ERASMUS JOURNAL FOR PHILOSOPHY AND ECONOMICS

VOLUME 1, ISSUE 1, AUTUMN 2008

The Erasmus Journal for Philosophy and Economics (EJPE) is a peer-reviewed bi-annual online academic journal run by the graduate students of the Erasmus Institute for Philosophy and Economics, Erasmus University Rotterdam. The Journal provides a forum for research across three domains: methodology of economics, history of economic thought, and the conceptual analysis of inter-disciplinary work relating economics to other fields. For additional information, see our website: <http://ejpe.org>. All submissions should be sent via e-mail to: <editors@ejpe.org>

PEER REVIEW

EJPE WOULD LIKE TO THANK THE REFEREES WHO ASSISTED IN THE PRESENT ISSUE:

Sonja Amadae, Kenneth L. Avio, Chris Berry, Mark Blaug, Hans Blom, François Claveau, Till Düppe, Joshua Graehl, D. Wade Hands, Frank Hindriks, Clemens Hirsch, Geoffrey Hodgson, Shashi Kant, Matthias Klaes, Bart Leeuwenburgh, Craig McLaren, Michiru Nagatsu, Malcolm Rutherford, Ana Cordeiro dos Santos, Esther-Mirjam Sent, André van Hoorn, Jack Vromen.

 ISSN: 1876-9098

ERASMUS JOURNAL FOR PHILOSOPHY AND ECONOMICS
VOLUME 1, ISSUE 1, AUTUMN 2008

TABLE OF CONTENTS

PHD THESIS SUMMARIES

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. iv-x.
http://ejpe.org/pdf/1-1-intro.pdf

Welcome to the inaugural issue of the EJPE

According to John Stuart Mill, "the definition of a science has almost invariably not preceded, but followed, the creation of the science itself" (Mill 1844, 86). The subject matter, the methods, the main topics and questions, the foremost goals and challenges, all such elements need to be properly understood before a suitable definition of the discipline can be established. Mill was referring to economics. Yet, today there is still no full agreement on what exactly economics is, what it is supposed to be about, nor on the methods that it should employ.

Since the early nineteenth century, such reflections on the nature of economics—its definition, subject matter, methodology, logic, epistemic and empirical basis, normative implications, as well as its relations to other fields—have slowly given rise to the new field of philosophy of economics. However, this development has been rather sporadic, possibly due to the disciplinary structures and boundaries that have, for example, limited enthusiasm for economic methodology among philosophers of science, and isolated historians of economic thought from their colleagues in economics departments interested in the relevance of the past for a better understanding of contemporary economic theory. Indeed, significant full-fledged attempts to organise, systematise, and recognise philosophy of economics as a distinct discipline only started some 30 years ago.[1]

Nonetheless in recent decades various international institutions and academic journals have begun to appear, including the International Network of Economic Method (INEM) with its specialised and highly regarded *Journal of Economic Methodology* and the prestigious and more general *Economics and Philosophy*.[2] These organisations have mainly served to shape and develop the exchange of ideas among people who were already knowledgeable and active in the philosophy of economics, thus bringing about a mostly highly specialised academic exchange.

[1] Some representative examples of this stage are Rosenberg 1976; Blaug 1980; Hausman 1981; Caldwell 1982; Boland 1982, 1989; McCloskey 1985; Dow 1985; Hoover 1988; Mirowski 1988. For more recent overviews see Backhouse 1994; Hands 2001; and Hausman 2008.

[2] Many more philosophy and economics resources can be found in the 'Related Links' section of the EJPE Website: <http://ejpe.org/related-links/>

This stage was followed in 1997 by the founding of the successful Erasmus Institute for Philosophy and Economics (EIPE) by a small international group of academics. At the time, EIPE was the only graduate institute entirely devoted to the training and research in the philosophy of economics. Since then, EIPE has played a significant role in the increasingly integrated state of the philosophy of economics.

In spite of lacking a precise definition for economics, it seems fair to say that almost everybody has at least an informal idea of what economics is all about: people in general have some notion about what economists might actually be doing. Things are entirely different for the philosophy of economics. As graduate students in the field, one of our aims with starting this new *Erasmus Journal for Philosophy and Economics* is to change all that. We want to spread the recognition and influence of the topics and the research being carried out in the philosophy of economics, and to persuade more academics from different disciplines to read and publish on crucial topics in the field.

The overall purpose of EJPE is therefore to provide a forum for the growing scholarly research that lies at the intersection of philosophy and economics. This intersection includes not only the philosophy of science applied to the special case of economics, but also research in the history of economic analysis which contributes to a better understanding of the contexts, ideologies, and culture behind the development of economic theory. Similarly, we want to include conceptual reflections on inter-disciplinary relations between economics and other disciplines, which have the capacity to enrich our understanding of all fields involved. EJPE's emphasis lies on publishing outstanding and original research, while also supporting the expansion and integration of the field by publishing critical survey papers by well regarded experts.

As a graduate student-run journal we recognise that EJPE has to demonstrate its academic credentials from the very beginning. Among our top priorities is depth in content (important, original, and rigorous research) supported by a formal and authoritative peer-review process. On this note, we are delighted that so many well-established academics have given generously of their time and expertise to assist in the selection and shaping of our submissions, and we have been no less impressed by the commitment and thoroughness of our young scholar referees.

In the same vein, we would like to stress some distinguishing features that we believe set EJPE apart from many other academic journals in a significant way. EJPE is particularly committed to supporting and encouraging young scholars (both graduate students and recent PhDs). This objective is built into the very structure of EJPE in a number of ways, including content, style, and participation. The breadth of our content is meant to be particularly relevant for young scholars who have not yet specialised in their research efforts and who may have bold inter-disciplinary perspectives. EJPE also contains a special section in which recent PhDs can publish short summaries of their theses in order to introduce their work to a wide and diverse audience, and to promote interaction among researchers working on related topics.

We have tried to give EJPE an open and supportive style that young scholar contributors will find particularly welcoming, for example, in designing our submission and peer-review processes so as to emphasise timely and constructive feedback to authors at every stage. We are committed to maintaining an efficient peer-review system that provides authors with initial decisions within 2 to 3 months, and this means that our publishing cycle is much faster than the norm. In addition, the journal encourages young scholars with relevant expertise to gain experience within the academic world by participating first-hand as referees and book reviewers. We are also proud that all EJPE issues will be free to access online, as a young-scholar friendly resource to everybody interested in the subject.

EJPE will also contribute by informing and raising critical debate among circles that are not yet so acquainted with philosophical discussions about economics. We hope thereby to support the development of interdisciplinary relations and conversations, not only directly between philosophers and economists, but also among a broader range of young practitioners and theorists from all existing schools of economics, the humanities, and social sciences.

EJPE also makes a concerted effort to follow important ongoing debates in philosophy and economics by commissioning expert articles and book reviews. In this issue, for instance, we took a closer look at the debate over Stephen Ziliak and Deirdre McCloskey's controversial work on statistical significance by inviting a book-review from Aris Spanos, and a reply to Spanos's review from Ziliak and McCloskey.

The initial response to our project has been impressive and encouraging. We are delighted to have received a large number of submissions for this inaugural issue from both well-established academics and young scholars all around the world, confirming that there is a great deal of interest in what this journal is meant to offer. In addition to several book reviews and PhD thesis summaries, this inaugural issue of EJPE contains five articles and one interview.

Opening the issue, Alessandro Lanteri examines the charge that students of economics are more selfish than students of other disciplines, as has been suggested by various economic experiments. Lanteri then explores in detail the alternative suggestion that economists are naturally selfish before their training begins and hence 'self-select' into studying economics. According to Lanteri, the self-selection explanation has been so readily accepted by economists because it requires no real self-examination of their teaching methods and contents; thus little effort has been made to properly corroborate it, while the roles of other plausible mechanisms, such as 'framing', have been neglected.

Next, Maurice Lagueux draws on the recent work of philosopher-economist Don Ross on microexplanation, and questions whether or not such work is contributing to a revolution in the methodology of economics by challenging the central pillars of the discipline: methodological individualism and the concept of rationality. More than an inquiry into the revolutionary status of economic methodology, however, Lagueux provides an in depth overview and analysis of the philosophical ideas that Ross has introduced to the economics literature, including Daniel Dennett's 'intentional stance'. Following Lagueux's article, in a short response, Don Ross himself elaborates on his points of contention with Lagueux's analysis and conclusions.

Following the pioneering work of Wade Hands and others, David Tyfield inquires into the potential for establishing an economics of scientific knowledge (ESK) from the sociology of scientific knowledge (SSK) literature. While acknowledging SSK's well-known problem of reflexivity, Tyfield argues that there are other, and more serious, philosophical problems with SSK that need to be rectified prior to developing an SSK-based ESK, requiring the introduction of a 'critical philosophy'.

In 'Bernard Mandeville and the 'economy' of the Dutch' Alexander Bick explores Mandeville's thoughts as elucidated in 'Remark Q' of *The*

Fable of the Bees. While historians of economics have traditionally focused on those elements of Mandeville's thought that figured prominently in the development of Adam Smith's thinking, Bick's paper sheds new light on how the development of political economy was informed by close examinations of actual economic practice: Bick explores Mandeville's first-hand account of England's economic possibilities in relation to the commercial experiences of the Dutch.

Cristina Marcuzzo's article is an invited adaptation of her presidential address to the 2007 annual conference of the European Society for the History of Economic Thought (ESHET), in which she considers the past, present, and future of the history of economic thought (HET). She describes a typology of four different techniques of HET and evaluates the roles they have played in the discipline; she also rejects the split between the good economist and the good historian, in relation to the required skills for a suitable HET. The demands of the subject require good historians of economics to be well-versed in both disciplines and able to toggle between deep context and (contemporary) economic theoretical frameworks.

As a special contribution to this inaugural issue of EJPE, we present the first of an envisaged series of interviews with well-established philosophers and economists. Uskali Mäki has been among the most important proponents, researchers, and institution builders of the discipline of philosophy of economics ever since its emergence three decades ago. Today he continues to contribute new and interesting projects to the field. In this interview, he offers his opinions on the current state of the philosophy and methodology of economics, a first-hand overview of the development of his own thought, as well as some detailed clarifications of his current philosophical ideas.

