirements of being reasons-responsive and in harmony with our deeper values if those decisions involve processing of the sort that we *are* typically conscious of, there is no reason to suppose that any sort of conscious awareness must cause or even temporally precede such processing in order for it to be reasons responsive or to harmonise with our values. And the empirical studies have no bearing at all on whether our choices are causally determined.

Finally, the non-naturalistic purported requirements of freedom do not meet criterion 3. If these are understood so as to genuinely meet criterion 1 (i.e. as involving a causal role for consciousness in itself and/or devoid of any correlation with prior neural events), this would require us to identify agents with something non-physical; e.g. a ghostly soul, which is an unmoved mover and is the ultimate source of choices and action. This is inconsistent with the thesis that the mental supervenes on the physical.

Such studies arguably ought to matter to someone with a classic Cartesian interactionist view. Since phenomenal consciousness, on this view, resides in an independent non-physical mind, and this causally influences brain activity, perhaps

phenomenal awareness must *precede* the initiation of choices. But Cartesian interactionism is hardly popular nowadays. There are reasons to reject it quite independently of the empirical case for epiphenomenalism; many of these reasons were identified immediately after its inception in the seventeenth century.

As noted earlier, this led other historical theorists to favour monism (idealist and materialist) or to favour parallelist or occasionalist analyses of the relation between the mental and the physical. The problem is, it's not obvious that proponents of alternatives to interactionism who endorse highly non-naturalistic accounts of free will have any good reason to suppose that neuroscientific research tells us anything *at all* about the mind. These views place the mind thoroughly outside of the reach of scientific investigation.

Moreover, contemporary philosophers overwhelmingly endorse physicalism. Those who take seriously non-naturalistic conditions of freedom are rare. Those who endorse those conditions *alongside* Cartesian interactionism are rarer still (an endangered species, if not extinct).

Libet-style studies present no evidence for supposing that we cannot meet naturalistic conditions of freedom. And as for non-naturalistic ones, anyone who takes neuroscientific research seriously had overwhelming reason to suppose nobody could meet *those* quite independently of these studies. In contrast, those who maintain that we do meet non-naturalistic conditions, for the most part, must regard such things as outside of the reach of neuroscientific research.

1.6 Conclusion

The landscape of the free will dispute is largely unaffected by the empirical case for epiphenomenalism. Surprisingly, it seems that empirical researchers have been prone to presupposing a picture of freedom, from the start, that would be completely inconsistent with sort of minimal physicalism upon which the whole enterprise of neuroscience is based. Somehow, they are still haunted by the ghost of Cartesian interactionism.

Insofar as we endorse a broadly naturalistic and physicalist picture of the mind (a picture that I think we have overwhelmingly good reason to embrace), it is not obvious that the experimental data provides any serious challenge to meeting the conditions of freedom that could plausibly be met consistent with that picture anyway. In contrast, for anyone who explicitly rejects that picture, the empirical data will be regarded as being of dubious relevance from the start.

There are, I believe, compelling reasons on the basis of which to embrace free will scepticism. But these reasons have little to do with the presence of unconscious precursors to the decisions we make.

References

- Atherton, Margaret, ed. 1994. *Women Philosophers of the Early Modern Period*. Indianapolis: Hackett Publishing Company.
- Ayer, Alfred J. 1954. Freedom and necessity. In his *Philosophical Essays*. New York: St Martin's Press.
- Baumeister, Roy F., E.J. Masicampo, and Kathleen D. Vohs. 2011. Do conscious thoughts cause behavior? *Annual Review of Psychology* 62 (1): 331–361.
- Berkeley, George. 1710, 1982. Winkler (ed.). *A Treatise Concerning the Principles of Human Knowledge*. Indianapolis: Hackett Publishing.
- Berofsky, Bernard. 2002. Ifs, cans, and free will: the issues. In *The Oxford Handbook of Free Will*, ed. Robert Kane. Oxford: Oxford University Press.
- Block, Ned. 1990. Inverted Earth. *Philosophical Perspectives* 4: 53–79.
- . 1995. On a confusion about a function of consciousness. *Behavioral and Brain Sciences* 18 (2): 227–247.
- . 1996. Mental paint and mental latex. In Villanueva, ed. *Perception*. Atascadero: Ridgeview.
- Block, Ned, and Jerry A. Fodor. 1972. What psychological states are not. *Philosophical Review* 81: 159–181.
- Cavendish, Margaret. 1664, 2017. *Philosophical Letters Or Modest Reflections upon some Opinions in Natural Philosophy*. Charleston: CreateSpace Independent Publishing Platform.
- Davidson, Donald. 1970. Mental events. In *Experience and Theory*, ed. Foster and Swanson. Amherst: University of Massachusetts Press.
- Dworkin, Gerald. 1970. Acting freely. *Noûs* 4 (4): 367–383.
- Ekstrom, Laura. 2003. Free will, chance, and mystery. *Philosophical Studies* 2 (1): 153–180.
- Elzein, Nadine. 2017. Frankfurt-style counterexamples and the importance of alternative possibilities. *Acta Analytica* 32: 169–191.
- Fischer, John Martin, and Mark Ravizza. 1998. *Responsibility and Control*. Cambridge: Cambridge University Press.
- Fodor, Jerry A. 1974. Special sciences: Or the disunity of science as a working hypothesis. *Synthese* 28: 97–115.
- Frankfurt, Harry G. 1969. Alternate possibilities and moral responsibility. *The Journal of Philosophy* 66 (23): 829–839.
- . 1971. Freedom of the will and the concept of a Person. *The Journal of Philosophy* 68 (1): 5–20.
- Franklin, Christopher E. 2018. *A Minimalist Libertarianism: Free Will and the Promise of Reduction*. New York: Oxford University Press.
- Gallagher, Shaun. 2006. Where's the action? Epiphenomenalism and the problem of free will. In Pockett, Banks, & Gallagher (eds.) *Does Consciousness Cause Behavior?*, Cambridge, MA: MIT Press.
- Ginet, Carl. 2002. Reasons explanations of action: causalist versus noncausalist accounts. In *The Oxford Handbook of Free Will*. Kane (ed.) New York: Oxford University Press.
- . 2007. An action can be both uncaused and up to the agent. In *Intentionality, Deliberation and Autonomy: The Action-Theoretic Basis of Practical Philosophy*, ed. Lumer & Nannini. Aldershot: Ashgate.
- . 2008. In defense of a non-causal account of reasons explanations. *The Journal of Ethics* 12: 229–237.
- . 2016. Reasons explanation: Further defense of a non-causal account. *The Journal of Ethics* 20: 219–228.
- Goetz, Stewart. 1988. A noncausal theory of agency. *Philosophy and Phenomenological Research* 49: 303–316.
- . 2008. *Freedom, Teleology, and Evil*. London: Continuum.
- Hobbes, Thomas. 1651, 1994. *Leviathan*. In *Leviathan, with selected variants from the Latin edition of 1668*, ed. Curley. Indianapolis: Hackett.

- Hodgson, Shadworth H. 1880. *The Theory of Practice: An Ethical Enquiry in Two Books*. London: Longmans, Green, Reader, & Dyer.
- Honderich, Ted. 1982. The argument for anomalous monism. *Analysis* 42: 59–64.
- Jackson, Frank. 1982. Epiphenomenal qualia. *The Philosophical Quarterly* 32: 127–136.
- . 1986. What Mary didn't know. *Journal of Philosophy* 83: 291–295.
- Kane, Robert. 2000. The dual regress of free will and the role of alternative possibilities. *Philosophical Perspectives* 14: 57–79.
- . 2002. Some neglected pathways in the free will labyrinth. In *The Oxford handbook of free will*, ed. Kane. New York: Oxford University Press.
- . 2004. Agency, responsibility, and indeterminism: reflections on libertarian theories of free will. In *Freedom and Determinism*, ed. O'Rourke Campbell and Shier. Cambridge, MA: MIT Press.
- Kim, Jaegwon. 1989. Mechanism, purpose, and explanatory exclusion. *Philosophical Perspectives* 3: 77–108.
- . 1993. The non-reductivist's troubles with mental causation. In *Mental Causation*, ed. Heil & Mele. Oxford: Clarendon Press.
- . 1998. *Mind in a Physical World*. Cambridge, MA: MIT Press.
- . 2005. *Physicalism, or Something Near Enough*. Princeton: Princeton University Press.
- Kittle, Simon. 2018. When is an alternative possibility robust? *European Journal of Philosophy*: 1–12.
- Leibniz, Gottfried W. 1695, 1989. A new system of nature. In *Philosophical Essays*, ed. Ariew Garber. Indianapolis: Hackett Publishing.
- Levy, Neil. 2011. Expressing who we are: moral responsibility and awareness of our reasons for action. *Analytic Philosophy* 51: 243–261.
- . 2013. The importance of awareness. *Australasian Journal of Philosophy* 91 (2): 211–229.
- . 2014a. *Consciousness and Moral Responsibility*. New York: Oxford University Press.
- . 2014b. Consciousness, Implicit Attitudes and Moral Responsibility. *Noûs* 48: 21–40.
- Levy, Neil, and Tim Bayne. 2004. Doing without deliberation: Automatism, automaticity, and moral accountability. *International Review of Psychiatry* 16 (3): 209–215.
- Lewis, David. 1981. Are we free to break the laws? *Theoria* 47: 113–121.
- Libet, Benjamin., Elwood W. Wright Jr, Bertram Feinstein, and Dennis K. Pearl. 1979. Subjective referral of the timing for a conscious sensory experience: A functional role for the somatosensory specific projection system in man. *Brain* 102: 193–224.
- Libet, Benjamin. 2003. Can conscious experience affect brain activity? *Journal of Consciousness Studies* 10 (12): 24–28.
- . 1982. Brain stimulation in the study of neuronal functions for conscious sensory experience. *Human Neurobiology* 1: 235–242.
- . 1985. Unconscious cerebral initiative and the role of conscious will in the initiation of action. *Behavioral and Brain Sciences* 8: 529–566.
- Malcolm, Norman. 1968. The conceivability of mechanism. *Philosophical Review* 77: 45–72.
- Malebranche, Nicolas. 1674, 1997. *The Search After Truth* ed. Lennon & Olscamp. Cambridge: Cambridge University Press.
- McCann, Hugh J. 1998. *The Works of Agency: On Human Action, Will, and Freedom*. Ithaca: Cornell University Press.
- . 2012. Making decisions. *Philosophical Issues* 22: 246–263.
- Maoz, Uri, Gideon Yaffe, Christof Koch & Liad Mudrik. forthcoming. Neural precursors of decisions that matter — an ERP study of deliberate and arbitrary choice. *BioRxiv* 097626. <https://doi.org/10.1101/097626>.
- Moore, George E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Moya, Carlos. 2006. *Moral Responsibility: The Ways of Scepticism*. New York: Routledge.
- . 2007. Moral responsibility without alternative possibilities? *The Journal of Philosophy* 104: 475–486.

- . 2011. On the very idea of a robust alternative. *Critica* 43 (128): 3–26.
- Nagel, Thomas. 1974. What is it like to be a bat? *Philosophical Review* 83: 435–456.
- Nahmias, Eddy. 2010. Scientific challenges to free will. In *A Companion to the Philosophy of Action*, ed. O'Connor & Sandis. New York: Wiley–Blackwell.
- O'Connor, Timothy., and John R. Churchill. 2010. Is non-reductive physicalism viable within a causal powers metaphysics? In *Emergence in Mind*, ed. Macdonald & Macdonald. Oxford: Oxford University Press.
- Peacocke, Christopher. 1983. *Sense and Content*. Oxford: Oxford University Press.
- Pereboom, Derk. 2000. Alternative possibilities and causal histories. *Philosophical Perspectives* 14: 119–137.
- . 2001. *Living Without Free Will*. Cambridge University Press.
- . 2003. Source incompatibilism and alternative possibilities. In *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*, ed. Widerker & McKenna. Aldershot: Ashgate.
- Putnam, Hilary. 1967. Psychological predicates. In *Art, Mind, and Religion*, ed. Capitan & Merrill. Pittsburgh: University of Pittsburgh Press.
- Pylyshyn, Zenon. 1984. *Computation and Cognition*. Cambridge, MA: MIT Press.
- Schlosser, Marcus E. 2013. Conscious will, reason-responsiveness, and moral responsibility. *The Journal of Ethics* 17: 205–232.
- Shabo, Seth. 2010. Uncompromising source incompatibilism. *Philosophy and Phenomenological Research* 80 (2): 349–383.
- Shepherd, Joshua. 2015a. Consciousness, free will, and moral responsibility: Taking the folk seriously. *Philosophical Psychology* 28 (7): 929–946.
- . 2015b. Conscious control over action. *Mind & Language* 30 (3): 320–344.
- Smart, John J.C. 1961. Free will, praise and blame. *Mind* 70: 291–306.
- Smithies, Declan. 2012. The mental lives of zombies. *Philosophical Perspectives* 26: 343–372.
- Soon, Chun Siong, Marcel Brass, Hans-Jochen Heinze, and John-Dylan Haynes. 2008. Unconscious determinants of free decisions in the human brain. *Nature Neuroscience* 11: 543–545.
- Sosa, Ernest. 1984. Mind-body interaction and supervenient causation. *Midwest Studies in Philosophy* 9: 271–281.
- Stoutland, Frederick. 1980. Oblique causation and reasons for action. *Synthese* 43: 351–367.
- Stump, Eleanor. 1999a. Alternative possibilities and moral responsibility: The flicker of freedom. *The Journal of Ethics* 3 (4): 299–324.
- . 1999b. Dust, determinism, and Frankfurt: A reply to Goetz. *Faith and Philosophy* 16 (3): 413–422.
- . 2003. Moral responsibility without alternative possibilities. In *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities*, ed. Widerker & McKenna. Aldershot: Ashgate.
- Timpe, Kevin. 2007. Source incompatibilism and its alternatives. *American Philosophical Quarterly* 44 (2): 143–155.
- . 2008. *Free will: Sourcehood and its alternatives*. New York: Continuum.
- Waller, Robyn. 2012. Beyond button presses: The neuroscience of free and morally appraisable actions. *The Monist* 95 (3): 441–462.
- Walter, Sven. 2014. Willusionism, epiphenomenalism, and the feeling of conscious will. *Synthese* 191: 2215–2238.
- Watson, Gary. 1975. Free agency. *The Journal of Philosophy* 72 (8): 205–220.
- Wegner, Daniel M. 2002. *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- . 2004. Précis of *The Illusion of Conscious Will*. *Behavioral and Brain Sciences* 27: 649–659.
- Widerker, David. 2006. Libertarianism and the philosophical significance of Frankfurt scenarios. *The Journal of Philosophy* 103 (4): 163–187.

- . 2009. A defence of Frankfurt-friendly libertarianism. *Philosophical Explorations* 12 (2): 87–108.
- Wolf, Susan. 1990. *Freedom within Reason*. Oxford: Oxford University Press.
- Zagzebski, Linda. 2000. Does libertarian freedom require alternate possibilities? *Noûs* 34: 231–248.
- . 2010. Foreknowledge and human freedom. In *A companion to Philosophy of Religion*, ed. Draper Taliaferro and Quinn. Chichester: Wiley-Blackwell.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 2

Causality, Agency and Change



Stephen Pratten

Abstract Mainstream economists intermittently recognise a dilemma at the core of their project. On occasion they note that the widely accepted intuition that people have real choice and agency is inconsistent with their objective of increasing the explanatory power of economic theory. They worry that as causal explanations of economic phenomena are extended the more agency and choice must be recognised as ultimately illusory. The dilemma once recognised is typically set aside and the conventional modelling practices of mainstream economics persisted in without further delay. It is argued in this paper that the noted supposed dilemma is false and arises primarily because formalistic methods (and notions of explanation that accommodate them) tend to be adopted in economics prior, and without sufficient attention being paid, to the sorts of objects that constitute the subject matter of social inquiry. It is argued that the methods and forms of explanation that mainstream economists recognise as legitimate presuppose an ontology that is unable to either accommodate agency within nature or recognise possibilities for genuine change. In the constructive part of the paper a strategy for examining the dilemma that mainstream economists note that places ontology up front and centre stage is adopted. It is argued an ontology of real active powers, capacities, tendencies, dispositions and potentials supports a thoroughly naturalistic conception of agency and genuine change that can serve to inform the choices economists make about how best to pursue explanatory projects without generating tension or incoherence.

2.1 Introduction

Mainstream economists periodically pose what they take to be a dilemma. They note that human intentional agency, including the capacity for choice, aspects of human behaviour that they wish in turn to acknowledge and promote, seem to be threatened, or fundamentally undermined, the more successful economic

S. Pratten (✉)

Kings Business School, Kings College London, London, UK

e-mail: stephen.pratten@kcl.ac.uk

explanation becomes. For example, Melvin Reder insists that economists must confront a dilemma between choice and the success of their explanatory projects:

Associated with the assumption of stable preferences, but logically distinct, is the ‘thrust for endogenization’. A leading manifestation of this tendency is Stigler’s attempt to explain – and constrain – the behaviour of political decision makers, but this is not the only one ... Successfully to endogenize a new variable is to enhance the explanatory power of economics, and there is much interest in such achievements. However, it must be noted that where variables are made ‘endogenous’, they can no longer serve as objects of social choice. To the extent that variables are endogenized ... choice is explained ... ‘society’s’ freedom of choice is seen as illusory. Freedom appears to consist not in power of choice, but (pace Hegel) in recognition of necessity. This is not a likely conclusion for followers of Adam Smith, and surely not one they desire, but one from which they can be saved only by the failure of this direction of research (Reder 1982: 34–35).

From a mainstream perspective, any progress in formulating acceptable economic explanations would, seemingly inevitably, be accompanied by people increasingly being revealed to be merely propelled along by external events and conditions and exposed as lacking agency and choice. Moreover, given the template adopted by mainstream economics for screening appropriate forms of explanation, the only notion of change that can be salvaged is a severely impoverished one. Social life is reduced to a flux of events while change relates merely to alterations in the patterning of events.

In this paper this apparent dilemma of only being able to recognise real agency, choice and genuine change by giving up on explanation is considered from an explicitly ontological perspective. The argument advanced is that the noted dilemma arises primarily because notions of explanation in economics tend to be adopted prior, and without sufficient attention being paid, to the sorts of objects that constitute the subject matter of social inquiry. It is argued that the forms of explanation that mainstream economists recognise as legitimate presuppose an ontology that is unable to either accommodate agency within nature or recognise possibilities for genuine social change.

In the constructive part of the paper a strategy for examining the dilemma that mainstream economists note (only to then ignore) that places ontology up front and centre stage is adopted. It is argued that a structured ontology of real active powers, capacities, tendencies, dispositions and potentials supports a thoroughly naturalistic conception of agency that can guide the choices economists make about how best to pursue explanatory projects.¹ Once this alternative ontological position is sketched and its adequacy defended it becomes clear that economists do not face any necessary dilemma between explanatory progress and the recognition of human agency and choice. It is further argued that this ontology can productively inform initiatives aimed at bringing about rational, intentional social transformation.

¹The kind of structured ontology being referred to here is associated with the work of the Cambridge Social Ontology Group, (see Lawson 1997, 2003, 2013, Pratten (2015), and Faulkner et al. 2017) which itself is aligned closely with the work of critical realists such as Archer et al. (1998), Bhaskar (1978, 1979) and Collier (1994, 2011).

2.2 Mainstream Economics, Ontological Neglect and the Denial of Agency

Given the central place that choice occupies in mainstream accounts of the contribution economics makes, why do mainstream economists often end up positing a dilemma between explanatory progress on the one hand and agency and choice on the other? In order to understand why mainstream economists hit upon this dilemma it is necessary to clarify the metaphysical or ontological presuppositions underpinning their insistence that only certain methods and forms of explanation are legitimate within an appropriately scientific economics.

If research practices are to count as proper economics at all, mainstream economists insist that formalistic modelling methods (and forms of explanation that can accommodate such methods) need to be adopted. This stipulation is made without any assessment being provided of the ability of such methods to illuminate the social domain. For any method to be able to illuminate a domain of reality, the nature of the phenomena of that domain must be of a sort to render that feasible. The problems of modern mainstream economics, including the posing of a dilemma between explanatory progress and agency and choice, stem from a failure to recognise this insight and more broadly from profound neglect of ontological issues.²

Mainstream economists start with a particular type of method and presume mistakenly that it must be appropriate to all social contexts. The result is that in their conceptions mainstream economists end up distorting social phenomena so as to render them open to treatment by their chosen method. At times the distortion is such that the results and implications of the modelling exercises grate with the unelaborated intuitions that mainstream economists themselves have about basic features of the social domain and this tension is then manifest in the positing of dilemmas or conundrums that can be stated but not resolved or transcended.

The procedures of formalistic modelling characteristic of mainstream economics typically involve a reliance upon functional relations. When mainstream economists address phenomena such as consumption, investment, production or human well-being they characteristically seek to formulate consumption, investment, production and utility functions. It is this emphasis upon formalistic modelling and functional relations that is inappropriate to the analysis of most social phenomena and leaves economists unable to accommodate agency, choice and genuine change.

If an approach to economics that utilizes mathematical functions is to be viable and provide real insight, then it can be shown that it must be assumed that social events relate to each other in very specific stable ways. If a reliance on functions is to be an appropriate way to proceed in economics event regularities or event correlations must be commonplace in the social realm. The insistence a priori upon formal

²The critique of mainstream economics briefly set out in this section is elaborated at length by Lawson (1997, 2003, 2015b).

mathematical methods can be seen to carry with it a set of misleading and largely unrecognised ontological commitments. Specifically, the presuppositions implied by formal methods, understood as essential by mainstream economists, include that the basic ontological units are events, or states of affairs, that causation is to be analysed in terms of necessary connections holding between events and that all events including actions are strictly deterministically caused by prior ones according to laws in an essentially passive way.

Before proceeding it is worth considering further the conceptions of choice/agency, causality and social change implied by the mainstream insistence on the universal application of formalistic modelling methods.

The mainstream economists' models are, of course, essentially deterministic in the specific sense that human agency and choice are effectively denied. In a standard consumption function, for example, consumers expenditure is depicted as a stable function of disposable income, if amongst other things, so that a certain change in income or in its rate of change, triggers an adjustment in consumption by an amount that is fixed even before actual consumers know about any such change and whatever might have brought it about. Here it is far from obvious how the 'agent' can be interpreted as being actively involved in their own actions at all. The agent on such an account appears merely as a site of automatic adjustment rather than a subject capable of actively bringing about change.

Among heterodox economists it has long been recognised that the mainstream treatment is hostile to agency and choice. Veblen, in a famous passage that retains much of its relevance today, writes:

The hedonistic conception of man is that of a lightening calculator of pleasure and pains, who oscillates like a homogenous globule of desire of happiness under the impulse of stimuli that shift him about the area, but leave him intact. He has neither antecedent nor consequent. He is an isolated, definitive human datum, in stable equilibrium except for the buffets of the impinging forces that displace him in one direction or another. Self-imposed in elemental space, he spins symmetrically about his own spiritual axis until the parallelogram of forces bears down upon him, where upon he follows the line of the resultant. When the force of the impact is spent, he comes to rest, a self-contained globule of desire as before. Spiritually, the hedonistic man is not a prime mover. He is not the seat of a process of living, except in the sense that he is subject to a series of permutations enforced upon him by circumstances external and alien to him (Veblen 1919: 73).

Shackle argues that, given the framework they deploy, mainstream (or conventional) economists ought really to drop entirely any reference to choice:

Conventional economics is not about choice, but about acting according to necessity. Economic man obeys the dictates of reason, follows the logic of choice. To call this conduct choice is surely a misuse of words, when we suppose that to him the ends amongst which he can select, and the criteria of selection are given, and the means to each end are known. The theory which describes conduct under these assumptions is a theory of structure, not of creation of history. Choice in such a theory is empty, and conventional economics should abandon the word (Shackle 1961: 272–3).

2.3 Humean Causality and Event Focussed Conceptions of Change

In terms of causality the conception most straightforwardly reconciled with the mainstream modelling orientation is that of Humean causality. From this perspective the analysis of causality relates to the tracing of connections holding between events, so that the cause of some change is always a prior chain of events from which it follows according to a rule. Hume denied the possibility of establishing the independent existence of things or the operation of natural necessity. In Humean philosophy, the only properties of which we can have any knowledge are those which give rise to distinct impressions. These properties include the perceptible qualities of bodies, such as their shape, size, colour, etc. We can observe the speeds and directions of, say, two colliding billiard balls immediately before and after they collide, and we may identify a regularity in the way these speeds and directions are connected. What we do not observe is something beyond this that constitutes the capacity of one billiard ball to move another. On Hume's account of perception and causality, as traditionally interpreted, it is experiences, constituting atomic events and their conjunctions, that are viewed as exhausting our knowledge of nature. On such a view, generalities of significance in science must take the form of event regularities for these are the only sort of generalities that such an ontological position can sustain.

Mainstream economists from time to time acknowledge their Humean heritage. The celebrated econometrician Hendry R.F, for example, writes:

I think 'causality' is only definable within a theory. I am a Humean in that I believe we cannot perceive necessary connections in reality. All we can do is to set up a theoretical model in which we define the word 'causality' precisely, as economists do with $y=f(x)$. What they mean by that in their theory is that if we change x (and it is possible to change x), y will change. And the way y will change is mapped by f , so we have a causal theory. They could give a precise or formal definition of the mapping $f(\cdot)$. Empirically, concepts such as causality are extraordinarily hard to pin down. In my methodology, at the empirical level, causality plays a small role. Nevertheless, one is looking for models which mimic causal properties so that we can implement in the empirical world what the theorist analyses; namely, if you change the inputs, the outputs behave exactly as expected over a range of interesting interventions on the inputs (Hendry et al. 1990, p. 184).

Hendry, like so many mainstream economists, is restricted by his Humeanism to searching out correlations at the level of events and to sophisticated forms of data analysis (see Pratten 2005).

Finally the notion of change sustained by the mainstream is similarly impoverished. Consider, for example, the way econometricians typically partition relationships into endogenous and exogenous components. In drawing this distinction the possibility of retaining an, albeit limited, conception of change seems to be opened up. If some of the exogenous variables are interpreted as choice variables or understood as instruments that can be directly manipulated by government policy makers, it initially seems that an element of choice is afforded to this elite group in shaping or controlling the configuration of final outcomes. Econometric models are valued

as tools that can assist policy makers as they seek to fix or control future patterns of events. Yet from the perspective of contemporary mainstream economics with its emphasis on rational, near omniscient, individuals the assumption adopted here that the variables treated as exogenous, the instruments of government policy, remain unpredictable is difficult to justify. There is a persistent tendency within the mainstream project to endogenise ever more variables and an underlying belief that everything can ultimately be endogenised.³ Within the framework adopted by the mainstream economist there is neither provision, nor in consequence scope, for transforming any structural modes of determination. Social life is reduced to a flux of events. At most, any conception of change that can be sustained is the set of policy maker instigated adjustments to events and states of affairs.

Thus by adopting the conventional modelling practices of mainstream economics a host of metaphysical or ontological positions are thereby assumed that mean that agency, choice and change are at best little more than illusory.

2.4 Defending a Depth Realism

The ontological presuppositions of most methods of mathematical economic modelling are hostile to agency, choice and genuine change. Mainstream economists cannot coherently sustain the reality of people experiencing their power of choice even though the latter is an aspect of human behaviour they claim explicitly to acknowledge. It is the implicit ontology presupposed by their insisted upon methods that generates a dilemma whereby nature construed in terms of chains of events is a block on agency. Subjectivists in economics recognise the difficulties associated with accommodating agency that mainstream economists face. However, their own failure to decisively break away from the event-based ontology encourages in reaction an essentially voluntarist conception of human agency to be adopted and the acknowledgement of only a highly restricted role for social structure.⁴

In order to appreciate that the dilemma is a false one it is necessary to recognise that human agents are active causers of change in a world of active causers of

³The powerful tendency within mainstream economics to endogenise ever more variables is reflected where the advocates of the rational expectations hypothesis assert that changes in government policy instruments are in the final analysis predictable. Thus, Sargent and Wallace note: "The conundrum facing the economist can be put as follows. In order to have normative implications, it must contain some parameters whose values can be chosen by the policy maker. But if these can be chosen, rational agents will not view them as fixed and will make use of schemes for predicting their values. If the economist models the economy taking these schemes into account, then these parameters become endogenous variables and no longer appear in the reduced form equations for the other endogenous variables. If he models the economy without taking the schemes into account he is not imposing rationality" (1976: 183). For discussion and criticism of such developments, see Lawson (1994, 1997).

⁴Lawson (1994), argues that Hayek at least in his famous 'Scientism and the Study of Society' essay adopts such a subjectivist position and proceeds to pinpoint the problems and tensions associated with it.

change. It is an error to view nature as a block on agency since we are part of nature and what is required is a thoroughly naturalistic understanding of agency. This, in turn, requires the elaboration of a quite different ontological framework. A broader range of basic ontological categories needs to be elaborated upon that extends beyond events and fixed relations between them to include centrally real natures or forms, powers, capacities, tendencies, potentialities, processes.

An ontological conception opposing the dominant Humean event causal theory and seeing agency instead as involving the irreducible exercise of powers has recently enjoyed something of a revival in the philosophy of science (Bhaskar 1978; Cartwright 1992; Harré and Madden 1975) and the philosophy of action (Groff *forthcoming*; Steward 2012). On this ontological conception it is real things and their powers or ways of acting that are considered to be knowable and are taken to endure. Specific kinds of things have powers to act in definite ways in appropriate circumstances by virtue of certain relatively constant intrinsic structures or constitutions, or more generally, natures - which are discerned *a posteriori* in the process of science and general experience. It is these essential natures that designate what things are. Moreover, once we know what a thing is then, if certain 'activating' or 'triggering' conditions hold, we know how it will behave.

Without some such structured ontology, the absence of strict event regularities necessarily threatens the search for scientific generalities. With a structured ontology persistence and generality can obtain at a different level. It is worth briefly considering how an ontology that acknowledges the reality of deeper structures and mechanisms can render significant aspects of scientific activity, i.e., experimental activity and the application of scientific knowledge outside the experimental set-up, intelligible in a way that an event based metaphysics is quite unable to.

It is important to note that experiments are actively brought about, they are not spontaneous natural occurrences. Yet what they are designed to reveal is how nature acts when and where such active interventions are not taking place. If experiments only revealed how nature acted in the confines of the laboratory they would be of limited use.

What must the world be like for the kind of practical intervening associated with experimental practice to be, in fact, of assistance in understanding what happens outside of experimental contexts? Hume and many empiricists assume that by simply observing nature constant conjunctions of events can be discovered. Outside astronomy this does not happen. It is not the case that every time an object falls to the ground it does so with a constant rate of acceleration - an object such as a feather might be affected by all sorts of counteracting forces. The concept of cause is nevertheless applied in such cases, yet there is no constant conjunction which for the Humean is all that causality amounts to.

Outside astronomy the only way to be sure of constant conjunctions occurring is to carefully design and practically configure experiments. Every time we drop a heavy object in a vacuum it falls with a constant rate of acceleration. When an object is dropped outside of a vacuum all sorts of forces are likely to be in play preventing it from falling with a constant rate of acceleration.

A successfully designed experiment ensures that irrelevant forces which would influence the outcome are absent. If Hume were correct experiments would be superfluous but they are not because cause is accompanied by constant conjunction *only* when other things are equal and in nature, other things are never equal. The natural world is (what is referred to in the relevant philosophy of science literature as) an *open system* and experiment establishes a *closed system*.

In order for nature to be discoverable in its openness by artificially establishing closed systems it must be governed by many causal mechanisms, conjointly producing events. It is an open system because there are many of these mechanisms; it can be studied experimentally because we can in certain contexts isolate one of them, either by preventing others from operating, or keeping their operation constant, or making allowances for their operation. Successful experiments reveal the real working of natural mechanisms one by one, but in the spontaneous course of nature they are working conjointly to produce outcomes that are not, like the results of an experiment, predictable.

With an augmented ontology acknowledged then a central aim of the experiment can be recognised as being to help us understand causal mechanisms not simply identify event regularities. In well-controlled experiments, stable underlying causal mechanisms are insulated from countervailing causes, so that their unimpeded effects can be straightforwardly identified. Thus objects fall with a constant rate of acceleration in an experimental vacuum, because aerodynamic and other causal forces are prevented from affecting the outcome. So experiments are primarily concerned with underlying causal factors. The point of the experiment is precisely to insulate and thereby empirically identify stable causal mechanisms.

Since actual events or states of affairs outside the laboratory are co-determined by numerous often countervailing mechanisms the action of any one mechanism though real may not be precisely manifest or actualised. Characteristic ways of acting or effects of mechanisms which may not be actualised because of the openness of the relevant system can be conceptualised as tendencies. Tendencies are potentialities which may be exercised or in play without being straightforwardly realised or manifest in any particular outcome.

This kind of more expansive ontology suggests that if there is an essential moment in natural science it involves identifying and understanding causes of phenomena of interest. The practice of successful event prediction, and any attendant mathematical reasoning, may aid this process where it is feasible but is not an essential feature. The essential mode of inference in science is neither induction nor deduction but one that can be termed retrodution or abduction and explanatory, rather than predictive, power becomes the dominant criterion of theory adequacy, while the objective of assessing the reality of the posited mechanism has to be explicitly acknowledged. Once committed to knowable deeper levels of reality a priority becomes elaborating ways of identifying underlying causes.⁵

⁵For elaborations and defences of the depth ontology sketched here see Bhaskar (1978), Collier (1994), and Lawson (1997).

2.5 Situating Agency and Choice Within Nature

Once some such augmented ontology is acknowledged does it remain the case that economists must accept the apparent dilemma that increasing explanatory power can be achieved only at the expense of denying the possibility of human agency and genuine change? The short answer, of course, is no. By refocusing the analysis onto things and their potencies the pressure to construe nature's order in terms of sequences of events is removed and the dilemma of how to accommodate human agents in a natural world conceived metaphysically in terms of event chains simply does not arise. For reconceived as a world of objects, systems, totalities, mechanisms and their potencies there is no obvious tension between causal nature and active agency. Agents take their place in such a metaphysical framework as active exercisers of powers and capacities that are distinctive of the kinds of objects they are. Nature is not a block to agency, rather human agents need to be understood as being part of nature.

With this structured ontology adopted agency is not only recognised as real and an irreducible primary cause of change in the world but, importantly, also understood as being constrained by social conditions. It is part of our nature that we are essentially community beings. Modern social conditions are such that they distort our relation to our nature and often harm us. Ours is a social world based in large part on disregard for living human beings in which people's needs and well-being remain peripheral concerns. Nature, including the social world, is composed of totalities some of which have human agents as components. The organisation of social totalities determine individuals and their actions not by negating their agency but by working through it – shaping agents powers and capacities. Societal determination does not stop at external constraints but enters into the constitution of the agent's practical capacities.

A particular position in social ontology is being evoked here.⁶ Social reality is that collection of phenomena whose existence depends necessarily on human beings, including human interactions. There are two extreme opposing views about the nature of social phenomena. The first is a social atomist view that insists there is nothing to the social realm apart from a collection of human agents. The second extreme view – social holism – is that there is nothing to human agency except the positions individuals occupy in social totalities. However, there are no de-socialised human agents, that is people always exist in communities, and equally neither are there any depopulated communities. Individuals exist and each of us have our own relatively unique nature, but social totalities are not just the plural of these individuals: they too have their own nature, in the arrangement or organisation of relations in which individuals stand. Individuals do not lose their identity by standing in relations to one another within communities, but their identity is partly constituted by these relations. One is an employer, an employee, self-employed, a husband, a

⁶See Lawson (2013) and Collier (2011) for summary accounts of this position in social ontology.

mother and so on and acquires rights and obligations and follows collective practices as such positions are moved into.

Both extreme views – that is both social atomism and social holism - are unsustainable accounts of the nature of the social realm – what is needed is a relational account of the nature of the social domain. Social totalities are neither bundles of separate individuals nor are they mysterious collective subjects. Communities are organised sets of relations between individuals and their environment, relations that pre-exist any given individual and partly constitute the character and the powers of the related individuals. Society and communities exist in the sense that they are not a mere plural of person. On this relational conception of the social all social forms – the economy, the state, international organisations, trade unions, universities, households – are communities that depend upon, or presuppose social relations. And of special concern are the relational positions into which individuals essentially slot with their associated relationally defined collective practices, rights, obligations, prerogatives, etc.

A further feature of the social ontology being drawn on here is its transformational account of social activity. *Voluntarists*, while observing that making history is undertaken by human agents, exaggerate their ability to create social structure. *Structural determinists*, while acknowledging that human agents operate in conditions not of their own choosing but enabled and constrained by social structure, tend to conceive of structure as a fixed constraint. The insights from both perspectives need to be retained in a more encompassing *transformational model of social activity*.

On the transformational model the existence of social structure is the often unacknowledged but necessary condition of an individual's intentional acts, as well as a typically unintended, but inevitable outcome of, individual actions taken in total. Social structure is the unmotivated condition of our motivated productions, the non created but drawn upon and reproduced/transformed condition for our daily economic/social activities. One works to earn a living and thereby contributes to the accumulation of capital and reproduction of capitalist relations. One pays into a pension scheme so as to cover expenditures in old age and thereby helps to reproduce the financial system. The transformational model of social activity highlights both that the course of human history consists in a series of intentional actions of individual agents and their often unintended consequences and underlying this history there are relatively enduring mechanisms, which are constituted by the structure of relations between human agents and between those agents and their natural environment.

Let us consider more directly the notion of human agency and choice opened up once this relational understanding of the social and transformational account of social activity are adopted. Agency and choice if they exist must be powers that are irreducible to their exercise or manifestation in specific concrete acts. Although it has been argued that the concept of human intentionality entails the pre-existence of social structures which facilitate intentional acts, there is no necessity to suppose that such acts must be completely determined. Thus, although the structure of language facilitates speech acts it does not fix what is said – the status of human agency

is maintained. Similarly, the highway code enables safe driving without determining the journey undertaken; the market system promotes buying and selling without forcing any specific purchase to be made. Such examples emphasise action within structures. However, with social structures dependent on human conceptions and actions, the possibility also exists for both unintended and intended structural transformation.

What forms of consciousness are presupposed by human choice and action? Real choice, the ability to act or to have acted otherwise, presupposes both the transformational capacity of being able to make a difference and that action is in some sense and to a degree controlled by the agent. A significant feature of human action is that it is intentional in the sense that it is caused by reasons i.e., beliefs grounded in the practical interests of life - it is always directed towards some end. However, as Lawson (1994: 21) notes the pursuit of ends cannot be understood as a “simple, unitary, always reflected upon, discursive activity”. Consideration of the common everyday activities people engage in indicates that this cannot be the case. The complexity of human beings is of such an order that human powers extend beyond initiating changes in purposeful ways and include the monitoring and controlling of performances. The social activity that agents reflexively monitor constitutes a continuous flow – the individuals own acts, the acts of others, the socially constituted appropriateness of forms of conduct in particular contexts, etc. The reflexive monitoring of activity must also occur on an ongoing basis rather than in a piecemeal fashion. Since the monitoring of conduct is continuous then it must be tacit – moment by moment, explicit commentary and reflection would not be possible. Our ability to reflexively monitor activity presupposes a level of tacit consciousness. Beyond tacit consciousness a level of unconscious motivation can also be recognised as bearing on human praxis. In their actions, humans draw upon not only discursive thought and tacit and unacknowledged skills etc., but also unconscious needs and motivations.

2.6 Causality, Change and Social Transformation

With the status of human agency and choice preserved let me now turn to the issue of what conception of causality is presupposed by the proposed structured social ontology. The kind of ontological perspective sketched above conceives of causality as referring not to connections between discrete events but to systems actively exercising or displaying their powers and capacities. Reference to efficient cause is seen as needing to be supplemented - final, formal and material cause are understood as equally fundamental categories. Causality pertains to the powers and capacities that objects including human agents possess in virtue of their natures. On this view causality is construed as being objective rather than subjective.

The causal pluralism facilitated by the adoption of this ontological perspective makes it possible to conceptualise social determination in terms of final, formal and material causation rather than efficient causality alone. To elaborate with regard

formal causality specifically, a system can be understood as a set of elements that have an integrity when considered together as a totality, where the latter is formed via an organisation of those basic elements. The organising structure of a system emerges concurrently with the emergent totality that comprises the system as a whole. It not only renders the (organised) basic elements distinct components of the system but also accounts for any emergent causal powers of the emergent system. This organising structure also connects a subset of components to features of the environment; a system always exists in some context. The influence that this organising structure has is a type of formal causality. Lawson provides a useful example:

Consider the construction of a bridge. Here various items or materials may be brought together to form components of a totality, including, perhaps, pieces of wood, brick, stone, cast and/or wrought iron, mild, high-tensile and/or alloy steel, aluminium, steel-reinforced and/or pre-stressed concrete, glass-reinforced plastic, and so forth. These are organised or assembled, in a specific environment, and in a manner such that the resulting totality allows the crossing of a space, perhaps containing a river (whilst the resulting totality itself can survive potential stress caused by such factors as bending, compression, impact, oscillation, pressure, tension, torsion, vibration; contraction, corrosion, erosion, expansion, fatigue, friction, rain, river flow, sea-water, scouring, temperature changes, tidal flow, turbulence, waves, wind erosion, wind gusts, wind pressure etc.). The totality that is the bridge clearly emerges simultaneously with the organising relational structure of the materials enlisted as components, and, significantly, the latter organising structure makes a (causal) difference to the emergent causal powers of the totality. Were the resulting bridge to be taken apart again and the various materials assembled blindly, it is unlikely that any resulting outcome would possess the causal properties of a bridge. The arrangement matters; it is a type of *formal* causation. (Lawson 2014: 25)

Most social systems are continually reproduced through the everyday human interactions which they facilitate. In these systems human individuals are amongst the components. It is through the sum total of their activities, qua components, that the system is reproduced. Consider as examples local communities, firms, markets, tutorial groups, reading clubs, financial centres and workplaces. Each is an emergent form of organisation possessing novel emergent causal powers at the level of the emergent totality, but these causal powers can only ever be realised through the actions of its organised members. Each such system possesses an organising structure that facilitates certain individual actions of system components and is subsequently reproduced (or transformed) through those actions. The organising structure conditions, facilitates and constrains the development of individual's powers. The individuals are irreducible active causers of change yet the kind of change they can and cannot bring about, the kinds of actions they can and cannot perform, are conditioned by the organising structure of the totality.

The notion of change entailed by this augmented ontological perspective is not reducible to the minimalist account of discrete and exogenous changes in events and states of affairs encouraged by the mainstream economics approach. Significantly this broader metaphysical position can distinguish different kinds of change, including accidental from necessary. Developing into a dog is in the nature of a puppy

while, say, breaking a leg is not. The latter is an event due to contingent external factors, while growing into a dog is a necessary tendency given the kind of thing it is. It is our understanding of the nature of things that allows us to make judgements of particular cases. We do not alter our description of what a dog is when we come across one with a broken leg. Instead we make a judgement of this individual and recognise that it is lacking something. When we make such judgements about individual cases we are not making empirical descriptions or statistical generalizations but are assessing what is proper or essential to a particular individual given the kind it belongs to. The assessment of the individual is made in relation to an appreciation of the powers and activities a dog *qua* dog has and does, the parts it has and the life phases it passes through. These assessments are normative. When we make the assessment that the particular individual lacks something that is proper to its nature or *life-form* we thereby recognise that in the relevant sense it exists privatively. The category of life-form here plays an explanatory role in accounting for what happens in the course of an individual's life to the extent that things proceed as they should,⁷ while it is when things go wrong that appeal must be made to chance and accident.

This kind of ontological framework, once supplemented with insights from critical social science, can serve to inform emancipatory initiatives that are focussed on, not the controlling of variables so as to guarantee predictable outcomes, but, the bringing about of rational, intentional, social transformations that facilitate human flourishing. Given the natures we have we are creatures capable of flourishing in certain conditions and the flourishing of each of us depends on the flourishing of others and ultimately all.⁸ If we do not flourish it is not because our interests are necessarily opposed, nor is it that our nature is inherently alien to us but is due to problematic social conditions existing that distort our relation to it. Contingently existing social factors foster a wrong form of life that fails to facilitate human well-being and flourishing. Agency is irreducibly real but it is not sufficient for our flourishing.⁹ The good society remains a real potential that can be realised if different social arrangements, requiring significant structural transformations, can be brought into existence.

⁷For elaboration on the category of life form and its significance, see Thompson (2012) and Reeves (2016a).

⁸For an extended defence of the thesis that the flourishing of each is dependent on the flourishing of all others, see Lawson (2015a).

⁹Reeves (2016a, b) elaborates on a distinction between agency and freedom in the context of competing interpretations of Adorno's contributions. He clarifies that there is no contradiction between insisting on someone's agency and yet also recognising their unfreedom due to currently obtaining social conditions.

2.7 Conclusion

Mainstream economists, due to their mistaken assumption that the scientific credentials of the discipline hinge upon the deployment of formal mathematical modelling methods, are constrained by a host of implicit metaphysical presuppositions. By assuming that these methods are always relevant they take the social world to be constituted in effect by passive, isolated atoms. An appropriately formulated economic model, stipulating well specified functional relations, serves to filter out agency, choice and genuine change. This is a most distressing result for mainstream economists to acknowledge since they typically see choice as being a central, even defining, disciplinary concern. A fundamental dilemma between explanatory progress and choice is then posited and seen as inevitable.

Once an explicitly ontological orientation is adopted this dilemma can be shown to be false. An alternative structured ontology of things, powers, potentialities can be identified that is not only coherent but able to render intelligible key aspects of modern scientific practice in a way the implicit ontology presupposed by mainstream economists is quite unable to do. Both agency and choice can be recognised without either constituting a block on explanatory progress. Economists can accept the goal of seeking to improve the explanatory power of their theory and acknowledge the reality of agency, human choice and intentionality. The recognition that the dilemma is false follows from a direct engagement with ontology. Adopting a structured social ontology also allows a more adequate conception of social change to be accommodated which can inform emancipatory initiatives. Given the state of contemporary economics a renewed focus on ontology is urgently needed if sustained explanatory progress is to be facilitated.

References

- Archer, M., R. Bhaskar, A. Collier, T. Lawson, and A. Norrie. 1998. *Critical Realism: Essential Readings*. London: Routledge.
- Bhaskar, R. 1978. *A Realist Theory of Science*. Hemel Hempstead: Harvester.
- . 1979. *The Possibility of Naturalism*. Hemel Hempstead: Harvester.
- Cartwright, N. 1992. Aristotelian natures and modern experimental method. In *Inference, Explanation and Other Frustrations: Essays in the Philosophy of Science*, ed. J. Earman. Berkeley: University of California Press.
- Collier, A. 1994. *Critical Realism: An Introduction to Roy Bhaskar's Philosophy*. London: Verso.
- . 2011. The social ontology of critical realism. In *Sociological Realism*, ed. A. Maccarini, E. Morandi, and Riccardo Prandini. London: Routledge.
- Faulkner, P., S. Pratten, and J. Runde. 2017. Cambridge Social Ontology: Clarification, Development and Deployment. *Cambridge Journal of Economics* 41 (5): 1265–1277.
- Harré, R., and E.H. Madden. 1975. *Causal Powers: A Theory of Natural Necessity*. Oxford: Blackwell.

- Hendry, D.F., E.E. Leamer, and D.J. Poirier. 1990. The ET dialogue: A conversation on econometric methodology. *Econometric Theory* 6 (2): 171–261.
- Groff, R. forthcoming. Sublating the free will problematic: Powers, agency and casual determination. *Synthese*.
- Lawson, T. 1994. Critical Realism and the Analysis of Choice, Explanation and Change. *Advances in Austrian Economics* 1 (1): 3–30.
- . 1997. *Economics and Reality*. London: Routledge.
- . 2003. *Reorienting Economics*. London: Routledge.
- . 2013. Emergence and Social Causation. In *Powers and Capacities in Philosophy: The New Aristotelianism*, ed. R. Groff and J. Creco. London: Routledge.
- . 2014. A speeding up of the rate of social change? technology, resistance, globalisation and the good society. In *Late Modernity: Trajectories towards Morphogenic Society*, ed. M.S. Archer. New York: Springer.
- . 2015a. Critical ethical naturalism: An orientation to ethics. In *Social Ontology and Modern Economics*, ed. S. Pratten. London: Routledge.
- . 2015b. *Essays on the Nature and State of Modern Economics*. London: Routledge.
- Pratten, 2005. Economics as progress: The LSE approach to econometric modelling and critical realism as programmes for research. *Cambridge Journal of Economics* 29: 179–205.
- Pratten, S., ed. 2015. *Social Ontology and Modern Economics*. London: Routledge.
- Reder, M. 1982. Chicago Economics: Permanence and Change. *Journal of Economic Literature* 35: 1–38.
- Reeves, C. 2016a. Beyond the postmetaphysical turn: Ethics and metaphysics in critical theory. *Journal of Critical Realism* 15 (3): 217–244.
- . 2016b. Adorno, Freedom and Criminal law: The ‘Determinist’ Challenge’ revisited. *Law and Critique* 27 (3): 323–348.
- Sargent, T.J., and N. Wallace. 1976. Rational Expectations and the theory of Economic Policy. *Journal of Monetary Economics* 2.
- Shackle, G.L.S. 1961. *Decision, Order and Time in Human Affairs*. Cambridge: Cambridge University Press.
- Steward, H. 2012. *A Metaphysics for Freedom*. Oxford: Oxford University Press.
- Thompson, M. 2012. *Life and Action: Elementary Structures of Practice and Practical Thought*. Cambridge: Harvard University Press.
- Veblen, T.B. 1919. *The Place of Science in Modern Civilisation and Other Essays*, as reprinted with a new introduction by W. J. Samuels. New Brunswick: Transaction (1990).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 3

How Economics Becomes Ideology: The Uses and Abuses of Rational Choice Theory



Jason Blakely

Abstract Rational choice theory is often presented as essential to the truly scientific study of economics. To the contrary, in this paper I argue that when rational choice is treated as the key to a science of human agency, it ensnares economics in certain intractable dilemmas. Drawing on hermeneutic philosophy, I argue that economists need to distinguish legitimate from illegitimate uses of rational choice theory. Failure to do so not only leads to self-defeating research goals, but also mires the discipline in an extreme form of free-market ideology. Economists thus have not only empirical but also ethical and political reasons for rejecting the mainstream deployment of rational choice. The paper concludes by briefly sketching a philosophically defensible use of rational choice in economics – one that is more sensitive to the expressive dimensions of human agency and the limits of mechanistic causality.

3.1 Introduction

Unbeknownst to many economists, much of their discipline has been absorbed into a worldview known to philosophers as “naturalism.” Naturalism can take a dizzying array of forms but at the most general level it holds that the study of human beings should adopt the basic conceptual features of the natural sciences.¹ In economics this sometimes occurs through the attempt to place human agency under supposedly

¹ Jason Blakely, *Alasdair MacIntyre, Charles Taylor, and the Demise of Naturalism* (Notre Dame: University of Notre Dame Press, 2016), ch. 1.

J. Blakely (✉)
Pepperdine University, Malibu, CA, USA
e-mail: jblakely@pepperdine.edu

value-neutral, formal causal laws akin to physics.² It is also evident in the way economics is often lauded as outpacing the other social sciences on the road to science.³

In this paper, I give a critical account of how a philosophically naturalist notion of agency and causality has shaped a central method of mainstream economics—rational choice theory.⁴ Of course, rational choice is far from being the whole of economics and naturalism does not comprise its only philosophical basis.⁵ Nonetheless the two working in tandem remain enormously influential. One of my major claims is that rational choice remains philosophically legitimate only insofar as it disentangles itself from naturalism. Rational choice has a role to play in social science but it needs to be fundamentally rethought.⁶ To advance my argument I draw on the resources of the philosophical traditions of hermeneutics and anti-naturalism.⁷ These traditions maintain that because human agency is expressive of meanings it is not compatible with the formal, mechanistic causality of naturalism.

My argument proceeds in three stages. First, I explain how rational choice's absorption into naturalism generates intractable dilemmas for economics. Hermeneutics offers a better account of these dilemmas than the leading proponents of naturalist rational choice. This leads to the second stage of my argument in which I explain how subsuming agency under naturalist notions of causality does not achieve scientific neutrality. Instead, rational choice becomes fused with highly contestable forms of technocratic and neoliberal ideology. This means economists have political reasons for abandoning a naturalistic use of this method. Finally, hermeneutics offers economists an alternative conception of rational choice—one consistent with the unique form of causality typical of human agency. A philosophically defensible use of rational choice must remain aware of certain limitations.⁸

²For a key early work in neoclassical economics that avows many naturalist goals, see: Lionel Robbins, *An Essay on the Nature and Significance of Economic Science*, second edition (London: Macmillan & Co., 1935).

³Cf., Daniel Hausman: "Introduction," in *The Philosophy of Economics*, 3rd edition (Cambridge: Cambridge University Press, 2008) 3; "Philosophy of Economics," *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition), Edward N. Zalta ed., <https://plato.stanford.edu/archives/win2013/entries/economics/>

⁴For the importance of rational choice to economics, see: Julian Reiss, *Philosophy of Economics* (New York: Routledge, 2013) 6; Daniel Hausman, "Philosophy of Economics," in *Routledge Encyclopedia of Philosophy*, vol. 3, ed. Edward Craig (London: Routledge, 1998) 211–222.

⁵Bruce Caldwell notes the pluralization of post-positivist theories: "Economic Methodology and Behavioral Economics: An Interpretive History," in Benjamin Gilad and Stanley Kaish, eds. *Handbook of Behavioral Economics*, vol. 2 (Greenwich, CT: JAI Press, 1986) 11–13.

⁶For why some critics advocate abandoning this method entirely, see: Norman Denzin, "The Long Good-Bye: Farewell to Rational Choice Theory," *Rationality and Society* 2:4 (1990): 504–506.

⁷For key volumes on the hermeneutic or interpretive turn, see: Michael Gibbons, *Interpreting Politics* (Oxford: Basil Blackwell, 1987); David Hiley, James Bohman, and Richard Shusterman, eds., *The Interpretive Turn* (Ithaca: Cornell University Press, 1991); Paul Rabinow and William Sullivan, eds., *Interpretive Social Science: A Second Look* (Berkeley: University of California Press, 1987).

⁸This runs against the qualitative-quantitative debate in the social sciences that treats methods as a binary choice between paradigms. For a sketch of this debate, see: Matthew B. Miles and A. Michael Huberman, *Qualitative Data Analysis: An Expanded Sourcebook, Second Edition* (Thousand Oaks, CA: SAGE Publications, 1994) 40.

Far from economics leading the way to science, the discipline should return to its historical roots in the humanities and become part of the interpretive turn.

3.2 Rational Choice and Scientific Causality

Rational choice is widely recognized as a key methodological resource for mainstream economics.⁹ At its most basic level, rational choice consists of the construction of an ideally rational agent by positing certain *a priori* postulates or axioms. Two axioms of particular importance are completeness and transitivity. The completeness axiom holds that an ideally rational actor always ranks preferences—though ties and indifference are allowed. While transitivity posits that these preferences are transferable from one object to another—such that someone who, say, prefers Giotto to Monet, and Monet to Warhol, also prefers Giotto over Warhol.¹⁰ What this vision of human agency gives economists is an idealized model that can be used to conceptualize an enormous variety of decision scenarios.

Today many economists implicitly ground rational choice in naturalist philosophical assumptions. This happens in two ways I wish to examine at some length. First, economists sometimes slide into treating rational choice as a quasi-behavioral, scientific account of the structure of human agency (that is, they *anthropologize* rational choice). Meanwhile other economists concede rational choice does not offer a quasi-behavioral science, but nonetheless insist on this method's scientific, predictive value. What both approaches share is the naturalist conviction that rational choice helps loft economics onto a higher plain of science than the other social sciences.

First let's examine the naturalist assumption that rational choice captures an essential behavioral structure of human agency. The idea that economics is premised on a true anthropology has important antecedents in classical and neoclassical economists who helped envision humans as a species of *homo economicus*—naturally calculative, haggling, and materially acquisitive. Such speculations go back at least as far as Adam Smith and John Locke.¹¹ Later William Stanley Jevons gave an influential utilitarian account of *homo economicus*.¹² Yet mainstream economics only took on its current form when economists abandoned these earlier psychological conjectures in favor of what some scholars have called a “pure logic of choice” (a move we will return to below).¹³ Nonetheless, the project of *homo economicus* has never fully vanished. One common way *homo economicus* has survived is through the assumption that rational choice captures some essential behavioral features of folk psychology.¹⁴

⁹ Julian Reiss, *Philosophy of Economics* (New York: Routledge, 2013) 6.

¹⁰ Itzhak Gilboa, *Rational Choice* (Cambridge, MA: MIT Press, 2010) 39–40.

¹¹ For discussion of the eighteenth-century myth of the bartering “savage,” see: Karl Polanyi, *The Great Transformation* (Boston: Beacon Street Press, 2001) 45–46.

¹² William Stanley Jevons, *The Theory of Political Economy* (New York: Sentry Press, 1957) 23, 21.

¹³ Caldwell, “Economic Methodology,” 6. This theory of rational choice originated in the strategic nuclear defense analysis of the Cold War: S. M. Amadae, *Rationalizing Capitalist Democracy: The Cold War Origins of Rational Choice Liberalism* (Chicago: University of Chicago Press, 2003).

¹⁴ See: Reiss, *Philosophy of Economics*, 29–31.

Folk psychology is the view that the reasons people hold are the causes of their actions.¹⁵ According to folk psychological approaches to economics, rational choice reveals the essential logical features of human reasoning, and so its formal models can provide predictable outcomes for how humans in aggregate will reason and behave in decision-making scenarios.

A prominent example of this approach is Nobel laureate Gary S. Becker's defense of what he called the "economic approach" to human behavior. Becker believed that rational choice assumptions offered researchers a "unified framework for understanding all human behavior" that was compatible with the empirical findings of the social sciences as well as evolutionary biology.¹⁶ Indeed, Becker insisted that "the economic approach" together with "modern psychology come to similar conclusions."¹⁷ Becker thus rendered rational choice assumptions into a quasi-behavioral claim that grounded a science of society more generally, explaining everything from marriage rituals to crime patterns.¹⁸

This quasi-behavioral account of the structure of human agency has had a large ideological impact in the United States.¹⁹ It has also been championed by prominent social scientists working in other disciplines. For instance, in political science the public choice approach to institutions (associated with the economist James Buchanan) has led to rational choice postulates being used to model the behavior of various political actors including voters, legislators, bureaucrats, and welfare recipients.²⁰ A leading example of this paradigm of study is David Mayhew's massively influential research on Congress.²¹

Despite its large presence in the social sciences, this quasi-behavioral approach to rational choice has never overcome certain intractable dilemmas. Specifically, psychological findings contradict this account of human agency showing that actual individuals both on the micro-level and in aggregate violate the axioms of rationality in myriad ways. For example, individuals sometimes treat the same choice dif-

¹⁵For a philosopher sometimes evoked to justify this view, see: Donald Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 2011).

¹⁶Gary S. Becker, *The Economic Approach to Human Behavior* (Chicago: University of Chicago Press, 1976) 14.

¹⁷Becker, *The Economic Approach*, 10.

¹⁸Becker, *The Economic Approach*. See also: Gary S. Becker and Julio Jorge Elías, "Introducing Incentives in the Market for Live and Cadaveric Organ Donations," *Journal of Economic Perspectives* 21:3 (2007): 3–24.

¹⁹This has partly occurred through popularizations of *homo economicus* on the fringes of professional economics: Steven Levitt and Stephen Dubner, *Freakonomics* (New York: Harper Perennial, 2005) 16; Tyler Cowen, *Discover Your Inner Economist: Use Incentives to Fall in Love, Survive your Next Meeting, and Motivate Your Dentist* (New York: Penguin, 2008).

²⁰Buchanan's Austrian School background made him less prone to describe economics as a full-blown science. Nonetheless, he often slid into treating *homo economicus* as reflecting "empirical presuppositions." See: Geoffrey Brennan and James Buchanan, *The Reason of Rules in The Collected Works of James M. Buchanan, vol. 10* (Indianapolis: Liberty Fund, 2000) 58–59.

²¹David R. Mayhew, *Congress: The Electoral Connection* (New Haven, CT: Yale University Press, 2004).

ferently (reversing their preferences) depending on how it is framed.²² They also exhibit circularity of preferences when holding conflicting criteria for deciding between options.²³ This implies that the quasi-behavioral grounding of rational choice theory is empirically false. While humans do act on the basis of reasons their reasoning is not formally fixed or structured by rational choice postulates.

This brings us to a second way economists have commonly grounded rational choice theory in naturalist philosophy. Namely, by arguing that although its axioms are empirically false, they nonetheless generate scientific predictions. An influential statement of this position remains Milton Friedman's controversial 1953 essay, "The Methodology of Positive Economics."²⁴ Like many economists working today Friedman saw economics as akin to physics since both relied on idealized models (for example, in physics falling bodies are imagined in a perfect vacuum).²⁵ Friedman inferred from this that neither economics nor physics depended on the realism of their theoretical assumptions but instead on their ability to yield reliable scientific predictions. As Friedman put it: "the relevant question to ask about the 'assumptions' of a theory is not whether they are descriptively 'realistic,' for they never are" but instead "whether it yields sufficiently accurate predictions."²⁶ In other words, Friedman offered economists a way out of the empirical problems facing rational choice: false or distortive assumptions might still have scientific value.²⁷

Yet once again naturalist rational choice was unable to succeed on its own terms. For rational choice has not granted economists predictive powers even approaching that of physics and the natural sciences. To the contrary, the most recent studies have found that the ability of economic experts to predict phenomena as diverse as GDP growth, inflation, unemployment, and entry or exit from membership in free-trade agreements is no better than dilettantes.²⁸ As Philip Tetlock summarized these findings: "people who devoted years of arduous study to a topic were as hard-pressed as colleagues casually dropping in from other fields to affix realistic prob-

²² Framing a choice in terms of gain led some individuals to risk aversion, while the equivalent choice framed in terms of loss led to more risk taking. Amos Tversky and Daniel Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science* 211 (1981): 453–458.

²³ For example, when selecting a spouse according to criteria of beauty, intelligence, and wealth: Kenneth O. May, "Intransitivity, Utility, and the Aggregation of Preference Patterns," *Econometrica* 22:1 (1954): 1–13. Other cases of intransitivity were observed when features were gradually added to an initial option (for example, accessories to a car purchase). Amos Tversky, "Intransitivity of Preferences," *Psychological Review* 76:1 (1969): 31–48. For an early list of violations of rational axioms, see: Amos Tversky and Daniel Kahneman, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5 (1992): 298.

²⁴ See: Hausman, "Philosophy of Economics," 218; Caldwell, "Economic Methodology," 10.

²⁵ Reiss notes the idealization of objects in physics differs from rational choice in that the former does not add substantive features to reality. Reiss, *Philosophy of Economics*, 131.

²⁶ Milton Friedman, "The Methodology of Positive Economics," in *The Philosophy of Economics*, 3rd edition, ed. Daniel Hausman (Cambridge: Cambridge University Press, 2008) 153.

²⁷ Cf., Caldwell, "Economic Methodology," 10.

²⁸ Philip Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* (Princeton, NJ: Princeton University Press, 2005) 59, 247.

abilities to possible futures.”²⁹ Indeed, economic experts were not even able to decisively best mindless, computer-generated predictions in either the form of “crude extrapolation algorithms” or “sophisticated statistical ones.”³⁰ Thus, rational choice cannot be philosophically grounded on its scientific predictive powers.

The leading naturalist uses of rational choice are therefore stuck in a bind where entire research programs are based on false philosophical assumptions.³¹ This brings us to the first place hermeneutics is useful to economists. Namely, hermeneutics avoids these dilemmas by clarifying why both *homo economicus* and strong prediction are not appropriate conceptions of rational choice. But seeing how this is the case requires articulating some basic philosophical concepts.

Hermeneutics holds that human agency is expressive of meanings, making the social sciences philosophically distinct from the natural ones. One way to understand this is through Charles Taylor’s influential claim that humans are “self-interpreting animals.”³² Self-interpretation refers to how in human beings any one belief or action is related to a wider web of meanings that are the result of creative, interpretive activity. Grasping human beliefs and actions therefore requires that economists interpret the interpretations of the agents in question. Human behavior is not an ahistorical structure but a set of historically local meanings (like written texts in need of decoding). This is what is meant by the claim that the human sciences—including economics—are interpretive disciplines. Namely, they must engage in the “hermeneutic circle,” relating a given action and belief to a wider context of meaning.³³ Economists must learn the art of interpreting meanings developed over centuries in the humanities (particularly literature and history).

Self-interpretation also implies a different kind of causality than that typical of the natural sciences. The natural sciences largely focus on necessary causality, as classically articulated by David Hume who theorized that a “cause” referred to the repeated experience of observing two events joined together where the antecedent event was producing the ensuing effect. As Hume put it: “we may define a cause to be an object, followed by another, and where all objects similar to the first are followed by objects similar to the second.”³⁴ This Humean line of thought inspired a certain vision of the natural sciences in which predictive power was made possible by identifying antecedent conditions “X” in constant conjunction with a consequent set of conditions “Y” (so long as there are not intervening factors). In this view,

²⁹Tetlock, *Expert Political Judgment*, 54.

³⁰Tetlock, *Expert Political Judgment*, 54.

³¹This dilemma is widely recognized in the philosophy of economics. For example, the leading textbook on the philosophy of economics candidly admits that the discipline is bedeviled by a cyclical set of debates and unresolved paradoxes concerning its assumptions about human agency: Reiss, *Philosophy of Economics*, 127, 141, 172.

³²Charles Taylor, “Self-Interpreting Animals,” in *Human Agency and Language* (Cambridge: Cambridge University Press, 1985) 45–76.

³³For an influential discussion, see: Charles Taylor, “Interpretation and the Sciences of Man” in *Philosophy and the Human Sciences* (Cambridge: Cambridge University Press, 1985), 14.

³⁴David Hume, *An Enquiry Concerning Human Understanding* (Chicago: The Open Court Publishing, 1912) 79.

predictive science is achieved when an ahistorical mechanism is discovered between two variables. When two variables have been causally linked a law of science might be inferred or formulated—such as whenever, X occurs, all things being equal, Y necessarily follows.

The problem is that in the case of human agency there is no set of antecedent conditions that necessitates or predictably generates the outcome of a consequent belief or action. To the contrary, because human beings are self-interpreting agents this implies that an individual is not locked into a fixed pattern of belief or action. The pattern of beliefs and actions is always contingent and never necessary. The formal structure of belief which rational choice posits is therefore unable to predict or explain human actions. Instead, beliefs and meanings must be described and explained in terms of a narrative or story about why a particular pattern of action and belief arose. This narrative is historically situated and not a formal, ahistorical structure.

The key lesson for economists is that they must always keep in mind that human agency is contingent and not mechanistic as in the natural sciences. This means economists might very well correlate variables and discover interesting patterns, but such variables never amount to causal discoveries or explanations (causal inference in this mode is impossible). Tightly associating variables while perhaps capturing a true description of the world still leaves all the explanatory work to be done. This is a key philosophical insight of hermeneutic theorists like Taylor as well as analytic philosophers like Donald Davidson.³⁵ Yet contemporary economists frequently make the philosophical mistake of assuming that non-agential variables are causally explanatory (e.g., inflation, the pricing system, environmental triggers, etc.). The quasi-behavioral appropriation of rational choice is therefore like the strongly predictive project also doomed to fail. This is because the self-interpretive nature of human agency does not universally conform to an *a priori* structure. Humans need not, for instance, prefer things in a predictably complete and transitive manner, because other beliefs might push them in the direction of circularity or lead them to reverse their preferences willy-nilly.

Hermeneutics also gives grounds for rejecting an argument often made by economists that anomalous behavior that breaks away from the parameters of rational choice is “noise” that can be ignored in aggregate by considering the overall statistical propensities and patterns of a population.³⁶ The problem with this argument is that rational choice does not in fact explain anything—not a particular individual’s actions nor those of a group in aggregate. This is because only a narrative capturing the contingent meanings and beliefs of an actor or group explain their actions. Of course, particular, historically situated agents might happen in some cases to approximate or match the structure, but it still is not the decision structure doing the explaining. What does the explaining is a story about the beliefs that were taken on

³⁵ See: Donald Davidson, “Actions, Reasons, and Causes,” in *Essays on Actions and Events*, 3–20.

³⁶ Leonard Lee, On Amir, and Dan Ariely, “In Search of Homo Economicus: Cognitive Noise and the Role of Emotion in Preference Consistency,” *Journal of Consumer Research* 36 (2009): 173; Reiss, *Philosophy of Economics*, 111.

board by particular individuals, groups, and societies. In other words, the common defense of rational choice as a folk psychology is far from a true folk psychology. To the contrary, the idea that humans act for reasons structured by technical axioms in fact mutes and ignores an individual's reasons for acting. The transcendental formalism of rational choice is incompatible with the thick psychological, cultural, and historical realities motivating self-interpreting agents.

3.3 Rational Choice and Neoliberal Ideology

So far I have argued that rational choice fails at the naturalist goal of a hard science. But there are other problems with the absorption of this method into naturalism. One of the primary doctrines of naturalism is the fact-value dichotomy, which holds that scientific theories are purely factual and logically distinct from ideological commitments. In economics this doctrine is echoed frequently as the claim that “positive” economics is “independent of any particular ethical position” and so “an ‘objective’ science, in precisely the same sense as any of the physical sciences.”³⁷ But do naturalist conceptions of rational choice achieve value neutrality? I now want to argue that naturalism links rational choice to a form of technocratic neoliberal ideology. This happens in at least two ways that parallel the above discussion—first by rendering neoliberalized notions of selfhood inescapable features of human agency; and second by buttressing neoliberal technocracy.

I must begin with a brief definition of my political terminology: neoliberalism emerged at the end of the twentieth century as an ideology maintaining markets are natural, spontaneous forms of human association and that the state is “inherently inefficient when compared with markets.”³⁸ Neoliberals advocate shrinking government through austerity, privatization, and revamping public institutions according to market logics. What is less frequently appreciated is that neoliberals present their arguments as justified by a value-neutral science of economics—that is, by a form of naturalism.³⁹

Hermeneutics reveals that when rational choice is presented as a quasi-behavioral psychology it reduces human agency to a highly contestable form of life—a neoliberalized selfhood in which all goods are subject to calculation within a market. Take the “completeness” maxim of rational choice, which assumes that a rational actor can always compare and rank preferences. On this view not being able to consistently rank two or more preferences is by definition irrational.⁴⁰ This assumption entails that all the goods in a human life are in principle reducible to the level of calculation

³⁷ Friedman, “Methodology of Positive Economics,” 146.

³⁸ Mark Bevir, *Democratic Governance* (Princeton: Princeton University Press, 2010) 30. For an overview of this ideological tradition, see: David Harvey, *A Brief History of Neoliberalism* (Oxford: Oxford University Press, 2005).

³⁹ See: Bevir, *Democratic Governance*.

⁴⁰ Gilboa, *Rational Choice*, 39–40.

and exchange—that they are finite and bounded. The aims of human life are analogous to commodities and ordered the same way as market goods in terms of a hierarchy of preferences.

But what is the nature of such a self when socially and psychologically embodied? The hermeneutic psychologist, Philip Cushman, has written extensively on the emergence and proliferation of a particular kind of “empty self” that he associates with a “neoliberal way of being.”⁴¹ This is a form of selfhood is expressive of a culture of mass consumerism, in which human problems are treated as resolvable through the right purchases on a market. According to Cushman, “the empty self is soothed and made cohesive by becoming ‘filled’ up with food, consumer products and celebrities.”⁴² This kind of self views human fulfillment as a form of preference maximizing. But Cushman found that such a self suffered various psychological pathologies as preference maximizing behavior is unable to meet deeper needs for what he calls “mutual recognition.”⁴³ Specifically, treating all human relations as subject to calculable rankings undermines unbounded, infinite sources of allegiance and value—be it to other people, creeds, or communities. In other words, Cushman argues that there is clinical evidence that the attempt to reduce all goods in life to a hierarchy of calculative preferences damages human psychological health.

Whether or not we accept all of Cushman’s conclusions, his psychological research suggests that a self who ranks all goods as-if on a market is only one possibility and not the universal state of human psychology. This fits with hermeneutic claims that there is no necessary, scientifically fixed way of holding preferences. Instead forms of self are expressive of rival self-interpretations and therefore also radically contestable. Indeed, Taylor has argued that moral selfhood is characterized by rival visions of what he calls “hypergoods” or sources of meaning that are of such ultimate, intrinsic value that they cannot be placed on a calculable value scale.⁴⁴ Hypergoods are sources of ultimate value—for example, a sense of human dignity; a role in a family or nation; adherence to a political or religious way of life—that in turn organize the other goods of a particular identity. According to Taylor there is no way to subject hypergoods to the market logic of exchange, pricing, and calculation. To the contrary, hypergoods are such that the very act of haggling corrupts and undermines them, potentially leading to betrayal or a compromised sense of self.⁴⁵ Humans thus live deep tensions between competing hypergoods; say, their love of a family member or friend and their sense of loyalty to a religious or political cause. Individuals can find themselves stuck in tragic binds over these unbounded goods, in which there is no simple way to calculate and rank them.

⁴¹ Philip Cushman: “Relational Psychoanalysis as Political Resistance,” *Contemporary Psychoanalysis* 51:3 (2015): 423; “Why the Self is Empty: Toward a Historically Situated Psychology,” *American Psychologist* 45:5 (1990): 599–611.

⁴² Cushman, “Why the Self is Empty,” 603.

⁴³ Cushman, “Relational Psychoanalysis,” 444.

⁴⁴ Charles Taylor, *Sources of the Self* (Cambridge, UK: Cambridge University Press, 1989) 66.

⁴⁵ Taylor, *Sources of the Self*, 31–32.

If many humans structure their lives around hypergoods then rational choice is a woefully limited and hobbled account of action. This is because the central and most important goods motivating action are often incommensurable, not subject to ranking, calculation, completeness, or transitivity. In fact, human life (including the market) is saturated by claims to ultimate meaning. These meanings are conceived as unbounded and potentially infinitely demanding. For this reason, the completeness maxim, which simply stipulates that there are no incommensurable, unbounded goods is highly problematic. Far from constituting an autonomous, separate zone of inquiry, economics is instead pervaded by a field of hypergoods. Economics, in other words, occurs inside of ethics and politics and cannot be abstracted from it without creating serious misunderstandings and distortions.

Hermeneutics opens the deep historicity of the human subject. The content of human psychology is not ahistorically one of market preference calculation. This fixes one possible life world and obfuscates all the rest. Rather, humans make choices against a backdrop of historical meanings that rational choice erases and mutes in favor of a supposedly logically necessary structure. And yet normally economists think of rational choice as making no substantive assumptions about human psychology. They do this by drawing a distinction between “formal” versus “substantive” theories of rationality—where the former only provides constraints of consistency, the latter imposes substantive values onto the agent.⁴⁶ But the foregoing analysis suggests that the self of rational choice taken as a quasi-behavioral thesis is a substantive subject—a hyper-marketized, neoliberal way of being. As S. M. Amadae has made clear this neoliberalism has in turn generated a game-theory view of society as profoundly non-cooperative and self-interested. This is not an uncovering of the natural, universal mechanics of society but rather a form of “neoliberal subjectivity.”⁴⁷ Rational choice taken as a science of action is thus entangled in a substantive psychology and political theory of society.

In other words, philosophical naturalism offers the exact opposite of a descriptive, value-neutral theory. Instead, a tendentious form of selfhood grounds the basis for neoliberals’ belief that markets are the best way to organize societies. This means naturalist rational choice is often creating a new kind of ideological world, not describing or discovering what is already given. Indeed, naturalist forms of rational choice are closer to purely normative accounts of what rational agents ought to be, akin to a transcendental Kantian subject (albeit a transcendental subject whose moral concerns are far less lofty). Axiomized, thin constructs of a rational subject might chiefly become normative and ideological interventions within the economic and political lives of ordinary citizens. Institutions are restructured, incentives calibrated, in order to form and discipline neoliberal subjects.

The interventionist possibilities of naturalism bring us to the second ideological dimension of these theories—namely their use of a purported predictive science to bolster technocracy. Here the descriptive claim that economics is a science is used

⁴⁶ Reiss, *Philosophy of Economics*, 51.

⁴⁷ S. M. Amadae, *Prisoners of Reason: Game Theory and Neoliberal Political Economy* (Cambridge: Cambridge University Press, 2015) 293.

to construct a form of political authority—that of the technocrat.⁴⁸ Historians have shown that rational choice was self-consciously developed during the twentieth century as a science of society meant to bolster liberal capitalism and fend off Marxist accounts of human agency as based in common class interests.⁴⁹ This led to the radical conceptual revamping of voting, legislation, collective action, consent, and even the social contract itself as all based on hyper-individualistic notions of self-interest.⁵⁰ Today rational choice is used as an interpretive lens for understanding human actions in everything from the high offices of government to the low-level haggling of the market.

This massive, restructuring interpretation of politics in North Atlantic democracies was largely carried out by a new expert class trained in economics departments and business schools, who came to see rational choice as a science.⁵¹ A cadre of neoliberal technocrats maintained a relentless public relations effort to persuade ordinary citizens that they were the masters of a science of wealth creation, innovation, and efficiency. In Britain and America this type of technocracy ascended in the 1970s with the welfare state as its chief foe. Neoliberal technocrats declared the state in crisis and the public sector inherently inefficient—the authority of Keynesian and civil service technocrats was challenged by the new neoliberal technocracy. These new technocrats demanded that the state needed to be rolled back and also engineered into quasi-markets. The social democracies built at midcentury were unremittingly disassembled via austerity, deregulation, contracting out, and privatization. In addition, market discipline through auditing and customerization were imposed on the public sector.⁵²

Public choice economists like James Buchanan proclaimed the discovery of a science of rational choice that also happened to attack social welfare and advance neoliberalism.⁵³ For example, Buchanan used rational choice assumptions to argue that modern citizens had become too “soft” and allowed “parasites” to take advantage of them through welfare programs.⁵⁴ He used a similar rationale to claim that civil servants and bureaucrats necessarily had self-interested incentives to inflate government budgets indefinitely.⁵⁵ His proposed reforms were meant to design institutions from the top-down—especially through constitutional amendments that

⁴⁸ Indeed, technocratic defense rationalists invented modern choice theory by modeling what strategic action looked like between two players competing in a zero-sum, nuclear weapons showdown. Amadae, *Rationalizing Capitalist Democracy*, 75–80, 176–189.

⁴⁹ Amadae, *Rationalizing Capitalist Democracy*, 2–5, 12–14.

⁵⁰ Amadae, *Prisoners of Reason*, Part II.

⁵¹ Amadae, *Rationalizing Capitalist Democracy*, 57–75; Bevir, *Democratic Governance*.

⁵² Bevir, *Democratic Governance*, 28–29, 30–31, 67–75.

⁵³ James Buchanan and Gordon Tullock, *The Calculus of Consent* (Ann Arbor, MI: University of Michigan Press, 1965) 12, 19, 35, 123, 291.

⁵⁴ James Buchanan, “The Samaritan’s Dilemma,” in *Altruism, Morality and Economic Theory*, ed. Edmund Phelps (New York: Russell Sage Foundation, 1975) 75.

⁵⁵ See: James Buchanan, “Why Does Government Grow?” in *Budgets and Bureaucrats*, ed. Thomas Borcherding (Durham, NC: Duke University Press, 1977) 3–18.

limited the will of the majority (Buchanan later worked with the Chilean dictator Augusto Pinochet on a neoliberal constitution).⁵⁶ Yet Buchanan tellingly insisted that his theories were a value-free “science” akin to “the physical scientist” making “progress toward uncovering the laws that govern the natural world.”⁵⁷ In this way austerity, marketization, privatization, and the dismantling of social democracy were all presented as dictates of economic science.⁵⁸

Yet hermeneutics shows how the technocratic authority of neoliberals like Buchanan rests on a misplaced application of ahistorical formalism and mechanistic causality onto the human cultural matrix. This means the kind of knowledge claimed by neoliberal technocrats is a chimera. Neoliberal technocracy is at best a misunderstanding at worst a pseudoscientific power grab. And where neoliberals often present markets and the minimal state as anti-elitist the current analysis suggests the reverse—errant majority decision-making must constantly be corrected from the top down by economic “science.” Neoliberalism shares affinities with a specific form of elite state control.⁵⁹

Economics imbued in naturalism thus cannot uphold a split between so-called “positive” and “normative” branches of the field. The philosophical approach to the positive side of rational choice has normative implications. This implies it is valid to reject naturalist appropriations of rational choice on ethical and political grounds. The values of ordinary citizens can and should be used to contest these supposedly inescapable, “scientific” facts.

3.4 An Alternative Rational Choice

So what is the alternative to rational choice conceived as a value-neutral predictive science? Hermeneutics encourages economists to conceive of rational choice as a heuristic. Heuristics, in this sense, are fruitful ways of generating insights or conclusions about social reality. Heuristics are not to be confused with explanations. Because humans are self-interpreting, explanation requires grasping the actual meanings and beliefs that contingently cause a belief or action. To explain social reality economists must construct narratives or stories. By contrast, a heuristic need neither describe nor explain economic life. Instead, it offers a way of thinking about

⁵⁶ Buchanan, “Samaritan’s Dilemma,” 71, 77–82. For Buchanan’s connection to authoritarianism and Pinochet see, Nancy MacLean, *Democracy in Chains* (New York: Penguin Random House, 2017) 155, 157–164, 168.

⁵⁷ Buchanan and Tullock, *The Calculus of Consent*, 295, see also: xvii, 28–29, 294–296.

⁵⁸ More recently neoliberal technocracy played a major role in the 2008 global financial crisis where idealized models fueled the housing crisis and financial deregulation. See: David Colander, et al., “The Financial Crisis and the Systematic Failure of Academic Economists,” in *The Economics of Economists*, eds. Alessandro Lanteri and Jack Vromen (Cambridge: Cambridge University Press, 2014): 344–360.

⁵⁹ Compare: Milton Friedman, *Capitalism and Freedom* (Chicago: University of Chicago Press, 2002) 15.

society that might help consider possible outcomes and hypothetical scenarios. This means that rational choice must never be confused as having generated an actual account of what particular people are doing or why they are doing it.

Colin Hay helpfully proposes the notion of rational choice as an effective method for considering as-if or hypothetical scenarios. Hay argues that rational choice is at best a tool for generating “hypothetical thought experiments” that pose the question “what if the world were like this?”⁶⁰ For example, rational choice can illuminate the vicious consequences of social situations where agents have a tendency to think in very limited strategic and self-interested ways. In an age of neoliberalization, rational choice scenarios might issue forth political warnings about the destructive consequences of treating shared goods like education and the environment in certain ways. As Hay notes, this use of rational choice does not claim a “predictive hypothesis” but only an admonition or “precautionary political warning.”⁶¹ Self-interpretation implies that the world might become more this way even if it is unlikely to ever become fully neoliberalized.

Clearly hermeneutics dethrones the centrality of rational choice in economics. Hermeneutics implies the use of this method depends on judgment in context and certain rules of thumb. First, economists must be sensitive to the way rational choice as a kind of formalism drops the cultural and ethnographic content of economic reality. This idealized formalism means rational choice is more useful when thinking about contexts where human beings are already engaged in strategic reasoning. In particular, heuristic rational choice is more likely to yield insights in cases where individuals have an opportunity at iterative, calculative repetition, weighing what they view as bounded, commensurable goods. One modern practice where agents regularly approximate this kind of strategic subjectivity is the mundane market—another is gaming and gambling. Of course, this does not mean humans even in these scenarios are ever simply rational choice beasts. It just means that certain cultural practices come closer to fitting the model than others and so are more likely to be relevant to a heuristic analysis. This echoes what early critics of rational choice like Donald Green and Ian Shapiro have referred to as a domain limited approach.⁶² It also goes a long way towards explaining the strange mix of successes and failures of rational choice economics as a discipline. Too often economists have jumped to naturalist conclusions about modeling, instead of seeing it as an only sometimes useful way of thinking about social contexts where strategic agency is being continually repeated and practiced.

For the same reasons rational choice can also be wildly misguided when considering human action that is heavily imbued in hypergoods or those ultimate goals and meanings of human life that are not reducible to rational calculation. Perhaps this is why naturalist rational choice has always seemed more plausible for researchers

⁶⁰ Colin Hay, “Theory, Stylized Heuristic, or Self-Fulfilling Prophecy? The Status of Rational Choice Theory in Public Administration,” *Public Administration* 82:1 (2004): 55.

⁶¹ Hay, “Theory, Stylized Heuristic, or Self-Fulfilling Prophecy?” 57.

⁶² Green and Shapiro, *Pathologies of Rational Choice Theory* (New Haven: Yale University Press, 1994) 54.

whose primary goal is thinking about mundane commodities, markets, and money than for those thinking about political realities like justice and identity or psychological ones like recognition and self-actualization.

This brings me to a second rule of thumb. Economists who actually wish to understand and explain economic reality will often need to reintegrate heuristic rational choice into a wider research program that incorporates the ethnographic, cultural, and historical knowledge of the other social sciences. To escape naturalism economists must take an interpretive, historicizing turn, transforming economics into a culturally aware discipline. This means either achieving dexterity in multi-methods or cooperating with other researchers whose specialty is, say, long-form interviews, languages, history, area-expertise, and ethnographic observation. Obviously there are practical obstacles to training cultural economists of this kind in the near future. Moreover, many of the most sophisticated methods require years of training. Hermeneutic, cultural economists might therefore try drawing on the enormous amount of existing research in the social sciences in order to repurpose it in light of new research questions and method expertise. In a sense, naturalist research might be harvested and reconfigured for hermeneutic ends.

Ultimately, hermeneutics encourages economists to become more aware of the way economies are always immersed in meanings, history, politics, and culture. Recognizing the historicity of human agency has the effect of de-naturalizing all versions of *homo economicus* and the transcendental pretensions of rational choice. This opens the door to understanding that there have been multiple, rival economies across history and not simply one universal economic mechanics or logic of action underlying all social arrangements.⁶³ Thus, hermeneutic economics opens the door to the great historical sociologies of the nineteenth century—albeit now free from naturalist distortions and recognizing more deeply the contingency of history. This means that historical sociology and economics might reunify. Indeed, the first seeds of such a transformation are evident in some prominent economists' frustration with the neoclassical craze for formal modeling free of greater historical awareness.⁶⁴

Finally hermeneutics implies that economists should renounce the use of rational choice as a technical jargon that props up technocratic authority and status while excluding ordinary citizens. Economists cannot legitimately claim to have discovered the single scientific discourse for capturing economic reality that somehow overrides the concerns of ordinary citizens. Once naturalism is rejected the role of the economic expert is to offer theories, facts, insights, and knowledge that aid in the deliberative process amid ordinary citizens. There is no scientifically necessary or inescapable feature of economy that must be adopted by a given society. Rather,

⁶³ Cf., Karl Polanyi, *The Great Transformation*, 48–52.

⁶⁴ Ha-Joon Chang's call for an "historical approach to economics" is an example even if he slides into naturalism when he uses history to construct inductive, midlevel generalizations and reifies institutions. Ha-Joon Chang, *Kicking Away the Ladder: Development Strategy in Historical Perspective* (London: Anthem Press, 2002) 5–8. A similar point could be made for the important work of Thomas Piketty in *Capital*, trans. Arthur Goldhammer (Cambridge, MA: Belknap Press, 2014).

ordinary citizens are perfectly philosophically justified in contesting the meaning of economic theories and the heuristic uses of rational choice. Economists have no special authority over the meaning or significance of the economy nor do they have any special predictive or prophetic powers. Rather, they must join fellow citizens in the process of contesting and debating the moral, ethical, and economic significance of, say, deregulating a market or enacting a particular policy. And they must come to terms with being much more on the same level as others vis-à-vis generating informed guesses about the economy.

Economists, in other words, should reconceive their role along a humanistic not a scientific authoritative model. Humanists advise and offer understanding and knowledge of a subject area, without making the naturalist move of claiming scientific status and epistemic authority over all other participants. Humanists are sensitive to the contingency and contestability of social reality, they recognize the creative agency of human beings, and they are aware of the radical changes and ruptures that mark human history. Where the naturalist technocrat claims science dictates how things must be, the humanist recognizes the role of creative agency and the need to dialogue over contested beliefs, practices, and institutions. Thus, economics becomes humanistic when it abandons its top-down view of the discipline, and sees itself instead in deliberative dialogue with the society it inhabits.

In sum, hermeneutics urges an end to philosophical naiveté in the use of rational choice. Any researcher who picks up this tool in light of the foregoing discussion must ask careful questions like: What are the ideological ramifications of modeling social, economic, and political reality in this way? What are the limits of this idealized formalism? What does it hide from the economic situation in question? What does it illuminate? How do the formal heuristic elements relate back to the world of thick cultural and historical selves? Is there a good fit with the self-interpretive situation of the particular individuals being studied? How does the articulation of economic theory fit inside a particular historical instantiation of economy and the meanings and practices of a given community?

Because economics is not conducted in isolation from its object of study, economists must remain aware of the role economic discourse has in shaping political reality. To study economics always has the potential to reshape economies in non-trivial ways. The case of Marxist economics is dramatic in this regard but at this point so should neoliberal uses (and the same point holds for Keynesianism, mercantilism, and every other kind of economic paradigm). The study of economics is itself a practice internal to economic and political reality.

This means that economics needs to be put back into politics. A particular danger unleashed by the influence of naturalism is that economics becomes an “imperial” discipline with no desire to learn from sociology, anthropology, or politics let alone the humanities, philosophy, literature, and history.⁶⁵ Unfortunately, at present the influence is very much in the other direction, with naturalist conceptions of rational choice colonizing disciplines as varied as political science and evolutionary biology.⁶⁶ Yet even to have spent a few hours thinking about economics philosophi-

⁶⁵ Cf.: Caldwell, “Economic Methodology,” 11.