As we have quickly learned, editing an academic journal involves a variety of challenges and rewards. There is the challenge of juggling opposing interests pulling in quite different directions: to publish high quality original research, while simultaneously ensuring a diversity of contributions across the journal's domains; to steer a reasonable path between opposing but well-argued referee reports; and so on. The first time an editor has to read a submission and decide whether it is a candidate for publication, they quickly realise the weight of the responsibility. But these real challenges are coupled with many rewards, from learning how certain aspects of academia work "behind the scenes" and developing valuable relationships with colleagues, to the sense of achievement derived from guiding a submission all the way through the peer-review process into a published issue of EJPE. We hope that we, as editors, have lived up to the expectations and responsibilities entrusted to us by our colleagues and friends in the academic community.

We are grateful to the Department of Philosophy at Erasmus University Rotterdam for generous funding, and to the many members and friends of EIPE who have provided advice and support. And in addition to the authors who made this issue possible, we would like to extend our thanks to the referees and the EJPE Advisory and Executive Boards for helping to transform EJPE from a lofty idea to a reality. We hope you all enjoy the outcome.

TYLER DESROCHES
LUIS MIRELES-FLORES
THOMAS WELLS

The EJPE Editors
<editors@ejpe.org>

REFERENCES

Backhouse, Roger E. (ed.). 1994. *New directions in economic methodology*. London: Routledge.

Blaug, Mark. 1980. *The methodology of economics: or how economists explain*. Cambridge: Cambridge University Press.

Boland, Lawrence A. 1982. *The foundations of economic method*. London: Allen & Unwin.

Boland, Lawrence A. 1989. *The methodology of economic model building: methodology after Samuelson*. London: Routledge.

Caldwell, Bruce J. 1982. *Beyond positivism: economic methodology in the twentieth century*. London: Allen & Unwin.

Dow, Sheila C. 1985. *Macroeconomic thought: a methodological approach*. Oxford: Basil Blackwell.

Hands, D. Wade. 2001. *Reflection without rules: economic methodology and contemporary science theory*. Cambridge: Cambridge University Press.

Hausman, Daniel. 1981. *Capital, profits, and prices: an essay in the philosophy of economics*. New York: Columbia University Press.

Hausman, Daniel M. 2008. Philosophy of Economics. The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), ed. Edward N. Zalta. SEP Website. http://plato.stanford.edu/archives/fall2008/entries/economics/ (accessed November 2008).

Hoover, Kevin D. 1988. *The new classical macroeconomics: a sceptical inquiry*. Oxford: Basil Blackwell.
McCloskey, D. N. 1985. *The rhetoric of economics*. Madison: University of Wisconsin Press.
Mill, John Stuart. 1844 [1836]. On the definition of political economy, and on the method of investigation proper to it. In *Essays on some unsettled questions of political economy*, J. S. Mill. London: Batoche Books, 86-114.
Mirowski, Philip. 1988. *Against mechanism: protecting economics from science*. Totowa (NJ): Rowman & Littlefield Publishers.
Rosenberg, Alexander. 1976. *Microeconomic laws: a philosophical analysis*. Pittsburgh: University of Pittsburgh Press.

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 1-23.
http://ejpe.org/pdf/1-1-art-1.pdf

(Why) do selfish people self-select in economics?

ALESSANDRO LANTERI
University of Piemonte Orientale

Abstract: Several game-theoretical lab experiments helped establish the belief that economists are more selfish than non-economists. Since differences in behaviour between experiment participants who are students of economics and those who are not may be observed among junior students as well, it is nowadays widely believed that the origin of the greater selfishness is not the training they undergo, but self-selection. In other words, selfish people voluntarily enrol in economics. Yet, I argue that such explanation is unsatisfactory for several reasons. I also suggest alternative explanations for the observed differences, which have been so far unduly disregarded.

Keywords: economics, experiments, moral trial, self-interest, self-selection

JEL Classification: A11, A13, A14, C70, C90

Do economists make bad citizens? (R. Frank, et al. 1996). Are political economists selfish and indoctrinated? (Frey and Meier 2003). How tempting is corruption? More bad news about economists (B. Frank and Schulze 1998). With so many obvious value judgments, these are questions one would not expect to read in an economics article and, with such a dismal depiction of their profession, one would expect economists to answer negatively. Both expectations would be wrong.

These unflattering questions are but a small sample of the titles of economics articles from a stream of literature that has waged against the economics profession a veritable 'moral trial' (Lanteri 2008a). Moreover, although some attempts have been made at defending economists (Yezer, et al. 1996; Laband and Beil 1999; Lanteri 2008a, 2008b; Hu and Liu 2003; Zsolnai 2003), affirmative answers to these questions prevail by far, both among economists and non-economists.

AUTHOR'S NOTE: I am grateful to two anonymous referees, and to Tom Wells, for very useful comments. Remaining mistakes are my responsibility.

To be under these types of attacks may not be altogether surprising for the practitioners of a discipline which, since its origin, has been criticized for countless reasons (Coleman 2004): because it is false, useless, or harmful, because its practice is conceited, biased, bidden, or methodologically inadequate, because its subject is overstretched in scope or overemphasized in value, and so forth. Such longstanding hostility is perhaps a reflection of economists having different opinions from the rest of the people, not only when it comes to strictly economic matters (McCloskey 1990; Caplan 2001; Rubin 2003; Klamer 2007), but also political (Kearl, et al. 1979; Fuller and Geide-Stevenson 2007) and moral ones (Frey, et al. 1993; Haucap and Just 2003).

Yet, where do these differences come from?

Quite simply, those who have faith in the overarching power of self-interest may gravitate towards economics so that, as Steven Rhoads (1985, 162) wrote in *The economist's view of the world*, the "[p]eople who think […] narrow self-interest makes sense are more likely to become economists". In other words, there is a process of *self-selection* into the discipline. On the other hand, *training* in economics may modify the view of the world of its students. Therefore, George Stigler (1959, 528) suggested that the origin of the differences "surely lies in the effect of scientific training the economist receives", because the typical economics student "is drilled in the problems of all economic systems and in the methods by which a price system solves these problems!"

The differences in economic, political, and moral opinions seem indeed to reflect the central features of economic theory, according to which individual agents behave as self-interested and rational 'economic men' and their voluntary interactions in free markets produce an optimal final state. Though this simplistic vision of the world is arguably no longer part of the mainstream, it is the kind of economic knowledge that is regularly fed to undergraduate students (Colander 2007).

Such an 'economist's worldview', moreover, may have consequences that reach farther than personal opinions. It has been shown in several lab experiments (Marwell and Ames 1981; Carter and Irons 1991; Frank, et al. 1993; Rubinstein 2006) that economics students behave in accordance with the predictions of economic theory—i.e., selfishly—whereas non-economics students behave contrary to those predictions. As a reaction to such evidence, it has become common to accept the claim that "economists are more selfish than other persons" as "a fact beyond doubt" (Frey and Meier 2002, 2). This literature, therefore, moves beyond the traditional critiques of the discipline of economics and focuses instead on the individual behaviour of economists in an attempt to uncover their morals.

On the whole, this literature can be regarded as a moral trial against economists on charges of selfishness. The trial is *moral* because of two unusual features of this literature compared with most economics articles: a clear focus on the personal character and behaviour of individual economists with respect to a normative moral standard; and the abundance of value judgments, which reflects a level of emotional involvement that surpasses the typical intellectual curiosity of a scientific inquiry.[1] What makes the moral trial a *trial*, instead, are the prescriptions for correcting this deviant moral conduct.

Yet, corrections may only be advocated and enforced after some clarity has been shed on the origin of the misdeed. For instance, one could suggest that the content or the method of economics teaching should be changed, but this would only be effective if economics teaching had an effect on economists' selfish conduct. If one subscribes to the self-selection explanation, on the other hand, this is no longer the case. Self-selection may be a moral accusation, to be sure, but it is not one that economics teachers are accountable for. If selfish people converge to the discipline, what can economists do?

Within this moral trial, as I collectively refer to the contributions to the literature on the morality of economists, the training explanation for the origins of the difference in selfishness has received very little support and most researchers have found indications of self-selection, which is therefore accepted as the leading interpretation for the evidence.

In the following, however, I will address two orders of problems: whether we can conclude that economists are more selfish than non-economists and, to the extent that this is the case, whether the difference is explained by self-selection. With regard to the first problem, I will argue that the experimental evidence is not strongly conclusive, so the difference between economists and non-economists is

[1] One criterion to define a moral issue versus a non-moral one is emotional involvement. If we disagree about whether red wine goes with fish, the extent of our passion in defending our opinion against the opposing view is most likely milder than the passion involved in a disagreement over paedophilia. The latter is thus a moral issue, because it passes a critical threshold on an "emotional staircase" (Blackburn 1998, 9ff.). This seems the case of the moral trial as well; Ariel Rubinstein for example thanks the many economists who confirmed that his work "hit a nerve" (2006, c1n.).

not as sharp as it is sometimes presented to be. With regard to self-selection, I will complain that it is poorly specified and that it does not truly amount to a satisfactory explanation. Moreover, other explanations for the observed differences in behaviour, which might deepen our understanding of the conduct of economics majors and non-majors, have been unduly disregarded.

THE MORAL TRIAL

It was in the early Eighties that the first lab experiments were published which pointed out a behavioural difference between economists and non-economists. In a study of the private provision of public goods, the psychologists Gerald Marwell and Ruth Ames (1981) endowed their subjects with some tokens. Each participant could put her tokens into either an individual investment (a private good) or a collective investment (a public good). For every token invested in the private good, each participant received a small amount of money. Each token invested in the public good was awarded a larger amount of money, which was pooled and then equally shared among all participants—including those who had not invested in the public good. Economic theory predicts that nobody voluntarily contributes to a public good—i.e., everybody free rides—in the hope that they may still reap the fruits of the collective investment, should other participants so allocate their tokens. This is regrettable because, thanks to the higher compensation, the social optimum obtains when every participant puts all her tokens in the public good. Marwell and Ames show that various samples of students put roughly 49% of their endowed tokens into the collective investment: behaviour that is far from being collectively optimal, but also far from the predictions of economic theory. On the other hand, graduate students of economics only invested 24% of their tokens in the public good: not the strong free-riding pattern that economists would predict, but much closer to that.

This line of inquiry was extended ten years later, when John Carter and Michael Irons (1991) conducted an ultimatum bargaining game among randomly recruited freshmen and senior students, majoring in either economics or non-economics (and none of whom had ever enrolled in or taken graduate economics courses). They, too, confirmed that “a behavioral difference [between economists and non-economists] does exist” (Carter and Irons 1991, 171).