⁶⁶ For example: Mayhew, *Congress*; Amadae, *Prisoners of Reason*, 247–268.

cally—opening a dialogue between philosophical reflection and economic thought—marks the small stirrings of a new discipline. Whether a new, hermeneutic economics emerges from such considerations depends solely on the creative energies of future thinkers.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 4

Economics, Agency, and Causal Explanation



William Child

Abstract The paper considers three questions. First, what is the connection between economics and agency? It is argued that causation and explanation in economics fundamentally depend on agency. So a philosophical understanding of economic explanation must be sensitive to an understanding of agency. Second, what is the connection between agency and causation? A causal view of agency-involving explanation is defended against a number of arguments from the resurgent noncausalist tradition in the literature on agency and action-explanation. If agency is fundamental to economic explanation, it is argued, then so is causation. Third, what is the connection between causal explanation and the natural sciences? It is argued that, though the explanations given in economics and other social sciences are causal explanations, they are different in kind from the causal explanations of the natural sciences. On the one hand, then, the causal explanations of the social sciences are irreducible to those found in the natural sciences. On the other hand, the causal relations described by the social sciences are not completely autonomous; they do not float free of, or operate independently from, the causal relations charted by the natural sciences.

4.1 Economics and Agency

I shall start by defending the thesis that explanation in economics always depends on, or is mediated by, phenomena that essentially involve agency. When we spell out the causal relations that are charted in economics, I shall argue, we can see that the mechanisms by which economic causes produce their effects are always, ultimately, dependent on human agency.

In some cases, the dependence of economic explanations on agency is very obvious, because what is being explained is itself an action or a set of actions. If we

W. Child (✉)
University College Oxford, Oxford, UK
e-mail: bill.child@univ.ox.ac.uk

explain why British people generally save less than Japanese people, for instance, we must explain why certain people make certain choices or act in certain ways. And we cannot do that without employing the concepts of agent and action.

In other cases, considerations of agency are not so close to the surface. Thus, there are economic explanations in which neither the explanandum nor the explanans directly mention actions or agents. When we spell out the mechanism by which the effect was produced in such cases, however, agency quickly comes into the picture. Consider, for instance, the claim that inflation in England in the 1580s was caused by an increase in the money supply. That claim makes no immediate reference to any human actions or any instance of agency. But when we ask how, or why, inflation was caused by an increase in the money supply it is easy to see that the explanation depends on agency-involving phenomena. To simplify: the influx of gold from the Americas led to an increase in the amount of gold in the hands of consumers and, thereby, increased the money supply; when shopkeepers tried to raise the prices of goods, therefore, customers were able to pay the higher prices rather than deciding not to purchase the goods at all, as they would previously have done; so the prices that people set and paid for goods generally increased; which is to say – there was inflation. Or again, take the claim that the UK economy shrank slightly in the second quarter of 2018 because there were severe storms in April. Again, that explanation does not itself mention actions or agency. But we see how the explanation depends on facts about actions and agency when we spell out how it was that the bad weather caused the economy to shrink. Because of the storms, many people decided to stay at home and not to buy the goods and services they would otherwise expect to buy in the spring: summer clothes, barbecues, sports equipment, outdoor holidays, and so on. So firms did not sell as much as they had planned, and the overall level of economic activity decreased: the economy shrank.

That is the general pattern. It is worth mentioning two apparent exceptions: cases where there might seem to be explanations in economics that are not underpinned by agency. The first is game theory – which is, of course, a major area of activity within the discipline of economics. Explanations within game theory do not mention human actions at all. Game theory is the study of pure rationality: what it would be rational for an agent to believe, or prefer, or choose in such-and-such circumstances, given specified information, aims, and preferences. That is a purely abstract, theoretical enterprise. Explaining why such-and-such an action would be rational in such-and-such circumstances is a genuine form of explanation. But it has nothing to do with causally explaining anything. Nor, *a fortiori*, is there any question of spelling out any mechanisms that underpin such explanations and revealing them as involving instances of agency. When we apply the theoretical structure of game theory to actual cases, however, assuming that real people are more or less rational, and using the theory to predict or explain their behaviour in situations of choice, then we are once more dealing with agency and our explanations will be underpinned by facts about agents and their actions. The claims of game theory itself, then, are not concerned with action or agency. But that is not a counter-example to the general claim that economic explanation essentially

depends on agency-involving phenomena. For, though game theory itself does not aim to explain any actual phenomena, when the conclusions of game theory are applied to real-world phenomena, the resultant explanations do invoke and depend on agency.

A second apparent exception to our general principle relates to automated buying and selling. Suppose that the stock market fell because there was a sudden worsening in the exchange rate. What was the mechanism that produced that effect; how did the worsening in the exchange rate bring about a fall in the stock market? Before the advent of automated trading, the mechanism would, as we have said, have involved human agency: people making decisions to buy and sell, in the light of their knowledge, expectations, and preferences. But in circumstances where changes in the stock market are produced by high volumes of automatic trades, a change in the exchange rate can trigger a fall in the market by triggering mass automated selling, without the involvement of any process of reasoning or agency at all. To that extent, some of the causal processes studied by economics can operate without any process of human agency. But that phenomenon does not undermine the general principle that causation and explanation in economics depend on agency. In the first place, there could not be an economy that worked entirely automatically, with every instance of buying and selling operating independently of any human agency. If it is to count as part of an economy at all, the structure of automated dealings must at some point make contact with real, human decisions and actions. In the second place, agency is involved in the design and implementation of the programmes that carry out automated deals. So part of the answer to the question, why the drop in the exchange rate caused a fall in the stock market, involves an appeal to agency: individuals and organizations programmed their trading systems to behave in exactly that way.

Explanation in economics, then, essentially depends on agency.

4.2 Agency and Causation

Many philosophers have been attracted by the idea that explanation in the social sciences is radically different from explanation in the natural sciences. Some of those philosophers have held that the essential difference between the two is that, whereas the natural sciences offer causal explanations, the social sciences offer interpretative or hermeneutic explanations which, it is said, are non-causal. And, in the light of Sect. 4.1, one way to develop that position is to argue that explanation that appeals to agency is a form of non-causal explanation. I shall argue, against that view, that there is no tension between agency and causation. Agency is a causal phenomenon. And explaining an action in terms of the agent's reasons for performing it is a kind of causal explanation. I shall defend that view against a number of arguments from the non-causalist tradition. If agency is fundamental to economic explanation, as I have argued in Sect. 4.1, then so is causation.

Many different positions have been associated with the phrase ‘causal theory of action’. And the claims that are rejected by critics of causalism often turn out to be different from those that are defended by causalists themselves. So it is important to be clear about what a causal view of action or agency involves. The central claim, as I understand it, is that reason explanation is a form of causal explanation; explaining an action by citing the agent’s reasons for performing it is a way of causally explaining the action. (The same applies, *mutatis mutandis*, for explaining an agent’s beliefs or intentions by citing her reasons for forming or retaining them.) Causalists about action are often taken to be making a different and more ambitious claim: that we can give a non-circular analysis of what it is for something to be an action in terms of the concepts of bodily movement, belief, desire or pro-attitude, and causation. But that is no part of the causal theory as I understand it. Suppose that Sarah raised her arm. The central claim is, as I have said, that explaining why Sarah raised her arm by citing her reason for doing so is a form of causal explanation. The aim of the causal theory is to understand the form of that kind of explanation; it is not concerned with analysing the phenomenon of action into simpler parts. So arguments against the possibility of a reductive account of what it is for something to be an action are not in themselves arguments against a causal view of reason-explanation. The causal theorist can accept that ‘Sarah raised her arm’ and ‘Sarah’s action of raising her arm’ are basic and unanalysable.

I shall defend the causal view of reason-explanation against two criticisms that are prominent in the non-causalist tradition. The first is that there is no good reason to accept the causal view and, in particular, that Donald Davidson’s classic argument for a causal view of action is ineffective.¹ The second is that the causal view cannot be true because the character of reason explanation is inconsistent with the requirements of causal explanation. I will respond to those criticisms in turn.

4.2.1 Defending the Basic Argument for a Causal View of Reason-Explanation

What is Davidson’s argument for a causal view of reason-explanation? Here are two famous passages from ‘Actions, Reasons and Causes’:

a person can have a reason for an action, and perform the action, and yet this reason not be the reason why he did it. Central to the relation between a reason and an action it explains is the idea that the agent performed the action *because* he had the reason (Davidson 1963, 9).

One way we can explain an event is by placing it in the context of its cause; cause and effect form the sort of pattern that explains the effect, in a sense of ‘explain’ that we understand as well as any. If reason and action illustrate a different pattern of explanation, that pattern must be identified (Davidson 1963, 10).

¹For Davidson’s argument, see Davidson 1963.

We can summarize Davidson's point like this. We can explain S's ϕ -ing by citing her reason for ϕ -ing. Suppose R was S's reason for ϕ -ing. Then we can say that S ϕ -d because R, or because she had R. To understand what kind of explanation that is, we need to understand the force of the 'because' that it contains. And reflection shows that there is no serious alternative to understanding the 'because' as a causal 'because'. Explaining an action by citing the agent's reason, then, is a form of causal explanation.

A natural response to Davidson's argument is to wonder whether he has given us any reason at all for thinking that the 'because' in reason-explanation has to be a causal 'because'. After all, 'because' is often used in ways that have nothing to do with causation. Consider the following cases: '7 is a prime number because it is divisible only by 1 and itself', 'The striker was offside because there were no defenders between her and the goalkeeper when the ball was passed', 'Mary is my sister-in-law because she's married to my brother'. In each of those sentences, 'because' occurs in an explanation. But none of those explanations involves causation, and none of those instances of 'because' is a causal 'because'. So what is it about the 'because' in reason-explanation that justifies the claim that it is a causal 'because'? To answer that question, we need to consider what it is that a reason-explanation explains. When I explain why someone did something, I am explaining a change or occurrence in the natural world: I am explaining why something happened; why an event occurred. And to explain why something happened, or came about, is in its nature to give a causal explanation. Non-causal uses of 'because' can explain many things. They can explain why a number – something that has no causes or effects – is a prime number, by showing how it meets the definition of a prime; they can explain why a player, whose existence and position on the field are taken as given, counts as being offside, by showing how she meets the conditions for being offside; they can explain why a person, whose existence and marital status are taken as given, counts as my sister-in-law, by showing how she meets the conditions for being my sister-in-law (she qualifies as my sister-in-law by being married to my brother, rather than by being my wife's sister). But a non-causal explanation can never explain why something happened or came into existence, or why something changed in the world. To explain those things, we need a causal explanation.²

With that in mind, we can consider a particular version of the current objection to the argument for causalism. It is sometimes claimed that Davidson's argument for causalism depends on the gratuitous assumption that there is only one kind of explanation: causal explanation. And, it is said, there is no reason to accept that. So we can respond to Davidson's argument by insisting that, contrary to what he claims, there is an alternative, non-causal way of understanding the 'because' in reason-

²There may be events or changes that have no causes. By they are not counterexamples to the general principle that explaining why something happened or occurred requires a causal explanation. If an event has no causes, then its occurrence cannot be explained in some non-causal way; on the contrary, it cannot be explained at all.

explanation; it can be understood in terms of the primitive, unanalysable relation of acting for a reason. Here are two examples of that kind of argument:

The most direct response to Davidson [is] that the difference between those reasons for which the agent did in fact act and those for which he might have acted but did not is not a difference in causal role at all. It is just the difference between the considerations in the light of which he acted and other considerations he took to favour acting as he did but which were not in fact ones in the light of which he decided to do what he did. (Dancy 2000, 163)

Pressed to state in other terms the difference between having a justification for acting and acting on it, one reply is that it can't be done and doesn't have to be. Thus Dancy takes the relation, *on the ground that*, as primitive. Davidson gives no argument against this. (Setiya 2009, 145)

The suggestion, then, is that we can simply treat the 'because' in 'She ϕ -d because she had R' as picking out a primitive relation. So Davidson's grounds for a causal view of reason explanation 'are radically inadequate' (Wilson 1985, 40); there is no case for causalism at all. But this line of response to Davidson's argument is completely ineffective. It misunderstands the structure and force of the argument.

The causalist about reason-explanation can agree that we can think of the 'because' in 'She ϕ -d *because* she had reason R' in terms of the relation of *acting for a reason* or *acting in the light of a reason*. If she ϕ -d because she had reason R, then it is indeed true to say that she ϕ -d *in the light of* reason R or, more simply, *for* reason R. The causalist can also agree that the relation of *acting for a reason*, or *in the light of a reason* is primitive; it cannot be reductively analysed in other terms. But none of that does anything to weaken the force of the argument for causalism. The fundamental point of the argument is this: reflection on the form of reason-explanation shows that the 'because' in 'She ϕ -d because she had R' is a causal 'because' because of what a reason-explanation is explaining: namely, a change or occurrence in the natural world. And that fundamental point is untouched by these arguments. The critics insist that the relation between a reason and an action that is explained by that reason is simply the relation of acting *for* the reason. But that is not an alternative to the causal view. For the force of Davidson's original argument still applies to the relation of *acting for a reason*. The argument shows that the notion of acting *for reason R* is itself a causal notion; and explaining an action by citing a reason *for which* the agent acted is itself a form of causal explanation.

Of course, as I have said, not every 'because' is a causal because, and not every kind of explanation is a form of causal explanation. So in some cases, it would be perfectly correct to respond to an argument with the superficial appearance of Davidson's argument by insisting that the 'because' in question marks a primitive, non-causal relation. Suppose I am considering possible reasons for the British Museum to return the Parthenon Marbles to Athens. Here are three such reasons:

- (i) The British Museum damaged the Marbles in the 1930s by inappropriate cleaning;
- (ii) The people who sold the Marbles to Lord Elgin were colonial rulers, not Greeks;
- (iii) Athens would be a culturally and physically superior site for displaying the Marbles.

Suppose that I have the following views. The Marbles should be returned to Athens. Each of (i)–(iii) is true. Each of (i)–(iii) gives a genuine reason for returning the Marbles; they all count in favour of returning the Marbles. But, all things considered, the reason why the Marbles should be returned is (iii), and not (i) or (ii): it would still be right to return the Marbles even if (i) and (ii) were false; and it would not be right to return the Marbles if (iii) were false, even if (i) and (ii) were still true.

Now, we can ask what the difference is between ‘The Marbles should be returned *and* Athens would be a better site for displaying them than London’ and ‘The Marbles should be returned *because* Athens would be a better site for displaying them than London’; what is the force of the ‘because’ in the latter proposition? In this case, it would plainly be wrong to conclude that the difference between ‘p *and* q’ and ‘p *because* q’ is a causal difference. Correspondingly, it is quite correct to say that the ‘because’ in ‘The Marbles should be returned *because* Athens would be a better site for displaying them’ expresses a primitive, non-causal relation: the relation of being *the reason why* something should be done. The reason for saying that the ‘because’ in this case is non-causal is, as before, the nature of what it is that is being explained. When I explain why the Parthenon Marbles should be returned to Athens I am not explaining why something happened; I am not explaining the occurrence of anything: a change or an event in the world. I am explaining why a particular course of action would be right. And to do that is not to give a causal explanation.

The intuition behind the response to Davidson’s argument that is given by Dancy and others is that the difference between ‘p *and* q’ and ‘p *because* q’ is not always a causal difference. But, as I have said, the defender of Davidson’s argument agrees with that point. Whether or not a particular ‘because’ is a causal one depends on the kind of explanation involved in the particular case. It seems clear that, when we explain someone’s doing something by citing her reason for doing it, we are explaining why something happened, or why an event occurred. And to do that is to give a causal explanation.

4.2.2 *The Many Faces of Causal Explanation*

Many writers in the non-causalist tradition have argued that reason-explanation cannot be a form of causal explanation because it does not fit the pattern of explaining something in terms of its causes. I give three examples of arguments of this kind.

Reasons for actions are not mental states, like beliefs, desires, intentions and so forth. Rather, they are facts about the non-mental world. So they are the wrong sorts of things to be the causes of action.³

³For the view that an agent’s reasons for acting as she does are typically facts about the external world, rather than mental states, see e.g. Alvarez 2010, ch 6. Alvarez herself remains neutral about the implications of her view of reasons for the thesis that reason explanation is a form of causal explanation (see Alvarez 2010, pp. 199–200). But others are less cautious, and take the ‘fact’ view of reasons to be incompatible with a causal view of reason-explanation.

Reason-explanation works by making actions intelligible by putting them in a context that makes sense of them. It does not work by citing the causes of actions.⁴

To explain S's ϕ -ing by citing her reason for ϕ -ing is to give the purpose with which she ϕ -d: to say what she was trying to do by ϕ -ing. But the purpose with which someone acted is not the cause of their acting. Suppose that Bob stood on his head in order to impress Claudia. Impressing Claudia was his purpose in standing on his head; he was trying to impress Claudia by standing on his head. But his trying to impress Claudia *just was* his standing on his head; it was not something else, which preceded his standing on his head. So it could not be the cause of his action.⁵

These anti-causalist objections, and others like them, I shall argue, work by assuming a very restricted conception of the form that a causal explanation can take. In particular, they take a very restricted view of what reason-explanation would have to be like in order to be a form of causal explanation. Very crudely, they assume that something will only qualify as a causal explanation of an action if it works by citing states or events within an agent that cause the movements of her body. But causal explanations are much more various than that. Once that point is appreciated, we can see that much of what the critics have urged about the ways in which reason-explanations work is in fact perfectly consistent with a causal view of reason-explanation.

Causal explanations in general can take many forms. To give a causal explanation is just to tell, or suggest, a causal story that helps to make something intelligible. But there are many ways in which we can tell a causal story. And something can be a causal story without explicitly citing the cause of an effect. We can illustrate that point with an example. A wine glass breaks. We ask why it broke. Here are some possible answers:

- (a) It was hit with a hammer
- (b) It was fragile
- (c) It was made of inferior material
- (d) The soprano's voice was so piercing
- (e) The washing-up water was too hot.

Those are all, in the relevant circumstances, legitimate explanations of the glass's breaking. And they are all causal explanations. But they take different forms. In (a), an event (the glass's breaking) is explained by citing its cause (its being hit by a hammer). In (b) and (c), we are not told exactly what it was that caused the breaking of the glass. Instead, we are told about a property of the glass (it was fragile; it was made of inferior glass). That explains the effect (the glass's breaking) by making it intelligible why some event (which is not specified, and may be unknown) caused the effect that it did. If the glass had not been fragile, or had not been made of inferior material, the event that actually caused it to break would not have done so. In (d) and (e) we are told about a feature of the cause, or of the

⁴For a recent statement of this position, see Tanney 2009. As Tanney says, this kind of view goes back at least to Ryle 1949, as well as to Anscombe 1959, Melden 1961, and other work in the Wittgensteinian tradition of the 1950s and 1960s.

⁵For this argument, see McLaughlin 2013.

surrounding circumstances, which, as before, explains the event by allowing us to see why the cause had the effect it did. In (d), it is implicit that the cause of the glass's breaking was the soprano's singing; but if her voice had not been so piercing, her singing would not have had the effect it did. In (e) it is implicit that the cause of the breaking was the glass's being put in the washing-up bowl; but if the water had not been as hot as it was, that cause would not have produced the effect it did.

In the same way, an explanation of an action may take many forms, consistent with its being a causal explanation. We can illustrate that point in connection with two of the suggestions about forms of reason-explanation that were mentioned above.

Suppose we agree that reason-explanations often explain actions by citing facts about the external world, rather than by mentioning the agent's mental states. Thus, for example, she climbed out of the window *because the door was locked* (not: because she believed that the door was locked). He chose the salad *because it was healthier than the other options* (not: because he believed it was healthier). The facts that are appealed to in those explanations are not facts about the agents' mental states. But that is no barrier to understanding these explanations as causal explanations. For facts about the non-mental world can causally explain agents' doing things. Take the case where someone climbed out of the window because the door was locked. That explanation does not tell us what the triggering cause of her action was. But it tells us something about the context that makes it intelligible why the triggering cause, whatever it was, produced the kind of effect it did: an act of climbing out of the window, rather than an act of leaving by the door. Suppose that the triggering cause of her leaving the room was, say, her noticing that it was 3 pm, or her receiving a message asking her to come to a meeting. We cite the fact that the door was locked to explain why she left the room through the window rather than the door. And that is a causal explanation. The fact that the door was locked only explains her climbing out of the window if she climbed out of the window *because the door was locked*. (Suppose she would have climbed out of the window anyway, whether or not the door was locked. Then it is not true that she climbed out of the window because the door was locked.) And there are the usual reasons for saying that that 'because' is a causal one.

Similarly, the causalist will accept that when we explain an action by giving the agent's reasons for performing it, we are explaining it in terms of the agent's aim or purpose. 'He was trying to impress Claudia', or 'In order to impress Claudia', are legitimate answers to the question, 'Why did he stand on his head?' They explain why he did what he did. But, the causalist will say, that is entirely compatible with the claim that reason-giving explanation is a form of causal explanation. As before, neither 'He was trying to impress Claudia' nor 'In order to impress Claudia' tells us the triggering cause of his action of standing on his head. But they do tell us something about the mental properties of the agent: that he intended to impress Claudia and believed that standing on his head would be a way of doing so. And that makes it intelligible why the triggering cause of his action, whatever it was (for instance, noticing Claudia approaching), produced the effect it did: an act of standing on his

head. That is one kind of causal story that makes the occurrence of an action intelligible; it is a way of causally explaining the action.

4.2.3 Conclusion

We have found no reason for abandoning the view that explaining an action by giving the agent's reasons for performing it is a form of causal explanation. In the first place, the suggestion that we can simply bypass Davidson's argument for causalism by treating the relation of *acting for a reason* as a basic, non-causal relation fails to understand the force of the argument. There may be good reasons for treating *acting for reason R* as a basic, unanalysable relation. But if we do, we must recognize that the notion of acting for a reason is itself a causal notion. In the second place, the suggestion that reason-giving explanations do not fit the template of causal explanation fails to appreciate the diversity of the kinds of causal explanation. I conclude that there is no prospect of distinguishing explanation in economics (and other social sciences) from explanation in the natural sciences by accepting (i) that economic explanation depends on agency, but holding (ii) that explanations that appeal to agency are not causal explanations. Social scientific explanation is indeed different in kind from natural scientific explanation. But the reason for that is not that social scientific explanation is non-causal. It is, rather, that social scientific explanation is grounded in *reason-giving* causal explanation rather than non-reason-giving, merely physical, causal explanation.

4.3 Causation in the Social Sciences and in the Natural Sciences

It is sometimes said that the causal relations of the natural sciences are a misconceived paradigm for the social sciences. If we treat the social sciences as a form of causal science, it is said, we reduce human beings to natural objects; we drain human agency of moral content. And that distorts and falsifies the ontology of agents and actions that is fundamental to the social sciences. I agree that explanation in the social sciences is indeed distinctive; it is different in kind from explanation in the natural sciences. But that does not prevent social-scientific explanation from being a form of causal explanation; that was the burden of Sect. 4.2, above. Nor does it mean that the phenomena and causal processes that are studied by the social sciences are completely independent of the phenomena and causal processes that are the subject-matter of the natural sciences. That is the position I defend in the current section.

On the one hand, then, the phenomena and the explanations of the social sciences are irreducible to the phenomena and explanations found in the natural sciences. On the other hand, there is a close relation between the causal phenomena studied in the

social sciences and the causal phenomena studied by natural sciences. We can put that relation in terms of determination. The low-level physical facts determine all the facts. If you completely determine the physical make-up of the world, and the natural scientific laws, then you completely determine all phenomena: including all the mental, moral, economic, sociological etc. properties of human beings. The same is true for causation. If you completely determine all the low-level causal facts, you determine all the causal facts: including the causal facts about agents and actions.

What should we say about the complaint that construing the social sciences as a form of causal science involves reducing human beings to natural objects? Well, human beings *are* natural objects. We are completely composed of matter, whose behaviour, at the micro level, is no different from the behaviour of other bits of natural stuff and can in principle be completely described and explained without appeal to the concepts of agency or reasons for acting. Our view of agency and reason-explanation must not conflict with that truism. But the truism does not imply that everything about human beings can be understood if we describe people at the level of their minute physical composition. It cannot be. Recognizing the distinctiveness of agency and reason-explanation, however, is consistent with accepting, first, that human beings are natural, material beings, and second, that fixing the low-level causal facts fixes the facts about agency. In that sense, the causal facts about agency are not independent of lower-level causal facts.

What should we say about the complaint that treating the social sciences as a kind of causal science distorts or falsifies the ontology of agents and actions? The mere idea that reason explanation is a form of causal explanation has no revisionary effects on the ontology of social science. The causal stories involved in accounts of agency, and in social science more generally, are stories about agents doing things and making choices in the light of their beliefs, preferences, circumstances, and so on. And, as I argued in Sect. 4.1, the causal mechanisms that underpin economic explanation essentially involve people. There is nothing distorting or revisionary in that. But will there be revisionary consequences for the ontology of social science if we also accept, as I have suggested, that all causal facts are determined by the low-level facts of physical causation? That depends on how exactly the relations between levels are claimed to work.

On one view, the ontology of the social sciences is the same as the ontology of the physical sciences. Causal reality, on that view, comprises a network of causally inter-related events, which constitute the common ontology of all human enquiry. Every causal story picks out and describes events that belong to this common, topic-neutral stock of events. But causal stories in different domains – the domain of the natural sciences and the domain of the social sciences, for instance – pick them out in different ways and organize them in different classes or patterns. So, in particular, though the descriptions and explanations of common-sense psychology or the social sciences cannot be reduced to those of the natural sciences, they pick out the very same events. In that sense, the social sciences and the natural sciences have a common ontology.⁶

⁶That is essentially the picture offered by Davidson's anomalous monism (see Davidson 1970, 1973, 1974).

But that is not a plausible view of the relation between the different domains. Consider the relation between the mental and the physical. It seems clear that individual mental events, of the kind described in the vocabulary of common-sense psychology, are not identical to individual microphysical events, of the kinds described by neuroscience, or chemistry, or physics. For particular microphysical events are much smaller, both spatially and temporarily, than any individual mental event. A natural response to that point is to say that mental events are *composed of* microphysical events; wherever there is a particular mental event, we could in principle assemble a collection of microphysical events that collectively occupies the same spatio-temporal region as that mental event. That is true. But it does not save the picture of reality as comprising a single, topic-neutral stock of events that are common to all areas of enquiry. For the collection of microphysical events that jointly compose a particular mental event is not itself an individual physical event: there is no natural way of picking it out in the vocabulary of the physical sciences; it is of no particular interest from the point of view of the physical sciences; it does not figure in those sciences as a cause or effect. The existence of mental events does not require the existence of anything more than the totality of microphysical events. But, for the reasons just given, that does not mean that common-sense psychology and the physical sciences have a common ontology of events. The same goes for the social and human sciences more generally.⁷

We need a looser picture of the relation between different kinds or levels of description and explanation. Here is how I think things work. There are numerous different levels of description that we can occupy when we describe the world: the level of fundamental physics; the level of chemistry; the levels of psychology, economics, and so on. And there are numerous different causal explanations, which we can give when we occupy these different levels. Suppose, for instance, that someone buys a ticket from a parking machine. We can tell a physical story about the causes of the movements of her body that are involved in her doing what she did. And we can tell a common-sense psychological story about her acting for a reason. Those are both causal explanations. But they explain different things. The physical story explains the motion of a bit of a human body. The psychological story explains a person's doing something. But though the two causal stories are different, the psychological story is not completely independent of the low-level causal processes described by the physical story. When someone buys a ticket in order to park their car, physical causal processes go on in their body. If those physical process did not take place, there would be no action of buying a ticket. Mental causation, as we have said, does not float free of physical causation.

But if we insist that all the causal facts are determined by the low-level physical causal facts, don't we have to say something more about the relation between high-level and lower-level causal stories: about how exactly the two stories marry up? Some have argued that there can only be genuine causal explanations in psychology and the social sciences if there are detailed, law-like correlations between the

⁷For the argument summarized in this paragraph, see Hornsby 1980–81.

higher-level phenomena identified by psychology or social science and the low-level, microphysical phenomena that determine the higher-level causal facts. If there are no such law-like correlations between the different levels, it is argued, then the explanations offered by psychology or the social sciences will be no better than the pseudo-explanations put forward by astrology or homeopathy. They will be stories that we tell ourselves: ways we have of trying to make sense of the world. But they will have nothing to do with why things really happen; they will not be true, or correct, causal explanations.

A response to that thought must address two questions. First, what reason is there to think that explanation in psychology and the social sciences is a genuine form of causal explanation? Second, what must be true of the relation between higher-level facts and lower-level facts in order for it to be true that the higher level facts are causally explanatory? I take those points in order.

We can have good reasons for thinking that psychology and the social sciences give genuine causal explanations, while astrology and homeopathy do not, without knowing anything in detail about any specific correlations between psychological or social scientific facts, on the one hand, and lower-level physical facts, on the other. In the first place, psychology and social science are successful explanatory practices, whose claims meet normal standards of explanatoriness. They put forward substantial causal claims whose truth is susceptible of confirmation or disconfirmation on the basis of experience. And importantly, explanations in psychology and social science do not compete with lower-level, physical causal explanations; they operate at a different level and have different explananda. So the phenomena they purport to explain, which have to do with agents' choices and actions, are not better explained in more basic, lower-level terms: on the contrary, they cannot be explained at that level at all. That contrasts with the purported explanations of astrology or homeopathy, which do compete with other explanations. Astrology aims to explain a person's characteristics and the success or failure of their relationships and career by reference to the positions of the planets. But the same phenomena can be fully explained in other ways without appeal to the positions of any heavenly bodies. Homeopathy aims to bring about and explain improvements in a person's state of health by reference, for instance, to the qualities of substances that are too dilute to contain any physical trace of their supposedly active ingredient. But the phenomena in question can be fully explained in other ways, without appeal to the principles of homeopathy.

So there are good reasons, on the basis of our ordinary knowledge, to regard psychological and social scientific explanations as genuine causal explanations. At the same time, I have suggested, we must accept a basic, background commitment to the idea that all causal facts are determined by low-level causal facts of the kind charted by the natural sciences. And if that is true, then in order for psychological and social scientific explanations to be genuinely causally explanatory, there must be some connection between the higher-level properties that we cite in giving such explanations and the underlying, lower-level causal facts that ultimately determine all the causal facts. If there were no connection at all, there would be no way in

which a thing's possession of high-level psychological or social-scientific properties could affect its causal behaviour. But the connection need not be tight or systematic. What is required is just that the higher-level properties supervene on the lower-level properties: that is to say, that if two worlds are alike with respect to all lower-level properties, they must also be alike with respect to all higher-level, psychological, economic, sociological etc. properties.

The fact of supervenience alone is not enough, in itself, to show that psychological properties, say, are genuinely causally explanatory. The case for thinking that they are genuinely causally explanatory comes from the point sketched in the previous paragraph: that common-sense psychology offers causal explanations that meet the normal standards of successful explanatory practice. But in order for the explanations of psychology to be genuinely causally explanatory, as we know them to be, the supervenience claim must be true. For consider how things would be if the supervenience claim were not true. Two worlds could differ in mental respects without differing with respect to any lower-level physical properties. But in that case, those mental differences could make no difference to people's behaviour. For, we have said, the causal facts are fully determined by the low-level physical facts. If mental properties are to have any causal implications, then, they must supervene on the low-level physical facts. But there need be no stronger or more systematic relation between the two levels than that. Accepting that all the causal facts are determined by lower-level physical facts is compatible with insisting on the distinctiveness of the social and human sciences and their irreducibility to the physical sciences.⁸

References

- Alvarez, Maria. 2010. *Kinds of Reasons: An Essay in the Philosophy of Action*. Oxford: Oxford University Press.
- Anscombe, G.E.M. 1959. *Intention*. Oxford: Basil Blackwell.
- Dancy, Jonathan. 2000. *Practical Reality*. Oxford: Oxford University Press.
- Davidson, Donald 1963. Actions, Reasons and Causes, in Davidson 1980.
- . 1970. Mental Events, in Davidson 1980.
- . 1973. The Material Mind, in Davidson 1980.
- . 1974. Psychology as Philosophy, in Davidson 1980.
- . 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Hornsby, Jennifer. 1980–81. Which physical events are mental events? *Proceedings of the Aristotelian Society*, 81: 73–92.
- McLaughlin, Brian. 2013. Why rationalization is not a species of causal explanation. In *Reasons and Causes: Causalism and Anti-Causalism in the Philosophy of Action*, ed. G. D'Oro and C. Sandis, 2013. Houndmills: Palgrave Macmillan.

⁸An earlier version of this paper was presented at the symposium, 'Causation, Agency, and Supervenience', at the Las Casas Institute, Blackfriars Hall, Oxford in July 2018. Some of the material was also presented at a workshop, 'Ascription, Causation, and the Mind', at the University of Utrecht in May 2016. I am grateful to the participants on both occasions for helpful comments and discussion.

- Melden, A.I. 1961. *Free Action*. London: Routledge and Kegan Paul.
- Ryle, Gilbert. 1949. *The Concept of Mind*. London: Hutchinson.
- Setiya, Kieran. 2009. Reasons and causes. *European Journal of Philosophy* 19: 129–157.
- Tanney, Julia. 2009. Reasons as non-causal, context-placing explanations. In *Rules, Reason, and Self-Knowledge*, ed. J. Tanney, 2013. Cambridge, MA: Harvard University Press.
- Wilson, George. 1985. Davidson on intentional action. In *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, ed. E. LePore and B. McLaughlin. Oxford: Basil Blackwell.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 5

Causation and Agency



Peter Róna

Abstract ‘Causation’ covers a variety of dependent relationships between and among objects and events. The axiom concerning the unicity of reality has been thought to warrant the assumption that causal relationships of social phenomena, including economics, share common properties with corporeal objects, that, in short, agency is a form of causation. This paper defends the opposite view, to wit, that causation based on the properties and powers of corporeal objects (be they natural or man-made) is unlike causation based on agency. Whereas causation among the former is a function of the properties and powers of the objects at play, agency ‘causation’ is the product of human intentionality. Any theory of agency must account for free will even where, as in the case of rule based roles, the instantiation of free will is qualified. An agent’s action may, of course, set in motion a causal law by instantiating the properties of the *object(s)* producing the intended effect (eg. pulling the trigger of a loaded gun), and the agent will be responsible for the consequences, but the discharge of the bullet is the result of the properties and powers of the gun and the bullet and not of the agent. The profound ontological difference between causation and agency cannot be overcome with resort to epistemological, logical or linguistic considerations.

In economic and social life the agent relies not on the causal properties of corporeal objects, but on her action authorised by *social rules*. The causal properties and powers of objects are fundamentally different from the causal powers of rules, because, unlike the former, the latter are designed for the achievement of an intended result and cannot operate without intentional application. The proper study of economics is not the study of the properties of objects contrived by economic theory, but the nature of purposeful human action.

P. Róna (✉)
Blackfriars Hall, University of Oxford, Oxford, UK

5.1 Introduction

Modern social science is built on the belief that social reality, much like nature, has its own objective, mind-independent laws. If these can be discovered, social life can be predicted and explained much as law-like generalisations can explain and predict natural phenomena. But just as a greater knowledge of nature has brought about nature's accelerating destruction, the formulation of law-like generalisations about social and economic processes has compromised the moral autonomy of human beings, transforming in profound ways who we are and what we do. Prediction and explanation, – the core objective of both the natural and the social sciences, rests on how dependent relations - conventionally grouped under the concepts of causation and agency - are understood.

The idea that every single fact of reality is determined by a law is as old as recorded philosophy, having been first proposed by Democritus and elaborated with insistence by the Stoics. Its opposite, first formulated by Epicurus and developed by Aristotle,¹ recognised the operations of spontaneous chance, a modern version of which is the theory of emergence. The corollary of Democritus' material determinism, –that facts for which no identifiable law-like cause can be shown are not facts - is reasserted in some versions of XXth century logical positivism. A third approach first set out by Hume² and restated by Russell,³ denies that there is such a thing as a cause, and, instead sees regular patterns, the 'constant conjunction' between cause and effect as a mental operation. A much refined version but still broadly in the Humean tradition is Mackie's INUS condition theory where INUS stands for an Insufficient, but Non-redundant part of an Unnecessary but Sufficient condition. In addition to his well-known example of singular causal statements of a short circuit causing a fire, – where the spark causing the fire was insufficient because other conditions were also necessary, but non-redundant, because in fact the spark was caused by the short circuit – Mackie illustrates the INUS conditions in a general causal statement with an example from economics. When an economist speaks of credit restrictions causing unemployment, the causal relationship will obtain only in the context of a series of other conditions and makes the important admission that all of these other conditions cannot be specified, but he does not draw the conclusion from this admission, namely, that without a specification of the conditions the generalisation has, at best, only approximate truth value⁴ A fourth approach, set out among others, by Davidson sees cause as having to do with the logical form of action sentences.⁵

¹ Aristotle, *Physics*, Book II, Chapters iv, v, and vi.

² Hume, D. ((1740) *A Treatise of Human Nature* (Analytical Index by L. A. Selby Bigge; Second edition with text revised and notes by P. H. Nidditch) (1978) Oxford University Press, Book I, Part III.

³ Russell, B. (1912–13) "On the Notion of Cause" *Proceedings of the Aristotelian Society*.

⁴ Mackie, J. L. "Causes and Conditions" *American Philosophical Quarterly*, 2/4 (October 1965) 244–55.

⁵ Davidson, D. "Causal Relations" *Journal of Philosophy*, 64 (1967), 691–703.

There is, therefore, the distinction between the view that sees all dependent relations as caused/necessitated⁶ and, accordingly, assumes that in the case of singular causal statements there is always a universal statement from which the singular statement can be derived and those who, following Anscombe do not see the need for the support of universal statements for the validity of singular statements. But all of these approaches, including Hume's, revolve around the logical ordering of sense experience. Another distinction is made between mind-dependent and mind-independent events on the basis that these two categories of events are ontologically different in a way that the dependent relations of mind-dependent events are quite unlike those that characterise mind-independent events. On this view causal relationships are between and among mind-independent events and agency is the product of mind-dependent ones. This paper is principally concerned with the exploration and defence of this last claim.

Science, understood as the knowledge of nature, is concerned with natural reality as it exists. Economics, understood as the knowledge of how material well-being is or might be generated and distributed, is concerned with how individual and social objectives may be achieved. Whereas the natural sciences are concerned with what is, economics simultaneously addresses what is and what should be; it is inherently a teleological and normative discipline. Its putative regularities are not derived from actual phenomena, but, rather, are adduced from synthesised mind-dependent objects intended to resemble as much as possible experienced events. These objects are what is needed to obtain the proposition of the model, such as equilibrium, marginal utility or Pareto efficiency. Inflation, for example, is the mind-dependent object of increased prices/decreased purchasing power. The relationships and interactions between and among these synthetic objects are meant to reveal generalisations not about what is, but what should be. Economic action, being by its nature purposeful, requires a degree of clarity about the nature and scope of its purpose. Economics is expected to provide recipes for the achievement of societal purposes, and it can satisfy this expectation only if it can illuminate or generate causal relationships that will bring about the desired end. Causal relationships, therefore, are at the heart of economics. But are those causal relationships to be discovered - because they are there in the law-like mechanisms of economic phenomena - or are they to be crafted through purposeful action to achieve desired ends? Do these mechanisms operate without the aid of conventions and legal sanctions? Should we follow Democritus and look for universal laws derived entirely from the operation of physical nature in our search for an understanding of economic processes, or should we, instead, base our notions on an understanding of the *nature* of the object of our attention, should we seek an understanding of economic phenomena through an understanding of the purposeful action of the economic agent, or should we abandon the search for causes altogether and base ourselves on probabilistic correlations? Should economics concern itself with the unravelling of some existing reality or should it formulate hypotheses of optimality and the means with which to

⁶The extensively studied distinction between causation and necessitation is not relevant to the claim defended in this paper.

achieve them? Are there invariables at the service of economists, waiting to be discovered much as the constants of nature provide the foundation for natural science, or is economic reality made? Can models reveal causal relationships in ways that mirror recurrent relationships in economic life?

Understandably, the social sciences turned to the startling achievements of the natural sciences for the means with which to organise and define the domain of their discipline after the Galilean-Newtonian revolution. The triumph of the Enlightenment, the consolidated confidence in the unicity of all reality, and, in the phrase of Eugene Wigner, the “unreasonable efficiency”⁷ of mathematics in the natural sciences persuaded social scientists - or moral philosophers, as they were then called - to follow the paradigm being developed by the natural scientists in their search for understanding social phenomena. As Condorcet said in his speech at the French Academy on 12 February 1792:

As mathematics and physics perfect the arts of supplying our simple needs, is it not part of the same order of nature that progress in the moral and political sciences should exercise the same effect on the motives which rule our actions and feelings?

More than a century later Edgeworth puts the same point in the following terms: “The invisible world of electricity is grasped by the marvellous methods of Lagrange; the invisible world of pleasure may admit of similar handling”,⁸ and William Stanley Jevons draws a causal relationship between the trade cycle and the sunspot cycle. All that this requires, Edgeworth exuberantly explains, is “the conception of Man as a pleasure machine (that) may justify and facilitate the employment of mechanical terms and Mathematical reasoning in social science”.⁹ With these words the great project of modern social science – the reformulation of individual and collective social behaviour to accord with that of matter - is announced.

Modern economics, said to be the queen of the social sciences, is built on the assumption that the reality it addresses is, in all material respects, much like the subject matter of physics. The operation of the second law of thermodynamics and the interaction of supply and demand in the formation of prices are thought to have sufficiently numerous and important commonalities to warrant the search for a conception of a ‘law’ of similar form. Bernoulli’s or Toricelli’s theorem about the inverse relationship between pressure and the speed of a fluid’s movement at a point instantiates, it is thought, the same sort of causation as the inverse relationship between inflation and unemployment represented by the Phillips Curve.

The foundation for ambitions of this sort is a belief in the unicity of reality and the unity of knowledge. Accordingly, economics portrays “human acts (as) clearly part of the order of nature causing and being caused by events outside of ourselves.”¹⁰

⁷Wigner, E. (1960) “The unreasonable efficiency of mathematics in the natural sciences” *Communications in Pure and Applied Mathematics*, 13, pp.1–14.

⁸Newman, P. (2003) *F.Y. Edgeworth’s Mathematical Psychics and Further Papers on Political Economy*, Oxford University Press, p. 13

⁹Ibid. p. 15. Compared to Edgeworth’s, Jevons’ and Walras’ brutal materialism. Marx was a starry-eyed idealist who thought that if reality did not accord with theory, we should change reality.

¹⁰Donaldson, D. “Psychology as Philosophy” in: Brown, S. C. (Ed.) *Philosophy of Psychology* (1974) The Macmillan Press and Barnes, Noble; reprinted in Donaldson, D. (2013) *Essays on Actions and Events*, Oxford University Press.

This paper denies the unicity of reality and also denies the possibility of unity of knowledge, except at a level of abstraction that does more harm than good. At the simplest level, there is no unity between analytic and synthetic knowledge. Accordingly, it rejects the presumed need in the social sciences to formulate generalised theories about causation that can be linked to the causation theories of the natural sciences.

As its point of departure, this paper takes the Aristotelian position that objects and events - 'substances' in Aristotelian terms - have their essential natures, and these essential natures are marked by powers and properties proper to them. All objects and events have properties and powers in relation to other objects and events, but these properties and powers cannot be subsumed under a generalised notion of causation. A further complication arises with respect to the causal powers, if any, of concepts. The valence of an element is not the same sort of property or power as inflation, unemployment or marginal utility. The former conceptualises the physically identifiable number, charge and movement of electrons, that is to say, of corporeal objects. The latter are statistically based generalisations where the components of the statistical information are themselves conceptual constructions.

Interest in agency and its relation to causation increased during the middle third of the XXth century with important contributions from Collingwood,¹¹ Gasking¹² and von Wright.¹³ Notwithstanding the important differences among these three authors, the intuition that causation in nature differs in some basic sense pervades all three, but the ontological as opposed to the epistemological aspects of this difference remained unexplored. As von Wright put it:

Independently of our stand on the metaphysical questions, it will be readily admitted, I think, that the idea of experimentalist or manipulative causation has important application in the natural sciences - and also that its applicability becomes debateable when we move to the human (including the social) sciences. If we wish to identify causation as such with manipulative causation, we could then say that the category of causation is primarily at home in the (experimental) natural sciences and basically alien to the human sciences.¹⁴

The paper follows¹⁵ the Davidsonian distinction between the constitutive elements of belief, desire, intention and action on the one hand and the constitutive elements of physical objects on the other, but unlike Davidson, claims that they do not have any causal properties in common. It follows Searle¹⁶ when he notes that:

In the philosophy of mind there is an uneasy relation between intentionality and causality. Causality is generally regarded as a natural relation between events in the world;

¹¹ Collingwood, R. G. (1938) "On the so-called idea of causation", *Proceedings of the Aristotelian Society* (New Series), 38; (1940) *An Essay in Metaphysics*. Oxford University Press.

¹² Gasking, D. (1955) "Causation and recipes", *Mind*, 64.

¹³ von Wright, G. (1975) *Causality and Determinism*. Columbia University Press.

¹⁴ Quoted in Illari, P. and Russo, F. (2014) *Causality*. Oxford University Press.

¹⁵ Davidson, D. (1970) "Mental Events", reprinted in Davidson, D. (2013) *Essays on Actions and Events*, Oxford University Press.

¹⁶ Searle, J. R. (1983) *Intentionality*. Cambridge University Press.

Intentionality is regarded in a variety of ways but not generally as a natural phenomenon...

but does not follow him when Searle proceeds to

take a step towards Intentionalizing causality and therefore, toward naturalising Intentionality.¹⁷

Instead, it defends the view that causation and agency do not share common properties and sees the effort to bring them together as a part of the general drift of XXth century philosophy away from ontological issues, away from the Aristotelian nature of events and substances, a drift with serious adverse consequences for the moral autonomy of the individual. This trend purports to coin a commonality among ontological incommensurables on the basis of epistemological, logical or linguistic harmony. But (*pace* Searle) intentionality brings about effects where those events would not occur without it; natural causation takes effect without the introduction of properties alien to the properties of cause and effect, and agency produces events that nature does not produce. Although it is readily agreed that the more causation is stuffed into agency the closer the social sciences get to Newtonian methodology, the scope and significance of intentionality should not be compromised by fudging that autonomy with the importation of natural causal relations into its domain. The claim here is that resemblance to Newtonian physics and a deceptive precision coupled with the illusion of predictive capability is achieved at the cost of serious violence to the ontology of social reality.

Most authors on the subject of causation in economics and in the social sciences in general build their theories of causation and agency on the basis of an assumed analogy with causation in nature, and particularly Newtonian physics. John Hicks, for example, in his seminal *Causality in Economics*¹⁸ raids astronomy, physics and mathematics from Copernicus to Newton for a suitable scientific paradigm and poses the question: “How does it all apply to economics?”¹⁹ without any showing - other than some sort of an urge to resemble the natural sciences - as to why any of it should apply at all. But, as Ludwig von Mises put it: “The study of economics has been again and again led astray by the vain idea that economics must proceed according to the pattern of other sciences.”²⁰

A great deal turns on this question. If causation in economics is analogous to causation in the natural sciences, as Hicks, Samuelson or Mankiw among others would have it, moral responsibility for individual economic action is diminished and the case for the regulation of group action and macroeconomic phenomena is undermined. If markets are given nature-like powers and properties it would be foolish to resist or regulate economic life. Economics may be seen as part of nature just like the components of Mendeleyev’s table, and, because nature knows best,

¹⁷ Ibid. p. 112

¹⁸ Hicks, J. (1979) *Causality in Economics*, Basil Blackwell, Oxford.

¹⁹ Ibid. p. 37.

²⁰ von Mises, L. (2006) *The Ultimate Foundation of Economic Science; an Essay on Method*, Liberty Fund, Indianapolis.

man must comply with the natural laws of economics. If the optimal use and distribution of economic resources can be determined on the basis of causal laws, fairness, justice, human dignity and the protection of the environment all require unnatural action, or, in the alternative, only those instances of fairness, justice and human dignity can be legitimate that conform to or can be derived from natural laws. Conversely, if the natural sciences do not provide a relevant source for an understanding of dependent relationships in economic life, modern economics must be seen not as a project for an understanding of reality, but as the construction of a surrogate reality for the regulation of human behaviour in aid of achieving postulated ends.

5.2 Causation(s)

Apart from being plainly wrong,²¹ there is something odd about Hume's billiard balls. The point of the example - that causation is not a *thing* but a mental construct - leads Hume to see the question the wrong way around. He first constructs a sequence of events that starts with pure intentionality. The billiard balls strike one another because somebody wants to strike the first ball so as to bring about the desired hit of the second. But instead of looking for intentionality in why and how the first ball was set in motion, he looks instead for something with autonomous existence that could be called a causation, that somehow brought about the collision on its own whereas it is perfectly clear that the collision of the two balls is the consequence of the intentionality of the player. Having suppressed the premise upon which his example is built, he looks instead for a thing, and not finding it, he stops and announces that causation is in the mind of the beholder, instead of looking for the properties of the two balls that produce the result once the first one is set in motion.²² The two billiard balls strike each other because the player wants them to. They strike each other in a certain way because that is the result their properties bring about. There is nothing metaphysical, mysterious or enigmatic about the sequence of events. It starts with an intention and proceeds in accordance with the known properties of the objects set in motion by the intention. Obscurity arises because Hume discards both of the possible two sources of causal relations - intentions on the one hand and properties and powers on the other - and looks instead for something that neither exists nor is necessary to explain causal relations. But causality resides in the effect being derived from the cause. Cause and effect then are both properties of objects

²¹ For a thorough and compelling refutation of Humean causation theory, see Harre, R. and Madden, E. H. (1975) *Causal Powers*, Basil Blackwell, Oxford.

²² Anscombe puts the point about Hume's causation in *Causality and Determination* (Reference in footnote 18) as follows: "it is argued that 'all we find' is such-and-such, and it turns out that the arguer has excluded from his idea of 'finding' the sort of thing he says we don't find".

and causation takes place when these properties meet under conditions for the operation of the properties.²³

Science, however, marched on and did what Hume thought could not be done. It found causation not in observation, but in the measurable properties of substances, and established the method with which natural causation could be determined. The social sciences then surrendered to the primacy of natural causation as the universal law for all phenomena, and, although there was a rebellion principally lead by Elisabeth Anscombe, the dominance of the natural causation paradigm in the social sciences prevailed.

The cultivation of law-like generalisations reached its apogee with the short-lived triumph of logical positivism, but, despite the collapse of this school of thought, economics has neither sought nor found a new foundation for its theories and doctrines apart from making some use of the refinements of analytic philosophy. At the same time, the natural sciences and mathematics have been moving away from this structure, built as it is, on Newtonian reductive materialism.

An obvious requirement for reaching scientific status on par with the natural sciences is the formulation of a causation theory capable of both predicting and explaining the behaviour of its chosen object(s). But a causation theory requires the demarcation of a domain within which the theory is expected to operate, and the identification of the domain of economics has been elusive. As Keynes noted, economics

deals with motives, expectations, psychological uncertainties. One has to be constantly on guard against treating the material as constant and homogeneous. It is as though the fall of the apple to the ground depended on the apple's motives, or whether it is worthwhile to fall to the ground, or whether the ground wanted the apple to fall, and on mistaken calculations on the part of the apple as to how far it was from the centre of the earth.²⁴

The subject matter of the natural sciences - whether already given or only intuited - is invariably corporeal with one or more verified or intuited properties. Science deals in natural kinds. They are proper 'objects', not, in the words of Thomas Hobbes, things "made with words". The objects of the natural sciences are impervious to words. It may be that their corporeality is undetectable with the means available at a given moment in history, but, invariably, the claim is that 'it is an *it*' and with the right tools and procedures it will be found, as indeed, historic experience has often justified this expectation.

There is no easy or obvious answer to the question: what is the subject matter, what are the objects of economics? The conventional definition of chemistry - "the investigation of the substances of which matter is composed and of the phenomena of combination and change which they display"²⁵ - finds no parallel in any of the definitions of economics offered by the dictionary or the subject's leading scholars:

²³ Anscombe, G. E. M. (1971) *Causality and Determination*. Cambridge University Press.

²⁴ Letter to Roy Harrod, 10th of July, 1938 (misdated 26th July) in *Collected Writings of John Maynard Keynes*. Macmillan/Cambridge University Press, for the Royal Economic Society, Vol. XIV, p. 292.

²⁵ The New Oxford Shorter Dictionary of English.

- “economics is the branch of knowledge that deals with the production and distribution of wealth” (Oxford Shorter Dictionary of English);
- “The science which studies human behaviour as a relationship between ends and scarce means which have alternative uses” (Lord Robbins)²⁶;
- “the study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of the material requisites of well-being” (Alfred Marshall);
- “Economics is the study of how men and society choose, with or without the use of money, to employ scarce productive resources which could have alternative uses, to produce various commodities over time and distribute them for consumption now and in the future among various persons and groups of society” (Paul Samuelson); or
- “Economics is a science of thinking in terms of models... relevant to the contemporary world.” (John Maynard Keynes).

If we reflect on how hypotheses of causation might be formulated in chemistry and in economics respectively so as to advance the understanding of the subject under study, it seems immediately obvious that the objects of economics and the objects of chemistry are unlike. The objects of chemistry exist independently of the theories chemistry builds about them. The objects of economics are contrived by economic theory itself. While there is no doubt that there is, in some general sense, ‘supply and demand’ Alfred Marshall’s two curves have yet to gain empirical confirmation. The sample field within which causation is to be found can be defined for chemistry but cannot be defined for economics. Economics does not have a subject it can call its own in the sense that the initial conditions of the domain can be specified. The determination of what causes what can only be made with the help of artificial assumptions and the *ceteris paribus* clause because, as John Stuart Mill had already noted the ‘disturbing causes’ have to be separated from the object of study. But neither he nor his successors have found a way to distinguish the former, also called ‘exogeneous’ from the latter, also designated as ‘endogeneous’. This separation of the exogeneous from the endogeneous is accomplished, first with the selection of the cause and then with the aid of assumptions and *ceteris paribus* clauses that may or may not have a basis in reality. For example, the Arrow-Debreu competitive equilibrium model, a cornerstone of modern economics modelling - for which its authors have received the Nobel Prize in economics - the causal relationships that secure equilibrium, require 14 assumptions nine of which cannot exist in the world as we know it. But the selection of the cause is an ideological choice, one made on the basis of the sort of economy the economist favours or wishes to test. There is no objective, value free basis for making this choice. Chemistry by contrast, does not need assumptions for the delineation of its subject matter because it is given in nature. Instead of assumptions, it posits conditions that obtain in the real world. It only studies the transformations of substances under conditions that do in

²⁶ Robbins’ tortuous phrasing reveals the problem: how can behaviour be “a relationship between ends and scarce means”?

fact exist or that can, in fact, be generated. Economics studies wealth, choices about the alternative uses of scarce resources, individual and social action connected with the attainment of material well-being, etc., subjects that cannot be given durable content. What amounts to wealth, material well-being, what are scarce resources, what are relevant individual and social actions are all matters that change over time and that require reference to markets, prices, notions of saving and consumption, etc. that, in turn, require supporting assumptions about the conditions of private property, the legal system, customs, and so on. These notions are, as the quote from Walras below shows, contrived with the help of various techniques, such as abstraction, isolation and idealisation and supported by reliance on a *ceteris paribus* clause. But the *ceteris paribus* clause is suspect not only because, as usually noted, things do not stay the same, but, more importantly, because the isolated subject itself is unstable. There are no invariables in economics. These devices may help produce the appearance of causal relations, but their conceptual success is purchased at the expense of descriptive adequacy.

The reason for this failure is that whatever economics may be about, it should be clear from the above that, in contrast with the natural sciences, its subject matter is not a natural kind.

In one of the great essays of XXth century philosophy Nelson Goodman showed that the confirmation of an inductive argument, such as a scientific generalisation, is not a matter of logic, that its validity must come from the subject matter itself.²⁷ It is logically possible to use data for one theory and to use the same data for another theory. There is, for example, no *logically unique way*, to connect data points on a graph in comparison with any other way. But the different curves²⁸ will lead to different conclusions and predictions where predictions are made on the basis of the shape of the curves.

With the abandonment of ontology as a proper subject for philosophy and the eventual arrival of analytic philosophy in its place, economics increasingly saw models as its proper subject as opposed to any reality those models purport to represent. Economics became a discipline concerned with a surrogate reality the discipline itself constructed, based on axioms rather than any sort of reality existing independently of its theories. As Alexander Rosenberg noted:

Much of the mystery surrounding the actual development of economic theory – its shifts in formalism, its insulation from empirical assessment, its interest in proving purely formal, abstract possibilities, its unchanged character over a period of centuries, the controversies of its cognitive status – can be comprehended and properly appreciated if we give up on the notion that economics any longer has the aims or makes the claims of an empirical science of human behaviour. Rather, we should view it as a branch of mathematics, one devoted to examining the formal properties of a set of assumptions about the transitivity of abstract relations: axioms that implicitly define a technical notion of ‘rationality’, just as geometry examines the formal properties of abstract points and lines.²⁹

²⁷ Goodman, N. (1983) *Fact, Fiction and Forecast* (fourth edition), Harvard University Press.

²⁸ This is known as Goodman’s „curve-fitting problem”.

²⁹ Rosenberg, A. (1992) *Economics – Mathematical Politics or Science of Diminishing Returns*, Chicago University Press, p. 247. Quoted in Backhouse, R. E. (1997) *Truth and Progress in Economic Knowledge*, Edward Elgar, p. 107–8.

As Keynes had already noted, the subject matter of modern economics is thinking about models. Accordingly, causation is posited or attributed to the objects contrived for the model. Causation and agency in this surrogate context – where the model represents a non-existent reality – are validated by nomological deductive means without empirical verification. The procedure was clearly set out by Leon Walras in 1874, and has not undergone any significant change since.³⁰ Walras writes:

...this pure theory of economics is a science which resembles the physico-mathematical sciences in every respect... The mathematical method is not an *experimental method*; it is a *rational method*... From real-type concepts, these sciences abstract ideal-type concepts which they define, and then on the basis of these definitions they construct *a priori* the whole framework of their theorems and proofs.... Following this same procedure, the pure theory of economics ought to take over from experience certain type concepts, like those of exchange, supply, demand, market, capital, income, productive services and products. From these real-type concepts the pure science of economics should then abstract and define ideal-type concepts in terms of which it carries on its reasoning. The return to reality should not take place until the science is completed and then only with a view to practical applications.³¹ (emphasis in the original).

Walras notes that “Reality confirms these definitions and demonstrations only approximately”, but goes on to claim that “and yet reality admits of a very wide and fruitful application of these propositions.” But the parallel he draws between the “physico-mathematical” sciences and economics breaks down (i) when we consider that social phenomena in general and economic phenomena in particular is ceaselessly reflexive, and (ii) the ‘real-type concepts’ of the physico-mathematical sciences are derived from real, corporeal objects as opposed to the incorporeal notions of ‘exchange’, ‘supply’, ‘demand’, ‘market’, etc.

The marked growth in model building since the middle of the XXth century is nevertheless built on these Walrasian premises. As Jakko Kuorikoski and Caterina Marchionni note:

Although economists talk a lot about economic *theory*, models are really the working units driving the acquisition of knowledge. Economics is first and foremost a modelling science. ... Theoretical models are constituted by a set of assumptions, a set of conclusions and rules for deriving conclusions from the assumptions.³²

Scientific status is achieved, according to the authors, if no errors occur in the derivation and if the assumptions are empirically supported. But how is empirical support obtained? Economic data, unlike scientific data, cannot be replicated, and, as Walras noted, economics is a rational as opposed to an empirical discipline.

³⁰ For a detailed summary and analysis of current methodology in economics see Morgan, M. S. (2012) *The World in the Model; How economists Work and Think*, Cambridge University Press; Morgan, M. S. and Morrison, M. (ed.), (1999) *Models as Mediators*, Cambridge University Press.

³¹ Walras, L. (1874) *Éléments d'Économie Pure*, Corbaz 9292Cie, Lausanne; translated as *Elements of Pure Economics*, of the 1926 definitive edition by William Jaffé, (1954), George Allen and Unwin Ltd. London.

³² Kuorikoski, J. and Marchionni, C. “Philosophy of Economics” in French, S. and Saatsi, J. (2014) *The Bloomsbury Companion to the Philosophy of Science*, Bloomsbury.

According to Mäki, assumptions in economics are similar to laboratory controls in experiments, but he fails to observe, in addition to the replication problem, that laboratory controls have to do with the removal or enhancement of some aspect of physical reality; what is removed as well as what is kept is real and not assumed in contrast with the assumptions of economists that are, as Walras notes, *a priori*. Mäki, however, sees assumptions as serving the theoretical isolation of causal factors from disturbing causes.³³

Two examples will show that, in fact, there is no reliable way to secure empirical support for the assumptions of economic models. A standard trope of neoliberal economics is the claim that labour-market rigidities constrain output and employment. This claim assumes that employers are more willing to invest and employ if there are few if any laws specifying the terms and conditions of employment. One man's rigidity is another's security permitting him to have confidence in his employer, a confidence that may have a dependent relationship to his productivity. Greater "flexibility" may increase output in a context of high unemployment but will not do much in an economy struggling with labour shortages. Whereas the removal of Mill's 'disturbing causes' has an empirically demonstrable basis, economics does not have any scientific foundation for deciding what causes should be deemed to be disturbing – and therefore eliminated from the model – when labour market flexibility is decided upon. Another one, advanced by Stephen Moore and Arthur Laffer, claims that tax cuts will bring down government deficits and generate greater private investment. These claims, first made some 30 years ago, still await empirical confirmation. In each case the factors affecting the outcome are numerous and constantly undergoing change. Lower taxes may have the propensity to increase investment as long as there are opportunities for profitable investment. But profitable investment requires, among other things, an educated and healthy labour force that cannot be had without the resources provided by taxes. .

Yet another view of the truth value of assumptions was famously expressed by Milton Friedman, who thought that the assumptions of models need not have any truth value at all so long as the model makes accurate predictions. The learned Nobel laureate, however, did not explain how such outcome might be possible, and the decades that have since passed have not favoured us with persuasive examples.³⁴

An example of conflating the properties of the objects with intentionality was described by Karl Popper in a much overlooked essay³⁵ in which he replaces causal relations with propensities. Whereas causality is an inherent property in an object, a propensity, according to Popper, is inherent in a situation, that is to say, in the

³³ Mäki, U. (1992) "On the method of isolation in economics" in Dilworth, C. (Ed.) *Intelligibility in Science*, Rodopi, Atalanta and Amsterdam pp. 319–54.

³⁴ Friedman, M. (1953) "The methodology of positive economics" in *Essays in Positive Economics*, University of Chicago Press, pp. 3–43.

³⁵ Popper, K. (1990) *A World of Propensities*, Thoemmes Antiquarian Books Ltd. Bristol. The essay was first delivered by Popper on August 24, 1988 before the World Congress of Philosophy in Bristol, UK.

interaction of known and unknown objects. While properties cause determined results, propensities are indeterminate and give rise to new propensities. Popper writes:

Just like a newly synthesised chemical compound, whose creation in turn creates new possibilities for new compounds to synthesise, so all new propensities create new possibilities. And new possibilities tend to realise themselves in order to create again new possibilities.

But *pace* Popper, whereas compounds do not synthesise ‘in order to create’ new compounds since compounds, like Keynes’ apple, do not form intentions. Intentional human action is undertaken by the agent in order to create a new situation. That, physical property innocent of intentionality, versus purposeful action, in a nutshell, is the difference between causation and agency.

The contrary view, advanced by John Hyman³⁶ dissolves intention under the more general concept of any “active causal power” where no distinction is made between the causal powers of substances and the intentionality of human agents:

Action in general is simply the exercise of an active causal power – i.e. the power or ability to cause some kind of change – the agent being the one that causes the change and the patient being the one that undergoes it. Far from applying exclusively to human beings, the concept of action applies to every substance able to cause change.

This very broad notion of action is unhelpful in the social sciences and particularly unhelpful in relation to the explanation of economic action where the subject matter under study is the *purposeful*, that is to say, intentional action of human beings. Hyman explicitly rejects any theory in which “will or intention.. play an essential role in *human agency*” (emphasis in the original) on the ground that intentionality does not distinguish between activity and passivity in human life. But the intentionality of agents rests not on this distinction but on the difference between *properties* and *purposefulness*.

5.3 Intention(s) and the Will

The causal theories built on Walras’ method cannot account for economic events because neither his real-type nor his ideal-type concepts have causal powers. Rather than taking place between and among objects with empirically demonstrable and fixed causal properties, economic causality is not object based; it runs through the agent of economic action. The agent disrupts or generates the purported causal mechanism because she reflects on the meaning of what is happening and acts not according to some causal property of the event, but in accordance with her assessment and interpretation of the meaning of the event in question. Agency is purposeful action *designed to* change what is into what the agent desires. Unlike causation, which takes place without a purpose, agency is all about the achievement of a *telos*, about the disruption of the existing state of affairs. The ontology of agency consists

³⁶ Hyman, J. (2015) *Action, Knowledge and Will*, Oxford University Press.

of Keynes' 'motives, expectations, psychological uncertainties' triggering the agent's action rather than Walras' neo-Kantian *a priori* concepts.

Therefore, as Jaegwon Kim notes "any discussion of causation must presuppose an ontological framework",³⁷ and, although the author of this observation adds that there needs be "an accompanying logical and semantical framework", it is unclear how either logical or semantical considerations would throw light on the processes that take place in economic life. Indeed, the search for the ontological framework is quickly abandoned in favour of a debate with Mackie about necessary and sufficient conditions. Without it economics, following Walras, replaces the ontology of action with 'real-type concepts' in order to „construct *a priori* the whole framework of their theorems and proofs".

The use of the 'ontological framework' leads to rather different results from those found in some of the most influential essays on causation. Consider, for example, Davidson's analysis of Mackie's "this short-circuit caused this fire" in terms of the presence or absence of a non-truth-functional causal connective where the cause is seen - incorrectly according to Davidson - as a condition, all in order for Davidson to reject the notion that causes are fully expressed only by sentences. From an ontological and scientific view all of this is meaningless. Causation in the natural sciences is not a matter of either logic, language or the completeness of sentences. Science is not interested in either „necessary" or „sufficient" causes, and least of all in Mackie's INUS conditions. What matters is that short-circuits produce sparks of a certain, measurable intensity that *has the property of setting to fire* combustible material of a certain kind. The properties of the sparks meet the properties of the combustible material. These properties can be stated in the form of universal laws from which the events can be calculated but need not be. The short circuit caused the fire if the requisite conditions were met, to wit, there were sparks of a certain intensity in the proximity of some flammable material, there was the requisite amount of oxygen, etc. If the universal laws cannot be formulated because the available evidence is insufficient, the deficiency cannot be made up by constructing causal hypotheses either on the basis of language or on the basis of logic and the evident causal connection between the sparks and the flammable material cannot be denied on the ground that there is no known universal law of which the event is an instance. Arguments about necessitation cannot overrule the physical properties at play in the causal relationship.

Davidson compounds the confusion when he conflates intentionality with causation. He writes:

If I poison someone's morning grapefruit with the intention of killing him, and I succeed, then I cause his death by putting poison in his food, and that is why I am the agent of his murder.³⁸

But death was caused not by the intention of the agent, but by the lethal properties of the poison. Death by poisoning would have occurred if the poison would have

³⁷ Jaegwon Kim "Causes and Events: Mackie on Causation" in Sosa, E. and Tooley, M. (1993) *Causation*, Oxford University Press.

³⁸ Davidson, D. (2001) *Essays on Actions and Events* Oxford University Press, p. 48.

ended up in the grapefruit without anyone's intention. The chemical process causing death is separate from the intention that sets in motion that chemical process. Conversely, the intent to kill would not have 'caused' death without the lethal properties of the poison. Causation in Davidson's account is a legal as opposed to a scientific concept. In a trial for murder the prosecution would have to prove that (i) the poison placed in the grapefruit killed the deceased (excluding, for example, the possibility that the deceased suffered a heart attack before the poison took effect), (ii) the accused placed the poison in the grapefruit, and (iii) the accused placed the poison in the grapefruit with the intent to kill the deceased.³⁹ Criminal law, beginning with the Romans, made the distinction between the first two and called it the *actus reus* while the third one was named *mens rea* in order to distinguish causation from agency.

Causation then is a matter of proving the operation of one or more mind independent properties, of experimentally provable evidence rather than of some sort of an epistemological theory. It is not the mental operation of the beholder, but the physical behaviour of the object(s) in question. It is the operation of natural kinds.

Placing ontology ahead of epistemology amounts to the insistence on the priority of the 'what' over the 'why' and the 'how'. The problem facing economics is that the facts with which it has to deal are unlike the facts of the natural sciences. They are incorporeal and unstable. Economics by-passed the admittedly very difficult question of the 'what' – the ontological question – in the hope and expectation that concentration on the why and the how – the epistemological, logical and linguistic issues – would render the 'what' question otiose. In doing so, it domesticated tools, that had been developed for the study of subjects where the subject matter of the inquiry was reasonably settled and the collection of facts could proceed. But in economics there are few if any facts to be faced directly, since the facts are not found, but made by the purposeful actions of individual and collective agents.⁴⁰ Aristotle was right: it is the nature of the object that determines all else rather than our sensory perceptions or inferences determining its nature. "For Aristotle, the perceiver is the means for the fullest activation of the perceptible properties of objects in the world – which are activated as properties of objects rather than experiences of perceivers."⁴¹ Generalisations based on perception, logic or language will not supervene the radical ontological difference between objects and actions. A discipline such as economics, that is theory laden and given to creating its own reality is particularly vulnerable to the misapplication of "idealisation", "abstraction" and "isolation" of the object practiced in the natural sciences and, instead, it easily succumbs to the contrivance of the objects of its study at the expense of hard evidence, believing, as it does, that it is practicing science when mimicking idealisation,

³⁹ Hart, H.L. A. and A. M. Honoré (1959) *Causation in the Law*, Oxford at the Clarendon Press.

⁴⁰ The special problems of corporate action were explored in Róna, P. "Ethics, Economics and the Corporation" in Róna, P. and Zsolnai, L. (2017) *Economics as a Moral Science*, Springer.

⁴¹ Marmodoro, A. (2014) *Aristotle on Perceiving Objects*, Oxford University Press, p.1.

abstraction and isolation. Mary Morgan,⁴² for example, does not even consider how these techniques of the natural sciences could apply to actions as opposed to objects, as indeed John Stuart Mill similarly assumed, that the actions of agents were objects of economics permitting the identification and elimination of ‘disturbing causes’.

As the quote above from Leon Walras shows, economics, lacking empirically verifiable objects, contrives its own with the help of a variety of techniques.⁴³ Its objects are not empirically verifiable things or events, but axioms, purportedly embodying some sort of a generalisation. As Keynes indicates in the quote above, its subject is “the models... relevant to the world” it creates. Models, not Aristotelian substances, relevant to, but not of the world. The immense complexity, the ceaseless change and turbulence, the fundamentally provisional and contingent nature of the material forces the economist to abstract, isolate or idealise not some fact or facts with a verifiably independent existence bearing verifiably independent properties and powers, but rather, to inductively contrive facts with the help of concepts of the economist’s devising, such as supply and demand, equilibrium, GDP, etc. that outside the contrivance do not exist. These concepts, embedded in the relevant axiom, operate as surrogates for the objects of the natural sciences in the hope and expectation that they can be shown to manifest constant properties and powers. They are rather like Platonic Forms standing in the place of the Heraclitan flux.

SOCRATES: But if it is always passing away, can we correctly say of it first that it is *this*, and then that it is *such and such*?...

Then if it never stays the same, how can it *be* something?

.....

Indeed, it isn’t even reasonable to say that there is such a thing as knowledge, Cratylus, if all things are passing on and none remain. For if that thing itself, knowledge, did not pass on from being knowledge, then knowledge would always remain, and there would *be* such a thing as knowledge. On the other hand, if the very form of knowledge passed on from being knowledge, the instant it passed on into a different form than that of knowledge, there would be no knowledge.... But if there is always that which knows and that which is known, if there are such things as the beautiful, the good, and each one of the things that are, it doesn’t appear to me that these things can be at all like the flowings or motions... (Emphasis in the original) Plato *Cratylus*, 439 d, e, 440a, b.

But the idea that we can reduce the Heraclitan flux to some cognate of the Platonic Form without fundamentally altering the properties and powers of the former is an intellectual sleight of hand made possible by the premature burial of ontology.⁴⁴ Economics, from its very beginnings, has been and remains a search for divining the operations of the Invisible Hand. It has consistently believed in some sort of a hidden mechanism, directed by a sublime power concealed and compromised by ‘disturbing causes’. While some of its practitioners, as, for example Ricardo, Marx and Mill believed in the discoverable hidden unity of economic life, over time economic concepts increasingly became idealised distillations, devised

⁴² Cited below.

⁴³ Morgan, M. (2012) *The World in the Model*, Cambridge University Press.

⁴⁴ A “stinking corpse” according to Hilary Putnam. Putnam, H. (2004) *Ethics without Ontology*, Harvard University Press, p. 81.

not so much to unmask the empirically unknowable hidden nature of the Heraclitan flux, but rather, to contrive – under cover of the claim to scientific status buttressed by purported application of scientific techniques – a surrogate reality to which the flux should be adjusted by fiat, its movement guided with the help of appropriate institutions, laws and policies all in the name of a posited optimum. If the recommended or coerced human action in aid of achieving the posited optimum has secured an extraordinary level of material well-being for much if by no means all of humanity, as indeed it has – the result is not due to the discovery of natural causal properties, but to the guidance or coercion of human behaviour towards the posited end. Much like Socrates, not believing that the “flowings or motions” can be known, economic theory holds that the material of economics must “*be something*”. But it isn’t, it is what we contrive it to be.

If a detailed examination of the process with which economic theory displaces the ceaseless becoming of the Heraclitan flux of economic reality with its Platonic Forms is beyond the scope of this paper, it seems reasonable to conclude on the basis of the above, that the strategy of reification, the transformation of the flux into objects does not work. Just as stock and flow in accounting cannot be reconciled in the sense of deriving the one from the other, Flux and Form remain, despite Searle’s efforts, incommensurable. The movement of the flux is the product of the agent’s action. Transforming the flux into objects so that the agent’s action can be disregarded – so that we have the Form without knowing what Flux brought it about – as the strategy for the discovery of causal dependencies is, therefore, a dead-end. There is no economics in nature and no nature in economics. It is all about conscious, purposeful human action.

5.4 Rule-Based Roles

Although John Stuart Mill committed economics to the tragicomedy of the utility maximising *homo oeconomicus* with his maxim that a greater gain is preferred over a lesser one, his more basic intuition to the effect that the purpose of economic action is to secure material goods seems reasonable even if Adam Smith’s notion, that the basic human motivation is gaining the esteem of fellow humans appears more convincing. But securing material goods is no simple matter. Apart from the need to develop suitable technologies, know-how and skills, there is a need to avoid unbridled chaos in its pursuit. Economic behaviour must be regulated in order to keep social peace. This need results in the construction of *rule-based roles*. *Economic agency is conditioned upon compliance with these roles.*

Human action arises from the agent’s dissatisfaction with some existing circumstance. The whole point of action is to change what exists or what is thought to exist to something more in line with the agent’s hopes and expectations. But society does not give free reign to any and all dissatisfactions and therefore does not permit any and all action. All economic action requires social authorisation and the authorisation is granted through the often highly specific and elaborate content of these roles.

A ‘seller’ is a person with legal title to the goods he sells and is held responsible for the truthfulness of the explicit as well as the implicit or attributed representations made in connection with the sale. A ‘buyer’ is an agent with the capacity to satisfy the conditions of the seller. Employers and employees have specified roles and non-performance of the role has serious consequences. In sharp contrast with the view of Rom Harré quoted and explored elsewhere in this volume⁴⁵ to the effect that “institutions are not ontologically basic...(that) an institution is an appearance, an illusion...(that) there is and was no banking system”, individual, group and institutional roles are formed and held together by powerful written and unwritten rules specifying authorised action and behaviour and sanctioning deviations from it. ‘Causation’ in agency is generated by the purposeful performance of a rule-based role. Economic roles have a clear purpose; intentionality is built into the role. The agent, by assuming the role, adopts the built-in intentionality. But there is a subtle cleavage between the actions of the self and the performance of a role by the self. Not being able to bring about economic action without adopting the role prescribed for the achievement of the desired end, she takes on the legislated intentionality of the role, but the gap between individual responsibility and responsibility for the performance of a role is evident. Accordingly, responsibility based on the operations of the conscience is replaced by the legally defined responsibility associated with the role. The most extreme form of this problem is the corporation and the question of corporate agency where the identity of the acting agent between the agents of the corporation and the corporation itself is disrupted by the corporate veil. Medieval Catholic theology was acutely aware of the problem,⁴⁶ saw roles as *ficta*,⁴⁷ (best translated as ‘made things’, or contrivances) and the concern dominated English jurisprudence well into the eighteenth century. Chief Baron Manwood, for example, remarks in *Tipping and Pexhal’s Cafe*⁴⁸: “Corporations are invisible and immortal and have no Soul; None can create Souls but God; but the King can create Corporations, therefore they have no Souls”.⁴⁹ Souls have virtues and vices. Immortal objects have or lack usefulness. Mixing the two leads to the loss of the Soul.

The hallmark of modernity is the resolution of this tension between the soul as the source of conscience and moral autonomy and the rule-based role of the self in favour of the latter. With it the traditional notion of agent intentionality recedes, and

⁴⁵ See Archer, M. p.

⁴⁶ See for example Pope Innocent IV’s decretal of 1246 in which he contemptuously dismisses the corporation as *persona ficta*.

⁴⁷ ‘Persona ficta’ is routinely translated in legal and other textbooks and monographs as ‘fictitious person, but this translation is entirely mistaken. Far from being a fictitious person, the modern corporation is the most powerful person in the modern world. The correct translation is based upon the root of ‘cicta’, namely ‘fingere’, and accordingly the term should be translated as ‘made or contrived person’.

⁴⁸ 2 Bulstr. 233; also in *The Laws of Corporations: containing the Laws and Customs of All the corporations and Inferior Courts of Record in England* (1702) London.

⁴⁹ Róna, P. “Ethics, economics and the corporation” in Róna, P. and Zsolnai, L. (2017) *Economics as a Moral Science*, Springer.

agent action is constrained to the rule-based role. Intentionality is turning a choice from a legislated menu rather than being freely formulated by the self.

This trend is further exacerbated by the use of models, algorithms and artificial intelligence in economic decision making, production, logistics, marketing and distribution. In these cases, the agent executes the tasks or adopts the results specified by these means without the ability to deviate from them. The ‘loan officer’ of the typical bank of today operates a computerised program for the determination of the applicant’s eligibility for a loan that does not permit the officer to proceed to the next question without an answer to the open one deemed satisfactory by the program. Whose intentionality is at play here? No one’s. The program determines the specifications of the role of the eligible borrower, and the applicant either fulfils this role or fails to do so.

5.5 Conclusion

Causation in the sense in which the term has meaning in the natural sciences does not obtain in economics because economic phenomena do not have the sort of causal properties that obtain in natural objects; the subject matter of economics is not a natural kind. The study of dependent relations in economics, therefore, is entirely the study of agency, where any inductively obtained generalisation about a dependent relationship is a consequence not a of the properties of a natural kind, but of human action. This remains to be the case even if agency and the sort of intentionality traditionally associated with agency implying the autonomy of the agent is being replaced by the rule-based role, and the rule-based role is increasingly replaced by algorithms and artificial intelligence. Dependent relationships secured through algorithms depend neither on the properties of the objects involved nor the intentionality of the agent, but, rather, are *prescribed relationships* based on the rule posited in the algorithm. The algorithm is given the capacity and power to override both Causation and agency and brings about consequences that could not be obtained without it. This phenomenon requires analysis going beyond both causation and agency, and therefore beyond the scope of this paper.

The triumph and the tragedy of modern economics rests not in its failure to discover law-like regularities resembling those found in nature, but in its imposition on society of a surrogate reality, built with concepts borrowed from the determinist world of Newtonian physics. But the application of Newtonian causality to economic life is illegitimate. Instead of discovering something about an existing reality, economics generates its own - in Hobbes’ phrase a ‘made with words’ reality - that conforms to its agenda. Economic optimality is just as much a subjective *telos*, the choice of which just as much the product of value judgment as any of the other value judgments economics purports to exile. Something is optimal in relation to something else the choice of which is not a matter of either science or logic. With its materialist determinism economics has compromised the moral autonomy of the

agent, expelled moral responsibility for the morally loaded choices it compels us to make, and transformed for the worse the physical world around us.

References

- Anscombe, G.E.M. 1971. *Causality and Determination*. London: Cambridge University Press.
- Aristotle, *Physics*, Book II, Chapters iv and v.
- Backhouse, R.E. 1997. *Truth and Progress in Economic Knowledge*. Cheltenham: Edward Igar.
- Collingwood, R.G. 1938. On the so-called idea of causation. *Proceedings of the Aristotelian Society (New Series)* 38: 85–112.
- . 1940. *An Essay in Metaphysics*. Oxford: Oxford University Press.
- Davidson, D. 1970. Mental Events. first given as a lecture at the university of Massachusetts, published in *Experience and Theory*, ed. L. Foster, and J.W. Swanson. The University of Massachusetts Press and Duckworth, reprinted in Davidson, D. (2013) *Essays on Actions and Events*, Oxford University Press; also reprinted in Sosa, E. and Tooley, M. (1993) *Causation*, Oxford University Press.
- . 1993. Causes and Conditions. *Journal of Philosophy* 64 (1967): 691–703.
- Friedman, M. 1953. The methodology of positive economics. In *Essays in Positive Economics*, ed. M. Friedman, 3–43. Chicago: University of Chicago Press.
- Gasking, D. Causation and recipes, 64, *Mind*.
- Goodman, N. 1983. *Fact, Fiction and Forecast*. Cambridge: Harvard University Press.
- Hart, H.L.A., and A.M. Honoré. 1959. *Causation and the Law*. Oxford: The Clarendon Press.
- Harré, R., and E.H. Madden. 1975. *Causal Powers*. Oxford: Basil Blackwell.
- Hicks, J. 1979. *Causality in Economics*. Oxford: Basil Blackwell.
- Hume, D. 1790. *A Treatise of Human Nature* (Analytical Index by L. A. Selby-Bigge); Second Edition with text revised and notes by P. H. Nidditch (1978) Oxford University Press.
- Hyman, J. 2015. *Action, Knowledge and Will*. Oxford: Oxford University Press.
- Illari, P., and F. Russo. 2014. *Causality*. Oxford: Oxford University Press.
- Keynes, J.M. 1971–1989. *The Collected Writings of John Maynard Keynes*. Macmillan/Cambridge University Press for the Royal Economic Society.
- Kuorikoski, J., and C. Marchionni. (eds.). 2014. Philosophy of economics. In *The Bloomsbury Companion to the Philosophy of Science*, ed. S. French and E. Saatsi. London: Bloomsbury.
- Mackie, J.L. 1965. Cause and conditions. *American Philosophical Quarterly* 2/4 (October): 244–255.
- Mäki, U. 1992. On the method of isolation in Economics. In *Intelligibility in Science*, ed. C. Dilworth, 319–354. Atalanta/Amsterdam: Rodopi.
- Marmodoro, A. 2014. *Aristotle on Perceiving Objects*. Oxford: Oxford University Press.
- Morgan, M. 2012. *The World in the Model*. Cambridge: Cambridge University Press.
- Newman, P. 2003. *F. Y. Edgeworth's Mathematical Psychics and Further Papers on Political Economy*. Oxford: Oxford University Press.
- Popper, K. 1990. *A World of Propensities*. Bristol: Thoemmes Antiquarian Books, Ltd.
- Putnam, H. 2004. *Ethics without Ontology*. Cambridge: Harvard University Press.
- Róna, P. 2017. Ethics, economics and the corporation. In *Economics as a Moral Science*, ed. P. Róna and L. Zsolnai. New York: Springer.
- Rosenberg, A. 1992. *Economics – Mathematical Politics or Science of Diminishing Returns*. Chicago: Chicago University Press.
- Russell, B. 2012. On the notion of cause. In *Proceedings of the Aristotelian Society*, vol. 13, 1–26.
- Searle, J.R. 1983. *Intentionality*. New York: Cambridge University Press

- von Mises, L. 2006. *The Ultimate Foundation of Economic Science; an Essay on Method*. Indianapolis: Liberty Fund.
- Von Wright, G. 1975. *Causality and Determinism*. New York: Columbia University Press.
- Walras, L. 1874. *Éléments d'Economie Pure*, Cobran 1313 Cie, Lausanne, translated by William Jaffé (1954) of the 1926 definitive edition. London: George Allen and Unwin Ltd.
- Wigner, E. 1960. The unreasonable efficiency of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics* 13: 1–14.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Part II

Praxis

Chapter 6

Why Aquinas Would Agree That Human Economic Behaviour Is Largely Predictable



Richard Conrad and Peter Hunter

Abstract Many people, from retailers and advertisers to the Chancellor of the Exchequer, work on the assumption that human economic behaviour is to a fair degree predictable, at least statistically. This paper asks how far the thirteenth-century Thomas Aquinas would agree that human behaviour (including economic behaviour) is predictable, both the behaviour of individuals and the behaviour of groups, and on what grounds. In doing so it also asks how any elements of predictability would square with Aquinas' conviction that human beings enjoy *liberum arbitrium*, "free decision". In the context of the present volume, exploring Aquinas' position may promote a nuanced and multivalent approach to the question of what causes human behaviour, and liberate us from the fear that if human behaviour is *caused*, it cannot be *free*.

Aquinas was aware of the extreme complexity of the human psyche and of the organic interactions among its components. In particular, *liberum arbitrium* is achieved in interaction between intellect and will. The human will is *the rational appetite*, the ability to *be attracted by* the good perceived by reason. Any predictability of behaviour is therefore not a statistical result of intrinsic arbitrary randomness, as if acts of will were a kind of mental coin-flipping. Truly free behaviour is rationally explicable in terms of the goals it is right for human beings to pursue; *final causality* operates, in a way appropriate to responsible agents.

In an ideal world, not marred by sin, this would make human nature predictable to a *limited* degree. People would behave sensibly, as individuals and as communities, avoiding anything harmful. But people *naturally* differ in talents and temperament; geographical and historical circumstances vary; and human thinking is open-ended – within the time available, we can examine a situation from different points of view. In an unfallen world, people would happily adopt different social roles, and leaders' decisions about how to apply the Natural Law to particular

R. Conrad (✉) · P. Hunter (✉)
Blackfriars Hall, Oxford University, Oxford, UK
e-mail: richard.conrad@bfriars.ox.ac.uk; peter.hunter@bfriars.ox.ac.uk

circumstances would be sensible, but also “prudential” as not determined by a rigid reasoning process such as geometry uses. Within a context of good citizenship, people would make varying choices about practical matters, to the extent that an unfallen race would be more interesting and vital – because *more* human – than a fallen one.

Our world is not ideal. The basic dynamics of intellect, will and emotions remain good, but individual temperaments can include propensities to vice as well as to virtue. Intellect and will are in important ways blank slates at birth, and embedded in a biological and social context. In the long growth towards moral maturity people are vulnerable to corrupt customs which can obscure even obvious points of the Natural Law. Though God’s grace is operative, not many people achieve the full moral freedom of an integrity in which emotions *enhance* a rational behaviour personally owned. When people do build up virtue, and thereby are partially liberated from the effects of the Fall, it becomes possible to predict in a general way that they will behave virtuously. But the open-endedness of thinking means that virtuous people will make varying choices about practical matters, *more* interesting and various than the dull and sadly predictable behaviour of people tied to various vices.

Since thorough-going vice is unnatural, most people tend towards a mediocre, partial moral consistency and *behave rationally enough*, obeying laws that carry sanctions, and listening to the more persuasive good advice; this will result in a certain predictability of behaviour on the part of the majority. They will also tend to follow bad laws and false persuasion without adequate reflection. Further, the failure to develop full rational control of their emotions (a control that must be “political” rather than “despotic”) leaves people vulnerable to emotional drives: in the here-and-now they often perceive lesser, but easy and immediate, goods as preferable to greater goods that they *know* are better, but which are more demanding and distant. At the personal level, individual temperaments lead to a certain predictability of behaviour; at the social level, predictability may result from the proportion of temperaments that is statistically likely. Aquinas saw these temperaments as due to inheritance and astrological influences; we would replace astrological explanations by genetic ones and a better understanding of children’s psychological development in its social context. For Aquinas, astrological influences remained potent throughout life, influencing the will *indirectly*, through the imagination and humours. We reject that form of predictability, but psychological experiments show that subconscious environmental factors, as well as fashion and peer-pressure, are potent. Aquinas also saw good and bad angels as influencing the human imagination and humours. Whether or not we agree with him on that, we recognise elements of unpredictability in social behaviour that are due to our vulnerability to unexpected mass movements, mass movements that we are inclined to label “demonic” – though there can also be *good* (“angelic”) mass-movements.

In conclusion: Aquinas would agree that human behaviour is predictable in some degree, but his perceptive *pre-modern* understanding of human psychology invites us to reflect afresh on the nature of freedom and on the forms and causes of predictability.

6.1 Introduction

Many people, from retailers and advertisers to the Chancellor of the Exchequer, assume that human economic behaviour is to a fair degree predictable, at least statistically. They expect forms of persuasion, and fashion, to *cause* enough people to buy certain goods to make advertising, or stocking up on certain goods, profitable. They *predict* that increasing the tax on tobacco will *cause* a worthwhile proportion of smokers to quit the habit. This paper examines how far Thomas Aquinas would agree that the behaviour of individuals and of groups is predictable, and on account of what factors. In doing so it will touch on how elements of predictability would square with his conviction that human beings enjoy *liberum arbitrium*, “free decision”.

Aquinas¹ says little on economic behaviour as such,² but is worth including in this volume because he worked before 1277. In that year certain “Aristotelian” propositions were condemned by Paris and Oxford Universities, an event David Luscombe describes as a “watershed”.³ Aquinas and Bonaventure disagreed about the relative priority of intellect and will, but agreed that free decisions arise within the interaction of these components of human nature.⁴ Scotus, working after 1277, roots freedom in an *affectio iustitiae* distinct from our natural intellectual appetite for happiness.⁵ A broad-brush survey of Western thought might see this dissociation of freedom from our “natural dynamics”, plus Ockham’s “voluntarism”, as profoundly affecting the way free will is perceived by many today. Arguably, Aquinas approach to issues of human agency is refreshingly pre-modern, and stimulates a nuanced, multivalent approach to the causes of human behaviour while allaying the fear that if our behaviour is *caused*, it cannot be *free*.