The ultimatum game consists of dividing a fixed amount of money among two players: a proposer, who makes the division, and a responder, who may accept or reject the division offered. If the responder accepts, the division takes place as proposed. If she rejects, both players earn zero. Economic theory would suggest that the proposer keep as much as possible for himself, say 99.99%, and offer a mere 0.01% to the responder. The responder, on the other hand, should accept even that minuscule share, as it is better than nothing. Carter and Irons, therefore, asked the participants to state both the minimum amount they would find acceptable if it were offered to them by another player and the amount they would offer to another player. On average, non-economics students declared that they would accept 24.4% or more of the original sum and that they would keep 54.4%. Economics students, on the other hand, would keep 61.5% of the initial endowment for themselves and would be happy to be offered a mere 17%. Once again, the evidence reveals conduct that places economics students closer than the others to the standards set by economic man. The second important finding is that this difference can already be observed between freshmen, who had not had enough time to be indoctrinated by economics teachers. This is interpreted to mean that the phenomenon must be explained by self-selection. In short: "[e]conomists are born, not made" (Carter and Irons 1991, 174).

In another famous experiment, Robert Frank, Thomas Gilovich, and Dennis Regan (1993) assembled groups of three students. Each participant played two simultaneous prisoner's dilemmas with each of the other participants. In such situations, each individual faces a decision in which one choice (i.e., defection) yields a higher payoff (i.e., it is a dominant strategy), regardless of the choice made by the other player. If both players make that choice, however, each participant achieves a poorer outcome than they would have if everybody had chosen otherwise (i.e., to cooperate). Although cooperation is advantageous for both parties, economic theory is clear that every rational agent will defect in a one-shot prisoner's dilemma.

In this experiment, the defection rates of economics students were 60.4%, compared with 38.8% for non-economics students. Economics students, as usual, get closer to the predictions of economic theory. Since, yet again, the differences are present even when comparing junior economics and non-economics students, there must be some self-selection at play. Frank and his colleagues, however, also show that the longer the participants had been trained in economics, the more they expected other people to be dishonest and therefore, probably, to defect. This fits together with the further observation that, whereas the progress of non-economics education is correlated with an increase in cooperative behaviour, the same pattern is not observed among economics students, suggesting that "the training in economics plays some causal role in the lower observed cooperation rates of economists" (Frank, et al. 1993, 168). This causal role, however, is not identified or explored elsewhere in the literature.

Before assessing the merits of the self-selection explanation, however, let me turn to some remarks on the kind of evidence offered by these experiments.

SOME METHODOLOGICAL OBSERVATIONS

In reaction to the moral accusations lingering around this experimental evidence, many scholars began producing counterevidence. For example, the results of several field experiments showed that more students of economics than students of other disciplines who found abandoned envelopes containing dollar bills, returned them to the owners instead of pocketing the cash (Yezer, et al. 1996); that at the University of Zurich more students of political economy than students of business administration, medicine, or veterinary science contributed to voluntary social funds to support needy students and foreigners (Frey and Meier 2003); and that more practicing economists than political scientists or sociologists declare their real income and pay their professional associations membership fees accordingly (Laband and Beil 1999).[2]

Besides handling these ambiguous results, an assessment of the moral trial has to come to terms with the problem of pinpointing vague concepts like 'selfishness' or even '*more* selfish' than someone else. Most authors seem unconcerned with specifying the charge of selfishness in much detail. Frank and his co-authors (1996, 192) observe that in the debate with Yezer and his co-authors, and also more generally among the contributors to this literature, there are three points of agreement: "that economics training encourages the view that people are motivated primarily by self-interest", since many economists maintain that "the average human being is about 95 percent selfish in the narrow sense of the term" (Tullock 1976, quoted in Frank, et al.

[2] For a more complete review, see Kirchgaessner 2005. For more detailed scrutiny of the moral trial, see Lanteri 2008b.

1993, 159). Moreover, “this view leads people to expect others to defect in social dilemmas” (Frank, et al. 1996, 192). Economics students’ ratings of the likelihood that a businessman would honestly report a mistaken bill (to his disadvantage) dropped after one term of training in microeconomics (Frank, et al. 1993, 168ff.). The third point of agreement is that people who hold such expectations “are overwhelmingly likely to defect themselves” (Frank, et al. 1996, 192). Indeed, “almost all respondents (30 of 31 economics students and 36 of 41 non-economics students) said they would defect if they knew their partner would defect” (Frank, et al. 1993, 167). In other words, the economics students may *behave* selfishly because they anticipate dishonesty. So they may be protecting themselves, in a display of mild self-interest, or they may be animated by a sense of justice in punishing the dishonest co-player, and in such ways come to behave exactly as a selfish person would.

Another interpretation of uncooperative behaviour, different from self-interest, has also been proposed. Marwell and Ames (1981) label the small contributions in their public good experiment as distinctly economic conduct, and explain it by the observation that “the meaning of ‘fairness’ in this context was somewhat alien” to economics students, who were half as likely as non-economics students to say they were “concerned with fairness” and who believed that tiny or zero contributions were fair. They admit that such a difference may be associated with the self-selection of economists “by virtue of their preoccupation with the ‘rational’ allocation of money and goods”, or be the result of their “behaving according to the general tenets of the theories they study” (Marwell and Ames 1981, 309), although they do not discriminate between these explanations. Uncooperative behaviour is thus seen as the outcome of a ‘different understanding’ of fairness, either pre-existing and associated with rationality or acquired through economics indoctrination.

Such interpretations notwithstanding, the subsequent articles simply took observations of uncooperative behaviour as evidence of selfishness. Though their study investigates cooperation in a prisoner’s dilemma and though defection as they characterise it may be ascribed to self-defence or to a sense of justice, Frank and his colleagues (1993, 163) seem to ultimately regard their game as “a rich opportunity to examine self-interested behavior”, and defection rates as a measure of self-interest.

The literature also employs a variety of other criteria for measuring self-interest besides uncooperative behaviour. Anthony Yezer and his colleagues use the return rates of envelopes containing dollar bills as a measure of cooperation, but it has been convincingly argued that their field experiment elicits observations of (dis)honest behaviour rather than (non)cooperation (Zsolnai 2003). Both not cooperating and behaving dishonestly can be seen as manifestations of self-interested conduct, and so can the lower accepted and lower proposed offers of economics students compared with non-economics students in Carter and Irons's ultimatum game (Carter and Irons 1991).

These numerous possible indications of self-interested conduct, however, may blur rather than advance the result of the moral trial. For instance, David Laband and Richard Beil (1999, 86) describe economists as "honest/cooperative", though the two are distinct concepts, and refer to free-riding on professional associations' fees as "cheating", though admittedly cheating may *not* be "the appropriate description for the behavior that [they report]" (99n.). It is indeed remarkable that they concur in such oversimplifications, since they are aware that "the terms 'selfish', 'uncooperative', 'dishonest', and 'cheater' […] are not perfect substitutes" (99n.). They nonetheless follow in the tradition of mixing these terms because they sense that there is a "general shared meaning", so that all the authors involved in the moral trial refer to "the same kind of behavior: selfish *versus* unselfish, cooperative *versus* uncooperative, etc." (99n.).

I doubt that this emphasis on these contrasts provides a clearer concept of self-interest; if anything, it complicates the definition because it suggests that not only does selfishness equal cheating and defection, but also that unselfishness equals honesty and cooperation. That these authors believe it to be clearer, however, makes the important point that, in order to make sense of the experiments, one needs to subscribe to two central conjectures. On the one hand, one must believe that a selfish subject defects in the prisoner's dilemma, makes and accepts small offers in the ultimatum game, free-rides in the provision of a public good, is dishonest when he finds an envelope full of money, and so on. On the other hand, one must also believe that someone who acts in these ways is selfish. Such actions are indeed compatible with selfishness, but it does not follow that they must always be exclusively motivated by selfishness. Though in theoretical research it may be both useful and plausible to assume self-interest as the sole motivation of human behaviour, the assumption is inadequate to address empirical questions about actual behaviour in a range of only slightly similar circumstances.

At any rate, for someone who accepts the two conjectures, experimental evidence gathered under the auspices of the standard economics assumption that individuals behave as *homines economici* may, depending on its interpretation, be used to address two types of questions. The empirical investigation of a descriptive theory examines whether the theory corresponds closely enough to observations, and so it tests whether the theory is accurate. On the other hand, the empirical testing of a normative theory can say nothing about its accuracy, since it can only investigate whether the observed behaviour fits with the standards set by the theory itself.

Therefore, if one considers standard economics as a descriptive account of individual behaviour, one may empirically test whether it is accurate, as Marwell and Ames (1981) did when they conducted their experiments precisely in order to prove that most subjects do not free-ride as economic theory would have them. The other contributors to the moral trial literature, however, have nothing to say about the descriptive accuracy of economic theory. If one instead considers microeconomics as a normative theory of how an individual ought to behave in order to be rational, one could empirically investigate how many subjects fail at behaving as they ought to, and therefore behave irrationally. A comparison of the different degrees of irrationality prevailing among various groups of subjects with different educations might be an interesting academic pursuit, especially if understanding the differences helped with correcting the failures, so that more people behaved in line with rationality and thus with economic theory. God forbid!

As mentioned above, economics and economists are not very much liked. The critics blame economics because it has twisted evil into looking good in a way that is "intellectually acceptable" (Lux 1990, 135) so that immorality now finds its "intellectual and theoretical justification in the name of economics" (Lux 1990, 129), with the consequence that this 'dismal' discipline has really become "pernicious" (Moffat 1878, 5) and it should "be simply swept away" (Henderson 1981, 12). Though not all critics of economics are so severe, there seems to be a presumption that the closer to the standards set by microeconomics, the less ethically one behaves. The moral trial thus seems to take a peculiar twist in considering microeconomics as a normative theory of what an individual should *not* do in order to be moral.[3]

Laszlo Zsolnai (2002, 40-41) seems to endorse this moral-normative interpretation of the moral trial when he mentions these experiments as those “famous studies [showing] that people are *moral beings* in their economic actions”, because they do not behave as economic theory predicts. It is the very same experiments that force Frank, et al. (1996) to respond to the allegations that ‘economists make bad citizens’, because they arguably *do* behave as economic theory predicts, and so on. These justified, but contradicting interpretations of the same experimental result amount to the problem of under-determination of theories by data (Quine 1951), a methodological problem according to which it is virtually impossible to find evidence which demonstrates that a theory is correct. Instead, the most we can strive for would be evidence that could disprove an interpretation, following Karl Popper’s (1969) falsificationism, according to which scientists should begin with a theory or a conjecture and then conduct experiments that could show the starting theories and conjectures wrong. A theory is then provisionally upheld, insofar as it is not falsified (although even falsifying a theory may not be straightforward).[4]

If the starting assumption is that economists behave the way economic theory predicts, however, we must acknowledge that “even economists sometimes fall short of the behaviour expected of all good homines economici” (Carter and Irons 1991, 177). For instance, in their ultimatum game experiment, the average amount kept ($6.15 out of $10) is fifteen standard deviations from the prediction of economic theory ($9.99) and 40% of the economics students offer the 50-50 division. In the standard prisoner’s dilemma, 40% of the economics majors cooperate and, when they are allowed to make a promise to cooperate, the figure rises to 71.4% (non-majors: 74.1%), although economic theory considers such unenforceable promises irrelevant and dismisses them as ‘cheap talk’.