¹References to Aquinas’ *Summa Theologiae* are given in the format Part Question, Article. 1a = *Prima Pars*, 1a2ae = *Prima Secundae*, etc.

²He condemns usury in 2a2ae 78 (cf. *De Malo* 13, 4) but in 78, 2 ad 5 recognises the legitimacy of making a profit (or loss) on a joint project in which one invests.

³*Medieval Thought. A History of Western Philosophy*, II. Oxford: OUP, 1997. 114–121.

⁴Bonaventure, *Commentary on Book II of the Sentences*, Dist. XXV, Part I, Qq. 3 & 5.

⁵Thomas Williams, “How Scotus Separates Morality from Happiness.” *American Catholic Philosophical Quarterly* 69 (1995) 425–445.

6.2 Free Decision *Within* a Complex Psyche

The essential background to our discussion is Aquinas' acute awareness of the extreme complexity of the human psyche. The salient points are summarised here.⁶

6.2.1 *Abilities of a Complex Form of Life*

Aquinas uses nouns like “intellect” and “will”. These “powers of soul” should not be reified as if they were departments in an office-block, each with its own decision-making power. A soul (of a plant, animal, or human) is a “form of life” unifying and animating the organism. Its powers (Aristotle’s “potentialities”) are *abilities of the whole organism* that result from its form of life. All living things are able to take in nutrition, to grow and to reproduce. Animals can also perceive and respond in subtle ways; human beings can perceive and respond both in “animal” and in specifically human ways.

6.2.2 *Animal Abilities to Interpret and Respond*

Some abilities are “active”. Digestion *works on* the food we ingest; our “agent intellect” *works on* the rich and complex material in the imagination. Some abilities are “passive” in the technical sense of “receptive”. The power of hearing is the animal’s ability *to be affected by* sound so as to hear things; the “emotions”⁷ are its abilities *to be attracted by* what it perceives as suitable, and *repelled by* what it perceives as noxious.

⁶Relevant texts include: 1a 77; 78; 79, 2–3; 80–86 (collected in R. Pasnau, *Thomas Aquinas on Human Nature: A Study of Summa Theologiae* 1a 75–89. Cambridge: CUP, 2002); 1a2ae 8–10; 22–23; 25. Secondary literature includes:

Diana Fritz Cates, *Aquinas on the Emotions: A Religious-Ethical Inquiry*. Washington DC: Georgetown University Press, 2009. Powers of soul are summarised on pp. 267–8.

R. Pasnau and C. Shields. *The Philosophy of Aquinas*. Boulder: Westview Press, 2004. Chapter 7 (on sensory and intellectual powers).

Nicholas E. Lombardo, *The Logic of Desire: Aquinas on Emotion*. Washington: Catholic University of America Press, 2010. Chapters 1–4.

Robert Miner, *Thomas Aquinas on the Passions: A Study of Summa Theologiae* 1a2ae 22–48. Cambridge: CUP, 2009.

⁷*Passiones animae* are basic patterns of attraction and repulsion. To call them “passions of soul” might imply too much *passion* to suit them all. To call them “emotions” risks importing a modern psychological concept, but does hint at the complex social life of the higher animals, and so remind us of what we share with them.

Both kinds of ability have inbuilt dynamics, partly due to a shared nature, partly due to “individual nature” (e.g. the biochemistry of digestion is common to all humans, but some have more robust digestive systems than others).

Besides the five senses of touch, sight, etc., animals coordinate the sense-data received, and recall it *via* the “imagination”. They also *make sense of* their world: by their “estimative sense” they perceive meanings like danger, and “affordances”.⁸ Hence they react to stimuli both internal (e.g. hunger) and external.

We share many abilities with other animals; much *human* consciousness is “animal consciousness”.

6.2.3 *Limited Conscious Control; Pre-conscious “Acts”*

Some abilities are not under conscious control. We can decide when and what to eat, but not what the body does with what we eat.

Much coordination and interpretation of sense-data goes on *pre-consciously*; we become aware of plates and food, not “raw” patches of colour.⁹ Reactions, too, are partly pre-conscious: when charged by a bull we *automatically* feel fear. Both Aquinas and we attribute such activities to the brain, plus other bodily structures – nerve impulses and hormones in our case, blood vessels and humours in Aquinas’.

6.2.4 *Rational Perception and Reaction*

Human beings have abilities other animals do not, namely *intellectus* and *voluntas*, “intellect” and “will”.

The intellect is our ability to draw out and grasp *universal truths*, including those of applied sciences such as ethics-politics. We *abstract* universal concepts from particular instances located within space and time, and organise them into bodies of knowledge.

As *the rational appetite* the will is, in the technical sense, “passive” – the ability to *be attracted by* the good that reason perceives. It can “rise above” particular drives to “higher goods”, even what Aquinas calls “universal good”. I can resist hunger for the sake of political protest, or choose a painful medical procedure for the sake of future health.

⁸A term coined by J. J. Gibson. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin Harcourt, 1979. 127.

⁹So good are we at seeing *things*, it is hard to become a good visual artist or cartoonist and isolate the shapes and patches of colour we need to focus on when painting or drawing.

6.2.5 *Co-operation Between Intellect and Sensory Abilities, Between Will and Emotions*

Since “materiality” ties things down in space and time, Aquinas sees intellect and will as spiritual powers, not *in themselves* the functioning of bodily organs.¹⁰ This leads to a nuanced account of what external factors can influence them. However, our active *use* of concepts requires close cooperation between intellect, on the one hand, and the imagination and the estimative sense, on the other. We apply general knowledge to particular instances (e.g. a vet uses her knowledge of diseases in diagnosis). Even in *abstract* thought we make use of mental pictures, examples, implicit symbols, etc. Hence whatever affects brain functioning can affect our ability to think – sleep, drunkenness, brain injury, strong emotion... Both intellect and the senses are *transformed* by this close cooperation¹¹: we are rational in an animal way, and animal in a rational way. Our imagination can “play” with memories; our estimative sense is transformed into the “cogitative sense”.

Will and emotions also influence each other. We are voluntary in an animal way, and animal in a voluntary way. Emotional “drives” can make things attractive or repugnant to us – to us as *responsible* beings. Will and emotions cooperate: many physical movements are under voluntary control, but it is through the emotions that the will puts them into effect, while our emotions wait upon the will’s command before initiating deliberate movement.¹²

6.2.6 *Co-operation Between Intellect and Will in Free Decision*

A key point is that *liberum arbitrium* – sometimes translated as “free will”, better translated as “free choice” or “free decision” – *is achieved in two-way interaction between intellect and will*. The cooperation between thinking and wanting is so close that Aristotle spoke both of “desirous reasoning” (*orektikos nous*) and “reasoning desire” (*orexis dianoetikē*)¹³; following NE 3.3 (1113a11) Aquinas considers “intellectual appetite”¹⁴ better than “appetitive intellect”.¹⁵ Thinking and

¹⁰ 1a, 75, 2; 1a2ae 9, 5.

¹¹ Candace Vogler, “The Intellectual Animal.” A lecture delivered at Blackfriars, Oxford, on 2 March, 2017. Available <https://www.youtube.com/watch?v=4lhNm1fa8GE>. *New Blackfriars* (forthcoming).

¹² 1a 81, 3.

¹³ *Nicomachean Ethics* (NE) 6.2 (1139b4).

¹⁴ To call the will “*intellectual appetite*” does not mean it is always “highbrow”. I have an intellectual appetite for dark chocolate, since I *know* I will enjoy it.

¹⁵ 1a 83, 3–4.

wanting are so entangled that Mark Jordan speaks of “the untellable circlings of will and intellect”.¹⁶

6.2.7 *Development of Habits and Virtues*

Many abilities are “open-ended” and can be “moulded” further by training and practice. We can “shape” our intellects with concepts of economics or engineering. We can be trained as children, or train ourselves later, to enjoy adult food.¹⁷ This channelling of an ability is a “habit”, not in the sense of a nervous tic that takes us over, but rather like a learned skill we can deploy. A virtue, a good habit, is an enabling life-skill. A bad habit, a vice, means some aspect of our behaviour is habitually mis-directed.¹⁸

We can develop “intellectual virtues”, “strengths of mind”, that empower us to use our minds readily in theoretical and practical ways.¹⁹ These do not make us good people. We can also develop “moral virtues”, “strengths of character”, that bring our emotions and wills into harmony with right reason, so that “by second nature” we desire what is truly good, truly fulfilling, and live it out with readiness and a sense of fulfilment.²⁰ Moral virtues are deployed by the intellectual virtue *prudentia* (“prudence” or, better, “good moral sense”) and in turn support it, so that, unlike sciences and crafts, it is inseparable from *being* good.

6.2.8 *Limited Conscious Self-Awareness*

Consciousness is not a core concept for Aquinas: animals are conscious in various and shifting ways, and so are we. Being conscious of toothache makes it difficult to be conscious of other things.

In particular, we are not conscious of the whole “contents” of our intellects *or of our wills*. The human psyche is not geared to introspection. I am not consciously aware of concepts I have learned until some situation prompts me to call upon them. Nor am I fully aware of the priorities I hold, the “habits” that “structure” my will, until they result in *acts* of will.²¹ Both concepts and willed choices emerge from “the

¹⁶ *Teaching Bodies: Moral Formation in the Summa of Thomas Aquinas*. New York: Fordham University Press, 2017. 102.

¹⁷ This reminds us to hear “passive ability” in a nuanced way.

¹⁸ Habits are treated in 1a2ae 49–54; virtue in 55–56.

¹⁹ 1a2ae 57.

²⁰ 1a2ae 57–61. For Aristotle and Aquinas, reason takes charge of emotions in “political”, not a “despotic” way (1a 81, 3 ad 2; 1a2ae 17, 7): emotions have dynamics we must work with sympathetically.

²¹ 1a 87; 1a2ae 112, 5.

habitual retention of knowledge and love”.²² We sometimes *discover* our wants; I may *surprise* myself by what I do – it may be different from what I supposed I wanted to do!²³

6.2.9 Influences Upon “Embedded” Free Decision

It should now be obvious that internal and external factors affecting our bodies affect both our outward senses and our imaginations and emotions. Through them they can impinge on our intellects and wills, both because thinking draws on and cooperates with the senses, and because we can perceive our internal state (of hunger, health, emotion, etc.) to some extent. Our free decisions involve perceptions and motives at various levels of our psyche, many of which have been shaped by past behaviour and interactions.

The human will does not spring into being as part of a fully-formed, self-sufficient, adult psyche able to select its choices within a landscape it surveys. We come into being (a) needing to grow towards the use of reason, (b) “embedded” in a biological and psycho-social situation on which we are highly dependent – and to which we are vulnerable – and (c) with both intellect and will as “blank slates”²⁴ even though they have intrinsic dynamics towards the true and the good. In our growth to maturity we remain highly dependent on human interaction; in developing virtue we must deal sympathetically and practically with the effects nature, nurture and earlier decisions have had on our emotions.

At no stage is the human will an “unmoved first mover”, spontaneously and independently bringing an act of will into being.²⁵ It is true that while Aquinas saw external factors as able to influence our senses and emotions, and to some extent our intellects, he held that *no external agent could directly influence the human will*. However, he insisted that whenever *any* human being’s will acts, it is attracted into “act” by God as the Unmoved First Mover. For we are “embedded in God” who is in fact *within* the will, (a) as the Source of all being who holds the will and its “acts” in being, and (b) as the Infinite Good, the ultimate attractive Goal, who attracts our will into seeking good.²⁶ Further, by Grace God can “enlarge our hearts” (Psalm 119:32) to embrace him as our *Friend* and to liberate us from false thinking

²² 1a 93, 8.

²³ Nicely described by Gareth Moore, *The Body in Context: Sex and Catholicism*. London: Continuum, 2001. 16.

²⁴ 1a 84, 3.

²⁵ 1a2ae 9, 4.

²⁶ 1a2ae 9, 6. Bonaventure, too, denied that a *liberum arbitrium* “presides over” intellect and will, initiating their directions of thinking and loving. *Commentary on Book II of the Sentences*, D. XXV, Part I, Q. 2.

and willing. Grace is typically mediated through the Sacraments God gave to the Christian community.²⁷

This outline of the contexts of our free decisions warns us not to expect a simple – let alone a reductionist – analysis of what might make for predictability in human behaviour.

6.3 Explicable But Open-Ended Freedom

To see the will as the ability to *be attracted by* the good perceived by reason is very different from a “voluntarism” which (to some extent caricatured) sees the will as a free-floating, pro-active, arbitrary deciding power surveying a landscape of options and saying, without motive or other such constraint, “I shall have X today.” This would make an act of will the mental equivalent of flipping a coin. Pinckaers labels the resultant view of freedom “freedom of indifference”: the will is faced with a range of options “on an even field”.²⁸ Arguably, it lies behind the presumption that increasing people’s freedom is a matter of maximising their choices. Modern discussions of freedom, and Benjamin Libet’s experiments,²⁹ do often focus on such arbitrary choices. If individual choices were *purely* arbitrary, the resulting randomness might translate up into statistical predictability of behaviour. If, as Libet’s experiments have suggested to some, choices are made prior to conscious awareness, they might become predictable once the biochemical or other factors that precipitate them are discovered.

However, investigations of arbitrary, random choices do not engage with Aquinas’ account of *truly human, deliberate* choice, for he would label unthinking choices “acts of human beings”; by contrast, “*human acts*” engage us as rational, responsible, goal-seeking agents.³⁰ They are free by what Pinckaers terms “freedom for excellence”,³¹ that is, by “responsible ownership” of decisions. Motives and reasons do not *constrain* us, but *contribute to* freedom: if I can *explain* my behaviour, this means I have *owned* it by reflection on what is good for me, i.e. what on truly contributes to my well-being.

This view of freedom goes with a “metaphysics of morals” in which characters, decisions and actions are susceptible of more or less goodness; and insofar as they have more goodness, they have more being or reality, more integrity and truth.³² Hence freedom is susceptible of varying degrees. If our hierarchy of values corresponds to

²⁷ 1a2ae 62; 109–112; 3a 62.

²⁸ Servais Pinckaers, *The Sources of Christian Ethics*. Edinburgh: T & T Clark, 1995. Chapter 14.

²⁹ Helpfully described and critiqued at http://www.informationphilosopher.com/freedom/libet_experiments.html

³⁰ 1a2ae 1, 1. As Mary Midgley remarked, “Randomness would not be freedom.” *Wickedness: A Philosophical Essay*. London: Routledge & Kegan Paul, 1984. 111.

³¹ *The Sources of Christian Ethics*, Chapter 15.

³² 1a2ae 18, 1; “deep down”, truth and unity “coalesce” with goodness and being: 1a 5, 1; 11, 1; 16, 3.

the truth of things, and if we have the integrity to pursue wholeheartedly what is really good for us, we are the more free. Truly free behaviour is *more* rationally explicable than less free behaviour – explicable in terms of the goals it is happy-making for human beings to pursue.

The causality involved in truly human behaviour is **final causality**. It operates in a way appropriate to rational agents who act in the light of goals held with (ideally) increasing degrees of coherence.³³

Insofar we behave in a human way, any predictability of our behaviour will not result from the statistics of random choices, or from *mere* biochemistry, but from the *rational explicability* of responsible choices. Given our psyche's "embedded complexity", one might hope to predict what an individual would choose in a given situation knowing her "personal chemistry", her acquired patterns of behaviour, and her account of her motives, her goals. However, besides our susceptibility to outside influences, there are (a) a "cascade effect" in "the untellable circlings of will and intellect" which seems to limit the predictability of particular choices, especially ones that do not engage any portentous moral judgment, and (b) elements of mystery in the human psyche:

- The will has an inner dynamic towards "the good" in the sense of *beatitudo* (equivalent to Aristotle's *eudaimonia*: happiness, flourishing, fulfilment). This is not a matter of free choice; we want to be happy *voluntarily*, but *cannot choose* whether or not to want to be happy.³⁴
- This inbuilt dynamic responds to a prior apprehension of "the good", and of the good as to be pursued, built into the intellect, which Aquinas calls *synderesis*.³⁵
- Our drive to happiness energises reflection on what will make us happy, which we thereby find attractive. We can perceive power, pleasure or wealth, etc., as our chief priority – or God.³⁶ In an important sense people must make this basic decision once they have the use of reason.³⁷ It can be an *implicit* decision; and we can change it. We can even change our priority from God to something else.³⁸
- In the pursuit of more basic ends, we choose means to them; *this* is where free decision comes in.³⁹
- Free decision is not needed when there is only one obvious means to a goal.⁴⁰
- In selecting a means, we exercise free choice by *comparing contingents*. Choice is free because our *thinking* is open-ended: we can note advantages and disadvantages of various means, and (since decisions cannot be deferred indefinitely)

³³ 1a 60, 2; 1a2ae 1, 1 & 6; 6, 1.

³⁴ 1a 82, 1; 1a2ae 10, 1.

³⁵ 1a 79 12; 82, 4 ad 3; 1a2ae 94, 2.

³⁶ 1a2ae 2 details these possibilities.

³⁷ 1a2ae 89, 6.

³⁸ Turning against God is rarely explicit; more often it is *implied* by some serious sin: 2a2ae 34, 2.

³⁹ 1a 82, 2; 1a2ae 8, 2–3; 14, 2.

⁴⁰ 1a2ae 10, 2.

a certain indeterminateness and spontaneity operates here – within the time available we are not determined to choose a particular means.⁴¹ We can fail to take something important into consideration, and *failure* can be *voluntary*.⁴²

- Having made a careful decision, we can still fail to carry it through owing to factors such as fear and laziness. That is, we can fail to follow our conscience. This failure, too, is *voluntary*.⁴³

An “economic” example illustrates how, in this “cascade”, a means becomes an end for a subordinate decision, and how difficulties at lower levels can prompt rethinking at higher levels. If a concert is arranged in Manchester, it might be predicted that many people will attend, so that special train services are laid on. If my musical interests make attending the concert a “medium-sized” goal, I investigate means to that end (e.g. train travel) and, with those means acting as a subordinate end, I weigh up further means (buying a ticket on a chartered train, an advance-purchase ticket on a specified train, or an open return). Which I choose depends on how I perceive the alternatives (flexible timing *versus* a cheaper fare; a quiet journey *versus* the “buzz” of a chartered train). Maybe *no* means is open (there are no advance-purchase tickets; I can’t afford a flexible ticket; I cannot cope with the crush on a chartered train), and I give up my “medium-sized” goal. A means might be open if I sacrifice a higher good: I could steal the money – then my *basic* goal comes into question, since theft is sinful, and (hopefully) I still give up the medium-sized goal. I may well not weigh up every option – but failing to consider a certain option *may* be a voluntary mistake (e.g. I don’t notice a very cheap plane fare, either because I couldn’t have realised a plane service had been established, or because I knew there might be one but didn’t look into it). Further, while pursuing the medium-sized goal, I might culpably fail to think of every relevant thing, and so, culpably, fail to respect a greater goal (e.g. I promised to meet an important deadline, and going to Manchester makes me break my promise).

Clearly, such thinking does not have the compelling force of a geometrical proof. Even if my friends could have predicted I would try to attend the concert, and even if the train company correctly predicted the number travelling by train, my decision was not determined. In retrospect I will be able to explain my actions, but in doing so I may realise I made bad decisions, or failed to execute good ones as I should. Rationality is compatible with elements of spontaneity, mystery and open-endedness; open-endedness is multiplied by the complexities both of the psyche, and of the situations among whose contingencies we negotiate practical decisions.

Many decisions in regard to the concert entailed no disobeying a serious moral law, hence one could not determine whether I would attend it simply on the basis of my goodness. Yet, Aquinas holds, no concrete human act is morally indifferent. Buying a first-class ticket might be no sin, but it would be “more perfect” to buy a cheaper ticket and give money to charity, and yet more perfect to miss the concert

⁴¹ 1a2ae 13, 5–6; 14, 2, 3 & 6.

⁴² 1a2ae 6, 3 & 8.

⁴³ 1a2ae 19, 5, cf. 6, 6–7.

and spend the time working in a hospice.⁴⁴ A great deal must fall into place for human decisions to be made, and yet more if our characters and behaviour are to have perfect integrity and truth, and thus perfect goodness.

6.4 In Humanity's Ideal State, Would Behaviour Be Predictable?

On Aquinas' account of the Fall, the first human beings were endowed with gifts such as moral integrity; if they had not sinned, all human beings would have been born with those gifts.⁴⁵ Children would still have grown towards the full use of reason, and acquired knowledge,⁴⁶ but their wills would have been formed by Charity.⁴⁷ Aquinas sometimes performs a "thought-experiment" about what an unfallen state would have been like to help distinguish what is natural to us both from what is the result of the Fall, and from what (being supernatural⁴⁸) must *always* be God-given.⁴⁹

Some of Aquinas' speculations have implications for the predictability of behaviour in an ideal state. People would naturally have behaved sensibly, as individuals and as communities, avoiding anything harmful and practising good citizenship.⁵⁰ They would have engaged in practical reasoning because of the complexities of the world and of society. People would have been unequal: owing to shifting patterns of the stars, and variations in climate, some would have been stronger and wiser than others, though there would have been no birth *defects*.⁵¹ By free choice, some would have advanced in knowledge and justice more than others.⁵² There would have been no slavery; people would have been governed for *their own* and for the *common* good, not *used*.⁵³ Within the context of good citizenship, people would have made varying choices about careers to pursue, where to live, whom to marry – "personal" but not irrational choices. It would have been possible to count on a balanced pattern of social roles. Lawgivers' decisions about how to apply the Natural Law to particular circumstances, and leaders' policies, would have been sensible yet also "prudential",

⁴⁴No act actually performed is morally indifferent (1a2ae 18, 9), but Aquinas distinguishes precepts from counsels (1a2ae 108, 4): we are not obliged under pain of sin to follow counsels.

⁴⁵ 1a 95; 100, 1.

⁴⁶ 1a 101.

⁴⁷ Charity is a love for God that empowers a journey into him: 1a2ae 62.

⁴⁸ The "supernatural" is not the "spooky", but the divine. Forgiving enemies is more supernatural than levitating.

⁴⁹ Besides 1a 94–101, see 1a2ae 109, 1–5.

⁵⁰ Hinted at in 1a 97, 2 ad 4.

⁵¹ 1a 96, 3.

⁵² *Ib.* This affirms the claim that, *without sin*, people can choose among more and less perfect options.

⁵³ 1a 96, 4.

not determined by rigid reasoning process such as geometry uses.⁵⁴ Clearly, an ideal society would be anything but a uniform, regimented army. An unfallen race would be more interesting, vibrant and vital – because *more* human – than a fallen one. People would have been free for excellence.

The thought-experiment reassures us that variety⁵⁵ and vibrancy are *natural*. A fallen world's warped perspective might suggest that vice is interesting and virtue boring, but the opposite is true; this meshes with Aristotle's conviction that while virtue is a *mean* between extremes, it is not mediocre,⁵⁶ and with Aquinas' implication that friendship with God grows into an "exhilarating resonance".⁵⁷ In the world as it is, people often are liberated from some effects of the Fall by God-given virtues and, I should argue, by the ways we help each other acquire virtues.⁵⁸ While we can predict that virtuous people will behave well, good citizenship does not make them clones of each other. Human variety, the open-endedness of practical thinking, and the strange beauty of grace, mean that virtuous people's choices will often be more interesting and various than the dull and sadly predictable behaviour of people tied to various vices.

6.5 Fallen, Vulnerable Humanity's Predictability

For Aquinas, the Fall deprived us of *supernatural* gifts such as Charity; these are restored through Christ. Along with these gifts, others, that came to be called "preternatural", were also lost⁵⁹; in God's providence, these are not restored in full in this life.⁶⁰ They remedied the physical and psychological vulnerability that are *natural* to us as complex and interdependent, hence their loss has wounded us: our darkened intellect no longer has a firm grasp of the principles of moral reasoning, a grasp that actively pervades our practical decisions; our will is weakened, and our emotions often run ahead of reason, or impede it.⁶¹ There remain natural, and basically good, dynamics within our faculties⁶²; Mary Midgley brought out – in a way similar to Aquinas – the dangers attendant on failing to integrate these with

⁵⁴ 1a2ae 96, 1 ad 3. Cf. the need for counsel, circumspection and caution: 1a2ae 14, 1; 2a2ae 49, 7–8.

⁵⁵ Naturally, in a fallen world there can also be immoral variations in preference.

⁵⁶ *Nicomachean Ethics* 2, 6 (1107a6–8).

⁵⁷ Andrew Pinsent, *The Second Person Perspective in Aquinas's Ethics: Virtues and Gifts*. London: Routledge, 2012. 96–98.

⁵⁸ To recognise how moral education, like medical techniques, can alleviate some effects of the Fall is not to deny our dependence on God for the restoration of Charity, and for total healing in the final resurrection.

⁵⁹ 1a2ae 82; 2a2ae 164. Moral/psychological integrity and immunity from sickness are "preternatural" since they could conceivably exist in people who loved God as Creator, but to whom he had not offered the *supernatural* goal of sharing his bliss.

⁶⁰ 3a 69, 3.

⁶¹ 1a2ae 17, 7; 24, 3 ad 1; 85, 3 & 5–6.

⁶² 1a2ae 24, 2 & 4; 63, 1.

rational considerations.⁶³ It is difficult to attend to all relevant aspects of a situation; and to carry through our decisions we must often overcome laziness, fear, or disordered desire.

Aquinas affirmed Aristotle's account of how the training of children, good laws, the advice of friends, and personal practice, promote "acquired" virtues that restore some degree of moral integrity.⁶⁴ All this is liberating, not coercive. It involves practical ways of dealing with emotional difficulties; for example sadness may require companionship, sleep, bathing, legitimate pleasures, or cathartic weeping⁶⁵ rather than "pulling yourself together" – we would add suitable drugs, when medically indicated, to the list.

Unless they are so strong as to dement us, emotions cannot override freedom, but they can affect it.⁶⁶ It is difficult to achieve a moral integrity in which emotions *enhance* a rational behaviour personally "owned". Since vice, in which both reason and the affective powers cooperate in evil, is unnatural,⁶⁷ thorough-going vice is relatively rare. Aquinas seems to share Aristotle's suspicion that many people are neither virtuous nor vicious, but "controlled" or "uncontrolled"⁶⁸; if, for example, the presence of onlookers shames them, they do what is good "through gritted teeth"; sometimes they do evil, but "with a bad conscience". Many people follow their emotions, and do not make the effort to rise above them⁶⁹; they often perceive lesser goods that are easy and immediate as preferable to greater goods that in principle they *know* are better, but which are more demanding and distant.⁷⁰

We can expect most people to tend to a mediocre level of morality in which they *behave rationally enough*, at least when "controlled", and especially as regards the most basic principles of moral/practical reasoning that remain innate and potent.⁷¹ We can predict that most humans will make efforts to keep themselves alive, will beget children and care for them, will be social, and will worship God (or "gods").⁷² In a reasonably well-run society, we can count on most people obeying laws that carry sanctions and listening to persuasive advice; for example, if certain foodstuffs are lauded as healthy by respected authorities, people will buy them, unless they are so costly that doing so would deprive them of basic necessities. If certain substances are known to be dangerous, or are prohibitively expensive, or their use carries severe sanctions, few people will be so irrational as to purchase them.

At the same time, a failure to develop full rational control of their lives will leave many people vulnerable, not only to emotional drives, but also to bad laws and cor-

⁶³ *Wickedness*, Chapters 1, 2, 4 & 9.

⁶⁴ E.g. 1a2ae 63, 2; 92, 1.

⁶⁵ 1a2ae 38.

⁶⁶ 1a2ae 6, 6–7.

⁶⁷ 1a2ae 71, 2.

⁶⁸ *Nicomachean Ethics* 7, 1–3 (1145a15–1147b19).

⁶⁹ 1a 115, 4 ad 3; 1a2ae 9, 5 ad 3.

⁷⁰ 1a2ae 63, 1 ad 4; 75, 1–2; 2a2ae 20, 2.

⁷¹ 1a2ae 100, 3.

⁷² Cf. the natural inclinations in 1a2ae 94, 2.

rupt customs, to threats and peer-pressure. Human and God-given courage, and often the support of friends, empower some people to resist, but coercive rulers can often count on a large proportion of people yielding to even unpopular policies.

If many people do not rise above their emotions, personal temperaments (due to nature and nurture) will lead to some predictability in their behaviour once we know them and their backgrounds. We will examine below the factors Aquinas saw as influencing personal temperament. It seems he would not reject a statistical predictability resulting from the proportion of temperaments that is likely, or observable, owing to genetic, geographical and social factors.⁷³

To those he graces, God imparts a panoply of virtues that gradually reintegrate our psyche, liberating us from Original and actual sin,⁷⁴ though in most cases moral reintegration is partial this side of the grave. Aquinas does not predict that those whom God graces will regularly behave in ways that strike society as odd: Charity does not typically snatch people away from family commitments, but typically affirms, purifies, orders, divinises and widens natural affections.⁷⁵ All the same, in a fallen world there is a war on against evil, so that Charity always leads to *some* dramatic gestures (e.g. fasting)⁷⁶ and makes people wary of following the herd uncritically. In *some* cases it leads people to adopt odd forms of life, e.g. by making religious vows.⁷⁷ In the Middle Ages this had an immense, and arguably beneficial, *economic* effect on society.⁷⁸ It would be interesting to assess the current impact of ways in which devotion prompts people to stand out from the herd.

6.6 Factors Causing Predictability, Especially of the Majority

When reason is truly free, our decisions have a vitality that *sometimes* make us stand out from the herd, and we have the insight and integrity to assess laws, peer-pressure and emotional urges rather than blindly follow them. But Aquinas held that many people do not take full rational control of their lives, with the result that factors operating at less personal levels to influence their emotions and perceptions will affect their behaviour unduly, and often in ways that make for some predictability.

⁷³ Aquinas saw *individual chance variations* as translating up into a *statistical pattern* in at least one area: it is by chance that a child is conceived male or female, but in this way “Nature” designedly produces the same number of males and females overall (1a 92, 1 ad 1).

⁷⁴ 1a2ae 62; 63, 3.

⁷⁵ 2a2ae 25, 6 & 8–9; 26, 6–12.

⁷⁶ 1a2ae 63, 4.

⁷⁷ 2a2ae 186.

⁷⁸ The growth of Cistercian life reclaimed land for farming; monasteries were empowering centres of local economy.

6.6.1 *Heavenly Bodies*

Following contemporary astronomy, Aquinas saw the heavenly bodies as having an *immense* influence on what happens here below: no reproduction was possible without an input of “energy” from the Sun. Their shifting patterns allowed the ongoing variety of earthly events, but – owing to the *chance* coincidence of causes, and the varying dispositions of matter – without imposing determinism on them.⁷⁹ For human beings, (a) the pattern of the heavenly bodies at conception influenced someone’s “personal chemistry”,⁸⁰ (b) the heavenly bodies had an on-going effect on our bodily organs, and through them on our imagination and emotions, and (c) thereby, indirectly, the heavenly bodies could influence our intellect and, to a lesser extent, attract our will. However, *they could not force the human will* so as to *cause* human acts, because the will does not necessarily follow the inclinations of the lower appetites.⁸¹ Nevertheless, since many people do not make the effort to rise above their emotions, particular patterns of the heavenly bodies could, *predictably*, provoke “mass movements” through their influence on people’s imaginations and emotions. Hence astrologers could make correct predictions of events like wars.⁸²

This view, though incorrect, did imply a recognition of (i) “personal chemistry”, the propensities to virtue or vice in us at birth; (ii) our susceptibility to physical influences; and (iii) the possibility that mass movements might be explicable. Aquinas would not reject modern studies of our susceptibility to environmental, subconscious, pre-conscious and “psychological” influences.⁸³

6.6.2 *Inheritance*

Aquinas accepted that human beings can inherit bodily defects, and characteristics like athletic ability or mental agility, but not personally acquired skills.⁸⁴ Since he attributed *propensities* to particular virtues or vices to innate bodily dispositions, he presumably thought these were heritable, and would not be averse to studies of how character traits and behaviours that are not consciously chosen can be inherited. If patterns of behaviour do run in families, the relative proportions of patterns in a stable community might last for many generations. Modern science attributes to genetic factors a great influence over people’s “character”, in a sense replacing Aquinas’ account of the heavenly bodies’ real and natural causal power.

⁷⁹ 1a 115, 3 ad 2, & 6.

⁸⁰ For being born with certain propensities, see 1a2ae 50, 1; 63, 1; for the influence of the heavenly bodies, 1a 96, 3.

⁸¹ 1a 115, 4.

⁸² *Ib.* ad 3.

⁸³ For some time, advertisers have used subliminal techniques; the influence of environmental factors, including scents, on people’s behaviour in shops has been widely studied.

⁸⁴ 1a2ae 81, 1–2.

Aquinas' recognition of environmental and social factors, of personal choice, and of God's grace, reminds us not to rely *only* on genetics when investigating what contributes to people's characters.⁸⁵

6.6.3 *Climate*

Aquinas recognised climatic factors⁸⁶; he might well not be surprised by seasonal affective disorder, and might expect the climate of a region to have some predictable influence on patterns of behaviour there.

6.6.4 *Corrupt or Worthy Customs*

Current sociology would probably accord more influence than Aquinas did to cultural factors, to society's "mood" and presuppositions, and would analyse more deeply why many people follow these. But Aquinas did hold that, while in some sense the basic and obvious points of the Natural Law are built into everyone, we are still vulnerable to corrupt customs which can obscure even things that should be obvious.⁸⁷ On the other hand, worthy customs can be a force for good behaviour and hence character-formation.

6.6.5 *Coercive Law*

In a fallen world, it is not easy to work out the fine details of the Natural Law; it requires a long-term effort on the part of the wise.⁸⁸ Governments have the task of applying the Natural Law to local circumstances, of revising laws, and of granting dispensations justly.⁸⁹ This requires a special kind of *prudentia*.⁹⁰ It can be assumed that most people will follow law, either because they are good, or because it carries sanctions.⁹¹ Given the power of corrupt customs, Aquinas might not be surprised to find that in certain countries and cultures honesty in filling in tax returns is easier to count on than in others. He would presumably agree that, out of fear, many people

⁸⁵ Good science is non-reductive in this regard; it also recognises "Lamarckian" as well as "Mendelian" inheritance, but judiciously: Edith Heard and Robert A. Martienssen, "Transgenerational Epigenetic Inheritance: Myths and Mechanisms." *Cell* 157 (2014) 95–109.

⁸⁶ 1a 96, 3

⁸⁷ 1a2ae 94, 4 & 6.

⁸⁸ 1a2ae 100, 1.

⁸⁹ 1a2ae 91, 3; 95, 1–2; 97, 4.

⁹⁰ 2a2ae 47, 10–11; 50, 1.

⁹¹ 1a2ae 92; 96, 5.

follow laws that should be disobeyed⁹²; we might want to explore more deeply what instincts lead people to do what they are told even when it is questionable or unpopular.

If the proportion of people who will follow a particular kind of law can be estimated for a particular society, on the basis of observation, those who practise economics, as well as legislators and policy-makers, can rely on this estimate. It would be useful to study the power of law to influence not only behaviour, but also moral *perceptions*; this is probably greater than is often supposed.⁹³

6.6.6 *Persuasion and Protreptic*

Perceptions and desires can be moulded by persuasion and protreptic on the part of preachers, teachers, experts and officials. Aquinas practised protreptic⁹⁴; his work was embedded in a Scriptural and Liturgical system of moral formation in which Christ was the great Moral Exemplar.⁹⁵ While modern psychology has studied more deeply children's development in its social context, and phenomena such as peer-pressure, Aquinas would expect educational systems to have a predictable effect on people's behaviour.

One thing he might have difficulty with is brainwashing, since the behaviour of a brain-washed person "comes from within" and so seems to count as voluntary⁹⁶; yet it is imposed from without, against the person's deliberate choice.

6.7 Angels, Demons and Grace: Causes of Unpredictability?

Thus in the Mediaeval trio "the World, the Flesh and the Devil",⁹⁷ the world (social influences) and the flesh (the emotions, and genetic and environmental factors that influence them) largely make for predictability. For Aquinas, angels and demons also had a strong influence.⁹⁸ He saw them as affecting many physical occurrences

⁹² He has a nuanced view about whether bad laws should be obeyed: 1a2ae 96, 4.

⁹³ E.g. to what extent did the Abortion Law make people see abortion as a morally legitimate option? One such study is Yuval Feldman and Oren Perez, "How Law Changes the Environmental Mind: An Experimental Study of the Effect of Legal Norms on Moral Perceptions and Civic Enforcement." *Journal of Law and Society* 36 (2009) 501–535.

⁹⁴ Adam Eitel, "The Protrepic of Summa Theologiae I-II, qq. 1–5." *The Thomist* 81 (2017) 183–212.

⁹⁵ Mark Jordan, *Teaching Bodies*. Chapters 1–4.

⁹⁶ 1a2ae 6, 4.

⁹⁷ 1a 114, 1 ad 3.

⁹⁸ 1a 111–114. Angels' existence seemed obvious, since the heavenly bodies kept revolving because angels pushed them: experience suggested that when you stop pushing, things grind to a halt. The fourteenth century saw develop the concept of a temporary impulse, later replaced by the concept of inertia: James A. Weisheipl, *Nature and Motion in the Middle Ages*. Washington DC: Catholic University of America Press, 1985. 31–73.

on earth; they could move humours in the body and influence imaginations and emotions, and so could present us persuasively with ideas. Angels could not, directly, put abstract concepts into our intellects, but they could *strengthen* them.⁹⁹ Our dependence on the imagination both for forming and for employing concepts meant our mental susceptibility to angelic influence was marked; however, as with the heavenly bodies, the human will was open to God alone – no angel could “get inside” it.¹⁰⁰ Rebel angels – demons – could also influence us, towards evil, but only under God’s providential permission.¹⁰¹ They could not influence the human will directly, any more than could the good angels or the heavenly bodies.

A healthy aspect of Aquinas’ account of moral evil is his resistance to the idea that all sins are due to the devil; often enough they are due to “the world” or “the flesh”, or simply to the mysterious lability of a created will.¹⁰² Many modern Christians accept the influence of angels and demons, but Aquinas remains a valuable conversation partner for people who do not: he was not reluctant to recognise many forms of vulnerability to non-voluntary, pre-conscious interior and exterior influences, and would not rule out in principle *psychological* accounts of mental and moral problems. If we do accept the existence of angels and demons, we should (with Aquinas) recognise that they can only work on and with the human psychological material accessible to their influence; we would expect them as a rule to do so in subtle ways.

The influence of angels and demons might result in *unpredictability*: one could imagine that (under God’s Providence!) demons might incite a people to evil in a way that could not have been foreseen – or (sent by God) good angels might excite a good social atmosphere. Recent centuries have seen more than their fair share of mass hysteria, genocidal hatred, revolutionary fervour that leads to a reign of terror, and similar events that might be labelled “demonic” even by people who do not believe in demons. We have also seen strikingly beneficent social movements; the fall of apartheid without a blood-bath springs to mind. If we believe in angels, we might surmise they have a role in such events, under divine providence, and without prejudice to the priority of God’s grace which alone can be at work *within* the human will.¹⁰³

Since angels and demons only work on and with the human psychological material accessible to their influence, their role might be, if not predictable, at least interpretable in retrospect. Certain studies have suggested that political history moves in stages which include times of marked vulnerability.¹⁰⁴ If so, and someone like Hitler comes on the scene at such a time, disaster can happen “out of the blue”;

⁹⁹ 1a 111, 1 & 3–4.

¹⁰⁰ 1a 111, 2.

¹⁰¹ 1a 114, 1.

¹⁰² 1a 63, 1; 114, 3. Moral evil, as intrinsically “messy”, has neither a fully satisfying rational explanation, nor a unitary cause: 1a2ae 73, 1.

¹⁰³ 1a 105, 4; 1a2ae 9, 6.

¹⁰⁴ The studies were helpfully and critically reviewed by Frances Hagopian, “Political Development, Revisited.” *Comparative Political Studies* 33 (2000) 880–911.

if someone with a charisma for good is providentially granted (e.g. Martin Luther King, Nelson Mandela), un-hoped-for improvements take place. In such periods the influence of angels and demons might be especially marked as they work on and with natural liabilities, including our susceptibility to “public opinion” as enunciated (or manipulated) by the media, by demagogues – or by orators of truth and goodness.

Finally, we must recall that God’s grace, “enlarging our hearts”, is at work. Typically, it does not cut across the natural, but brings it to perfection in a way that promotes moral beauty.¹⁰⁵ Hence it does not annul the natural elements of predictability and unpredictability we have examined. The striking things some of God’s friends are inspired to do are not its most usual activities. Grace may be doing more than we realise – we dare not tame or corral it – to cause much more to go right in human behaviour than would go right if such a complex creature were left to its natural resources alone. On the other hand, it would not be true to Aquinas to deny the value of those resources which, being natural, we retain in a fallen world.

6.8 Conclusion

Aquinas’ *pre-modern* view of the complexity of human psychology, of the decision-making process, and of the situations calling for decisions, has invited us to reflect on the nature of freedom and on the forms and causes of predictability. Freedom involves self-possession, not randomness of decisions, hence the statistics of randomness are not the correct way to approach human predictability. Final causality operates, in a distinctively human way, so that we can give a rational account of any free act of ours, even if in doing so we come to realise flaws in our thinking. At a first level of analysis, we might expect our inner dynamics towards truth and goodness to lead to predictability: people will tend to do the sensible thing. This *is* the case to some extent: most people make adequate provision for themselves and their families through work and household management; they generally follow law and custom. In an ideal world, they would do so through good citizenship; in a fallen world it is either through good citizenship or through fear of sanctions. People are also liable to be misled (in thinking and behaving) through corrupt laws and customs. However, *natural* human variety means that personal choices would make an unfallen society vibrant with variety; in a fallen world, grace leads God’s close friends to buck the social trend in at least some ways.

Much of our psyche is *responsive* rather than pro-active; much of what goes on in it is pre-conscious. Free decisions involve interaction between thinking and wanting, and thinking is open-ended, hence individual decisions are not determined, *because* of being rational. A mysterious element of spontaneity is present regarding which components of a situation we attend to, or fail to consider, and how we perceive each one; this goes with our being embedded in a biological, social, historical

¹⁰⁵ 1a2ae 110, 2; 2a2ae 23, 2.

and geographical context, within which virtuous people take control of their wants and behaviour and, though good citizens, become in some ways *less* predictable than the majority. Most people settle for moral mediocrity, and fail fully to integrate their emotions, which leaves them susceptible to a wide range of interior and exterior influences, most of which make their behaviour predictable in a dull way (a) as *rational enough*, and (b) as influenced by factors not personally chosen, nor adequately recognised, evaluated or corrected for:

- “personal chemistry” due to inheritance and upbringing;
- environmental factors;
- law and custom;
- persuasion (now including psychologically astute advertising).

Individual variations in personal chemistry *may* translate up into statistical regularities.

However, our susceptibility to external influences leaves us vulnerable to factors that are to some extent unpredictable such as mass-movements, whether these are due to psycho-social patterns and forces, or to demons, or to the latter working with the former. Under God’s providence, we can also be swayed by forces for the good, whether these are due to charismatic leaders, or to angels, or to the latter working with the former.

In short:

- Human behaviour is, ideally, rationally explicable within a personal variety that conduces to the common good;
- In the majority of cases it is rational enough, so that well-made law and public policy are fairly successful;
- The majority of people will behave in statistically predictable ways owing to their limited success in taking free, personal, rational control of the many less personal factors to which we are susceptible;
- And people at large are liable to be swayed by unpredictable mass-movements for good or ill.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 7

Agency, Time and Morality: An Argument from Social and Economic Anthropology



Paul Clough

Abstract This paper reflects on the ways in which the concepts of agency, time, and morality can be linked together. It does so by considering how the kind of thinking which goes on in social anthropology and economic anthropology can contribute to our understanding of ‘agency’. It then relates the discussion of agency to the concepts of time and morality. The mode of reasoning in social anthropology can be described as cyclical: social theory generates general questions that can be asked of a human society in particular times and places. Those questions stimulate empirical investigation, which results in a descriptive analysis of human social reality – known in social anthropology as ethnography. In turn, ethnography often challenges existing paradigms in social theory, because it throws up puzzles about human social interaction that can only be resolved by revising existing theory. Thus, there is a cycle in which theory stimulates ethnography, which throws up puzzles that can only be resolved by revising or changing theory. The same cycle can be discerned in economic anthropology, that sub-branch of social anthropology which understands ‘economy’ as a holistic social process interweaving production, exchange, and consumption (Narotzky 1997). The ethnography of micro-economic reality has provided rich evidence about human decision-making that can be used to address general theoretical questions. From its origin in studies of gift exchange (eg., Malinowski 1922) to contemporary studies of money and debt (eg., Graeber 2011), economic anthropology, like its parent discipline, has been preoccupied with theoretical questions like the tension between ‘agency’ and ‘structure’, ‘individual’ and ‘society’.

P. Clough (deceased)
University of Malta, Msida, Malta
e-mail: paul.clough@um.edu.mt

© The Author(s) 2020
P. Róna, L. Zsolnai (eds.), *Agency and Causal Explanation in Economics*,
Virtues and Economics 5, https://doi.org/10.1007/978-3-030-26114-6_7

115

7.1 Agency in Social Anthropology

In a lucid short essay on the history of agency in anthropological theory, Nigel Rapport shows how ideas of agency have emerged from debates over the connection between individuals and social structure. These debates explored the nature of individual consciousness and its freedom from external determination. He defines agency as the ‘capability... to be the source and originator of acts’ (Rapport and Overing 2007: 3). Discussions go back at least as far as the classic differences between Durkheim and Weber. Durkheim’s account of human action focused on the norms whereby choices are guided, and those structures in which norms are formed. Leaving little or no room for self-creation, Durkheim’s theory contrasted strongly with that of Weber, in which the distinctively human aspect of action consists of the various features of ‘rationality’ – consciousness, reflection, intention, purpose, meaning. The subsequent literature has tried to overcome the opposition between Durkheim and Weber, by exploring the limits on the individual’s capacity to act independently of structural constraints. In Rapport’s view, all attempts have ended up privileging either structure or agency (Rapport and Overing 2007: 4).

The most influential recent attempt in anthropology to resolve the opposition between structure and agency has been Pierre Bourdieu’s theory of practice (Bourdieu 1977, 1980). For Bourdieu, most consciousness is habitual. The body acts as a mnemonic device for communicating pre-existing collective codes of gender, normalcy, propriety, and domination. Thus, children encode the collective culture in their bodies through early training. Perhaps most strikingly, Bourdieu argues that competency in the different kinds of social interaction operates by individuals *not* thinking about how they act. Hence his key concept of habitus – ‘the system of durable dispositions, predisposed to act as ... principles [that] generate ... practices and representations ... *without presupposing a conscious aiming at ends*’ (Bourdieu [1980] 1990: 53, emphasis added). In his various precise ethnographic analyses, the room for individual manoeuvre is strictly limited by prevailing dispositions and practices. Despite Bourdieu’s continuous claims to have produced a unique synthesis from the analytical opposites of structure and agency, none is detectable (Clough 2016: 202–204).

There have been notable efforts in social anthropology to outline theories that do stress the importance of agency. For example, Gregory Bateson emphasized *individual creativity* in generating social change (Bateson 1973). Kenelm Burridge’s outline of the process whereby idiosyncratic, ‘heroic’ individuals persuade others, or get copied by others, demonstrates how ‘eccentrics’ inaugurate patterns of action that eventually become the new norm (Burridge 1979). However, in this essay I will concentrate on the arguments of Terry Evens in his 2008 volume, *Anthropology as Ethics*, because he ranges so widely across philosophy and social theory. Evens links agency to intention and to moral responsibility. He stresses as well the need to allow room for the unprecedented, for the act of creative imagination (Evens 2008:

116). For Evens, a full description of the attributes of human agency entails the idea of ‘self-conduction’, which in turn implies the activity of genuine choice. Furthermore, reflection by the self on the self contains the awareness of a tension between the self and its surroundings, and equally, between self-interest and the demands of others. Awareness of this tension between the self and others includes belief in a personal responsibility to resolve that tension. Thus, the awareness of responsibility generates awareness of the reality of self-change (Evens 2008: 128–130). I contrast Bourdieu’s weak idea of agency as limited, structured change in a social system, with Evens’ strong concept of agency as self-conduction in responsibility to others.

7.2 The Argument

Through my ethnography of one farmer-trader in West Africa, I will argue that this man’s short-term economic decision-making and long-term planning implied the operation of a particular kind of agency. His decision-making involved imagination - the ability to visualize a future different from his present. The practical application of imagination entailed the qualities of intentionality and purpose, and the ability to plan for the imagined future. Furthermore, in the very framing of the goals used in his planning, he had to make choices in the ethical dilemma between expenditure on self and expenditure on family, kin, or clients. Thus, my argument favours ‘strong’ agency as self-conduction (Evens 2008) over ‘weak’ agency as the ability to slightly alter existing structures (Bourdieu 1977, 1990). Moreover, agency implicates time, because the very meaning of agency shifts as the analytical focus changes from the ‘short-term’ to the ‘long-term’. If, in a ‘short term’, the question often asked with regard to agency is whether individuals can act differently than they do at particular points in time and space, the more appropriate question in a ‘long term’ is whether individuals can strive over time to act differently in the future than they do in the present – can install new goals, or even change some of their habits, over a period that is open to the future. Equally, reference to the long-term dimension of agency implicates morality. Here I draw on Levinas’s philosophy of ethics as the inexhaustible or unending responsibility of the self to the other (in Levinas’s language, the ‘infinite’ responsibility of Self to Other). Levinas exposes the tension between self-interest and other-regard. Awareness of tension between the demands of the self and the demands of the other includes belief in personal responsibility for the resolution of that tension. Thus, the awareness of responsibility – ethics – generates awareness of the reality of self-change *over a long term* (Evens 2008: 128–130). In summary, *where* a person persists with a particular intention, *then* consistency of purpose can lead to self-change over the trajectory of a life.

7.3 An Ethnography of Economic Action

These various dimensions of agency can be seen in my ethnography of an individual farmer-trader among the Muslim Hausa of northern Nigeria, West Africa, to which I now turn. My research on this trader, whom I will call Abdulkadiri, was part of broader fieldwork on the relations of production and exchange in one village of Hausaland in northern Nigeria; and equally, on the trading networks for agricultural products that linked my village to other economic regions of Nigeria. Both the economic and cultural contexts of research are crucial, because they enable me to posit connections between the nature of agency and the real nature of global markets at the end of this essay. With regard to the economic context, the region where I did my research has long had a highly commercialized rural economy. By the time that I began fieldwork in Marmara village in 1976, the process of commodification had extended from products to land and finally to labour power. In Marmara, while half of all farmland had been inherited by the users, half had been acquired primarily by purchase or by borrowing the land in return for a money loan of its market value. Moreover, a half of farmland was farmed by household heads who combined family labour with hired labour. The top fifth of landowners combined farming with trading in various lines, using the profits from trade to invest in farmland and hired labour. Hence I use the term 'farmer-traders' for the social stratum of which Abdulkadiri formed a part. With regard to the cultural context, Hausa rural society was piously Muslim, and polygyny was the norm: the majority of family heads had from two to four wives. The marriage system accords wives an important degree of autonomy. Divorce is easy. Most men and women have been through several marriages.

In the complex ethnic composition of Hausa society, Abdulkadiri's ethnic background was actually Fulani, his original language – an ethnicity particular to pastoralists throughout West Africa – though he had migrated to the Hausa village of Marmara as a result of acute land scarcity in his home region. In his new home, he combined farming with trading in livestock, due to his close relationships with Fulani pastoral kinsmen who owned herds of cattle. This ethnography analyses 1 year of accumulation in Adulkadiri's life.

At the beginning of the year, in January 1978, he faced the need to make crucial decisions. He was the leader of a large family farming enterprise that included his two sons and two of his younger brothers. He combined trading in grains with trade in livestock. And he had 1600 Naira of personal savings – no mean sum in the currency of the time – which he looked upon as his working capital. But he had little land and only one wife. By the end of that year, in December 1978, Kadiri had invested all of his savings in farmland, which he had increased from 6 to over 20 acres. By combining his family work force with large amounts of hired labour, he used the additional farmland to more than quadruple his output of grains. He used part of his savings to contract successful marriages for his first son and for his unmarried brother. And he managed to marry first one, and after divorce, another woman who thus became his second wife. Over the year, Abulkadiri had successfully engaged in the accumulation of capital – investing profits from previous trade

and production in the increased control of land and hired labour. But moreover, he had accomplished the prestigious cultural transition from a monogamous to a polygynous family head. It was – by any local measure – an astonishing transformation. How had he pulled it off?

In the course of my fieldwork, I very gradually learnt more about Abulkadiri from his occasional statements made in the midst of work. What follows is thus my analytical reconstruction based partly on his observed actions, and partly on his discourse. Years before, when he had first emigrated to Marmara, his intention had always been to eventually return to his home region to ‘embrace’ his parents. But they had died, leaving him without a compelling reason to return. Moreover, though his mother’s family were well-connected to the hereditary rulers, his father had left very little farmland and few cattle. Thus, he weighed the absence of reasons to return against the opportunities in Marmara. By the beginning of 1978, he had come to the decision to remain in the economic region of southern Katsina, to abandon trading in livestock, and to focus his energies on agricultural production and trading. The timing was pushed because his eldest son and younger brother had reached maturity and needed to be married. In Hausa culture, the institution of family farming (called *gandu*) is based on a reciprocal exchange: sons and other working kinsmen receive from the household head the bridewealth needed for marriage, and food for themselves, their wives, and offspring, in return for their farm labour during the ‘rainy’ season and help with the household head’s off-farm occupation during the ‘dry’ season. Hence, Kadiri’s decision to abandon his residual traces of pastoral Fulani identity entailed large family responsibilities in his new focus, Hausa farming. I could see that this change in economic identity was accompanied by a certain change in his cultural identity. At the start of the year he could sometimes be seen in the customary Fulani dress, whereas by the year’s end he was only wearing the Hausa robe and cap. This is not to deny that he retained close links with his Fulani kinsmen having herds of cattle. They kept for him his few remaining cattle, and being wealthy in terms of livestock, were a source of loans in emergencies (Clough 2014:274–326, 2016: 201–207).

Kadiri could not expand his landholding without finding the funds to hire the labour needed to farm it. In his trading he bought grains from local farmers for sale at a distant weekly marketplace where prices fluctuated at a much higher level than in Marmara. He continually emphasized to me that the weekly profits derived from the relatively large scale of his trading were necessary to pay for agricultural labour and other inputs. At the same time, since he had sunk his own savings into farmland and family marriages, he became increasingly dependent during the year on loans to finance his farming. Essentially, he was immersed in social relations of credit that were as important as the institution of family farming. These relations were of two kinds – patronage and trading friendship. His patron was a much wealthier trader in the same line, crop trading, who loaned him money for trade in return for Kadiri’s long-term services in finding the patron new areas of supply, and selling his patron’s traded grain in distant markets. His ‘trading friends’ (in Hausa, *abokanan ciniki*) were men in other lines of trade, with incomes much closer to that of Kadiri, whose marketing had different periodicities than Kadiri’s crop trading. They were thus in a

position to loan him money at times when they had realized their own commercial proceeds.

All trading loans were without interest in rural Hausa society. This had less to do with the Islamic ban on usury than it had to do with the fact that all traders were involved in such tangled and variable claims for assistance from diverse kinsmen and friends, that they could not guarantee the regular money growth necessary for stable rates of interest to settle. Here, too, as in farming, economy depended on an ethic of reciprocity. As a trader sought to obtain loans, he found himself under pressure to give them. He knew, in borrowing, that part of the credit would be loaned out again. In the local circle of traders, potential borrowers were not only being assessed for their honesty, but equally, for their 'character' (in Hausa, *hali*). A person with 'character' had many men from whom he attracted credit. 'Character' in Hausa incorporates the ideas of redistribution and reciprocal justice. Since the success of most traders depended on the willingness of others to loan them commercial funds, reciprocal justice required that they lend to others still struggling to succeed.

Given the demands of kinsfolk and poorer traders on his funds, it may seem surprising that Abdulkadiri was able to achieve any capital accumulation at all. His singular feat was to borrow more from the circle of patrons and trading friends than he loaned out to others. Secondly, he was able during that year to convert many short-term loans into medium-term credit. He was even able to transform the medium-term loan from his patron into a long-term loan. In these ways he generated working capital for farming in the short term.

This summary gives a more seamless sense of the logic of the rural economy that is true of the human reality. Farmer-traders are managers of multi-purpose family enterprises. In consequence, the boundaries between domestic and productive expenditure are blurred. So I turn briefly to his discourse. Over the year, he was continually praising to me the beauty of women at the distant weekly marketplace where he sold his produce. It is from there that he found the two women whom he sought for marriage. In order to ensure that the second lady remained in his household, his mind fastened on a dramatic scheme – to completely rebuild his compound in order to better house, not only his two wives, but also those of his son and younger brother. The building took time, the result impressed the whole village. To finance the construction, he converted one of his largest trading loans into building money, and as far as I could tell, was only able to eventually repay that loan by tapping the funds of distant pastoral kinsmen. We have here a whole man, endowed with an aesthetic and with desires, who sought emotional security from the company of wives and friends. Reviewing the entire year, he actually spent more on household marriages and building than his considerable expenditure on farms and farming. It was larger, too, than his weekly investment in trading. He offered a revealing summary of his recent past: 'For four years, I have not postponed marriages in my house – for myself, for my son, and for my brother. *If not for this, I would have gone east* [ie., to Mecca, on the Muslim pilgrimage]!' Note that he did not say, 'I would have had more money to buy farms or hire labourers.' Expenditure on farmland and hired labour were embedded in the expansion of his household and the reciprocal obligations of patronage and trading friendship in a Muslim circle of traders.

7.4 Theoretical Implications of Ethnography for a Theory of Agency

Can we generalize from the specificity of micro-ethnography to philosophical or economic theory? Does a micro-study have implications for our understanding of agency? More particularly, how can meaning, purpose, intention, or self-reflection in the consciousness of the main character here - and, moreover, the changes in these - be seen to interconnect with the sequence of his actions over the period studied? I begin by asking whether a structural analysis of 'practices' in the manner of Pierre Bourdieu - 'skilled performances' based on indigenous practical knowledge in particular socio-economic niches - can explain Abdulkadiri's decision-making. It must be admitted that a structural analysis in the style of Bourdieu, based on the material necessity of certain actions in a particular economic niche, does indeed explain certain contours of his action over time. Having come from a pastoral family with limited livestock, Abdulkadiri almost inevitably gravitated from cattle herding to cattle trading. At the same time, he gradually realized that an expansion in his material income and social status required a certain oblique movement from being a trader focusing on cattle to being a trader focused on farming and agricultural trade. Analyzed thus, his change from a Fulani pastoral to a Hausa farming identity was certainly predictable.

Nevertheless, what a structural analysis cannot do is explain three elements of his consciousness. First, *the unpredictability of his search for beauty*. Being a man of middle age searching for women of similar age, material necessity cannot explain his search for a second wife. Moreover, a structural analysis does not construct a satisfying mental picture of his contrary urge to extravagance, as when he threw the loan of a major trading friend into house building rather than farming or trade. Secondly, *his capacity to dream of a future different from his past or present*. Put differently, while Bourdieu's structuralism can explain *how* a conscious being responds to the material pressures facing him, it cannot explain the existence of a dreaming, imaginative being *in the first place*. Thirdly, *his sense of a disjunction between dream and reality, between present planning and its future realization, between a relatively certain today and a merely possible tomorrow*. Thus, Abdulkadiri was aware of being caught between and among:

The claims from his wives, junior relatives and other close kin.

The claims from the Muslim trading community to which he belonged.

The pressures of personal desire in a polygynous context.

Such a being, being 'between', is somehow aware that he is distinct from the present spheres he seeks to manage, is disjunct from the future that he tries to turn into a reality. Human action is the putting into physical effect of wishes and dreams. Human action is a process of self-change that sometimes, as in this case, involves self-expansion.

In a sense, Kadiri was turning himself into a cause that initiates effects. Thus, he became by his actions a part of the 'structure' that constrains, or puts pressure on,

other people, and which they in turn take into account when they initiate actions. In other words, the structure is mobile. It is continuously changing because it is full of initiatory individual actions. In contrast to Bourdieu, who claimed to have transcended the analytical opposition between structure and agency, my reflections on this one farmer-trader suggest, not some synthesis, but rather a re-worded solution in favour of agency. This is not weak agency in Bourdieu's sense of 'social change', whereby in any society one action induces an array of consequent actions that can be specified predictively given adequate knowledge. Rather, it is strong agency, what Terry Evens calls 'self-conduction' (Evens 2008: 114). Self-conduction by numerous inter-actors generates a continuously changing structure - either slowly or rapidly. 'Structure' is the assemblage of interacting actions, where by 'action' is meant the human translation of dreams and wishes into a previously non-existent reality. Since structural fixity is a chimera, it makes no sense to analytically counterpose structure and agency in the first place. 'Structure' as the assemblage of interacting actions is shot through with agency and hence mobile.

Lastly, what is highlighted by the attention given to 'planning' in the actions of Abdulkadiri, is the analytical distinction between short-term and long-term in our understanding of agency. In the short-term, the *internal* compulsions to act in one way rather than another - prior convictions, fixed views, strong feelings, vivid memories - combine with the *external* constraints of political power and economic risk to limit the subject's possibility of choice. Therefore, if 'time' is glossed as a sequence of discrete 'points', then each point becomes a short-term in which internal compulsions and external constraints shape the consciousness and actions of the subject. Where, however, time is construed as a duration of indeterminate length, open to the future, then our perspective changes. We start to invoke different kinds of verb in our reflection - 'trying', 'attempting', 'striving', 'struggling', 'searching for' (in the manner of Gilbert Ryle experimenting with different verbs for 'thought' in his *Concept of Mind*). Perhaps, then, what the complex and variable language of 'choosing' or 'deciding' does, is to gloss an awareness of time *either* as a sequence of momentary points *or* as a temporal stretch of indefinite duration, open to the future. If, and when, a particular intention is persistent over time, then consistency of purpose leads to gradual self-change over a long period of time. The trajectory of a person's life cannot be described without reference to his or her intentions.

Although the temporal dimension of choice is specified in variable ways, the sense of self-conduction is ineluctable. Because I am aware of being a being suspended among present pressures and future possibilities, I am conscious that the self translates possible futures into an embodied present. Pervasive in this sense of self-suspension is the tension between self-desire and other-regard (Evens 2008: 128, 296-297). In a sentence: it is through my sense of the difference between present and future, and through my consciousness of responsibility to the other, that I become aware of the meaning of agency.

7.5 Implications for the Study of Global Markets

These arguments for strong agency as a focus on individuals *striving over time* to change their condition or even their character, allow us to more clearly see the nature of market formation and change in different parts of the world. Markets are like any other ‘structure’ in my argument: they are assemblages and sub-assemblages of initiatory human interaction. As such, markets are shot through with agency, and hence mobile, constantly changing.

First, whereas a short-term perspective on markets assumes that most individuals are seeking to maximize current net satisfaction across the range of their prioritized desires, a long-term perspective opens up the possibility of a change in the very nature of their satisfactions and in individuals’ hierarchies of desire. Such changes can alter the entire system of relative prices for different products and services. In the long-term, for example, changes in the consciousness of gender and of gender difference increasingly inflect consumer choice. New desires can lead to the emergence of new markets.

Secondly, focus on the intrinsically moral nature of agency enables us to shift attention from individual to collectively shared aspects of morality. I refer to the way in which social movements mobilize the collective force of their members to change public understandings of human obligation. Examples include the voluntary embargo on purchasing products from racist regimes, or the efforts by environmental movements to alter the consumer preferences of their members worldwide.

Thirdly, my Hausa micro-study points to a very particular structure in global markets – their cultural structure. The very different ways in which people in different parts of the world stratify their spending between social or public goods and individual private goods, demonstrate that markets are far more than the passive responses of anonymous, atomized individuals. Choices concerning what to produce or buy are made after public discussion and debate. In non-Western parts of the world, these choices are often made after community debates which can have recourse to religious beliefs, and – certainly in the case of Muslim societies – deliberation on the sacred texts of the religious group.

Fourthly, the cultural consequence over time of the intersection of different local and regional markets is *not* a universally shared concept of value. For example, despite a century of commercialization of the forces of production, all the evidence from my research area in West Africa reveals the continued priority given to household expansion through polygynous marriage, and to circles of trading friends through local systems of credit and debt, rather than to the capitalist accumulation of land and labour. As products and money circulate globally, their uses and meanings are altered in local contexts so that they express and sustain local systems of value (see also Parry and Bloch 1989).

Lastly, these diverse intersections between, on the one hand, local schemes of desire informed by political and religious movements, and, on the other, the global circulation of capital and search for profits, can only be given their full meaning if we see the entire assemblage of actions as shot through with agency in the strong

sense. While, at any moment in time, individuals are constrained internally by fixed tastes, views and values and externally by pressures on their income, a long and open duration of time gives more scope for persistence of purpose to become a force that gradually works its way into individual decision-making.

References

- Bateson, Gregory. 1973. *Steps to an Ecology of Mind*. London: Paladin.
- Bourdieu, Pierre. 1977. *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.
- . 1990 [1980]. *The Logic of Practice*. Trans. Richard Nice. Cambridge: Polity Press.
- Burridge, Kenelm. 1979. *Someone, No One: An Essay on Individuality*. Princeton: Princeton University Press.
- Clough, Paul. 2014. *Morality and Economic Growth in Rural West Africa: Indigenous Accumulation in Hausaland*. New York/Oxford: Berghahn Books.
- . 2016. Tension, Reflection, and Agency in the Life of a Hausa Grain Trader. In *Reflecting on Reflexivity: The Human Condition as an Ontological Surprise*, ed. T. Evens, D. Handelman, and C. Roberts. New York/Oxford: Berghahn Books.
- Evens, T.M.S. 2008. *Anthropology as Ethics: Nondualism and the Conduct of Sacrifice*. New York/Oxford: Berghahn Books.
- Graeber, David. 2011. *Debt: The First 5,000 Years*. Booklyn/London: Melville House.
- Malinowski, Bronislaw. 1922. *Argonauts of the Western Pacific*. Prospect Heights: Waveland Press.
- Narotzky, Susana. 1997. *New Directions in Economic Anthropology*. London/Chicago: Pluto Press.
- Parry, J., and M. Bloch. 1989. *Money and the Morality of Exchange*. Cambridge: Cambridge University Press.
- Rapport, Nigel, and Joanna Overing. 2007. *Social and Cultural Anthropology: The Key Concepts*. London/New York: Routledge.
- Ryle, G. 1949. *The Concept of Mind*. London: Hutchinson.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 8

The Switch from Agency to Causation in Marx



Scott Meikle

Abstract This paper traces the shift from Marx’s labour theory of value to his later forms of value theory and attributes the change to Marx’s shift from agency to causation as the dominant dependent relationship in economic phenomena.

8.1 First Edition Versus Second Edition

From the time of Engels and Kautsky, everyone believed that Marx held only Ricardo’s view that exchange value was labour time, known as the Labour Theory of Value (LTV). This began to change in 1972 when Pilling drew attention to the ‘forms of value’ in chapter one of Capital.¹ Everyone had known about these, but nobody ever saw anything in them, and Ronald Meek summed things up in his book on the LTV: ‘There is no need for us to follow Marx’s rather complex analysis of the “elementary”, “expanded” and “money” forms of value in any detail’.² But Pilling had inspired others to look again at chapter one, and a non-Ricardian side to Marx’s thought started to emerge. Since then there has been disagreement about what this is, but agreement that the second edition of Capital is where it must be found.

But the second edition cannot be trumps. Authors issue second editions for a reason, and if Marx issued his because he had a change of mind we need to know; second thoughts are not always best. Checking Marx’s earlier work we find he developed his own theory of exchange value through the Contribution of 1859, the notes on Bailey, and the first edition of 1867. But he suppressed it and gave the job to Ricardo’s theory instead, and the second edition was where he made the switch.

¹G. Pilling, ‘The Law of Value in Ricardo and Marx’, Economy and Society, 1972, vol. 1, no. 3, 281–307.

²Ronald L. Meek, Studies in the Labour Theory of Value, (London, 1973), 173.

S. Meikle (✉)
Research Fellow, Glasgow University, Glasgow, UK
e-mail: scott.meikle@glasgow.ac.uk

Let us see what that original theory was, and let's call it the 'theory of commodities and money' or TCM for short.³

8.2 The Theory of Commodities and Money (TCM)

Marx asks two questions in the Contribution: what is the commodity, and what is money? He starts by noting that in the simple relationship of commodities, '20 yards linen = 1 coat', the coat represents the exchange value of the linen. That is Marx's first result: the relationship contains an act of representation. The coat is performing a social office, in representing the linen's exchange value, which is a social property with no natural presence in the world of things, until the coat provides it.

The linen has the same relation to other useful things too, so that:

20 yards linen = 1 coat, 1 quarter tea, 2 lbs. iron, 1 ell yarn, etc.

Putting these relationships in this horizontal series gives us a sense of the range of the linen's exchange value.

We also see that a thing becomes a commodity when it enters a network of exchangeability relationships with other things. That is Marx's second result: being a commodity is being in a set of social relationships. And his third result is that exchange value is a social property that a thing acquires when it enters such a set of social relationships, and loses when it leaves it.

In that horizontal series, each thing represents the exchange value of the first. But what if only one thing does the representing? Then we would have a vertical arrangement like this:

20 yards linen = 1 coat
 = 1 quarter tea
 = 2 lbs iron, etc

Here only one thing does the representing, and the others are excluded from that job. This is Marx's final result, because this is money.

Marx gives names to the social roles played in '20 yards linen = 1 coat'. The representing role the coat plays he calls 'the equivalent', and the role the linen plays he calls 'the relative'. The role money plays in doing all the representing he calls the 'universal equivalent'. The entire explanation is social: social relationships entered into; social roles performed; functions fulfilled. These things are established in our behaviour, and rest on agreements. Marx makes the point when he says, as he often does, that such and such a function is performed 'with social validity', like gold doing all the representing for instance.

When Marx first asked what money was, he recalled that 'Gladstone, speaking in a Parliamentary debate on Sir Robert Peel's Bank Act of 1844 and 1845, observed

³Book I Part 1 of Capital is entitled 'Commodities and Money' so the title 'TCM' fits.

that even love had not turned more men into fools than meditation upon the nature of money'.⁴ He adds that you can't get anywhere until you realize that 'the commodity is the origin of money'. And at the heart of the commodity lie those social agreements, and the main one is the coat's act of representation, the equivalent. Samuel Bailey made a lot of the commodity relationship, or 'exchange relation' as he called it, '20 yards linen = 1 coat', yet he never managed to say anything interesting about it because, says Marx, he never asked the crucial question: 'how is it possible to express the value in exchange of A in the value in use of B?'⁵

Formulating the theory obviously depends on the logic of predicates, particularly quality, quantity, and relation, and on the nature-convention distinction. Marx knew the Greeks; had studied Aristotle's *Categories*; and had a Rhineland education and cultural background. But this sort of thought was vexed in Britain, where logic and metaphysics had been purged as Papist in the seventeenth century, and decried by native philosophers from Hobbes to Hume as 'School Metaphysics'. This cultural loss handicapped British political economists trying to understand the meaning of '20 yards linen = 1 coat', depriving them of requisite tools.

Marx's kind of social thinking was very unusual for the mid nineteenth century. British intellectual fashion didn't turn towards social things in this way until Wittgenstein in the mid twentieth century, and before then even thought and language were thought of non-socially in individualistic Humean ways. This may be partly why the TCM was so little appreciated. Fashion then looked to the physical sciences. And that was partly why economists were impressed by Ricardo's theory identifying exchange value with a natural quantitative property, viz. time taken making things. It may also be partly why Marx lost confidence in the TCM, or in its favourable reception.

When the coat represents the linen's exchange value, Marx says it 'counts as' the linen's exchange value. But this is a different way of construing '20 yards linen = 1 coat' from Ricardo's way, which says it is an arithmetical equation equating two quantities. This requires finding a property that each side of the equation has. Bailey's objection to Ricardo's theory was precisely that it construes the relation as an equation, and then looks for a property for each side to have equal amounts of, and there isn't one, so they make one up, a hidden one called 'value', to fill the vacancy created by supposing the relation to be an equation in the first place.

Marx likens '20 yards linen = 1 coat' to an encounter, and he asks what the parties see in each other when they meet? He appeals to the image of kingship, so he obviously thinks some kind of social status lies behind it: 'This man here is only King, because other men behave towards him like subjects. They believe, however, that they are subjects because he is King' (Dragstedt, 24).⁶

⁴ Karl Marx, *A Contribution to the Critique of Political Economy*, in Marx and Engels *Collected Works*, vol. 29, 303.

⁵ K. Marx, *Theories of Surplus Value*, vol. 3, (Progress Publishers, Moscow, 1971), 149.

⁶ Page references to the first edition of *Capital* are to the translation (the first into English) by Albert Dragstedt, in *Value: Studies by Marx*, New Park, London 1976, hereafter referred to as 'Dragstedt'. Dragstedt translates the original version of chapter one, and the Appendix entitled 'The Form of Value', both of which Marx suppressed in the second edition.

But how does status come into it? Marx suggests that commodity status exists, as other statuses like kingship do, in virtue of agreements conferring what he calls 'social validity'. This applies across social reality as Searle has argued.⁷ Social statuses exist by people agreeing to 'count' something as being something else, and they are often conferred in special procedures, or ceremonies. A church gets the status of a cathedral only if a bishop is appointed to it under proper authority, and similarly for someone 'counting as' a barrister, or some land 'counting as' someone's private property. Searle's formula for it is that over a certain domain C, X things count as Y things, or 'X counts as Y in C'. For example, 'these bits of authorised paper count as legal tender in this country'. But commodity status is unusual, because things acquire it without any procedure. They simply slip into the network of exchangeability relations and acquire the property of exchange value and the capacity to play the roles of equivalent and relative.

Marx continued with this theory into the first edition. The planned chapter six extended the critique of economics by introducing the concept of valorisation to expose how economics builds market features into its supposedly universal account of production. But the chapter was withdrawn. It appears today as the essay 'Results of the Immediate Process of Production' where, in the last few pages, Marx introduces a set of ideas that go dramatically against all his previous work (Penguin, 1044ff).⁸ He moves away from the theory of the commodity, and towards a theory of capital. He drops 'exchange value' and adopts the term 'value'. The surplus, which had been called 'surplus labour', is now called 'surplus value'. He now speaks of labour 'creating value', breaching the distinction between the labour and valorisation processes. 'Valorisation' goes, and 'accumulation' and 'self expansion' arrive. The two 'aspects' of labour appear. Capital becomes an agency with a tendency to expand. The mover is capital, not people, whom capital now makes its instruments. The critique of economics seems to shift towards a radical version of economics.

8.3 The Combined Theory: TCM & LTV

Marx finds himself holding both the TCM and Ricardo's theory at the same time. This means that he has to hold both constructions of '20 yards linen = 1 coat': the construction that says it means that the coat 'counts as' the exchange value of the linen, and the construction that says it means that the coat and the linen contain the same quantity of labour time.

Chapter one in the first edition sets out the TCM. But in the second edition it becomes a complicated attempt to marry parts of the TCM with Ricardo's theory. This is the source of the Byzantine complexities of chapter one that have bothered

⁷ John R. Searle, *The Construction of Social Reality*, Allen Lane, London, 1995.

⁸ K. Marx, *Capital* vol. 1, trs. Ben Fowkes, (Penguin, London, 1976), henceforth 'Penguin'.

readers for generations, and of most of the notorious conceptual problems faced by interpreters.

- (i) First, the problem of making labours homogeneous. Ricardo's theory logically requires this, because it treats labour as a single quantitative measurable, which presupposes that all its instances are homogeneous and uniform with all others. But labours (or 'activities' in ordinary language) are heterogeneous, belonging to different kinds. The obvious solution is the Utilitarian theory of action, in which there is only one kind of action, and one end that all actions aim at, utility. This has found few supporters except among economists. Marx inherits the problem from Ricardo in the second edition. He didn't face it in the first edition because the TCM does not say that labours are homogeneous, only that labours 'count as' homogeneous (abstract labour) when in the commodity relation.
- (ii) Secondly, it gives Marx the transformation problem. In the first edition there is only exchange value. But in the second edition there is also 'value', aka Ricardo's labour time. Marx now has to explain how 'value' is turned into exchange value. And that is the transformation problem, or turning labour time into prices. Generations have sweated blood over this, defending the second edition.
- (iii) Thirdly it makes abstract labour a quagmire. In the first edition abstract labour is what different labours 'count as' in commodity relationships. People may work together in a common effort, or privately for themselves. In this case they need to get what others make, and the things they make themselves become means of acquisition, and this is exchange value and things behaving as commodities. When things are exchanged, the labours that made them are still different, but that doesn't matter, because the labours 'count as' being alike. What Marx originally meant by 'abstract labour' is labours 'counting as' being alike, and so being exchange value. But in the second edition, incorporating Ricardo's theory means finding a way to connect labour time (n.b., not labour, but labour time) with exchange value. He adopts the idea that labour 'creates value'. But 'value' is ambiguous between labour time and commodity-hood, so Marx seems to say that labour creates a social status. He can't leave it like that, so he says it is 'an aspect of labour' that makes the value, and another 'aspect' that makes the thing. The first 'aspect' is abstract labour, so abstract labour is no longer an abstraction but a kind of activity with duration in time.⁹ Marx does this sort of thing to the TCM systematically in the second edition to make room for Ricardo.

In the process Marx introduces an ambiguous use of 'value' to mean the labour time embodied in a thing, and also to mean things in their social capacity as commodities, which are now referred to as 'values'.¹⁰ So the word 'value' now means

⁹Paul Sweezy builds his book on the notion of 'abstract labour time'. He writes 'abstract labour is susceptible to measurement in terms of time units', *The Theory of Capitalist Development* (Oxford University Press, 1942, [1946]), 33.

¹⁰As in Marx's phrase: 'the equal objectivity of the products of labour as values', *Capital* vol.1, (Penguin, 164).

both labour time embodied, and the social status of commodity-hood. This ambiguous usage eases Ricardo into the structure of the second edition, and affirms the combined theory. But it also merges a natural property (labour time) with a social status (commodity-hood), and this is commodity fetishism as Marx defines it. He does not want to advertise this, so in the discussion of fetishism in the second edition, he is forced to play down much of the TCM. The parts most affected are the four ‘peculiarities of the equivalent form’.

8.4 The ‘Four Peculiarities’

These are the analytical core of the TCM (Dragstedt, 55–60). The ‘first peculiarity’ is that it is the coat, the physical thing, and not its exchange value, that acts as representative of the linen’s exchange value. But it does this ‘(note well) . . . only within the value relationship . . . since no commodity can relate itself to itself as equivalent and thus can also not make its own natural skin into the expression of its own value, therefore it must . . . make the natural skin of another commodity-body into its own value-form’ (Dragstedt, 55, original italics). So the coat performs its social capacity in its natural capacity. It is already getting tricky to keep clearly in mind what is social and what is natural. Muddling them is the essence of the fetish of the commodity, and it is insinuating itself from the start.

The ‘second peculiarity’ is that ‘concrete labour becomes the appearance-form of its opposite, abstractly human labour’ (Dragstedt, 55). The different kinds of labour are passed around through the exchange of their products, and they all count as bits of one homogeneous uniform or ‘abstract’ labour, exchange value. But once the social fact of exchange value is established, the new fact comes to have overwhelming importance, and the activity that made the thing pales into insignificance in the face of the social property it counts as being. It is a case of ‘X counts as Y in C’, but one in which Y has become so dominant that X ceases to signify except insofar as it counts as a bit of Y.

Exchange value appears to exist as thing A, thing B, etc., and this has the effect of mystifying the useful thing, its social property, and the relation between them: ‘If I say: Roman Law and German Law are both law, that is obvious. But if I say, on the other hand, the Law, (this abstract entity) realizes itself in Roman Law and German Law, (these concrete laws), then the connection becomes mystical’ (Dragstedt, 57). It is all right to say ‘this thing counts as a commodity in the market system’ (‘X counts as Y in C’), but it is a mystification to say ‘this thing is a commodity’, because it does not make clear that a social status is involved, and even obscures it. This means that economic thought contains a mystification in its conceptual foundation in the notion of the commodity. From his first brush with economic thought, in the Comments on James Mill in 1844, Marx thought it was a distorted and disturbed vision that saturated reality with commerce. In the first edition of Capital, in the ‘four peculiarities’, he finally showed the deep foundation of it.

The ‘fourth peculiarity’ is: ‘the fetishism of the commodity form is more striking in the equivalent form than in the relative form of value’ (Dragstedt, 59). The property of exchange value inheres in things ‘only in our traffic with one another’ and is ‘not derived from nature’ as the property of weight is, though it appears natural (loc.

cit.). Physical analogies like sight involve causal relations between physical things, but commodity-hood is not a physical nature, and commodity relationships are not causal. Marx finds an analogy in the religious world, where things of our own devising give the appearance of being independent of us, and of having a life of their own in which they enter into relationships with us. ‘That is the way it is in the commodity world . . . [and] this is what I call the fetishism which clings to labour-products as soon as they are produced as commodities, and which is thus inseparable from commodity-production’ (Dragstedt, 60).

This particularly affects the equivalent. The exchange value of the ‘relative’ (the linen) becomes external to it, because the coat represents it. But with the ‘equivalent’ the natural thing and the social functionary are one and the same, or as Marx puts it, ‘the corporeal- or natural-form of a commodity counts immediately as social form, as value-form for another commodity . . . it thus appears as a social natural property of a thing’ (Dragstedt, 60).

8.5 Fetishism

The treatment of the equivalent in the first edition (Dragstedt, 53–60) is an imaginative logical and metaphysical analysis, and a step in explaining exchange value as a social property. It is set out under seven sub-headings. This care is necessary because the problem is delicate, and the prejudices and illusions surrounding the commodity are entrenched and persistent.

The second edition treatment (Penguin, 147–52) is incoherent, lacks detail, explanation, and organization. All the sub-headings have been removed. ‘20 yards of linen = 1 coat’, is now called ‘the value equation’ (147), thus installing the first construction to the exclusion of the second, and Ricardo’s LTV over the TCM. The social roles and functions played out in it are given no explanatory weight, and the reader gets no real understanding of the equivalent, or its place in the theory of the commodity. So the commodity and fetishism are not explained. Immediately after mentioning the ‘value equation’ Marx gets down to Ricardian quantitative considerations of ‘the proportion in which the two are exchanged’.

The poor presentation in the second edition has left fetishism out of focus ever since. Even the writer who became the authority on fetishism, Lukács, could not focus it sharply enough to explore it properly. Other authors were even less sure of their footing. Sweezy relies heavily on quotations from Lukács, as if he lacked the confidence to deal with it himself, and even Lukács relies unusually heavily on quotations from Marx.¹¹

Commodity fetishism appears, in the first edition, as the fourth of the ‘four peculiarities’, and is fully integrated into the TCM with the other parts: ‘20 yards linen = 1 coat’; the social roles things play in these relationships; the network of these relationships that forms the market; and the weird abstractions that this ‘commodity world’ throws up, such as abstract labour (compared by Marx to the ‘abstract animal’). The theory is a coherent whole, which, once stated, illuminates all its parts.

¹¹ Sweezy, op. cit., 34–40.

This is especially true of fetishism, because the notion arises from this theoretical structure, and cannot be described without it, as can be seen in the second edition where Marx tries and fails. Even Lukács could not recover the idea with any precision from the second edition, and without the TCM the best reconstruction he manages is 'reification', which is a poor approximation telling us only that 'a relation between people takes on the character of a thing', with no account of the social role-playing to explain how this is possible and what it means.¹²

Imposing Ricardo's theory forced Marx to consider which bits of the TCM he can mention and which bits he must avoid. Construing '20 yards linen = 1 coat' as an equation means that the commensurability of coat and linen must be accounted for by a common property, and not by the relation of 'counting as' as in the TCM. But it also rules out any serious use of the social roles and functions, because that would imperil Ricardo's account of commodity-hood, which being naturalistic and weaker, is vulnerable in a comparison. However, he cannot explain fetishism without those social roles, so he has to present it as best he can. Hence the problems faced by Lukács, Rubin, and Sweezy. The loss shows up in many ways. It is often said, for instance, that 'money is a social relation', which invites the economist's stock rejoinder that 'you can't buy anything with a social relation', and without the TCM there is no comeback.

Marx constantly tries to use the results of the TCM without stating the TCM. The analysis of the commodity and its social roles is replaced by vague descriptions and attempts at persuasion. He says, for instance, that the commodity also reflects the social relations of the producers to the sum total of labour as a social relation between objects, a relation which exists apart from and outside the producers. Through this substitution, the products of labour become commodities, sensuous things which are at the same time suprasensible or social (Penguin, 165).

This is not a reasonable summary of the explanation in the first edition. There is nothing explanatory in saying that when things become commodities, 'sensuous things' become 'suprasensible and social'. The TCM explained it with a theory, but here there are only impressionistic descriptions, and the passage is typical.

Lukács thinks fetishism is misperceiving a social relation as a thing, 'reification' as he calls it, but he does not explain. It seems that under capitalism it just happens that 'the relations between men that lie hidden in the immediate commodity relation, as well as the relations between men and the objects that should really gratify their needs, have faded to the point where they can be neither recognized nor even perceived', so that 'the reified mind has come to regard them as the true representation of his societal existence'.¹³

But fetishism is a failure of understanding not of perception. We do not understand that the commodity world is a social construction where things play social roles in a network of social relationships constituting the market. Marx leaves this out of the second edition, so Lukács was trying to recover the concept of fetishism

¹²Georg Lukács, *History and Class Consciousness: Studies in Marxist Dialectics*, trs. Rodney Livingstone, (Merlin Press, London, 1990 (1968)), 83–4. The work was originally published in 1923.

¹³Lukács, *op. cit.*, 93.

from fragments and the rhetorical flourish that material relations between people take on the fantastic form of social relations between things. Marx is forced to play down so much of the TCM that commodity fetishism is left without enough connection to its infrastructure to be intelligible. It is not surprising that for so long readers of Capital should have felt in their bones that fetishism was an important idea, but without feeling confident about where it belonged and what exactly it was.

8.6 Why Did Marx Impose Ricardo?

Events in economics may be part of the story. Marx kept up with economics, and attempts to make it resemble physics were growing. Ricardian economists had long hoped ‘to ground value in the absolute invariants they thought were provided by physical reality’.¹⁴ But in Marx’s time this ambition was being taken much further. Jevons was explicitly looking to physics to provide an analogy with exchange, and Walras was looking for a Newtonian model for understanding price and markets.¹⁵ It would be a nice irony if Marx were sacrificing the TCM for some similar pretension. Up until the end of ‘Results’, Marx’s work had been a critique of economics for putting capital at the centre and crediting it with the creative forces that really lie in the cooperative activity of people. Marx’s new picture seems to break with that line of criticism, because it too is putting capital at the centre.¹⁶ After ‘Results’ Marx has a different aim: he endorses the view of a Russian reviewer that ‘The scientific value of such an inquiry lies in the illumination of the special laws that regulate the origin, existence, development and death of a given social organism and its replacement by another, higher one. And in fact this is the value of Marx’s book.’ (Penguin, 102). Over some 20 years there was no suggestion that Marx saw his work in anything like these terms.

Marx loses a lot by the change. He takes on board a lot of conceptual incoherence, such as the idea that labour makes exchange value, and he has somehow to connect exchange value with labour time. By doing things in terms of ‘surplus labour’ and ‘rate of exploitation’, as he had before, he gained in transparency over the euphemism and fetishism of economic thought. By switching to ‘surplus value’ that transparency is lost behind a veil of technicality drawn from economics itself. He also loses the TCM, the social explanation of exchange value, and the theory of commodity fetishism.

Nonetheless, adopting Ricardo had advantages for Marx. He was nervous about not being taken seriously, and being thought a dilettante with strong and eccentric

¹⁴P. Mirowski, More Heat than Light (Economics as Social Physics, Physics as Nature’s Economics), (Cambridge 1989), 188.

¹⁵Mirowski, *ibid.*, 254–257.

¹⁶The title of Marx’s Contribution of 1859 had been ‘A Critique of Political Economy’, but by 1867 this was demoted to the sub-title of his new book, beneath the leading title Capital.

opinions.¹⁷ Standing on Ricardo gave him locus in the economic debate, and helped make him a player. Furthermore, ditching Ricardo would have contradicted his settled conviction, affirmed in both editions of *Capital* and in the notes on Bailey, that beneath the list of current prices a realm of hidden quantities ‘shines through’. But perhaps the greatest advantage Marx saw is that Ricardo’s labour time promised the single measurable quantitative variable needed to realise the Newtonian ambition of unifying all the laws into a single quantitative science of the ‘laws of motion’ of a market society: the concentration, centralisation, organic composition, and accumulation of capital etc. The ambition was more to do with appearance than substance, because there are no quantities in either edition beside sums of money and physical quantities (ells, tons).

Also, Marx’s constant struggle to keep the quantitative side involved, even as he developed the formal theory of commodity-hood, was a perennial albatross (e.g. Dragstedt, 16–20). He evidently concluded that he couldn’t have both. Choosing the quantities and Ricardo removed the problem.

Subsequent readers usually thought *Capital* was economics and Marx an economist. Not all sympathetic readers saw things this way. Orwell saw other things in Marx, and distanced himself from the then prevailing progressive orthodoxy. He defended Dickens from the charge of being merely a moral critic of society, suggesting that it is not at all certain that a ‘merely moral criticism of society may not be just as “revolutionary” ... as the politico-economic criticism which is fashionable at this moment’.¹⁸ But the mainstream opinion was that Marx had laid out the laws of motion of the market system, and some believed this gave socialists a superior ability to predict its future behaviour unavailable to others. It also conceded a lot to the scientific self-image economics had made up for itself in the second half of the nineteenth century. It also did something to absorb Marx further into the modern liberal tradition that had produced economic thought in the first place, which early Marxists were often keen to do.