Given that about half of the sample in Carter and Irons’s experiment behaves in plain contradiction with economic theory’s ‘canons of immorality’ and that many subjects deviate from its predictions, if one

[3] I shall on this occasion overlook the numerous problems that arise when equating—as mentioned above—certain observed behaviours with selfishness, selfishness with immorality, and then those behaviours with immorality.

[4] This methodological orientation was rather common among the early experimental economists (e.g., Smith 1982), however, it was later abandoned (Santos 2006, ch. 6).

hypothesised that this is how they behave the evidence of the moral trial should be perhaps best interpreted as falsifying the hypothesis. It would follow that the assumption that individuals—and specifically economists—are immoral like a *homo economicus* should be abandoned. Of course, what matters in the moral trial are not the absolute levels, but the differentials between economists and non-economists. The real origin of these differentials, however, is seldom explored and most authors account for their evidence by essentially suggesting that, unlike the rest, economics students are *homines economici* who self-select in the discipline.

SELF-SELECTION

Although the methodological weaknesses of the moral trial experiments just reviewed already cast some legitimate doubts over the charge that economists are selfish, there are also other specific problems concerning the prescriptions for changing the teaching of economics.

Teaching economics students that non-selfish motivations may play a major role in the conduct of economic transactions (as suggested by Frank, et al. 1996) would not work. Employing less mathematics and more case studies in economics classes (as suggested by Rubinstein 2006) would make no difference. Why? Because self-selection blames the observed differences in behaviour on differences that existed before the students were indoctrinated. The differences are not brought about by economists teaching formal models of rational self-interested choice, so there is no need for change. This conclusion overlooks the possibility that, while standard economics classes may not stimulate selfish behaviour, different teaching methods may mitigate it and so override the original differences and 'correct' economics students into less selfish people. For example, Harvey James and Jeffrey Cohen (2004) showed that students who attended an ethics module had higher rates of cooperation in a prisoner's dilemma experiment than those who did not attend. Yet, one might retort, selfish people would never enrol in such courses.

The alleged ineffectiveness of the proposed solutions, however, invites closer scrutiny of the standing of the self-selection explanation and its exact meaning. It was Marwell and Ames (1981) who first suggested the two most obvious accounts for the observed behavioural differences between economics and non-economics students: learning, which refers to the outcome of economics training, and selection, which refers to a pre-existing individual inclination. As noted above, the observation that the behavioural differences are already present between freshmen, is seen by most authors as sufficient ground to dismiss the training hypothesis and, ipso facto, "support the selection hypothesis" (Frey and Meier, 2003, 452), without allowing for any alternative explanation. Their evidence does challenge the learning hypothesis, but it achieves very little in the way of corroborating self-selection, which is but one of many possible alternatives.

The concept of self-selection itself is far from clearly spelled out. In the experimental literature, self-selection refers to the morally neutral problem that individuals with certain characteristics are especially likely to belong to certain groups, so that these groups are not truly representative of the population at large and therefore constitute a poor sample. As a consequence, the results obtained from such groups are not externally valid, which means that they cannot be extended to the entire population. In the moral trial, the target characteristics are not morally neutral: Carter and Irons (1991, 175) run regressions that substantiate the conclusion that economics majors differ from others not because "they are more skilled at the sort of deductive logic required to recognize and determine opportunities for economic gain", but rather differ "in terms of sentiments". Roughly speaking, the idea is that economics students display uncooperative conduct because "selfish persons choose to study economics" (Frey and Meier, 2003, 448), so that they must also have been selfish individuals before enrolling in economics. None of the experiments, however, attempts to track such inclinations to high school students, which might corroborate the hypothesis.

In the moral trial, self-selection is regarded not as a methodological shortcoming to avoid, but as a phenomenon to exploit. The comparison between the group of economics students and the group of non-economics ones is aimed precisely at 'measuring' the separation between the two groups. Yet these experiments do not attempt to show that economists differ from the general population, thus interpreting self-selection as above, but to show that they differ from specific subgroups. This attempt, however, appears to overlook a logical fallacy: that someone is not like everyone else does not entail that he is different from everyone else.

Perhaps if we had a truly complete pool of subjects—ideally comprising everyone—we would find that they collectively behave just like economists, because those who are more selfish and those who are less selfish than economists average out, so that economists turn out to be a representative sample after all. In the moral trial experiments (with the exception of Frey and Meier's, and Laband and Beil's), the conduct of economics students is nonetheless contrasted with that of very heterogeneous groups of non-economists that are far from being complete. This comparison seems to rest on the quite unlikely simplifications that non-economists are all alike, and that they are all different from economists in roughly the same way. It therefore seems that what should be an hypothesis to test—namely, that economists are different from the population at large—already constitutes a tacit assumption embedded in the setup of these experiments, and a tacit assumption of the kind that causes the methodological problem of theory-ladenness, that a scientist's prior theoretical assumptions affect the observations she elicits (Kuhn 1962). This also clarifies why Frank et al. (1996) complain against Yezer et al. (1996) that the students of biology used as a control group are trained with principles of natural selection founded on self-regarding behaviour that do not distinguish them sufficiently from economics students. They also complain against Marwell and Ames (1981) that the graduate students of economics they target and the high school students they use as a control group differ in several respects and therefore the observed differences in conduct may be caused by other factors, such as gender and age. Beside age and gender, there are other individual factors that remain overlooked in these experiments and the behavioural differences are accounted for only by means of a selective comparison between educational choices.

Given the virtual impossibility of conducting experiments on a complete pool of subjects, a more convincing case could hinge on the proof that economics students have some selfish personality trait, which explains why they behave selfishly, and that this trait is at work across all situations. Such a generalization, however, is contradicted by several findings from social psychology research, in which it is shown that situational factors affect individual behaviour to such a major extent (Darley and Batson 1973; Milgram 1974) that it is implausible that individuals consistently behave in accordance with some fixed personality. For instance, in one famous study (Milgram 1974), 65% of the participants actively murder an associate of the experimenters after he fails to answer some simple questions. Whereas the death was staged, the killing felt real. This dramatic observation, however, is not usually interpreted as meaning that the majority of Americans have a murderous personality, but that—under the rule of certain institutions, in some circumstances—even ordinary people may be pushed to behave in ways that are totally alien to their nature, that even a great evil may ultimately be caused not by the utmost cruelty, but more banally by following to the extreme the rules of one's institutions and situation (e.g., Arendt 1963).

The upshot of these findings is usually that "there is no empirical basis for the existence of character traits" (Harman 1999, 316) and this suggests that they can still be instrumentally employed as a tool for explanatory or predictive purposes (in other words, people behave 'as if' they possess certain character traits) or that they should be eliminated from theory as a misguided illusion. Regardless of how one defines the trait of selfishness, neither of these suggestions would seem to sustain a moral charge against economists. It also seems that the allegations about the type of character traits, with which nature endowed economists and which guide them to Econ–101, should be softened.

On the other hand, it may be conceded (Miller 2003, 381ff.) that there exist 'local character traits', which are activated in connection with *narrowly* defined situations of a certain kind. Someone may be a cheater when it comes to certain school tests but not a cheater all-around (Hartshorne and May 1928), or he could be talkative at lunch, but not on other social occasions (Newcomb 1929). Admittedly, narrowly defined situations that are different along a variety of dimensions may nonetheless elicit the same character trait. Therefore, one may simply show that the conditions encountered in the moral trial experiments are largely similar, if not to all, then to some everyday situations that elicit the trait of selfishness, so that economists' conduct can be generalised to a broader pattern. Alternatively, a narrower, but sufficient and more meaningful, ground for the case at hand would be to argue that the game theoretical experimental setups reproduce the central features of the decision to enrol in economics or in other majors.

This condition is reminiscent of the classical interpretation of game theory, according to which games capture the physical and institutional features of real world situations. In practice, however, this is not what happens (Janssen 1998): a game is *not* a full description of the elements of a situation, but rather a "description of the relevant factors involved in a specific situation as perceived by the players" (Rubinstein 1991, 917).

In the very moment that a game is embedded in a lab, however, it becomes very hard to predict what factors will become salient to each individual player. Norms of fairness, competitiveness, reputation effects, curiosity, intrinsic motivation, and the like, all seem to play a role together with or beyond the nominal payoff, and to do so in a highly idiosyncratic manner. Outside of the lab, in the complex real world, these matters become yet more difficult to capture, so that a direct connection between degree and career choices and conduct in the experiments may be very hard to establish.

The charge that economists are more selfish seems to rest on another tacit assumption: that the choice of different majors is associated with individual differences of some kind and that—conversely—the choice of the same major reflects some personal affinity. Christopher Boone, Woody van Olffen, and Nadine Roijakkers (2004) propose evidence supporting this intuition: different personalities are associated with various degrees of rationality in the choice process of selecting a major, and with the final choice itself.[5] This very evidence, however, also poses a challenge because the four different disciplines included in this study were economics, business administration, business education, and international economics and business studies. If large differences are present among the students who study these disciplines, which on the surface seem to be quite similar, then perhaps there exist even larger differences between them and students in disciplines such as chemistry and fine arts, although such evidence has not yet been produced.