8.7 Conclusion

The most effective weapon in the arsenal of economic thought, today as in Marx’s time, is the idea that the commodity and its market forces are things of nature, inescapable parts of the human condition, that they are ‘the real world’, and that however harsh it may seem sometimes, and however wide the gap between the needs of ‘the economy’ and the health and wellbeing of those living under it, this is

¹⁷In the *Preface* to the *Contribution* he writes: ‘This sketch of the course of my studies in the domain of political economy is intended merely to show that my views ... are the outcome of conscientious research carried on over many years.’ *Marx Engels Collected Works*, (Lawrence & Wishart, London, 1987), vol. 29, 265.

¹⁸George Orwell, *Critical Essays*, (Secker & Warburg, London, 1946), 18.

inescapable reality, and we had better learn to live with it, and stop deluding ourselves with heart-warming stories about being able to do things another way.

Marx's point is that when we find our lives blighted by 'market forces', we are not suffering the effects of natural laws operating; we are suffering the consequences of our own agreements of collective intentionality that constitute the commodity. These agreements have resulted eventually in the full-blown commodity system, the laws and forces of the market, and the trouble caused by the loss of control over our affairs entailed by this is something we have brought on ourselves. Until we come to understand that, and are ready to take control of our affairs back into our own hands, we shall just have to put up with the consequences whatever they are.

Writing in 1946 Orwell reflected that considering 'how things have gone since 1930 or thereabouts, it is not easy to believe in the survival of civilisation', but he did not think this meant abjuring politics. Rather: 'I think one must continue the political struggle, just as a doctor must try to save the life of a patient who is probably going to die. But I do suggest that we shall get nowhere unless we start by recognizing that political behaviour is largely non-rational, that the world is suffering from some kind of mental disease which must be diagnosed before it can be cured'.¹⁹

This essay is based on a forthcoming book Marx and the Commodity.

¹⁹The Collected Essays, Journalism and Letters of George Orwell, edd. Sonia Orwell and Ian Angus, Vol. 4, 248–9.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 9

The Morphogenetic Approach; Critical Realism's Explanatory Framework Approach



Margaret Scotford Archer

Abstract Critical Realism, the philosophy of the social sciences used here is equally applicable to all such disciplines and accords no special place to economics. In fact there has been disappointingly little take-up of it by heterodox economists (notable exceptions in Britain being Tony Lawson and Jamie Morgan). The generality with which Roy Bhaskar advanced CR means that necessarily theorists in each discipline must develop their own explanations, although these will share the same philosophical ‘under-labouring’ as Bhaskar characterized his own contribution. Specifically, this involves common endorsement of the following:-

- (a) A rejection of Humean ‘constant conjunctions’ as a deficient, because empiricist basis for conceptualizing social reality and causality.
- (b) A stratified ontology of the social order, endorsing emergence and the causal consequences of the second or third-order interplay between emergent properties and powers. In turn this spells acceptance of upwards and downwards causality between strata.
- (c) A refusal to assign automatic priority to structure (or culture) versus agency when accounting for causation in the social domain.
- (d) A reliance upon CR’s ‘three pillars’ for explanatory adequacy:-
 - Ontological realism
 - Epistemic relativism
 - Judgemental rationality

I will attempt to show how my own ‘Morphogenetic/Morphostatic’ explanatory programme usefully supplements the above with an interdisciplinary approach to accounting for change and stability in all social forms and institutions. This framework is obedient both to the four above principles but also provides a toolkit for those seeking to theorize about the development of particular social processes,

M. S. Archer (✉)

Department of Sociology, University of Warwick, G.B, Coventry, UK

e-mail: margaret.archer@warwick.ac.uk

practices and policies (and resistance to them) at any given $S^1 T^1$, wherever this is situated historically and geographically. In itself, the M/M approach is not a theory. Should some prefer to assign it to ‘methodology’, I would not object.

9.1 Philosophical Under-Labouring and the Need for an Explanatory Toolkit

Critical realists accept that the nature of social reality is such that its explanation requires the identification of the distinctive causal powers exercised at some given place or date. *This is the case for those processes that account for its social contours at any particular time; those that maintain a particular social configuration in being for some time; and those that transform its particular kind over time.* The difference between these mechanisms and those found in the natural order derives from the nature of the fundamentally different constituents of these two orders of reality.

Their difference also explains why ‘morphogenesis’ means something very different from in biology – where it is an entirely non-conscious process – than it does in social science. (The only thing they share is a common etymology). What is distinctive about social reality – or any section of it – is its being intrinsically, inherently and ineluctably ‘peopled’. Its ontological constitution is utterly *activity-dependent*, despite the fact that people’s thoughts and actions give rise to factors that are ‘not people’ – the most important of these being culture and structure.

Because of this I have argued along with Bhaskar that for any process to merit consideration as a generator of social change it must necessarily incorporate (i) structured human relations (*context-dependence*) because there is no such thing as ‘context-less action’ and calling it ‘situated’ makes no difference; (ii) human actions (*activity-dependence*) because even the most distant outcomes, such as GDP or Climate change in the Anthropocene would not exist without the continuous actions of people, and (iii) human ideas (*concept-dependence*) because activities like ‘voting’, ‘paying rent’ or ‘opening a bank account’ require that actors have some notion of what they are doing, however vague or misguided.¹ Necessarily, these three requirements make social theorising non-naturalistic, that is different from natural science. (As Roy Bhaskar often said, his *Possibility of Naturalism* could equally well have been called *The Impossibility of Naturalism*.)

A more familiar way of putting the above is that every theory about the social order necessarily has to come in a sack, SAC: it must incorporate Structure, Agency and Culture. The problem in hand will govern which of the three is accorded most attention and the acronym SAC is thus *not* a rank ordering of priority between the three elements. All are always indispensable.

¹ Roy, Bhaskar, 1979, *The Possibility of Naturalism*, Harvester Wheatsheaf, Hemel Hempstead.

During the last quarter of a century, the ranks of SAC deniers have swollen dramatically. As Porpora maintained in his recent book *Reconstructing Sociology: The Critical Realist Approach*,² the meta-theoretical stance of denial rests upon conflation of these three elements. Instead of distinctive properties and powers pertaining to structure, culture and agency, any pair is conflated with one another, thus ruling out examination of the (changing) interplay between them and its theorization.

When conflation rules the components of SAC are collapsed into only one ontological factor, which is given a proper name, whilst the existence of others is flatly denied. A relevant example in relation to the economy is Rom Harré's account of the 2008 financial crisis from his 'conversational model'.

Institutions are not ontologically basic, nor are any other seemingly structured entity-like beings. From the point of the conversational source model an institution is an appearance, an illusion presented by the relative stability of the flux of social acts that are constitutive of the then and there social reality. The need for a clear ontological viewpoint has been illustrated dramatically in the "collapse" of part of the "banking system" Talking that way distracts our attention from the reality, the flux of social acts performed by a loosely bounded group of active agents, following discourse rules that proved in the end to be incoherent. *There is and was no "banking system"*. (2009 my ital.)³

Instead, the guiding metaphor is of 'flows' or 'liquidity' – which depends upon a prior dissolution of all three components of SAC. Thus, the leading trope of 'liquid modernity' explicitly depends on an eclectic combination of denials of 'structure' (replaced, for example, by theoretical assertions about 'de-structuration' in the work of Ulrich Beck), denials of 'culture' as anything more than what people carry in their heads, (endorsed by Dave Elder-Vass), and of 'agency', rendered fluid by notions of serial self-reinvention (Beck and Anthony Giddens), thus severing ties with personal and group 'identity', 'interests' and 'commitments' (if anti-humanism does not make all such notions irrelevant by reducing agents and actors to Latour's 'actants'). In consequence, the picture of the social order being shaped and re-shaped by groups seeking to advance their material interests, their ideal interests and who they are is obliterated by the imagery of fluidity. In turn, the liquid society cuts loose from struggles for domination and control (societal or sectional) and becomes literally ungovernable and uncontrollable, as conveyed by the images of the 'runaway', 'juggernaut' or 'risk' society (Beck and Giddens). This spells the demise of both central 'Command and Control', but also of the robust Social Movements dominant in the macroscopic changes of the last two centuries.

Perhaps it is helpful to delineate the structure of realist social theory as in the diagram below, in order to be clear about the nature of explanation in realism. Let me start with two bold statements. First, Social Ontology (Bhaskar's contribution) explains nothing. Second, neither does my Morphogenetic Approach explain anything. That does not mean they do no work in social theorizing. Indeed, we need

²Douglas, V, Porpora, 2015, Cambridge, Cambridge University Press.

³Rom Harré, (2009) 'Saving Critical Realism', J. Theor Soc Behav 39(2), 129–143.

both of them. To begin with everyone has a social ontology (including lay actors) consisting of those things they hold to exist and not to do so. We know that lay actors differ – some hold that ghosts exist, some that exorcism can eliminate them and most neither. The same is the case for social theorists.

Nevertheless, all theorists have a social ontology, whether implicit or explicit, which effectively defines the constituents of the social world. Therefore, the SO performs a role of conceptual regulation because it governs those concepts that are deemed admissible in description as in explanation – just as an atheist cannot attribute his well-being to divine providence. Although a social ontology explains nothing, it frequently excludes certain explanations, cast in ‘improper’ terms.

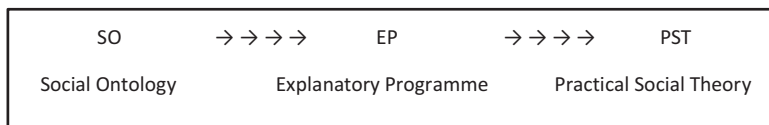
This was Bhaskar’s great contribution as he flags up in his posthumous book⁴; namely the task of rescuing ontology (how the world is) from epistemology (how we take it to be). He never for a moment claims that we have direct and infallible access to ‘how the world is’, including the social order that is ultimately of human making. Instead, his great philosophical contribution was to deny that we can answer this question in any science on the basis of empiricism (that which is available to us though data coming from our five senses). On that criterion, we will ‘perceive’ (smell, hear, etc.) something of what exists, but far from all that is in existence at the *actual or real levels*. If present at an event (a football match) a good deal of what happens will be inaccessible from anyone’s angle of vision. Yet, put these different perspectives together into some kind of generalized spectators’ view and such ‘actualism’ cannot reveal the unobservable but nonetheless real causal factors at work (for example, that key players have been bribed to lose the match). This is the level of the ‘real’ which exists and can be known only by its causal influence rather than by direct observation.

Matters are even more complex. Something may exist (the bribe) but its causal powers might not be exercised (the team cares more about winning than it fears subsequent retribution); powers may be exercised but not detected (perhaps a key shooter successfully fakes an injury), or exercised but not affect the intended outcome (the other team plays so badly that the bribe cannot prevent the bribed from winning). As researchers we have to deal with all of these ontologically variable scenarios. One consequence is that, unlike empiricism, we can rarely make predictions in the open system that is the social order – unlike the sterile laboratory – because nothing can prevent the intrusion of contingencies.

In itself, an SO tells no-one how to go about explaining anything. For this an explanatory programme is needed. That is what the Morphogenetic Approach is; the methodological complement of Critical Realism’s social ontology. The basic M/M diagram, supplies guidelines about how analytically to break up the material in hand to form the three temporal phases making-up a single morphogenetic cycle, which ends in either change or stasis and represents the start of the next cycle, viz <Structural and Cultural Conditioning → Social Interaction → Structural and/or Cultural Elaboration>. It is the investigator who contributes the material and problem to be explained and, if successful, produces what I have called a Practical Social Theory. The EP will have assisted in

⁴Roy Bhaskar, (2016), *Enlightened Common Sense, The Philosophy of Critical Realism*, Abingdon, Routledge.

marshalling the SAC components to account for the 'who', 'when', 'why' and 'what' of change, but it is the PST that does the explaining.



This means that PST plays a crucial role in the *development* of Critical Realism through *using* it in a substantive area of interest to the investigator. There are always such areas that have not previously attracted Realist research and this is necessarily the case because tomorrow's social changes (morphogenetic or morphostatic) cannot yet have been investigated. Similarly, though contingently, there is generally a large tract of issues that have simply not been investigated at all. This is why the framing of a research question is one of the most and often the most important element of a research project; it is what fosters the development of Realist theorizing. It can do so by prompting modifications in the SO because our ontology cannot remain immune to how reality is discovered to be. (As illustrated every day in medical research). Alternatively, research findings can require adjustments in the Explanatory Programme.

Although all structural and cultural properties found to be salient in any society are continuously activity-dependent, it is possible through analytical dualism to separate 'structure', 'culture' and 'agency' to examine their interplay in order to account for the structuring and re-structuring of the social order. Fundamentally, this is possible for two reasons. Firstly, 'structure', 'culture' and 'agency' are *different kinds of emergent entities*⁵ as is shown by the differences in their properties and powers, *despite* the fact that they are crucial for each other's formation, continuation and development. Thus, an educational system can be 'centralised', whilst a person cannot, and humans are 'reflexive', which cannot be the case for structures. Secondly, and fundamental to the workability of this explanatory methodology, 'structure', 'culture' and 'agency' operate diachronically over different time periods because (i) structure and culture necessarily pre-date the action(s) that transform them and, (ii) structural and cultural elaboration necessarily post-date those actions, as represented in the following diagram. It is this that distinguishes *analytical* dualism from philosophical dualism (Fig. 9.1).

Full significance is accorded to the timescale through which structure, culture and agency themselves *emerge, intertwine and redefine* one another, since this is the bedrock of the explanatory format employed in accounting for any substantive change in social forms.

It is entirely in the hands of the substantive researcher to judge where to start their explanatory project in time (and their T¹ may differ historically from that of previous research, which might have stimulated their own work). In a similar man-

⁵Roy Bhaskar, *Reclaiming Reality*, Verso, London, 1989, 'People and society [...] do not constitute two moments of the same process. Rather they refer to radically different things.' p. 76.

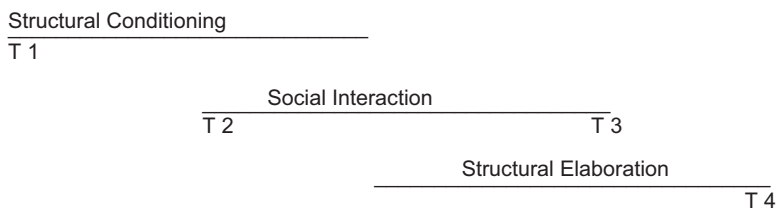


Fig. 9.1 The basic morphogenetic sequence

Margaret S. Archer, 1995, *Realist Social Theory; The Morphogenetic Approach*, Cambridge University Press, p.157

ner, the period between $T^2 - T^3$ in which groups vie with one another to shape some social form, organization, or practice can vary considerably from place to place depending on local circumstances. The EP needs to be treated as a flexible template. For example, one post-graduate had done a good study of the intricacies of local government decision-making over the 3 years of his registration, but came to me and complained that no notable changes had eventuated in this period despite considerable conflict between the groups involved. So be it, the struggles went on because the groups in question had locked in unresolved conflict, at least *pro tem*. Conflictual dynamics do not conveniently accommodate themselves to the duration of our grants! Conversely, it cannot be assumed that current forms and practices always pertained to the same nominal group, such as a profession throughout the individual careers of its members, which anyway would be of variable length. For instance, the referents of 'Company Director', 'Sales Manager' and 'Shop Floor worker' would have changed considerable, if still used, just as for Doctors and Teachers.

Most obviously, the same point pertains to T^4 , the explanandum in any theory or research focussing upon it. Given its activity dependence, every aspect of the social order is susceptible of change. Temporality is intrinsic to its constitution and so it must be in Realist social theorizing. Consequently, we must always be ready to recognize that ontologically change has occurred (although the 'when' will most likely vary with our research interests) and if it has done so significantly, it is time to start another morphogenetic cycle. The explanandum has altered, thanks to its 'activity dependence', so have the structural and/or cultural conditioning effects upon people and groups and consequently upon their concerns, vested interests and courses of action and most likely upon their outcomes. For instance, having spent many pages analysing the emergence of English State education as a decentralized system in my 1979 book⁶ (using two analytical cycles to do so), that structure morphed into something more centrally governed in the last three decades of the Twentieth century.

It follows that no social theory is eternal because what it purports to explain is often changing in some respect just as theories are advanced. The trope of 'liquid-

⁶Margaret S. Archer (1979), *Social Origins of Educational Systems*, London, Sage, reprinted 2013 by Routledge, Abingdon.

ity' had many defects, especially in relation to the global distribution of society's scarce material resources, but in addition it was presented as if it had eternal life. That is, no conditions were specified on which its continuation depended. Conversely, in the Centre for Social Ontology's current series of 5 volumes on *Social Morphogenesis*,⁷ we have been extremely cautious *not* to announce the advent of a Morphogenic Society, despite the empirical evidence of intensified morphogenesis in most social institutions almost everywhere. This was because the generative mechanism driving its manifestations (in my view, the synergy between financialized capitalism and digital technology⁸) can combine with other mechanisms to produce a variety of new social formations. Moreover, even that statement has a proviso. None of those possible changes will eventuate unless the global emissions from fossil fuels are cut back to fewer than 2% as per the 2015 Paris agreement – for which it may not already be too late – nor will any concrete utopia be realized because the game will be over. 'Morphonecrosis', the extinction of social forms and processes is a very important extension that Ismael Al-Amoudi has added to Critical Realist theorizing and the total extinction of the human race would be its ultimate expression.

To realists, nothing social, whatever its origins, is self-sustaining, which is what *inter alia* distinguishes the social from the natural world. Only a myriad of agential 'doings' (including thinking, believing, and imagining) keep any given higher level social entity in being and render it relatively enduring. In other words, whilst ever something like the centralised French educational system lasts, then move a marker, second-by-second, from the system's inception until today, and each and every moment of its 'centralisation' depends upon agential doings (including intentional inaction).

However, this is not equivalent to some Giddensian notion that *every* such doing on the part of *everyone* somehow contributes to maintaining the whole (in this case, an institution).⁹ On the contrary, some doings are entirely irrelevant to sustaining centralisation (keeping a dog), some are more important than others, and it is only because further 'doings' exist in tension with one another that things remain the way they are (Catholic and now Muslim religious practices 'provoke' centralisation to exercise its powers in defence of the *laïcité* of education in the French Republic). Still further doings are intended to change the status quo, but have not yet succeeded in doing so.

⁷Margaret S. Archer (ed.) (2013), *Social Morphogenesis*, (ed.), (2014), *Late Modernity: Trajectories Towards Morphogenic Society*; Margaret S. Archer, (ed.), (2015), *Generative Mechanisms Transforming the Social Order*; Margaret S. Archer (Ed.) (2016), *Morphogenesis and the Crisis of Normativity*; Margaret S. Archer (ed.) (2017), *Morphogenesis and Human Flourishing*, all published by Springer, Dordrecht.

⁸Margaret S. Archer, (2015), 'The Generative Mechanism Re-configuring Late Modernity', In Archer (ed). *Generative Mechanisms*, Ibid.

⁹For example, as in his paradigmatic case of language: 'when I utter a grammatical English sentence in casual conversation, I contribute to the reproduction of the English language as a whole.' Anthony Giddens, *Central Problems in Social Theory*, Macmillan, London, 1979, pp. 77–8.

What the morphogenetic approach allows us to do is to avoid the synchronic banality and futility of asserting that if a relational property endures, this must be because of some net balance of sustaining agential doings at each moment in time (reminiscent of Merton's 'net balance of functional consequences'). Instead, in completing a morphogenetic cycle, by issuing in structural elaboration, not only is structure transformed but so is agency, as part and parcel of the same process – as the *double morphogenesis*.¹⁰ (This point entirely fails to be understood in Dépelteau's misleading discussion of what he calls 'co-determination theories').¹¹ As it re-shapes structural and/or cultural relations at any given T^4 , agency is ineluctably re-shaping itself: in terms of domination and subordination, of organisation, combination and articulation; in terms of its vested interests and these in relation to those of other agents; in terms of the new roles and positions that some occupy and others do not; and in terms of the novel situations in which all agents now find themselves, constraining to the projects of some and enabling to the projects of others,¹² yet of significance for the motivation of all.

In other words, at any given T^4 something radical happens, not only to structure but also to agency. In cases of macroscopic change this affects the 'people' through transforming four 'parts' or levels of the social order: the systemic, the institutional, the role array and the positional (the life-chances of different sections of the population). Where the emergence of a significant economic change (at T^4) is concerned, one of its immediate effects consists in re-dividing the population, not necessarily exhaustively, into those with vested interests in (economic) maintenance and change respectively, according to the situations in which they now find themselves – involuntarily for the majority of people. To characterise an interest as a 'vested' one is to associate it with a particular position, the implication being that if positions (roles, institutions) change, then so do interests. As Porpora puts it, 'among the causal powers that are deposited in social positions are interests. Interests are built into a social position by the relationship of that position to other positions in the system [...] actors are motivated to act in their interests, which are a function of their social position. Again, this doesn't mean that actors always with necessity act in their interests, but if they don't they are likely to suffer.'¹³ Thus, 'opportunity costs' are differentially distributed to different groups of actors for the same course of action – hence providing directional guidance *vis à vis* the course of action each group adopts. (These are termed 'situational logics of action' in the M/M approach).

However, equal attention needs to be accorded to Morphostasis, that is simultaneously to account for the *relatively* enduring nature of structures and cultures,

¹⁰ See Margaret S. Archer. (1995), *Realist Social Theory: the Morphogenetic Approach*, Cambridge, Cambridge University Press p. 74 and Ch. 8.

¹¹ Dépelteau, François, 2008, 'Relational Thinking: A Critique of Co-Deterministic Theories of Structure and Agency', *Sociological Theory*, 26:1.

¹² See *Structure, Agency and the Internal Conversation*, Cambridge, Cambridge University Press, 2003, pp. 1–16.

¹³ Douglas V. Porpora, 'Four concepts of social structure', *Journal for the Theory of Social Behaviour*, 19:2, 1989, p. 208.

which, in turn, again serves to highlight the importance of the '*double morphogenesis*', or, in this case, its relative absence. A frequent difficulty with persuasive synchronic accounts – and I believe Bourdieu's concept of '*habitus*' to be an instance in point – is how to explain that a given relationship between parts is *ever* susceptible of transformation (indeed, he himself relied upon *external* intrusions, in common with normative functionalism). Instead, the threat of the double morphogenesis reveals how the synchronic '*forces*' (re-producing morphostasis) are an agential achievement, which is constantly threatened, rather than being ones conducive to eternal life for any social form.

To begin with, the losers in a struggle for institutional change do not quietly fade away but tend to fight on and may win concessions. Paradoxical as this might seem, morphostatic analysis cannot remain the same from one time interval to another. This is because the explanation of why something endures has to accommodate such changes in its constitution – changes that '*punctuate*' morphostasis diachronically. In other words, an emergent entity (such as a capitalist system) can retain its key relational properties and causal powers (those still making it a version of capitalism), without it remaining unchanged, as with the advent of multinational production, then financialization and, finally, digitalization.

Similarly, to simplify greatly, these new characteristics of familiar institutions also *define new groups of losers*; those with more limited economic opportunities but aware of the widening income/wealth divides. All groups in these new situations have vested interests in bringing about transformation, though not of precisely the same kind. With even greater over-simplification, the crucial question for *endurance versus change* is: '*Can these groups work together?*' This is an empirical question. What it means, however, is that we know where to look – and this is only contingently '*outside*' – to explain why time is eventually up for that which was only *relatively* enduring. When we then address the break-up of the tense balance of forces that had consistently maintained morphostasis, we also know what to do next, and that is to examine the next (potentially) morphogenetic cycle. What happens to the economy if the combination of the political '*populism*' engendered by those proclaiming themselves to be '*the 99%*' succeeds in re-structuring the previous pit-props that politics had provided for economics since the advent of capitalism, despite the gradual redefinition that the Parties of constitutional democracy had undergone?

Throughout this account it has been maintained that structural '*conditioning*' is necessarily *mediated by* (variable) agential responses to their circumstances. Without allowing for the personal powers of agents, it is impossible to explain the variability of their actions in the same circumstances. However, some question the notion of mediation itself. Thus, Manicas asks '*why postulate the existence of structure or culture as causally relevant if, to be causally effective, these must be mediated by social actors?*'¹⁴ Since he leaves the question there, it is presumably held to be unanswerable. However, structure and culture could only be deemed causally irrelevant if what was being mediated was, in fact, invented then and there by actors whose own

¹⁴ Peter T. Manicas, *A Realist Theory of Social Science*, Cambridge, Cambridge University Press, 2006, p. 72.

personal powers were entirely responsible for it. This ‘ban’ upon ‘mediation’ seems as untenable as holding that the wires bringing electricity into my house are entirely responsible for the working of my lights and electrical appliances and that the existence of a national grid and electricity generators are causally irrelevant.

This reflects a tendency among ‘weak’ realists to require some kind of instantiation of structure properties by agents before they are accorded any role in an explanation. In other words, far from their *impinging upon* agents, it is human subjects who literally bring them into play. Such a voluntaristic bias obviously provides rather better protection against being charged with reification. Examples would include John Searle’s¹⁵ notion of ‘the Background’, to which back-reference is made, for example, by listeners to disambiguate statements that require contextualization. Similarly, Manicas¹⁶ relegates structural and cultural properties to being ‘materials at hand’, without the capacity to exert causal powers but also, from his standpoint, without any explanation of why some are within easy reach of certain actors but out of reach for others. (It is thus unsurprising that Searle’s favourite social theorist appears to be Bourdieu, whilst Manicas’s book is a virtual repetition of Giddens: these two authors thus favouring the theoretical stance I have termed central conflation).¹⁷

9.2 Impatient ‘Innovative’ Responses and Their Deficiencies

Applying Critical Realism takes time. Some prefer to make a quick bid for fame, by advancing what are purported to be novel theoretical breakthroughs or announcing that they have ‘transcended’ some stubborn, unyielding problem (such as ‘structure and agency’, as bearers of different properties and powers, or ‘subjectivity and objectivity’, as indispensable for explanation but incommensurate). Frequently this takes the banal form that I term ‘Beyondism’ (coined and ridiculed by Bertrand Russell), because everything is in some sense beyond something else, including the Big Bang.

One of the most irritating is the supposed advent of new ‘turns’. When confronted with yet another last summer, I asked my Research Fellow how many he had encountered. Obliging he told me that he had listed them and his total was 48. In the next 2 weeks, I added two more, bringing us up to 50. Since such ‘turns’ imply theoretical advances and not merely that particular empirical phenomena acquire sudden popularity, it seems to me that these generally entail one of the following fallacies in theorizing.

¹⁵ John Searle, (1995), *The Construction of Social Reality*, London, Penguin, pp. 127–147.

¹⁶ Peter T. Manicas, *A Realist Philosophy of Social Science*, Cambridge, Cambridge University Press, 2006. ‘[P]ersons are the dominant causal agents in society – even while, of course, they work with materials at hand’. p. 75.

¹⁷ For a discussion of ‘what I have termed ‘central conflation’, see my *Culture and Agency*, Cambridge, Cambridge University Press, 1988, chs. 2, 3 and 4. Also *Realist Social Theory*: Ibid., Chs. 3 and 4.

First, they assign automatic causal prominence to some otherwise uncontroversial social feature, a manoeuvre that may well traduce SAC (as the 'cultural turn' either dismissed structure or conflated it with culture). Conflation is always mistaken if it conceals real differences in properties and powers, as in that case. Elements of the ideational cultural corpus can be shared without loss of value,¹⁸ unlike structural properties dependent upon the division of society's scarce material resources. Scarcity can only be characteristic of things ideational if it is artificially imposed (through, for example 'intellectual property rights' or patents or the deliberate withholding of education from certain groups).

Second, what assures any factor or feature of necessary dominance in the social order? 'Discourse' can indeed exert causal powers (note the replacement of 'gender' for 'sex' in United Nations documents and most media broadcasting). Usually, again, the answer is only conflation – as in 'knowledge constitutive interests' or the 'power-knowledge' amalgam. Their components do indeed exist but eliding them conceptually does not explain how they came to be conjoined. Certainly, the material basis of many interests encourages their bearers to scour the cultural archive for legitimatory resources, but it does not account for what ideas have been lodged there or which are selected from that array nor allow for human fallibility. Moreover, their elision occludes what was once usefully called 'ideological conflict' – useful because it retained the nexus between 'interests' and the 'ideas' deployed in the attempt to legitimate them in defensive or assertive action.

Third, some are simply tautological at best and imply a degenerative research paradigm at worst (as in the so-called 'relational turn' and its manifesto). What, a Realist would ask, is not relational? It is interesting and revealing simply to count the number of times that Bhaskar underlined how many of his key terms (in Chapter 2 'Societies' of *The Possibility of Naturalism*) were qualified as 'relational'. But to him, 'relationality' denoted real relations whose interplay generated emergent properties. This is light years from Emirbayer's¹⁹ use of the term – as an emergence denier – where the link between human relations and social change (or stability) merely points to an endless list of indeterminate 'transactions' without the most basic specification of the conditions under which a 'transaction' was likely to be successful, on whose part and to what end. For this reason, when Pierpaolo Donati and I co-authored *The Relational Subject* (2015),²⁰ we distinguished ourselves from

¹⁸ As Thomas Jefferson put it "If nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself; but the moment it is divulged, it forces itself into the possession of everybody...Its peculiar character, too, is that no one possesses the less, because every other possesses the whole of it. He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me." (August 13th 1813 Letter to Isaac McPherson). http://presspubs.uchicago.edu/founders/documents/a1_8_8s12.html

¹⁹ M. Emirbayer, (1997) 'Manifesto for a Relational Sociology', *American Journal of Sociology*, 103.

²⁰ Pierpaolo Donati and Margaret S. Archer, (2015), *The Relational Subject*, Cambridge, Cambridge University Press.

‘relationism’ and insisted on our being ‘relational realists’, basing our whole argument upon properties and powers emergent from relationality.

Overall, it seems to me that alongside the above attempts to substitute these forms of ‘central conflation’ for Realist efforts to reclaim reality, there is a generalized move to reinstate epistemology over ontology, one whose philosophical implications are often not appreciated by their exponents. It is revealing to examine the list of contents of social science Journals or just to glance at the list of staff seminars offered in any semester, since both give persuasive evidence of the pervasiveness of dealing with ‘representations’ rather than tackling reality itself. Undoubtedly, this is easier and also complements the preference for visual media over written sources (with the average hours given to television viewing in the EU now approaching those of the working week). Imagery can be influential, but only if it is never forgotten that these are images of something real and react back upon it. Insisting to young women that Size 0 pictured an ideal body-image was eventually accepted by many fashion houses as objectively damaging to the health of female teenagers by contributing to anorexia and bulimia.

These are symptoms of epistemology regaining ground but, if I am correct, they also work in combination with real changes in academia and augment negative consequences for the social sciences. Entry to the profession has itself become vastly more competitive as more graduates with (ever) better first degrees enter into an objective contest with one another to continue in academic life. One laudable but defensive consequence is that they appear to act supportively and non-adversarially towards one another. It is rare in the seminars I regularly run in various countries to find participants challenging the views of their peers. Since the unveiled objective competition continues in the probationary years of those succeeding in gaining a university post, so this contradiction persists. Because they know the rules of the game (for example, the importance of citations), it is not infrequent to encounter articles in which sentences are almost unreadable given the number of brackets opened to reference the work of their peers.

Superficially, this appears charitable but I believe it has a very negative unintended consequence. Argument is out and assertion is in. Yet argument is indispensable in Critical Realism. In a nutshell it is central to its ‘third pillar’ – the exercise of Judgemental Rationality. Since we have no direct access to the ‘real’ in any domain then reality cannot arbitrate on our theoretical propositions. Hence we are condemned to live with the most convincing contender at any given time and the arbitrator between competing claims is the strength of one argument against others.

9.2.1 The Effect of Anti-realist Evasions in the Current Global Crisis

When the Berlin wall came down, Peter Berger rightly complained about the silence of the sociologists; mute in advance of it and remaining so after it in terms of retrodictive explanation. Exactly the same reproaches can be levelled at our discipline for their

responses to today's crises, which are of even greater import and magnitude. Amongst the reasons for this temerity are the tendencies presented in the last section.

Climate Change is an obvious illustration of tardiness. In 1986 a NASA sponsored report that first published the Bretherton model, the part played by social activities was confined to a small black box on the outer periphery. The implication was not that 'we humans had created this problem; so we must solve it' and neither did the general public in the developed world take it that way. More depressing was again the silence of the sociologists. The ISSC's bibliographical analysis revealed in 2013 that a mere 3 percent of items dealing with global environmental change had been produced by sociologists in the preceding 20 years. Credit goes to the ASA's Task Force on Sociology and Global Climate Change for the publication of *Climate Change and Society: Sociological Perspectives*, but we had to wait until 2015 for Dunlap and Brulle's excellent volume.

Timorousness marked the sociological response to the whole gamut of interrelated crises, stemming from the economic crisis of 2008, through the application of austerity measures, the growth of jihadism and the closely linked migrant crisis, to the ultra-right's resurgence with nationalism and xenophobia threatening the institutions of the Western world (the EU, NATO and democracy itself); each one a threat to world peace and exacerbating increases in both social and system mal-integration at the most macroscopic level.

If one played at empiricism for a moment, I would hazard that more articles appeared about the representational significance of 'the veil' for Muslim women than about the social origins of terrorism. If one turned phenomenologist for a minute, I would cite an experience in the Vatican *Sala Stampa* in 2015, when a journalist asked me what to do about the thousands of migrants (those who managed to land in Greece). My response, namely that we needed a consolidated European fund for relief because this was not Greece's problem and neither was Lampedusa the graveyard of the Mediterranean. This was treated as unworthy of serious consideration. If one looked to the heirs of the old-left, they were still riveted upon the patrimony of historic social protests in modern dress, such as the 'indignation' of the ephemeral 'Occupy' movements. For example, Castells' exhaustive treatment in 2012 of their national manifestations, also lacked any account of how these could engage and integrate non-activists in the developed world, let alone in 'the rest'. What was missed by taking this focus were the disgruntled stirrings of populism and that the ultra-right wing political parties were already beginning to harness them.

Did Critical Realism in tandem with the Morphogenic Approach fare any better? In a generic sense, they did, meaning that owing to maintaining the distinction between structure, culture and agency and upholding SAC, some of us had consistently warned that when low system integration and low social integration coincided in time, this constituted the prime condition for explosive social change. Yet, we could not claim originality for that valuable proposition, which was first advanced by David Lockwood in 1964.²¹ However, expecting an explosion and detailing its fall-out are different things and in the latter we certainly failed and

²¹ David Lockwood, 'Social Integration and System Integration', in G.K. Zollschan and W.Hirsch (eds), *Explorations in Social Change*, Boston, Houghton Mifflin, pp. 244–257.

continue to do so. That can neither be understood nor forgiven (*tout comprendre n'est pas tout pardonner*) by pleading the impossibility of prediction in the open system that is the current global social (dis)order. Some were hanging on to the hope that a robust Third Sector could moderate the excesses of Market and State; some remained assured about the eventual evolutionary victory of collaboration and cooperation; and all were convinced that although morphogenesis would predominate over morphostasis, this could take a variety of different forms – given variations in SAC components and constitution in different parts of the globe. In a precautionary sense, we could hardly fail to be wrong there!

Yet why were we so prudential or positively pusillanimous? In part, because we had been brought up on the founding fathers, most of whose concrete utopias and dystopias had been not been realized (revolution, re-integration, and secularisation). In part, given we were more than a century of academic specialization further on also meant that none could feel confident of commanding global mastery of that portion of the literature we could even read. In part too, because of a sedulous Eurocentricism, which glued most of us to our familiar social institutions and their transformation, led us to discount the relevance of populism in, say, Latino politics. But, above all, we, like most in the Western world, simply did not believe that a globally explosive conjuncture could bring the whole social order, as we knew it, to its knees.

These are all plausible and probably contributory factors. But for Critical Realists they remain excuses. Had we clung to our macroscopic generative mechanisms and had we seriously explored SAC on the global canvas our contribution would not need excusing. The redemption of our theoretical approach can only consist in how, realistically, we conceptualise putting the pieces of our broken world back together.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Chapter 10

‘God Created Man *ἀντεξούσιον*’: Grotius’s Theological Anthropology and Modern Contract Doctrine



Jonathan Price

Abstract Hugo Grotius, the seventeenth century Dutch scholar, is most famous for his contributions to modern international law, particularly the law of the free seas. Yet, he has had just as lasting an effect on the formation of modern contract doctrine, originating in the same text that produced his maritime law. Grotius instigates a change in the theological anthropology implicit in late scholastic contract doctrine by importing a radical sense of God-given, unfettered freedom. This he calls ‘natural liberty’. He thereby renders contractual freedom freer, as it is now liberated: from its divine *telos* as part of man’s salvation; from the constraints of moral philosophy; and from the need of any ultimate end. Yet, he does not set the will completely at liberty in its contractual relations; certain formal and moral constraints remain. But it is no longer required that, in order to be well used, contractual liberty must be busy building the New Jerusalem. For, its divine purposes are now more mundane: peace and order on earth, beginning with oneself. For the moral theologians who developed ‘freedom of contract’, our natural promissory powers, as manifested in contractual relations, were ordained as means contributing to our salvation. For Grotius, they had been reduced to the providential means of our survival.

The great sense of natural liberty that Grotius vested in the freedom of contract would eventually open the door to radical contractual liberalization of the kind that was seen in North Atlantic nations in the nineteenth century, the primary causes of which are still disputed by scholars. In this chapter I argue that modern contract doctrine has never lost its theology of the radically free will since Grotius installed it there. If true, this theology of the will does some work in explaining the persistent liberal tendency in modern law and in modern economics (albeit not in economic theory), reducing the need to resort to naturalistic, materialistic, or ideological accounts.

J. Price (✉)

Faculty of Philosophy and Sociology, University of Warsaw, Warsaw, Poland

Blackfriars Hall, University of Oxford, Oxford, UK

© The Author(s) 2020

P. Róna, L. Zsolnai (eds.), *Agency and Causal Explanation in Economics, Virtues and Economics* 5, https://doi.org/10.1007/978-3-030-26114-6_10

151

10.1 Introduction

In describing the re-emergence of ‘freedom of contract’ as the result of the long development of contract doctrine by the late scholastics, Wim Decock says contract became ‘the instrument of a self-conscious *dominus* who [could] decide to do whatever he wants with his private property’.¹

True enough, but only within necessary limits. Those very scholastics who had freed *Christians* to arrange their affairs contractually—who had recognized and granted them ‘freedom of contract’—did not allow for the rapacity of such newly-liberated *domini* to destroy their liberty with license. There were moral brakes, side constraints, and substantive limitations on just what could be contracted, as well as in what manner. There were considerations of justice, such as equilibrium and fairness in exchange that curbed ‘freedom of contract’ and thus limited how lawmakers might extend the freedom to contract in the law.²

However, Hugo de Groot, the seventeenth century Dutch scholar better known by the Latin name Grotius, instigates a change in the theological anthropology implicit in late scholastic contract doctrine. He imports a radical sense of God-given, unfettered freedom. This he calls ‘natural liberty’, which he describes, in a passage also discussed below, as the normal activity of the *dominus*’ will. He thereby renders contractual liberty freer, in a sense, as it is now liberated from: its divine *telos* as part of man’s salvation; from the constraints of moral philosophy (but not wholly loosed from all morality); and from the need of any ultimate end. Grotius does not set the will *completely* at liberty in its contractual relations: one must never *knowingly* trespass revealed religion (orthodox Christian faith), or commit any of a few truly egregious sins, even in war (poisoning of water wells, rape, use of assassin who betray loyalty).³ But it is no longer required that in order to be well used contractual liberty must be busy building the New Jerusalem. For, its divine purposes are now more mundane: peace and order on earth, beginning with oneself. Peace and order have many possible ways, none being necessarily better or worse, according to Grotius. For the moral theologians who developed ‘freedom of contract’, our natural promissory powers as manifested in contractual relations had been ordained as means contributing to our salvation. For Grotius, they had been reduced to the providential means of our survival.⁴

The great sense of natural liberty that he vested in the ‘freedom of contract’ would eventually open the door to radical contractual liberalization of the kind that

¹Wim Decock, *Theologians and Contract law: The Moral Transformation of the Ius Commune (ca. 1500–1650)* (Leiden 2013) 213.

²George Gardner, ‘An Inquiry into the Principles of the Law of Contracts’, *Harvard L. Rev.* 1 (1932) details several of the moral principles implicit in contract law.

³Hugo Grotius, *De jure belli ac pacis / On the Law of War and Peace* (1625) III: XV–XIX. ‘DIBP’, henceforth.

⁴What then comes to the fore is the virtue of ‘moderation’, as it is translated. Cf. *DIBP* III.11–16. Each chapter attempts to curb excesses in war not by means of moral theology but by ‘moderation’.

was seen in North Atlantic nations in the nineteenth century, the primary causes of which are disputed by scholars. Below I put forward an idea, the argument for which I am developing as a monograph: modern contract doctrine has never lost its theology of the radically free will since Grotius installed it there. If true, this theology of the will does some work in explaining the persistent liberal tendency in modern law and modern economics, reducing the need to resort to naturalistic, materialistic, or ideological accounts. Here as elsewhere, despite the long-held line on secularity in legal development, 'evidence is mounting that many legal concepts are derived from theological traditions'.⁵ All the more, some theological concepts or doctrines remain present in or *as* legal doctrines after they have been so derived. As the origins of contractual liberalism are increasingly an object of study, the influence of theology and theologians on the emergence of it and other modern institutions is moving from grudging to open acceptance.⁶

10.2 Freeing 'Freedom of Contract' from Moral Theology

'Freedom of contract' is what Pedro de Oñate (1567–1646) celebrated as '*libertas contrahentibus restituta*'.⁷ This phrase should not be confused with the English-language phrase 'freedom of contract', which came into general use only in the nineteenth century as a term of contempt. That was during debates on granting limited liability to joint stock corporations, when there were great disruptions thought to be caused by hyper-liberal contractual freedom. Although, if I am right, a theological doctrine provided by Grotius ties these two 'freedoms of contract' together as two of a kind. The more ancient 'sacred history of 'freedom of contract'' is related by Decock in his *Theologians and Contract law: The Moral Transformation of the Ius Commune*. I pick up just after he leaves off, in the Protestant-led Low Countries that had by the early seventeenth century permanently severed their political ties with the Spanish king and with the Roman Church. However, they had not

⁵*Ibid.* 22–23, 22n75, 23n76 and following detail literature correcting the secular(ist) narrative in various areas of law. The formative influence of canon law is nearly nowhere denied.

⁶On the open acceptance side is the work of the late Harold J. Berman, in the two volumes of *Law and Revolution*. Volume 1 is *The Formation of the Western Legal Tradition*. Cambridge: Harvard University Press. (1983); and volume 2 is *The Impact of the Protestant Reformation in the Western Legal Tradition* (2003).

⁷Decock (2013) prologue; 212. Oñate put the result of taking promising seriously succinctly: '...natural law, canon law, and Hispanic law entirely agree and innumerable difficulties, frauds, litigations and disputes have been removed thanks to such great consensus and clarity in the laws. To the contracting parties, liberty had very wisely been restored [*contrahentibus libertas restituta*], so that whenever they want to bind themselves through concluding a contract about their goods, this contract will be recognized by whichever of both courts before which they will have brought their case and it will be upheld as being sacrosanct and inviolable. Therefore, canon law and Hispanic law correct the *ius commune*, since the former grant an action and civil obligation to all bare agreements, while the latter denied them just that.' *De Contractibus* (Romae 1646) 40 (I.I.2.5.166). Translation at Decock (2013) 163–164.

wholly rejected scholastic learning, at least not in the law. It would be the new theology, brought to bear on inherited legal doctrines, that accounts for the new trajectory taken towards liberalization.

To understand what that means for contractual liberalization, it must be remembered that the theologians who provided much of the apparatus of modern contract doctrine were more concerned with the salvation of souls than with economics or history or law or anything else *sub luna*.⁸ They were not shy to enter into every area of human life where morals or God were wont to go. That was seen as the mandate of a *science* with so grand a *telos* as theology. Francisco de Vitoria noted that ‘the office and calling of the theologian are so wide, that no argument or controversy on any subject can be considered foreign to his profession.’⁹

Besides the general mandate, particular interest in contract law related to the development of moral theology in the Western church in the later middle ages. Promising and promise-keeping, with its formal effects including contractual and quasi-contractual obligations, were natural fodder for theological reflection, especially within a Church built with Roman law. The result was the development by theologian(–jurists) of a general law of contract: a theorization of Roman law from its ancient action-based beginnings. Decock calls their doctrine ‘early modern scholastic contract doctrine’. It revolves round notions of freedom, the will, and mutual consent.¹⁰ Freedom, correctly understood and *effected*, participates in man’s salvation.¹¹

Yet, this ‘freedom’ would be extended far beyond the bounds of what scholastic moral theologians understand to be ‘free’—to the point where liberty becomes license, and thus where, if it is still to be called ‘freedom’ it has been redefined. Compare ‘freedom’ as the ability to do what is good (virtue), to ‘freedom’ as a lack of any or all constraint. The latter is what nineteenth century will theories permitted in their volitional understanding of ‘freedom of contract’.¹² ‘Licentious liberty’ came to be the meaning of ‘freedom of contract’ by the nineteenth century.