At any rate, choosing a major is but the first step in one's professional life. At a later stage, young graduates must also make a choice between either continuing studying or entering a job, and then among several career opportunities. For example, less than 50% of economics majors continue their education beyond the bachelor, and only about 3% pursue an advanced degree in the same field, while those who do not become non-economists. Therefore, very few graduates call themselves "economists" when they enter a job (Siegfried, et al. 1991,

[5] Following Julian Rotter's studies on personality (1954, 1966), in which students were classified either as having an internal locus of control (i.e., they had confidence in their capacity of affecting the events in their lives) or an external locus (i.e., they considered the events in their lives as driven mainly by forces beyond their control, such as chance or other people). Those with an internal locus were later found to be more likely to have actively searched for information prior to enrolment, and to have chosen study programs leading to more uncertain professional environments (e.g., international business, as opposed to teaching economics).

198). The U. S. Bureau of Labor (2007) estimates that there are 13,000 practitioners of economics presently active in the U. S. (a surprisingly small figure for a country in which over 30,000 students major in the field *every year*). These economists are variously employed in public administration, in politics, in international organizations, in public and private research institutes, in different types of teaching engagements, in consulting firms, in the media (Coats 1981, 1986, 1989; Frey 2000; Mandel 1999), though most economists still consider academia their career of choice (Scott and Siegfried 2002), where they are joined by the many PhDs in economics and econometrics, who originally followed bachelors in 'non-economics' (45% of the total).

The self-selection/training dichotomy, therefore, not only unduly rules out other plausible explanations for the observed behavioural differences between economics majors and non-majors (more on this below), but it also overlooks the heterogeneity of both training and self-selection: high school students self-select into economics majors, are thus trained in economics, later some of them self-select into graduate students and are again trained, and then all self-select into a variety of professions.

At each stage, some (and different) self-selection and learning take place. Neither economics and non-economics students, nor economists proper and non-economists, are tightly isolated and many students who are made into economists at some stage are, so to speak, unmade at a later stage and vice versa. Nonetheless, economics doctoral students contribute more to the University of Zurich social funds than doctoral students of other disciplines, despite the fact that these are the people who have both "absorbed the largest amount of economics teaching" (Frey and Meier 2005, 168) and self-selected the most times. And this observation questions, once again, the charge of selfishness that has been levied against economists.

CONCLUDING REMARKS

Are economists different? Probably. There exist numerous surveys, like the sources I briefly mentioned in the introduction, which expose the differences in moral and political opinions between economists and non-economists. I also reviewed many experimental observations that economists behave differently from non-economists in a stream of literature that puts economists on trial on allegations of selfishness, although it is not always easy to identify precisely either the charge of selfishness or who are the indicted in this moral trial. Such differences can be tracked to two possible and related problems: economists fail at distancing themselves from a normative standard to be avoided and/or economists do not regard such a normative standard as a standard to be avoided. The first problem would seem to indicate that economists are to a large extent morally incompetent. Though such a question could certainly be investigated with existing and well-known tools (Kohlberg 1984; Lind 1987), I am not aware of any findings produced along this line of research. On the other hand, economics students may believe that the decision-context within which they behave selfishly is one in which self-interest is not necessarily blameworthy. Perhaps they are truly born selfish and therefore think all opportunities to earn a buck must be seized, or perhaps they have been indoctrinated to expect others to be unreliable egoists and therefore to be guarded against. Both explanations seem plausible and are probably true of some economists. Yet, there exist other possible accounts.

When a given situation is framed as a market, most people behave more selfishly (Liebermann, et al. 2004), because there is broad acceptance of self-interest in market-like contexts or it may even constitute the social norm to follow (Bicchieri 2006). A candidate explanation for the evidence presented above would thus be that economics students frame decision-contexts differently from students of other disciplines, and specifically in a way that is more in line with the subject they study. Undergraduate economics classes put a strong emphasis on the so-called 'economic way of thinking', according to which each decision is best seen as a trade-off and each choice as a price to pay. It would therefore not be surprising if economics majors framed decision-contexts as markets more often than non-majors, and therefore also behaved selfishly more often, while believing they were simply doing the normal thing for the occasion at hand (Lanteri 2008a).

A related interpretation would be that the application of economic theory in the experiments reflects economics students' perception of the lab tasks as "an IQ test of sorts" (Frank, 1988, 226), in which they ought to apply the theories they had been taught. Such accounts could explain the observed behavioural differences between students of economics and non-economics, but they presume the acquisition of (at least some) economics knowledge.

The training explanation, after all, should not be too easily dismissed. The "logical implications" of Frank and colleagues' three points (1996, 192)—economics promotes cynical views of other people, such views support anticipations of defection from other people, such anticipations encourage defection—"place a heavy burden on those who insist that economics training does not inhibit cooperation". Showing that a difference in behaviour between economics majors and non-majors already existed is not enough to prove that training is irrelevant. After economics training, the observed behavioural differences remain, but they may be caused by different motives. Uncooperative behaviour may be the result of self-interest, but also compliance with social norms, a sense of justice, a lack of concern with fairness, and so on. Perhaps, economics seniors defect in the prisoner's dilemma because they had learned that framing, which they had not known when they were freshmen, and yet they also defected as freshmen, but for a different reason. The lab, field, and natural experiments employed in the literature so far are unfit to discriminate among different motivations for a targeted behaviour. Admittedly, when learning has not yet been possible but the difference can already be observed, as during the early weeks of college, the framing explanation I sketched becomes insufficient. Another plausible account of the early differences, however, can be proposed to side with framing.

Individuals behave in consonance with their identity, which is largely shaped by their current role and by the social expectations that role carries. It seems plausible that freshmen play the way they believe an economist should behave. Such a belief, moreover, probably follows some stereotypical idea of economists. Do such stereotypes exist, and what are they like? These stereotypes do exist, also outside of economics, and they are not very flattering. For example, unless they are given special instructions, students of occupational therapy do not defect very much in a prisoner's dilemma experiment. When they play with a student of economics or when they are asked to play 'as if' they were students of economics, however, they immediately start defecting and they do so slightly *more* than actual economics students (Lanteri and Rizzello 2008). If there were no stereotype of an economist, the students should be puzzled by the instructions, but they are not. They adjust both their decisions and their expectations very quickly. On the other hand, this evidence may hint that students of occupational therapy are not intrinsically less selfish or more moral because they do not defect: it is enough to let them see the situation under a different perspective to critically alter their responses. In other words, there is no need to posit a difference in character dispositions between economics and non-economics majors, but simply examine the combination of the individual's perceptions of her circumstances with her self-image. This way we do not necessarily reject self-selection, but we make our explanations less dependent on it.

Since its inception, the contributors to the moral trial literature have been concerned with discriminating between but two possible explanations: the self-selection of selfish individuals, and the indoctrination of cynical expectations or rational choice. There is, however, no need to stick to this dichotomy or to only one side of it. From high school onwards, there are a plurality of explanations that may capture the observed differences in behaviour between economists and non-economists: some economists may be selfish and self-select into the discipline; upon joining its ranks, some may adjust their decisions to those of the stereotypical economist; systematic exposure to the concepts of self-interest and trade-offs may make those concepts especially salient and therefore more likely to characterise one's framing of a situation; and, over time, the repeated exposure to the focus on material individual incentives may induce the expectation that other people are greedy or the belief that fairness need not be a major concern. These explanations are not mutually exclusive and it may very well be the case that different explanations are appropriate for the behaviour observed in different experimental tasks and for economists of different seniority.

In spite of the broad support for the self-selection explanation, any good description of what it amounts to or of the ways in which it plays out is regrettably lacking. My contention is that there remains ample room for further inquiries. The outcome of such inquiries will hopefully clarify whether the moral trial stands, and so also whether corrections are necessary and possible.

REFERENCES

Arendt, Hanna. 1963. *Eichmann in Jerusalem: a report on the banality of evil*. New York: Viking Press.

Bicchieri, Cristina. 2006. *The grammar of society: the nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Boone Christopher, Woody van Olffen, and Nadine Roijakkers. 2002. Locus of control and study program choice: evidence of personality sorting in educational choice. *Research Memoranda 005*, Maastricht: METEOR.

Caplan, Bruce. 2001. What makes people think like economists? Evidence on economic cognition from the survey of Americans and economists on the economy. *Journal of Law and Economics*, 44 (2): 395-426.

Carter, John, and Michael Irons. 1991. Are economists different, and if so, why? *Journal of Economic Perspectives*, 5 (2): 171-177.

Coats, Alfred. 1981. *Economists in government: an international comparative study*. Durham (NC): Duke University Press.

Coats, Alfred. 1986. *Economists in international agencies: an exploratory study*. New York: Praeger.

Coats, Alfred. 1989. Economic ideas and economists in government: accomplishments and frustrations. In *The spread of economic ideas,* eds. David Colander, and Alfred Coats. Cambridge: Cambridge University Press, 109-118.

Colander, David. 2007. *The stories economists tell: essays on the art of teaching economics.* New York: McGraw-Hill.

Coleman, William. 2004 [2002]. *Economics and its enemies: two centuries of anti-economics*. New York: Palgrave MacMillan.

Darley, John, and Daniel Batson. 1973. 'From Jerusalem to Jericho': a study of situational and dispositional variables in helping behaviour. *Journal of Personality and Social Psychology*, 27 (1): 100-108.

Ehrenberg, Ronald. 1999. The changing distribution of new Ph.D. economists and their employment: implications for the future. *Journal of Economic Perspectives*, 13 (3): 135-138.

Frank, Bjorn, and Gunther Schulze. 1998. How tempting is corruption? More bad news about economists. *Diskussionsbeiträge aus dem Institut für Volkswirtschaftslehre*, Universität Hohenheim, 164/98.

Frank, Robert. 1988. *Passions within Reason*. New York: Norton.

Frank, Robert, Thomas Gilovich, and Dennis Regan. 1993. Does studying economics inhibit cooperation? *Journal of Economic Perspectives*, 7 (2): 159-171.

Frank, Robert, Thomas Gilovich, and Dennis Regan. 1996. Do economists make bad citizens? *Journal of Economic Perspectives*, 10 (1): 187-192.

Frey, Bruno. 2000. Does economics have an effect? Towards an economics of economics. *Wirtschaftspolitik*, 1 (1): S. 5-33.

Frey, Bruno, and Stephen Meier. 2005. Selfish and indoctrinated economists? *European Journal of Law and Economics*, 19 (2): 165-171.

Frey, Bruno, and Stephen Meier. 2000. Political economists are neither selfish nor indoctrinated. *IEW-Working Paper*, No. 69. Institute for Empirical Research in Economics (IEW), Zurich.

Frey, Bruno, and Stephen Meier. 2003. Are political economists selfish and indoctrinated? Evidence from a natural experiment. *Economic Inquiry*, 41(3): 448-462.

Frey, Bruno, Werner Pommerehne, and Beat Gygi. 1993. Economics indoctrination or selection? Some empirical results. *Journal of Economic Education*, 24 (3): 271-281.

Fuller, Dan, and Doris Geide-Stevenson. 2007. Consensus on economic issues: a survey of republicans, democrats, and economists. *Eastern Economic Journal*, 33 (1): 81-94.

Harman, Gilbert. 1999. Moral philosophy meets social psychology: virtue ethics and the fundamental attribution error. *Proceedings of the Aristotelian Society*, 99: 315-331.

Hartshorne, H., and M. May. 1928. *Studies in the nature of character: I. studies in deceit*. New York: Macmillan.

Haucap, Justus, and Tobias Just. 2003. Not guilty? Another look at the nature and nurture of economics students. *University of the Federal Armed Forces Economics Working Paper*, No. 8. Hamburg.

Henderson, Hazel. 1981. *The politics of the solar age: alternatives to economics*. New York: Anchor Press.

Hu, Yung-An, and Day-Yang Liu. 2003. Altruism versus egoism in human behaviour of mixed motives. *American Journal of Economics and Sociology*, 62 (4): 677-705.

James, Harvey, and Jeffrey Cohen. 2004. Does ethics training neutralize the incentives of prisoner's dilemma? *Journal of Business Ethics*, 50 (1): 53-61.

Janssen, Marten. 1998. Individualism and equilibrium coordination games. In *Economics methodology: crossing boundaries. Proceedings of the IEA conference*, eds. R. Backhouse, D. Hausman, U. Mäki, and A. Salanti. London: MacMillan, 1-35.

Kearl, J. R., Clayne Pope, Gordon Whiting, and Larry Wimmer. 1979. A confusion of economists? *American Economic Review*, 69 (2): 28-37.

Kirchgaessner, Gebhard. 2005. (Why) are economists different? *European Journal of Political Economy*, 21 (3): 543-562.

Klamer, Arjo. 2007. *Speaking of economics: how to get in the conversation*. London: Routledge.

Kohlberg, Lawrence. 1984. *The psychology of moral development*. San Francisco: Harper and Row.

Kuhn, Thomas S. 1962. *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Laband, David, and Richard Beil. 1999. Are economists more selfish than other 'social' scientists? *Public Choice*, 100 (1-2): 85-101.

Lanteri, Alessandro. 2008a. *The moral trial: on ethics and economics*. Doctoral Dissertation. Rotterdam: Erasmus Universiteit Rotterdam.

Lanteri, Alessandro. 2008b. Guilty until proven innocent: economists and the moral trial. *mimeo*.

Lanteri, Alessandro, and Salvatore Rizzello. 2008. Ought (only) economists to defect? Stereotypes, identity, and the prisoner's dilemma. *Quaderni SEMeQ* 21/07.

Liberman, Varda, Steven Samuels, and Lee Ross. 2004. The name of the game: predictive power of reputations versus situational labels in determining prisoner's dilemma game moves. *Personality and Social Psychology Bulletin*, 30 (9): 1175-1185.

Lind, G. 1987. Moral competence and education in democratic society. In *Conscience: an interdisciplinary approach*, eds. G. Zecha, and P. Weingartner. Dordrecht: Reidel, 91-122.

Lux, Kenneth. 1990. *Adam Smith's mistake: how a moral philosopher invented economics and ended morality*. Boston: Shambhala Publications.

Mandel, Michael. 1999. Going for the gold: economists as expert witnesses. *Journal of Economic Perspectives*, 13 (2): 113-120.

Marwell, Gerald, and Ruth Ames. 1981. Economists free ride, does anyone else? Experiments on the provision of public goods, IV. *Journal of Public Economics*, 15 (3): 295-310.

McCloskey, Deirdre. 1990. *If you're so smart: the narrative of economics expertise*. Chicago: University of Chicago Press.

Milgram, Stanley. 1974. *Obedience to authority*. San Francisco: Harper & Row.

Miller, Christian. 2003. Social psychology and virtue ethics. *The Journal of Ethics*, 7 (4): 365-392.

Moffat, Robert. 1878. *The economy of consumption: an omitted chapter in political economy*. London: Kegan Paul.

Newcomb, Theodore M. 1929. *The consistency of certain extrovert-introvert behavior patterns in 51 problem boys*. New York: Columbia University College Bureau of Publications.

Popper, Karl R. 1969 [1963]. *Conjectures and Refutations*. London: Routledge & Kegan Paul.

Quine, Willard V. O. 1951. Two dogmas of empiricism. *Philosophical Review*, 60 (1): 20-43.

Rhoads, Steven. 1985. *The economist's view of the world: government, markets and public policy*. Cambridge: Cambridge University Press.

Rotter, Julian. 1954. *Social learning and clinical psychology*. Englewood Cliffs (NJ): Prentice Hall.

Rotter, Julian. 1966. General expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80 (609): 1-28.

Rubin, Paul. 2003. Folk economics. *Southern Economics Journal*, 70 (1): 157-171.

Rubinstein, Ariel. 1991. Comments on the interpretation of game theory. *Econometrica*, 59 (4): 909-924.

Rubinstein, Ariel. 2006. A sceptic's comment on studying economics. *Economic Journal*, 116: c1-c9.

Santos, Ana. 2006. *The social epistemology of experimental economics*. Doctoral Dissertation. Rotterdam: Erasmus Universiteit Rotterdam.

Scott, Charles, and John Siegfried. 2007. American association universal academic questionnaire summary statistics. *American Economic Review, Papers and Proceedings*, 97 (2): 588-591.

Siegfried, John, Robin Bartlett, Lee Hansen, Allen Kelley, Donald McCloskey, and Thomas Tietenberg. 1991. The status and prospect of the economics major. *Journal of Economic Education*, 22 (3): 197-224.

Siegfried, John, and Wendy Stock. 1999. The labor market for new Ph.D. economists. *Journal of Economic Perspectives*, 13 (3): 115-134.

Smith, Vernon. 1976. Experimental economics: induced value theory. *American Economic Review*, 66 (2): 274-279.

Smith, Vernon. 1982. Microeconomic systems as an experimental science. *American Economic Review*, 72 (5): 923-925.

Stigler, George J. 1959. The politics of political economists. *Quarterly Journal of Economics*, 73 (4): 522-532.

U. S. Bureau of Labor Statistics. 2007. Economists. *Occupational Outlook Handbook*, 2006/2007 Edition, at the United States Department of Labor Website. www.bls.gov/oco/ocos055.htm (accessed July 2007).

Yezer, Anthony, Robert Goldfarb, and Paul Poppen. 1996. Does studying economics discourage cooperation? Watch what we do, not what we say or how we play. *Journal of Economic Perspectives*, 10 (1): 177-186.

Zsolnai, Laszlo. 2002. The moral economic man. In *Ethics in the economy: handbook of business ethics*, eds. Laszlo Zsolnai. Bern: Peter Lang Academic Publishers.
Zsolnai, Laszlo. 2003. Honesty versus cooperation. *American Journal of Economics and Sociology*, 62 (4): 707-712.

Alessandro Lanteri is a post-doctoral fellow at the Department of public policy and public choice (POLIS), Faculty of political science, University of Piemonte Orientale (Alessandria, Italy). He holds a MA in economics, from Bocconi University (Milan, Italy); and an MPhil and a PhD in philosophy and economics, from EIPE at Erasmus Universiteit Rotterdam (The Netherlands). His main research interests rest at the meeting points between economics, moral philosophy, and psychology.
Contact e-mail: <alessandro.lanteri@sp.unipmn.it>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 24-55.
http://ejpe.org/pdf/1-1-art-2.pdf

Are we witnessing a revolution in methodology of economics? About Don Ross's recent book on microexplanation

MAURICE LAGUEUX
Université de Montréal

Abstract: The paper aims to assess whether the ideas developed by Don Ross in his recent book *Economic theory and cognitive science: microexplanation*, which relates neoclassical economics to recent developments in cognitive science, might revolutionize the methodology of economics. Since Ross challenges a conception of economics associated with what is pejoratively called "Folk psychology", the paper discusses ideas of the philosopher Daniel Dennett on which this challenge is largely based. This discussion could not avoid bearing on questions such as the nature of consciousness, the interpretation of ontological realism, the relations between agency and selfhood, and the nature and scope of economics. The paper attempts to rehabilitate the two pieces of the traditional conception of economics that were most radically contested by Ross, namely methodological individualism and the foundational role of (human) rationality in economics. A relatively nuanced judgment on Ross's bold enterprise is proposed in conclusion.

Keywords: methodology, intentional stance, consciousness, selfhood, individualism, rationality

JEL Classification: A12, B41, B52, D01, D03

It is remarkable how various contributions in methodology of economics that champion new orientations for this discipline have been published since the beginning of the new century. Recent books like *Modeling rational agents* by Nicola Giocoli (2003) and *Machine dreams* by Philip Mirowski (2002) were self-defined as contributions to the *history* of

AUTHOR'S NOTE: A first draft of this paper was presented at an EIPE Research Seminar in March 2007. For their comments, I wish to thank Menno Rol, Jack Vromen, and the other participants of the seminar, as well as to William Colish, Robert Nadeau, Christopher Pitchon, Christian Schmidt, Bernard Walliser, Don Ross himself, and two anonymous referees of the EJPE. I am also grateful to the Social Science and Humanities Research Council of Canada (Ottawa) for financial support.

economics, but they raise many questions that might have considerable repercussions for methodology of economics itself. These books were followed by two more explicitly philosophical essays, *The theory of the individual in economics,* by John Davis (2003), and *Economic theory and cognitive science: microexplanation*, by Don Ross (2005), which radically put into question some of the more respected tenets of traditional economic methodology

Many other important methodological contributions, especially papers in the main methodological and historical journals, have also been published during these few years, but my goal, here, is not to provide a survey of the work done in methodology of economics during this short period, but to inquire whether the methodological ideas proposed in such contributions can be described as revolutionary. Since it would not be possible in a short paper to seriously discuss each of them, I will focus on Ross's book, which Alex Rosenberg—who is himself one of the most respected methodologists of economics and has intensively published in the area throughout the last quarter of the 20th century—has presented as "the most important new work in the philosophy of economics in years" (jacket of Ross 2005).