Popularly, Grotius gets credited with forming modern contract doctrine. Yet, the further claim, that in order to arrive at modern contract doctrine, he alters the anthropology of late scholastic contract doctrine *theologically*, has not to my mind been laid at his door. Oñate wanted ‘freedom of contract’ to be, in Decock’s words, ‘the juridical principle that best fosters peace and moral comfort’ amidst scarcity.¹³ A canonical understanding of freedom is not far from this.¹⁴ Contract had ascended to the prominent place in the moral theological tradition, ‘as the principal tool for the

⁸Decock (2013) 5.

⁹*On Civil Power*, prologue, in Pagden and Lawrence (eds), *Francisco de Vitoria, Political Writings* (Cambridge 2001) 3; Decock (2013) 43.

¹⁰Decock (2013) 9.

¹¹James Gordley, *The Philosophical Origins of Modern Contract Doctrine* (Oxford: Clarendon 1991) 30ff.

¹²*Ibid.*, ‘Liberalism and Nineteenth-Century Contract Law’, 214ff.

¹³Decock (2013) 6.

¹⁴Helmholz, *The Spirit of Classical Canon Law* (London 1996) 49.

regulation of all human affairs, including international relations and the relations between citizens and the public authorities'; but this contractual freedom had erst-while been, 'the freedom to develop virtuousness, to express moral responsibility, and to strengthen mutual trust amongst human beings.'¹⁵

With the canonists, the moral theologians were concerned with *cura animarum*. But connecting salvation of the soul to contractual liberty is almost never something that those who inherit contract doctrine from Grotius do.¹⁶ Soteriology was already nearly absent in Grotius's own treatment of contract. This can be explained by the one glaring absence when compared to his Calvinist co-religionists and to the Catholic jurists who preceded him: Grotius seems not to have a robust doctrine of (original) sin and its (lasting) effects. In his wildly popular apologetical work written to be accessible to the common reader, *De Veritate Religionis Christiane* /The Truth of the Christian Religion (1627), neither the concept of sin nor its cognates nor even individual serious sins appear very often. And when they do, the topic is swiftly swept aside. Salvation is not a pressing psychological or social worry, if sin is viewed more as error or treatable disease than as terminal illness.

There are both practical and theological reasons that Grotius extends 'freedom of contract' in the way he does. Practically, many thinkers in his age had been chastened by theological conflicts. They were prevented, whether by others or by self-censorship, from developing theories that touched the exposed religio-political nerves.¹⁷ By removing contract doctrine from a web of notably Catholic and Jesuit theology, thereby generalizing it, Grotius made it both palatable to liberal Dutch Protestants and Remonstrants; and less offensive to the strict Calvinists that governed his land with a watchful eye for 'popery'.

Nevertheless, there would have been many ways to accomplish the same end – some of which could have usefully retained the moral constraints on contract, especially, if he expected it to gain broad acceptance. Grotius could have, for instance, built a general law of contract using uncontroversial or under-utilized stories from the Bible including some already appealed to by the late scholastics: 'Let your yes be a yes', and stories from Judges, respectively. The contextual explanation might eventually account for why Grotius avoided specific paths, but it falls short of explaining why he developed his thought on contract *just as he did*. As a writer of several theological works,¹⁸ Grotius's legal writings also benefit from his theological learning, at times explicitly. I thus look for the theological in the legal, to what Grotius actually believed about God and His creation, positing that without

¹⁵Decock (2013) 7; Zimmermann, *The Law of Obligations: Roman foundations of the civilian tradition* (Clarendon 1996) 544.

¹⁶*Ibid* 6.

¹⁷Mortimer, *Reason and Religion in the English Revolution: The Challenge of Socinianism* (Cambridge 2010) 9.

¹⁸Some modern editions of his theological works include, Grotius, H. (1990). *Defensio Fidei Catholicae de Satisfactione Christi, adversus Faustum Socinum Senensem*. Assen/Maastricht, the Netherlands, Van Gorcum; Grotius, H. (2001). *De Imperio Summarum Potestatum circa Sacra*. Studies in the history of Christian thought, v. 102. H.-J. v. Dam. Leiden, Brill; and Grotius, H. (2012). *The Truth of the Christian Religion*. Indianapolis, Liberty Fund (2012).

accounting for his theology of the will – or ‘natural liberty’¹⁹ as he calls it, and explains it using three carefully chosen terms – the question surrounding the origin of his contract doctrine remains only partially answerable.

10.3 The (Free) Will & Law

Mary Ann Glendon says modern law ‘touches nearly every aspect of human life, and different areas of the law typically emphasize different aspects of the person.’²⁰ Contract law emphasizes the will and its activity. This becomes apparent in cases of mistake, coercion, and duress, as well as being implicit in English law doctrines of offer and acceptance. But ‘the will’ often serves as a placeholder; a name without a face, or a name with too many faces. In discussions about contract, it is variously: choice, wish, desire, that which is chosen (or would have been chosen if not coerced), that which chooses; evidence of intention to form legal relations, or to be bound contractually, or of consent. However construed, the will either refers to a true psychological will, or to a presumed or attributed will, for instance, that of the reasonable man or of the diligent man in such-and-such circumstances. It serves so many functions as the point of justification of, brief explanation of, or some ultimate place of origin for, contractual obligation. Yet, the term ‘will’ neither refers to a well-developed concept, nor usually does it assume anything that could be called a philosophy of the will.

Except, that is, when it does. The late scholastics knew what was meant by ‘the will’. They knew its activities, its freedom and virtues, its limits and vices; and they could divide it out from the other parts of the soul with some precision. Although some nineteenth century will-theories also benefitted from a clear-cut doctrine of the will (particularly amongst German jurists), modern contract doctrine generally does not.²¹ If there is a coherent philosophy of the will operating, a rational reconstruction would need to be done before conceptual analysis is possible. The absence of a clear *concept* does not leave modern contract doctrine bereft of doctrines relating to the will.

Grotius flirts with a few ideas of the will, often implicit in notions such as ‘natural liberty’ and ‘*sui juris*’, and importantly with the use of the term ‘ἀυτεξουσιον’, all within a theological context that is descriptively reactionary (which I detail below). It was anti-Calvinist regarding determinism and predestination, and thus regarding the doctrine of the free will. This may have led him to cleave to freedom wherever it could be found, and to freedom in its most accessible form. It also could have led to producing less well constructed concepts than a constructive theory might have. It is also the case that Grotius is not a careful philosopher, who seeks conceptual clarity and logical rigour above all else. He is too historically-minded,

¹⁹ *De Iure Praedae Commentarius*, (Williams and Zeydel trs, Oxford 1950) 1:18.

²⁰ Mary Ann Glendon, ‘*Conceptualization of the Person in American Law*’. (Vatican City 2006).

²¹ Gordley (1991), 162–164.

such that more than a third of his universally famous tome consists of quotes from long-dead authorities; as well as at times too much of a jurist. He does, however, manage to present both a notion of the meaning of the free will and clearly indicate its centrality to his contract doctrine.

10.4 The 'Person of Law'

The vessel that carries Grotius's doctrine of the free will forward is his 'person of law', namely, he to whom law is addressed or for whom it is written. This logically pre-legal entity (substance) is a natural person (a moral entity), *who* also might be given legal personality, as a 'person in law'. There is then a shared vision between legal philosophy and law that can thereafter be used to justify the law. I here offer – and in the following sections start at defending – a rationally reconstructed definition of Grotius's 'person of law' as: he who is owner of his own liberty as property.²² When consenting and promising he does so with that very 'property', referring back to the 'self-conscious *dominus* who [could] decide to do whatever he wants with his private property', but now adding Grotian self-ownership to the mix.²³

Grotius' substantially libertarian vision of man, which I endeavour to show in his own words below, is a specific form of volitional personhood. In a system designed for a libertarian 'person of law', one would expect the 'person in law' to be given broad contracting powers. Said otherwise: Where there is a positing of *natural* freedom of contract, as in Grotius, the legally-sanctioned freedom to contract should be great (barring system-specific reasons curbing freedom), or greater than it would otherwise have been with a more limited notion of freedom active. Now, Grotius was neither a legislator nor a judge. And the life of the Grotian person in law, say, in the European legal codifications lies outside of the scope of this paper. Below, however, I do suggest just how free he might permit his 'person in law' to be by way of implication.

In contrast to the Grotian 'person of law', the contract doctrine of the late scholastics had as its 'person of law' a moral agent with a particular hierarchy of faculties, bound to a particular hierarchy of goods. This moral agent, a '*persona*', was not far from what one finds in Thomas Aquinas. Put crudely: reason (or understanding, intellect) was meant to rule the passions (or appetites, desires, whether for good or bad) by way of the will (the 'intellectual appetite', which is also the 'power of choice').²⁴ This was fundamentally in line with earlier scholasticism, where the Boethian inheritance affirmed that 'person' is '*naturæ rationalis individua substantia* /the individual substance of a rational nature.'²⁵ Any freedom that was to be

²² Partially derived from Richard Tuck, *Natural Rights Theories: Their origin and development* (Cambridge 1998) 60, 69.

²³ Decock (2013) 213.

²⁴ *Summa Theologiae* Ia.83.4 co (free-will); Ia81 (sensuality); Ia.79.8 (reason).

²⁵ Boethius, *Contra Eutychen et Nestorium*.

enjoyed could not be detached from reason, and reason was to be ordered to the good. With this ‘person of law’ in place, Christian moral philosophy transformed the *ius commune*, resulting in the restoration of ‘freedom of contract’. Eventually, the ‘person in law’ was made to bend to the reality that was the *persona* of the later scholastic tradition, resulting in the legal revolution described by Decock.

10.5 Contract as Promise

Grotius’s point of departure in thinking about contract doctrine is the conclusion of the late scholastics: ‘The moral theologians reorganized the *ius commune* tradition on contracts around the meeting of individual wills as the natural, necessary and sufficient cause to create contractual obligation.’²⁶ The result is that legal contract is at bottom a juridical form of natural promissory powers, given social form; exercised in as much freedom as conscience, convention, and (material) scarcity allow. Grotius deals with promise immediately before contract in *The Law of War and Peace*, and substantially follows the logic and themes that are mentioned in the above quotation.

The intellectual background to the elevation of the will in contractual obligations is referred to as ‘consensualism’. A doctrine of consensualism developed by the late scholastics is taken up by Grotius and folded into his natural rights theory. The result of consensualism in modern law is that it is now ‘normal to call an agreement between two or more persons a contract or convention and use these words as synonyms.’²⁷ It is true that jurists of the sixteenth century school at Leuven professed *pacta nuda sunt servanda*,²⁸ but this had not yet been translated into a *general* rule in legal discourse (even if it had already been present in canon law, and was enforced there). It also deviated from the long-held Roman law principle that clearly separates a formless agreement from contracts as recognized agreements, enforced by actions.²⁹ Pacts lacked such enforcement and were therefore ‘naked’. Even stipulation, a promise to do what the other party asks, required a formality in order to be enforceable. ‘Consideration’ in English contract law is a small example of the sort of extra assurances or formalities that the Roman law once required to ‘clothe’ contracts.³⁰ At times specific words were what the Romans wanted, and one could be bound to them, *even if one did not promise or intend to be bound*. Consensualism

²⁶Decock (2013) 10.

²⁷‘Change of paradigm in *contractus*’, in Boudewijn Sirks (ed), *Nova ratione: change of paradigms in Roman law* (Wiesbaden: Harrassowitz 2014).

²⁸Waelkens, ‘Was er in de zestiende eeuw een Leuvense invloed op het Europease contractrecht?’ in Tilleman and Verbeke (eds) *Actualia vermogensrecht* (Brugge 2005) 3–16; Decock (2013) 42.

²⁹The consensualist principle is also in the *Decretum Gratiani*.

³⁰Note that consideration was not a Roman law doctrine.

relies on promises being at the heart of contracts; whereas many forms of contract did not require it or look for evidence of it.³¹

Decock calls the result 'the victory of consensualism' as it offers a 'voluntarist account of contractual obligation'.³² Freedom of contract grants 'contracting parties the possibility to enter into whatever agreement they [want] on the basis of their mutual consent'. Moreover, '[t]hey could have the contract enforced before the tribunal of their choice'.³³ Again, this did not result in the ability to contract anything at any time. It emerged first amongst jurists many of whom were also moral theologians. Whilst nineteenth century versions of will theories of contract were 'characterized by the absence of moral considerations', so much so that fairness in exchange was seen to contest 'freedom of contract', earlier versions were safely ensconced in the world of the traditional pagan and Christian virtues and vices, divine revelation, natural law, and practical reason.³⁴

This was a dramatic change. It was not only that 'Roman contract law did not universally recognize the principle that agreements are enforceable by virtue of mutual agreements alone'.³⁵ Although actions and remedies of various kinds were scattered about the law, there was not even a general category of contract in Roman law until the moral theologians developed one. As far as actionability based solely on consensualist grounds, for instance sale, lease, mandate, and partnership, those formed a small set.³⁶ The natural law principle used to turn contract law toward consensualism was that all agreements are binding, supported by Holy Scripture: 'Let your yes be a yes' (Matthew 5:37).

With the victory of consensualism, it became necessary to have only three things for valid contractual obligation: *Animus obligandi* / the promiser's will to be bound, *promissio externa* / communication of the promise externally, and *promissio acceptata* / the promisee's offer of acceptance. All accepted offers are binding.³⁷ Fictitious promises, such as a case in which one party was not in possession of an *animus obligandi* were considered by many not to be binding, since that is the essence of promise.³⁸ Early modern scholastics contributed by consecrating and systematizing this new paradigm.³⁹

³¹ Consensualism in D 2,14,1–3 relies on the parties being of one in mind about what they want, because only then do they agree. Such agreement is at the heart of all contracts (according to Pedius and Ulpian). Stipulation is a crude (and partial) form of consensualism because the answer 'yes' to the question is a formal and evidentiary way of consenting. In Republican times error was not allowed in stipulation, suggesting the consensualism of Pedius did not lie at heart of it.

³² Decock (2013) prologue and 163; 142 (victory of consensualism), 153 (*pacta nuda*), 339 (general principle of contract).

³³ *Ibid* prologue.

³⁴ *Ibid* 1.

³⁵ *Ibid* 2–3.

³⁶ Gordley (1991) 3–4, 69–71 notes Soto, Molina and Lessius; Decock (2013) 3.

³⁷ Decock (2013) 163, 178.

³⁸ *Ibid* 193.

³⁹ *Ibid* 162–163.

10.6 Contract as Private Legislation

It is hard to underestimate just how invested certain Jesuits in the generation just before Grotius were in this consensualist approach to contract, with the obvious caveat that enforcement had to take place in order for consensualism to work. But with universal enforceability of agreements in place, they had secured a guarantee for freedom (*libertas*), a value that they esteemed to be priceless.⁴⁰ Leonardus Lessius (1554–1623) is the spiritual and intellectual progeny of that tradition and one of the sources from which Grotius collects his esteem for liberty.⁴¹ In those who follow Lessius, such freedom was envisioned by the return of the old metaphor of contract as a form of private legislation.

The near-contemporary of Grotius, Tomás Sánchez, says: ‘Every obligation which does not ensue from a law comes into existence through the private will of man’, adding, ‘so where the will is absent, the obligation is absent.’ And further: ‘promissary [*sic.*] obligation arises out of a private law which the promisor imposes upon himself, but no law is binding unless the legislator intends it to be binding [*nulla lex obligat nisi legislator obligare intendat*].’⁴² That was published in Antwerp in 1620, only 5 years before Grotius’s *De iure belli*.⁴³ All of which relies on ‘*voluntas libertatem possidens*’, that the will is controlled and controllable, either self-regulating or regulable by some other power,⁴⁴ such as reason or the ‘self’ in Charles Fried’s conception.⁴⁵ Self-ownership of a kind is being sought by way of the will; the will *possesses* freedom.

Nevertheless, it is not yet called ‘ownership of liberty’ or construed as a form of *dominium*. Yet, one can see that the leap to shore is now not far. A curious fact is that there were other signs of ownership, more evidentiary ones, that could have been used, but were not. Decock notes that ‘Lessius thinks it is the very sign of ownership that he who owns goods has the arbitrary power also to destroy them even out of pure lust, such as killing for pleasure [*perimere voluptatis causa*].’⁴⁶ The careful

⁴⁰ Decock (2013) 163.

⁴¹ At least 24 overt references to Lessius occur in *DIBP*.

⁴² Tomás Sánchez, *Disputationes de sancto matrimonii sacramento* (Antverpiae 1620) 30 (I.I.9.5) translation at Decock (2013) 193–194, where he discusses this locus and similar positions of early modern scholastics on contract as private legislation. Error and vices of the will are where volitional contract doctrine easily approaches difficulties, since each only ever has access to his own ‘*animus*’.

⁴³ Whether Grotius ever read Sánchez’s *Disputationes* is unknown to me. But these themes return below in Grotius’s notions of ἀντεξούσιος and natural liberty.

⁴⁴ Decock (2013) 163.

⁴⁵ Charles Fried, *Contract as Promise: A theory of contractual obligation* (Harvard UP 1981) 21; at 137n9 Fried commends, and I second, P. S. Atiyah, *The Rise and Fall of the Freedom of Contract* (1979) 41–60, 649–659 for Anglo-American writing on promise from Hobbes to modern times; For the obligation of promise from the element of reliance, see Neil MacCormick, ‘Voluntary Obligation and Normative Powers’ *Proceedings of the Aristotelian Society*, supp. Vol. 46 (1972) 59.

⁴⁶ Decock (2013) 166; Lessius, *De iustitia et iure* II,3,3,8.

construal of certain forms of *dominium* that man could have, would certainly make *dominium* of liberty arguable on such evidentiary grounds: it is only the man in question who can destroy his liberty by contractually binding himself with it; ergo, it is his *dominium*. It is not clear if Grotius knowingly tracked this route by way of the *ius abutendi*. However, I am inclined to believe he did. It is now time to allow him to speak for himself.

10.7 Grotius on 'Natural Liberty'

Grotius's *De Iure Praedae Commentarius* (1604–1608)⁴⁷ gives us the major elements of the phrase I am using as his 'person of law': he who is owner of his own liberty as property.⁴⁸ He writes:

Fecit enim Deus hominem αὐτεξούσιον, liberum suique juris, ita ut actiones uniuscujusque et rerum suarum usus ipsius, non alieno arbitrio subjacerent, idemque gentium omnium consensu approbatur. Quid enim est aliud naturalis illa libertas, quam id quod cuique libitum est faciendi facultas? Et quod *Libertas* in actionibus idem est *Dominium* in rebus. Unde illud: *Suae quisque rei moderator et arbiter*.

God created man αὐτεξούσιον,⁴⁹ free and *sui iuris*, so that the actions of each individual and the use of his possessions were made subject not to another's will but to his own. Moreover, this view is sanctioned by the common consent of all nations. For what is that well-known concept, "natural liberty," other than the power of the individual to act in accordance with his own will? And liberty in regard to actions is equivalent to ownership in regard to property. Hence the saying: 'every man is the governor and arbiter of affairs relative to his own property'.⁵⁰

Citing Aristotle immediately before and Plato immediately after this passage, he sandwiches this divine authority as the heart of the matter: 'God created man... free', Grotius says. Theologically there is no determinism, but freedom of the will, which he suggests is synonymous with 'natural liberty'. He communicates that *faith* into a legal reality in contract doctrine, first by proceeding to derive basic principles of justice: 'From the foregoing considerations the rule of good faith is derived: What each individual has indicated to be his will, that is law with respect to him. /

⁴⁷ For dating of the manuscript in relation to the development of his doctrines of law, see: Martine Julia Van-Ittersum, 'Dating the Manuscript of De Jure Praedae (1604–1608): What Watermarks, Foliation and Quire Divisions can tell us about Hugo Grotius' Development as a Natural Rights and Natural Law Theorist'. In: *History of European Ideas*. 2009; Vol. 35, No. 2. pp. 125–193.

⁴⁸ Although *DIP* (1604–5) remained in limited circulation until its publication in 1868, its influence on Grotius's *Mare Liberum* (1609) – which was a published section of *DIP* – and *DIBP* (1625) are discernible.

⁴⁹ For a use of 'exousia' as authority, see Romans 13:1ff.

⁵⁰ Translation, slightly corrected by me, from Grotius (1950), 1:18. Cf. 4.35.21 for partial; substantially repeated in *DIBP* II.5.6, II.20.48.2n6, II.21.12.

Hinc illa fidei regula, *Quod se quisque velle significaverit, id in eum jus est*.⁵¹ As is evident in the longer passage, Grotius understands natural right and the original acquisition of a right of ownership to be based on control (persistent ‘occupation’ by way of the will).⁵² This volitional route to (original) ownership shows up regularly in his treatises. It is thought by Grotius to be a logical extension of ‘natural liberty’ plus natural right: the will in the service of self-defence (broadly construed to include ‘chastity’) and self-preservation.⁵³

What about the two terms he relies on: ἀυτεξούσιον and *sui juris*? Regarding ‘*sui juris*’ there is a juridical use that would not necessarily include natural liberalism. For the Romans it meant ‘self-determination’, ‘being one’s own *pater familias*’, which would come with the end of paternal control, usually after the current *pater familias* dies.⁵⁴ In the middle ages ‘*sui juris*’ status would be recognized at the age of majority. But Grotius’s great legal treatises are not about the law of majority. They are about the law of nature in specific instances, especially in places where law has not or refuses to take cognizance, various untamed wildernesses of land and sea, as well as the ‘moral desert’ of war. There, law is only provided by the private legislator. ‘*Sui juris*’ is thus best understood to be another way of presenting the workings of ‘natural liberty’ in ways that would cause agreements to be forged, promises to be made, and contracts to provide order. Grotius teaches that man is naturally sociable.⁵⁵ Robinson Crusoe has the capacities implied by *sui juris* even while alone in nature; but he is not meant to remain alone in nature. Being *sui juris* among others entails exercising the will as a source of order.

In *DIBP*, ἀυτεξούσιος occurs three times, with Grotius seeming to use it in the same sense as in the passage from *DIP*. In one instance, a child which is no longer living at home and is grown is ‘altogether ἀυτεξούσιος, at his own disposal’ or ‘on his own’.⁵⁶ It, however, would seem to embolden the claim of *sui juris* but shade its meaning a little differently, appealing to being one’s own *authority*, rather than a self-legislator.⁵⁷ One might even see a logical priority, placing authority prior to

⁵¹ Concerning just how ‘free’ Grotius thought our divine endowment made us: he endorsed the freedom to act, but did not advocate the abolition of guilds, for instance.

⁵² *DIBP* II.III.1ff.

⁵³ *Ibid* II.I.5–7.

⁵⁴ Once ‘*major*’, you are no longer under the guardianship. With the Romans it meant ‘the state of self-determination’, since one had to be emancipated or wait until one’s *pater* died.

⁵⁵ *DIBP* II.I.9: ‘...we are drawn to friendship spontaneously, and by our own nature’.

⁵⁶ *DIBP*, II.6, and also II.XX.48.2n6 and II.XXI.12.

⁵⁷ The famous ‘ἀυτεξούσιος’ statement of *DIP*, I.18, is repeated in *DIBP* II.5.6, II.20.48.2 n.6, II.21.12. Grotius relies on the concept of ‘*sui juris*’ which he allows to retain the meaning of the late scholastics. Yet, he might have changed its legal use. [I must investigate how it was used by, e.g. Donellus. Was it more restricted (i.e., related to the Germanic *Vormundschaft* [guardianship] and *voogdij* [ditto])? Grotius’s *Inleidinge tot de Hollandsche Rechts-Geleerdheid* / Introduction to the Jurisprudence of Holland (The Hague 1631) is rather disorderly: II.1.47 & III.1.12 offer a broad sense. But in I.4.1 he merely distinguishes between ‘*mondigen*’ [with a voice] and ‘*onmondigen*’. The first are defined narrowly as ‘*qui personam habent standi in iudicio*’. Thomas Hobbes notably inherits Grotius and declares, to paraphrase him, ‘*Veritas non facit legem. Auctoritas facit legem*’.

legislation, such that it is the more basic fact (right) of nature. However, there is not much textual evidence for just how he wants it to be interpreted, and so this remains speculative. Below I dig into a possible source of the term in order to find some indication of how it would have been understood by Grotius's contemporaries, and thereby to suggest how it might be employed by him.

10.8 *Liberum* & αὐτεξούσιος

Grotius was allied with the theologians and statesmen in Holland supporting a doctrine of the radically free will against doctrines of divine predestination. These libertarians include the protestant theologian Jakob Arminius and the Remonstrants Brotherhood that was, and is still, associated with his theological legacy. Against them was Franciscus Gomarus and his Gomarists, called '*contra-remonstranten*' in Dutch, which included the strict Calvinists.

Grotius, although not known to be a member of the Remonstrants, did travel to London in 1613 in order to defend to His Majesty, the King of England, the orthodoxy of the Remonstrants Brotherhood.⁵⁸ Remonstrants differed from orthodox Calvinism on points affirming the free will: conditional (rather than absolute) predestination; universal atonement (anyone can choose God, for God has elected all); the possibility that one can resist divine grace (with the will); and the possibility of relapsing from grace (again with the will).⁵⁹ Their position was finally condemned, with political consequences to follow, at the Synod of Dort in 1618–19.

At the time, it might have been thought that belief in the doctrines of Arminius, along with other supposedly 'Catholic' or Jesuit-friendly teachings, meshed with foreign sympathies.⁶⁰ The Spanish crown had been expelled finally from the northern Low Countries within living memory. As noted above, the Jesuits had had an intellectual love affair with the restoration of the freedom of the Christian, and freedom as a divine good. Lessius was spreading a love of liberty – at the same time as Arminius was teaching – only a few cities south of the border in Antwerp. And Grotius's enemies were all too quick to smell 'popery' anywhere that their doctrines were denied. Moreover, any strong teaching on freedom seemed to lend itself to a Pelagian interpretation of salvation: that man has the power within himself to save himself (albeit only since God had done all the work already through Christ), which Augustine had been at great pains to reject, and which the Calvinists, purportedly following Augustine, condemned at Dort.

From either side Grotius was a proponent of the free will: as a theological claim about man as created by God '*liberum*' and αὐτεξούσιον; and as a residual claim

⁵⁸ Th. Marius van Leeuwen, Keith D. Stanglin, and Marijke Tolsma, eds. *Arminius, Arminianism, and Europe: Jacobus Arminius (1559/60–1609)* (Brill 2009) XVIII.

⁵⁹ 'Remonstrants' (2013) *The Columbia Electronic Encyclopedia* (Columbia UP) <encyclopedia2.thefreedictionary.com/Remonstrants>.

⁶⁰ Th. Marius van Leeuwen et al. (Brill 2009) 84.

about man's current state in the world, remaining *sui juris*, despite the effects of sin. But the most interesting and startling of Grotius's terms still has not been contextualized.

10.9 De libero arbitrio

These debates about the will were waged in terms of *voluntas* and *arbitrium*, with the opponents not always carefully distinguishing the two concepts. Grotius takes the debate up in very different terms, within his treatise on the law of plunder. And it is especially his use of ἀντεξούσιος that interests us. Where does it come from, and why that term?

The answer might be from Jakob Arminius himself. There exists a list of theses from a debate held in Basel between Jakob Arminius, while he was still a student, and Johann Jakob Grynaeus, who was also his teacher. It is not known (to me) whether the terms of the debate were chosen by the proponent or the opponent, something for which more research would need to be done. I have found no evidence that anyone has formerly presented these theses as the clue to a origin of Grotius's use of ἀντεξούσιον:

De libero arbitrio disputatio theologica

Johann Jakob Grynaeus

Basileae, Anno 1583⁶¹

1. The mind shows for the praise of God and our edification, that God indeed is *autexousios*, and a most freely acting agent, but that man is *hupexousios*, such that his liberty is circumscribed by his position and place in the universe (*centro spacioque*), namely by the law of God.
2. To him alone is rightly ascribed *he autexousiotēs*, who since he is supremely good, understands, wills, and makes – immutably and most perfectly – only that which is good; and he is supremely alien to every evil thing. But God alone is of this kind. Therefore, he alone is most perfectly *autexousios*.
3. On the other hand, to him is rightly ascribed *he hupexousiotēs*, who although in the beginning had been established both pure and good (but changeably / *sed mutabiliter*) unto the image of God, such that *ek prohaireseos* (from choice or decision) could obey God if he wished, by a willful disobedience made himself a slave of sin and freed himself from righteousness: but freed by the Son, becomes a slave of righteousness, and freed of sin, by the healed powers of understanding and choice within him. When once he has been perfectly restored (*restitutus*), he shall not be able to sin. Man is this kind of thing, by whose salvation the glory of the grace of God is manifested. He therefore is truly *hupexousios*.
4. ...

⁶¹ Translation made possible with help of Brian Lapsa.

5. A dutiful freedom (*pia libertas*), when it was joined by servitude to justice and righteousness, was gloriously manifested once in the state of creation, when man was first made right, and he was able not to sin, but he was also able to sin. Now, it shines forth in the grace of the one who regenerates by a state in which flesh and sin are mortified, and in man is established *eutaxia* (right ordering) of his powers, and someday *energia* will shine forth perfectly and perpetually in the state of our full redemption (*instauratio*): in this flourishing state, man will not be able, by choice (*ek prohaireseos*), not to obey the Lord his God, and indeed (to do so) perfectly.

Grotius's words were: '*Fecit enim Deus hominem αὐτεξούσιον, liberum sui que juris*'. The accusation against Pelagius by Augustine, and the tradition that followed him, was that he was making God's work in salvation unnecessary. It would seem that Grotius does the same, by applying a category, formerly reserved for God, to man. Such a man has little need of grace in order to do the right thing. This could also explain why Grotius says so very little about the (lasting) effects of (original) sin. Said differently, Grotius merges a doctrine of creation with a doctrine of salvation: man remains more or less as God created him, free, αὐτεξούσιον, and also capable of being *sui juris*.⁶²

10.10 The Limits of Freedom

Yet, one should not be tempted into thinking that Grotian 'natural liberty' is naked liberalism or even antinomianism. But the radical – novel even – claim of *perpetual* natural liberty must be addressed.

Originally one's duties and responsibilities, such as the familial responsibility to keep the property for the next generation, the *patrimonium*, were nothing that could be the subject of contract and could only be alienated from one by, say, moral turpitude (illegal forms of) or dire necessity (war). Grotius does not deny those obligations. But neither does he enumerate them. Moral considerations are notably absent in Grotius's understanding of legal obligation (if understood as perfect obligation). Where they return are in pious advice as to what a Christian *should* do (rather than what he *must* do, say, to avoid sinning). Following the Apostle, all things are permissible, if not all are profitable. As with the Apostle 'all things' for Grotius did not mean 'everything', but 'many things', or anything not strictly forbidden by natural liberty or divine law, particularly the New Testament.

Something libertarian present in Grotius later becomes definitional of modern contract doctrine. This is especially notable relating to his theological liberalism, which stands against both determinism and predestination: 'God created man... free.' Moreover, there is also liberalism in the role and priority of individual persons in Grotius, in the formation of legal orders and the derivation of the powers of the

⁶² See Henk Nellen, Hugo de Groot. *Een leven in strijd om de vrede*.

state from the individual's natural liberty (as natural right).⁶³ The logical conclusions of this will be guarded against by later jurists, since self-slavery is one obvious eventuality under a doctrine of liberty as *dominium*. Modern contract doctrine tends to stand against total property of one's liberty, as the scholastics did, in order to protect that very liberty.

Secondly, 'natural liberty' in its sundry philosophical cognates is almost as old as political philosophy itself. Grotius is correct when saying he is using a 'well-known concept'.⁶⁴ However, placing man in a state of freedom before society – defining it as 'the power of the individual to act in accordance with his own will' – or allowing that men retain natural liberty even after being members of society (meaning, an implicit right of revolt), is new. As Richard Tuck points out: 'In Aquinas, men do not have a *prima facie* natural right to liberty any more than they have a *prima facie* natural right to dominate other men.'⁶⁵ Being born free and possessing Grotian 'natural liberty' are different things altogether.

That men were born free, and what its implications might be for society, were always objects of discussion. A surprising number of ancient sources agree with *in principle* free birth – or they lean in that direction – which implies a (limited) doctrine of natural liberty. None, however, imagine man to enjoy this perpetually, or as an inviolable part of his nature, like possessing a free will. In that case it could always be invoked as a right against society, not only for revolt but also for reform. It is thus no accident that natural rights become just that sort of tool following Grotius.

Free will is a concept that emerges in Augustine's thought, but 'freedom' has been with law for much longer, both as a value and a goal. It is noteworthy that *The Institutes of Justinian* indicate that freedom should be striven for. There is, however, no *principle* of freedom (such as a right) as the starting point or judge of the justice of law. There is, moreover, no extra-legal 'freedom' (*qua* principle) on which the law rests or which it guarantees, even if men are, as it says, 'born free'. A formal *bias* in favour of natural liberty or basic freedom *within the law* can easily be discerned in the *Institutes*, for instance, in its manifold laws against the unnecessary perpetuation of slavery. It also claims to be law 'for persons'. Since those 'persons of law' are not defined within it, logically they must pre-exist the law, and the law guarantees them civil protections that declare:

A freeborn man is one free from his birth...whether both [parents] be free born or both made free, or one made free and the other free born. He is also free born if his mother be free even though his father be a slave, and so also is he whose paternity is uncertain, being the offspring of promiscuous intercourse, but whose mother is free.⁶⁶

Some favour is shown to those who are of 'mixed' birth'. But what is more curious is how the passage continues to hope for freedom, right up to the last minute: 'It is

⁶³ *DIBP* II.XI.

⁶⁴ Grotius, *DIP*, 1:18.

⁶⁵ Tuck (1998) 20.

⁶⁶ *Institutes* (Moyle tr, 5th ed., Oxford, 1913) I,4 'On Free Birth' and following.

enough if the mother be free at the moment of [the child's] birth, though a slave at that of conception'. All the more, the rule even allows that a child be born free if the woman was free at the time of conception, and enslaved before the birth of the child. This is held, 'on the ground that an unborn child ought not to be prejudiced by the mother's misfortune.' We are finally told in this remarkable passage that Marcellus thinks, and Justinian agrees, that 'it is enough if the mother of an unborn infant is free at any moment between conception and delivery' for the offspring to be born free.⁶⁷

Thus, it is not uncommon in legal systems for the prejudice to be on the side of freedom. Yet I know of no assumption of perpetual 'natural liberty', as Grotius sketches it, to be found in the ancient sources or in the Canonists and theologians who contributed to the development of freedom of contract. What are some other implications of the liberty of perpetual natural liberty? It proves to be both the blessing and the bane of the Grotian legacy in modern contract doctrine. For, it also frees man from the moral constraints on his liberty which formerly were thought to be just as natural as the liberty itself now is thought to be.

10.11 Natural Liberty and Conscience

What is often missed about Grotius's concept of natural liberty in legal or political analysis is its distance from any spiritual authority outside of rights exercised in nature. The theologians who incubated the freedom of contract could not imagine contractual obligation operating outside of the bounds of that which involves both regulation of the body and of the soul. In the world of man, law regulated both internal and external fora. Here the 'Protestant' part of Grotius's thought comes to the fore. The spiritual jurisdiction of law over the soul was removed from the Protestant church with the abolition of confession.⁶⁸ With the removal of an external check on the *forum internum*, conscience was gradually 'personalized, privatized and subjectivized'; yet, 'the rules of conscience were originally thought to be almost as objective as legal rules'⁶⁹ At the height of the influence of moral theology, '[a] theologian claiming to be able to solve a case of conscience without the support of the civilian and canon law tradition was considered to be arrogant.'⁷⁰

Grotius's notion that man is created 'free', and especially the appeal to authority in αὐτεξούσιος, includes a spiritual freedom from external authorities (i.e., non-divine, for the Bible is still the word of God) as the condition. If convention has placed (or places) a church or body of law over oneself, which does not contravene

⁶⁷ *Ibid.*

⁶⁸ Decock (2013) 27–28.

⁶⁹ *Ibid* 27–28. Recall that the English Court of Chancery is also called the Court of Conscience, which means little when compared to the modern notion of 'conscience'.

⁷⁰ Decock (2013) 40.

natural right (i.e, protection of life, limb and the things necessary for life⁷¹) and allows natural liberty (self-regulation in practice), then it is to be granted the authority that commands obedience in us. Moral and legal authority begin in ‘he who is the owner of his own liberty as property’. Grotius’s defence of the state’s power of punishment on the private right of defensive war illustrates this ably.⁷² The tradition of increasingly isolating considerations of liberty from considerations of morality in modern contract law and its attendant disciplines, is a faithful adherence to the residual belief that God creates each of us ἀντεξούσιον.

⁷¹ Summary of Grotius’s natural right in *DIBP* II.I.

⁷² *DIBP* II.I.16.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



Index

A

Actions, x, 5, 24, 40, 53, 70, 101, 116, 129, 138, 158
 Agency, ix, 8, 21, 37, 53, 70, 95, 116, 128, 138
 Algorithms, 42, 87
 Angels, 110–113
 Anscombe, G.E.M., ix, 60, 71, 75, 76
 Aquinas, T., 95–113, 157, 166
 Aristotle, 70, 83, 96, 98, 102, 105, 106, 127, 161
 Arrow-Debreu competitive equilibrium model, 77
 Artificial intelligence, 87
 Augustine, 163, 165, 166

B

Bateson, G., 116
 Beck, U., 139
 Becker, G.S., 40
 Berger, P., 148
 Bhaskar, R., 22, 27, 28, 138–141, 147
 Blackfriars Hall, Oxford, vii, 66, 70–88, 95–113, 152–168
 Bonaventure, G.F., 95, 100
 Bourdieu, P., 116, 117, 121, 122, 145, 146
 Buchanan, J., 40, 47, 48

C

Causal explanations, 4, 53–66
 Causation, ix, x, 3, 4, 13, 14, 24, 31, 32, 55–66, 70–88, 125–135

Cause and effect, 56, 70, 74, 75
 Changes, 21–34, 51, 55, 57–59, 72, 76, 78–81, 84, 85, 102, 116, 117, 119, 121–123, 125, 138–145, 147–149, 152, 158, 159, 162
 Charity, 103–105, 107
 Chemistry, 64, 76, 77, 102, 108, 113
 Choices, x, 3, 21, 38, 54, 77, 116, 156
 Climate change, 138, 149
 Commodities, 45, 50, 77, 126–135
 Condorcet, M. de, 72
 Consciousness, 3, 6, 7, 9–13, 15, 31, 97, 99, 116, 121–123
 Contract, 47, 118, 152–168
 “Contribution” (Karl Marx), 125, 126
 Corporeality, 76
 Critical realism, 138–150
 Cushman, P., 45

D

Davidson, D., 4, 40, 43, 56–59, 62, 63, 70, 73, 82, 83
 Democritus, 70, 71
 Demons, 110–113
 Descartes, R., 4
 Desires, 7, 10, 12, 22, 24, 51, 56, 59, 71, 73, 75, 81, 86, 98, 99, 106, 110, 120–123, 156, 157
 Determinism, 7–9, 12, 13, 70, 87, 108, 156, 161, 165
 Durkheim, E., 116

E

Economic anthropology, 39, 116–124
 Economic behaviour, 85, 95–113
 Economic explanations, 22, 53–55, 62, 63
 Economics, vii, ix, 22, 37, 53, 70, 99, 128, 145, 153
 Economics as a moral science, vii, 83, 86
 Edgeworth, F., 72
 Elder-Vass, D., 139
 Enlightenment, x, 72
 Epicurus, 70
 Epiphenomenalism, 3–16
 Equilibrium, 24, 71, 77, 84, 152
 Ethnography, 49, 50, 116–122
 Evens, T., 116, 117, 122

F

Fetishism, 130–133
 Final causality, 102, 112
 Flourishing, 33, 102, 165
 Free choices, 14, 98, 102, 104
 Freedom, 4, 7, 8, 10, 11, 13–16, 22, 33, 95, 101, 106, 112, 116, 152–158, 160, 161, 163, 165–167
 Freedom of contract, 152–159, 167
 Free will, x, 3–16, 95, 98, 153, 156, 157, 163, 166
 Friedman, M., 41, 44, 48, 80

G

Giddens, A., 139, 143, 146
 God, 86, 100, 102, 104–107, 109, 111–113, 152–168
 Goodman, N., 78
 Graces, 100, 105, 107, 109, 110, 112, 163–165
 Grotius, H., 152–168
 Grynaeus, J.J., 164

H

Harré, R., 27, 86, 139
 Hausman, D.M., 38, 41
 Hendry, R.F., 25
 Hermeneutics, 38, 42–46, 48–52, 55
 Hicks, J., 74
 Hobbes, T., 4, 76, 87, 127, 160, 162
Homo oeconomicus, x, 85
 Human behavior, 40, 42
 Human beings, 29, 31, 37, 42, 43, 49, 51, 62, 63, 70, 81, 95–97, 100–102, 104, 108, 155

Hume, D., 25, 27, 28, 42, 70, 71, 75, 76, 127
 Hyman, J., 81

I

Ideology, 37–52
 INUS condition theory, 70, 82

J

Jevons, W.S., 39, 72, 133

K

Kant, I, ix, x
 Keynes, J.M., 76, 77, 79, 81, 82, 84
 King, M.L., 86, 112

L

Labour, 80, 118–120, 123, 125, 128–134
 Labour theory of value (LTV), 125, 128–130
 Las Casas Institute, vii, 66
 Law-like generalisations, 70, 76
 Lawson, T., 22, 23, 26, 28, 29, 31–33
 Levinas, E., 117
Liberum arbitrium, 95, 98, 100
 Libet, B., 4, 5, 11, 12, 14, 16, 101
 Locke, J., 39
 Lord Robbins, L., 77
 Lukács, G., 131, 132

M

Mainstream economics, vii, 22–24, 26, 32, 38, 39
 Mäki, U., 80
 Mallinckrodt Foundation, vii
 Mandela, N., 112
 Mankiw, N.G., 74
 Marginal utility, 71, 73
 Marshall, A., 77
 Marx, K., 72, 84, 125–135
 Meikle, S., 125–135
 Mendeleyev's table, 74
 Mill, J.S., 77, 80, 84, 85, 130
 Modelling, 23–26, 34, 77, 79
 Moral autonomy, 70, 74, 86, 87
 Morality, ix, 106, 116–124, 152, 168
 Moral responsibility, 4, 11, 12, 74, 88, 116, 155

N

Nagel, T., 6
 Naturalism, 37, 38, 44, 46, 48, 50, 51, 138, 147
 Natural law, 75, 104, 109, 135, 153, 159, 161
 Natural liberty, 152, 156, 160, 161, 163, 165–168
 Natural sciences, ix, x, 28, 37, 41–43, 55, 62–66, 71–76, 78, 82–84, 87, 138
 Neoliberalism, 44, 46–48
 Newtonian physics, 74, 87

O

Ontology, x, 22, 26–31, 34, 62–64, 74, 78, 81–84, 139–141, 143, 148
 Orwell, G., 134, 135

P

Pareto efficiency, 71
 Physicalism, 4, 16
 Plato, 84, 161
 Popper, K., 80, 81
 Practical Social Theory (PST), 140, 141
 Predestination, 156, 163, 165

R

Rational choice theory, x, 37–52
 Rationality, 26, 40, 46, 54, 78, 103, 116, 148
 Reasons, x, 5, 24, 38, 55, 78, 97, 119, 141, 155
 Responsiveness, 11–12
 Ricardo, D., 84, 125, 127–134
 Robinson, J., 162
 Roman law, 130, 154, 158, 159
 Rule-based roles, 85–87
 Russell, B., 70, 146
 Ryle, G., 60, 122

S

Samuelson, P., 74, 77
 Scotus, D., 95
 Searle, J., 73, 74, 85, 128, 146
 Smith, A., 22, 39, 85

Social anthropology, 116, 117
 Social atomism, 30
 Social holism, 29, 30
 Social morphogenesis, 143
 Social sciences, ix, x, 33, 38–40, 42, 50, 55, 62–66, 70, 72–74, 76, 81, 138, 148
 Society, 22, 26, 30, 33, 40, 46, 47, 49–51, 77, 85, 87, 104–107, 109, 110, 112, 118, 120, 122, 125, 134, 139, 141, 143, 146, 147, 166
 Socrates, 85
 Soul, 13–15, 86, 96, 154–156
 Structure, Agency and Culture (SAC), 138, 139, 141, 147, 149, 150

T

Taylor, C., 42, 43, 45
Telos, 81, 87, 152, 154
 Theological anthropology, 152–168
 Theology, 86, 152–156, 167
 Theory of commodities and money (TCM), 126–128, 130–133
 Times, 4, 23, 25, 78, 84, 97, 98, 103, 104, 108, 111, 116–125, 127, 129, 133, 134, 141, 144–146, 148, 149, 155, 157, 158, 160, 162, 167

V

Values, 7, 12, 15, 26, 39, 41, 44–46, 48, 70, 77, 80, 87, 101, 112, 118, 123–131, 133, 147, 160, 166
 Veblen, T., 24
 Virtues, x, 5, 9–11, 15, 27, 31, 86, 99, 100, 105–108, 128, 152, 154, 156, 159
Voluntas, 97, 164
 von Mises, L., 74
 von Wright, G.H., 73

W

Walras, L., 78, 79, 82, 84, 133
 Weber, M., 116
 Wigner, E., 72
 Wittgenstein, L., 127