Instead of developing a systematic discussion of Ross's ideas, I will emphasise what I consider illuminative and potentially "revolutionary" in them, but I will conclude that we should not exaggerate the consequences of these rich contributions nor disqualify too swiftly the more traditional approaches to methodology of economics that was developed in the last three decades of the 20th century. In order to substantiate this view, I will devote the last two parts of this paper to a discussion of the two questions which are most at risk of being affected by these new ideas, namely: methodological individualism (a favoured target of Ross), and the notion of rationality (so closely linked to the radically redefined concept of an economic agent).

FOLK PSYCHOLOGY AS A PREDICAMENT OF ECONOMICS

A trait which will be immediately noted is the fact that economic questions are integrated in much larger considerations: the developments and theoretical debates in the 20th century mathematics, the rapid progresses realised in neuropsychology, in cognitive sciences, in artificial intelligence, and in biological evolution, are considered and discussed in order to characterise what should be the proper place and status of economics. Economics here is no longer treated as a separate science,[1] and we are witnessing, from this point of view at least, if not a revolution, at least a massive illustration of an approach that could not easily be put aside in future work in methodology. This trait is obvious in Ross's contribution, but it is his crucial theses rather than views on the scope of economics that will be mostly considered in the present paper.

To start with, one must acknowledge that Ross's book will have to be treated with circumspection, as it will be followed by a second volume, which is to focus on *macroexplanation*, and to draw on the economist Ken Binmore's contribution, just as his present volume on *microexplanation* draws on the philosopher Daniel Dennett's theses. At various places in his book,[2] Ross announces that he will come back to some discussed questions in his second volume, which are frequently the most typically economic ones. Nonetheless, the first volume clearly builds up the theoretical frame in which a microeconomics should be developed, and it is precisely this aspect that I would like to discuss.

To understand in what sense this frame differs so radically from the one that economists are used to, it is appropriate to recall the central problem that Alexander Rosenberg has so frequently raised throughout his various publications.[3] For Rosenberg, economics suffers from a predicament that has impeded it from making the kind of progress expected of a science born more than two centuries ago. Normally, such a progress should have been observed in the accuracy of its predictions, but nothing significant has been noted from this point of view. According to Rosenberg, the predicament is that concepts like "beliefs" and "desires" (or "expectations" and "preferences" as economists prefer to say) do not "describe 'natural kinds'" characterised as "sets of items that behave in the same way, that share the same manageably small set of causes and effects". Therefore, these concepts, typical of a folk psychology, "cannot be brought together in causal generalizations that improve on our ordinary level of prediction and control of human actions, let alone attain the sort of continuing improvement characteristic of science" (Rosenberg 1994, 224). Contrary to concepts like "gene" and "acid", they have not been "carved" by a rigorous scientific analysis; they are rather inherited from a popular way of

[1] For the notion of a *separate* science as "concerned with a domain in which a small number of causal factors predominate", see Hausman 1992, 224-225, 90-97.
[2] For example, at Ross 2005, 291, 303, 313, 316, 320, 345, 353, 373, 381, 386, 393.
[3] See, for example, Rosenberg 1988, 15-20; Rosenberg 1992, 129-131, 148-151; Rosenberg 1994, 217.

speaking, like the concept of "fish". Such concepts are useful in business life but not in scientific analysis. A consequence of this is that the terms of theories built with them "do not correlate in a manageable way with the vocabulary of other successful scientific theories" (Rosenberg 1994, 224). Note that this does not mean that folk psychology is inefficient in prediction. On the contrary, most philosophers who aim to replace it acknowledge its astonishing efficiency,[4] which allows us to predict with a remarkable accuracy so many decisions and actions typical of the daily life of human beings. The problem is that these predictions have not progressed like scientific ones should.

Ross and an increasing number of philosophers agree with Rosenberg's analysis and insistently look for a way to avoid relying on folk psychology concepts judged inappropriate for scientific investigation. The conviction that this goal will be reached is nurtured by the following argument, which is repeated on various forms by all of them. There is no reason why this psychology based on folk concepts like "beliefs" and "desires" should not be progressively replaced by a scientific psychology based on neurobiological data, just like the folk astrophysics which held Ancients to believe that the sun and the whole sky turned around the Earth was totally replaced by a scientific astrophysics according to which it is rather the Earth that is moving. Similarly, a folk biology—according to which the intervention of an Intelligent Designer is necessarily required to explain the remarkable adaptation of most organisms and of specialised organs, such as eyes and hearts—was progressively replaced (while not yet in every circle) by a scientific theory based on natural selection.

Given the continuous scientific progress that has been made over the past four centuries, the folk psychology of economics should be replaced in turn. However, this kind of replacement should not be interpreted as a straightforward elimination of the so-called folk concepts in order to replace them by scientific concepts, despite the program of those who are known as "eliminativists". Indeed, in the two paradigmatic cases evoked above, the folk concepts were designating phenomena or experiences which have not been eliminated but explained in a much more satisfactory scientific fashion. The diurnal movement of the sun and the whole sky from East to West is a phenomenon still experienced by everybody, but it was explained in a

[4] For example, Rosenberg 1988, 15-16; Dennett 1991a, 29, 42, 43; Churchland 1984, 58-59.

much more satisfactory fashion by Copernicus and modern astronomers than by the Ancients. Similarly, the functional adaptation of living organisms is a remarkable fact that is explained in a compelling fashion by natural selection. Therefore, it would be ill advised to drastically eliminate concepts like beliefs and preferences simply because their prospect as scientific concepts is rather poor. They refer to intentional states which are constantly experienced by human beings and that could not be treated as nonexistent, even if it is judged essential that one develop an alternative and more scientific explanation of them.[5] Such an explanation might be based on concepts related in some fashion to neurobiology, but (1) it must "save" the phenomena that we used to characterise as intentional, and (2) it must be really compelling as an explanation of their occurrence, whether as illusions, misconceptions, or whatever.

It is for having provided such compelling explanations of what was experienced that Copernican astronomy and natural selection have respectively replaced folk astronomy and folk biology. It is doubtful that the learned community would have massively rejected folk theories on these matters if, without proposing compelling alternatives to explain phenomena such as diurnal movement and functional adaptation of organisms, Copernicus and Darwin, and their respective successors, had contented themselves with claiming that, according to sound principles of dynamics, it is not reasonable to think that the whole sky turns around the Earth, or that Creationism does not fit well with the rest of biological discoveries. Put otherwise, given the irrepressible need for explanation of phenomena whose significance is perceived as major, the onus of proof lies on those who want to dislodge folk theories.

From this point of view, Paul Churchland's attitude is somewhat irresponsible when he defends his materialist bottom up methodology with the following argument: "If the thumb-worn categories of folk psychology (belief, desire, consciousness, and so on) really do possess

[5] John Searle (1997, 111) made an observation of this kind. He also developed a more radical argument aiming to show that the case of intentional states is crucially different from the two others I have mentioned, because "where consciousness is concerned the existence of the appearance is the reality". He illustrates this view with the following comment: "the experience of feeling the pain is identical with the pain in a way that the experience of seeing a sunset is not identical with a sunset" (Searle 1997, 112, and the point is revisited in 121-122, 124). I think that this is an interesting point which should be considered seriously, but which can be contested by one who would claim that, even though the pain and the feeling of the pain are non separable (or even identical), the point is to determine the very nature and the origin of what is called a feeling and not to decide whether experiencing it is a real experience or not.

objective integrity, then the bottom up approach will eventually lead us back to them" (Churchland 1984, 97). How can one proceed to research by assuming that phenomena experienced by everybody like beliefs, desires, and consciousness may be forgotten and a priori reputed inexistent if they do not happen to be rediscovered by a particular type of materialist analysis?

In any case, it is the merit of Daniel Dennett, together with a few other philosophers, to have devoted his own career to the difficult task of explaining beliefs, desires and even consciousness, while constantly remaining faithful to a commitment requiring that such an explanation be derived from an analysis of the brain and of its environment. And it is the merit of Don Ross to have boldly attempted the equally difficult task of adapting such an approach to economics, a science that has almost always assumed the full validity of the traditional conception of beliefs, desires, and consciousness. One must acknowledge indeed that Dennett and Ross have emphatically rejected eliminativist theses such as Churchland's. However, one must inquire about the ontological status granted to these intentional states, which are rescued from elimination in such a highly theoretical process.

THE ROLE OF THE INTENTIONAL STANCE

When it comes to explain why people *have* beliefs, the fundamental step in Ross's argumentation is based on Dennett's notion of "intentional stance". Since the role played by the intentional stance in Dennett's philosophy may easily be a source of confusion, it is worthwhile to recall what is involved in it. According to Dennett, it is an intellectual attitude corresponding to a strategy of interpretation that presupposes that an object (not necessarily people) has intentions and act rationally in such a way that it becomes possible for us to predict its behaviour.[6]

At first glance, the notion seems to be a very simple one, since it describes what we are so frequently doing when we say, for example, that our computer wrongly *believes* that we ask it to do an operation which, in fact, does not really interest us, or that it does not *want* to do a particular operation, or when we explain to a child that a frog *believes* that a fly is good to eat and *wants* to eat it. We may be convinced that such a parlance is metaphorical, but, for Dennett and Ross, it is not really so. Dennett admits that the metaphorical view is "immensely persuasive", but rejects this interpretation as deceptive in favour of his

[6] See, for example, Dennett 1987, 15; and Ross 2005, 38.

own position according to which "there is *nothing more* to *our* having beliefs and desires than our being voluminously predictable (like the frog, but more so) from the intentional stance" (Dennett 1987, 108, first emphasis added, second one is in the original). Taking an intentional stance when we consider human beings—which means granting human beings intentional states—allow us to voluminously predict their behaviour. But the behaviour of a computer or of a frog can be voluminously predicted as well when we grant them these intentional states by taking an intentional stance; therefore Dennett concludes that the difference between them and us turns out to be a simple matter of degree.

Now, the intentional stance is not the only way we have to predict the behaviour of what Dennett and Ross call a "system" (a computer, a human being, a frog, or any other animal). One may take a design stance, which consists in treating the system the way we are used to considering a machine that has been designed by an engineer. People unfamiliar with mechanics are frequently tempted to take an intentional stance toward their car which does not "want" to work properly, but, in this case, it is normally more appropriate to take a design stance and predict or explain the behaviour of the car from the examination of the functions that the engineer has reserved for the various parts of the engine. Dennett adds that a third stance which must be distinguished from these first two is the physical stance, the one we take when we predict or explain the behaviour of a physical body with the help of the causal laws of physics without assuming the intervention of any designer.[7]

Thus, in order to predict or explain the behaviour of a thermostat, one may take either a physical stance, looking at the causal laws affecting its material parts, a design stance, looking at the functional parts put into interrelation by engineers or an intentional stance looking at the thermostat as believing (rightly or wrongly) that the temperature is at such a level and wanting to avoid further heating. Dennett insists on the fact that, while the physical stance remains the most fundamental one, each of these three stances allows us to understand many phenomena that would be non-accessible from the other stances. This can be easily illustrated by an example familiar to economists. A Martian who would like to understand what is going on in an economic exchange by following, with the help of physical laws, the movements of

[7] For a presentation of these three types of stance, see Dennett 1987, 16-18, 38-40.

commodities and of pieces of money involved in the process would totally miss the economic exchange itself. Consequently, this Martian would have a lot of trouble to predict the developments of the phenomenon, in contrast with economists who quite appropriately take an intentional stance (they grant intentional states to the traders) in order to explain it.

None of these interesting considerations sounds very offensive for traditional economics, except, maybe, for the idea that associates one's beliefs and desires with the very fact of taking an intentional stance (taken by oneself or by someone else). That does not mean that every time that one takes an intentional stance toward something, let us say a car for example, one is committed to admitting that the car actually has beliefs and desires. To avoid this misinterpretation, Dennett and Ross insistently claim that a system's beliefs and desires can be associated with an intentional stance taken toward them only when the intentional stance is the *only way* to predict or explain the behaviour of this system. This proviso about the intentional stance may sound a little odd, but the idea is that, no more than the Martian evoked above, anyone can capture what is involved in an event such as an economic exchange without taking an intentional stance. Yet, no one will deny that an economic exchange really exists; it is a *real pattern* that must be explained as anything else. If the seller and the buyer's actions cannot be explained otherwise than with the help of beliefs and desires, Dennett and Ross will say that it is precisely because these actions cannot be explained otherwise that their beliefs and desires can be said to be real.

Similarly, if the moves of a sophisticated chess-playing computer cannot be predicted otherwise than by taking an intentional stance toward it, they will say that this computer *really* has beliefs and desires. In contrast, my car has no such intentional states: the simple fact that I frequently say that it believes or wants so and so does not allow me to explain or predict anything that I cannot explain or predict much better by taking a design stance toward it or by taking an intentional stance toward myself, the driver. But let us consider this view more closely.

BETWEEN SCYLLA AND CHARYBDIS

According to Ross, the "core thesis" of intentional-stance functionalism, which is the name he gives to Dennett's philosophical approach that he equally adopts, consists of the following claim:

> What it is to have intentional states—real ones, in the only sense of 'real' that attaches to any intentional states—is to exhibit behavioural patterns that can't be predicted or explained without recognition of the patterns indexed by the intentional states in question (Ross 2005, 63).

As seen above, Dennett, for his part, presents the same thesis in the following terms: "there is nothing more to *our* having beliefs and desires than our being voluminously predictable […] from the intentional stance" (Dennett 1987, 108). But what is the exact meaning of this sentence? If it is interpreted as simply saying that one's real beliefs and desires are inferred from the fact that we can predict one's behaviour with their help, in such a way that their real existence is confirmed, it is difficult to see in what sense this view significantly differs from folk psychology. Economists and historians who rely on folk psychology do not pretend to directly experience beliefs and desires of the agents they are studying; they assume that these agents have beliefs and desires and, since they can predict or explain their behaviour on this basis, they conclude that their assumption was well grounded. To avoid such a traditional interpretation and fully appreciate the originality of intentional-stance functionalism, one has to take seriously the words "nothing more" in Dennett's sentence, and construe the latter as *denying that the existence of beliefs and desires means anything more* than the fact that they are a necessary condition to predict a specific behaviour.

This is the interpretation that explains why Dennett's view has frequently been pejoratively characterised as "instrumentalist" (beliefs and desires being nothing but instruments to predict). According to Ross, one should instead characterise it as "behaviourist" (beliefs and desires being nothing but the fact that they are *required* to explain a particular behaviour), and admit that it is a consistent first step in a materialist attempt to "save" intentional phenomena from being purely eliminated. However, even if one is fully happy with this "behaviourist" way to characterise intentional states, one may raise an objection, that Ross does not consider explicitly, about the capacities of the "systems" that take such intentional stances. Who takes intentional stances? Clearly human beings take them, whether or not other systems also do. But if human beings have the capacity to take intentional stances, which means to interpret the behaviour of something by attributing it intentional states such as beliefs and desires, it is because they *already* *have* beliefs and desires themselves (they desire to predict behaviour and they believe that taking an intentional stance is the proper way to satisfy this goal). Now the point of invoking intentional stances was to solve the problem raised by the questionable existence of beliefs and desires. Explaining beliefs and desires by the fact that an intentional stance was taken toward the person who has them and explaining the capacity to take this intentional stance by beliefs and desires of the intentional-stance taker clearly looks like a case of chicken-and-egg type explanation.

A possible way out of this objection might be that even if human beings take intentional stances with the help of their beliefs and desires, this does not imply that intentional stances cannot be taken in quite different ways. I will consider this view below, but first I would like to examine how Dennett, according to Ross, answered to an objection of this type, put forward mostly by John Searle, according to which "our attribution of intentional meaning to states of artifacts is parasitic on the fact that we are already intentional interpreters" (Ross 2005, 43).[8] According to Ross, the strategy of Dennett's answer to Searle was to explain how people can have intentional states, just like computers! (Ross 2005, 44; see Dennett 1987, chap. 8). After all, human beings have been designed by natural selection with a brain that may be described as an exceptionally versatile computer. Note that this could hardly be considered a direct answer to the question I raised, which does not directly concern the possibility for people or computers of *having* intentional *states* or not, but the fact that human beings can *take* intentional *stances* to start with.

In the chapter 8 referred to by Ross, Dennett uses a few highly ingenious mental experiments to argue that if, as easily admitted, the "intentional states" of the device of a soft drink vending machine that accepts quarters and rejects slugs are granted to it metaphorically, so is the case for an extremely sophisticated robot and for human beings (Dennett 1987, 294, see also 290-298). The idea is that, even though the device, the robot and their "states of mind" are just artefacts, people would be artefacts as well, "artefacts designed by natural selection"

[8] With this sentence, Ross intends to capture the gist of an objection raised by Searle (in Searle 1980; see, for example 418, left column), which was implied by Searle's defence of his famous Chinese room argument; but, in contrast with the one I proposed above, this is not an objection addressed to Dennett's thesis on the intentional stance.

(Dennett 1987, 300). Natural selection would have designed not only our brain, which is a generally accepted view, but also the meaning which is circulating in it, just like engineers have designed not only the hardware of the robot but the software through which meaning is transmitted. Dennett does not hesitate to locate some intentionality in our genes in order to explain intentionality in our minds: "So our intentionality is derived from the intentionality of our 'selfish' genes! They are the Unmeant Meaners, not us!" (Dennett 1987, 298). This is granting quite a bit to natural selection and genetics; but if we accept this view, Dennett's answer to the question I raised above might be that *natural selection* has designed human brains that have beliefs, desires, and other intentional states including the capacity to interpret and, therefore, to *take* intentional stances.

Such a view, however, would introduce another problem for intentional-stance functionalism. If such human interpreters are produced ready-made by natural selection, with intentional states and the capacity to take intentional stances, the behaviourist interpretation of Dennett's view, according to which "there is nothing more to *our* having beliefs and desires than our being voluminously predictable […] from the intentional stance" (Dennett 1987, 108), no longer holds. Indeed, there would be *something more* in our having beliefs and desires, namely what natural selection would have provided to us, and which was acquired without the help of any intentional stance taken toward us, since natural selection does not take intentional stances. What meaning is left of this insistent behaviourist claim that beliefs and desires are *nothing more* than the fact of being predictable if it is admitted that natural selection has made people with desires, beliefs, and the capacity to take any stance they like? Should we conclude that we are back to a "folk psychology" interpretation according to which beliefs and desires fully exist, with the additional precision that this is due to the work of natural selection?

CAN CONSCIOUSNESS BE EXPLAINED?

For his part, Ross does not rely on such an alleged genetic basis of intentionality, even though he does not hesitate to grant to human systems the capacity to take an intentional stance not only toward other systems but also toward themselves, which means to interpret reflexively their own intentional states. Thus, he maintains that "[t]he main (relevant to present issues) difference between existing chess-playing machines and human chess-players is that the latter do, and the former don't, take the intentional stance toward *themselves*" (Ross 2005, 63). Ross, however, is fully aware of the difficulty involved in the notion of one adopting the intentional stance toward oneself. Since the subject that takes such a stance cannot "be a *part* of the system", when this reflexive intentional stance is evoked, the "most immediate and vicious sort of circularity thus seems to threaten" (Ross 2005, 286).

To avoid this threatening circle, Ross—who estimates with Dennett that deciding whether a chess machine can take an intentional stance toward itself has much to do with deciding whether it has consciousness or not—turns toward the thesis that Dennett developed in *Consciousness explained*. In this book, Dennett no longer relies on genes to explain intentionality. Instead, it is cultural selection that is invoked in order to account for human specific capacities and especially for consciousness (Dennett 1991a, 199-207). Thus, consciousness must be seen as "a product of cultural evolution that gets imparted to brain in early training" (Dennett 1991a, 219; see also Ross 2005, 160). After having rejected as a dead-end attempts to explain consciousness by looking *inside* the human brain, Dennett and Ross had little choice other than turning toward cultural and social factors to explain it (Ross 2005, 44-52); therefore, intentional-stance functionalism is a resolutely *externalist* approach.

Far from being explained by the very structure of the brain, consciousness, according to Dennett, must not be conceived as a solidly unified entity that would survey the activity of a person. It would be rather a result of multiple drafts (the word "draft" being understood here as the successive drafts of a paper written by a perfectionist author) in such a way that, each of these drafts being potentially operative at one point in time, there is no such thing as a final draft, which could be considered as "the moment of consciousness" (Dennett 1991a, 126; see also 113, 125-126) It is the reason why Dennett has named "Multiple draft model (MDM)" his model of consciousness. But how can these drafts be developed? Essentially as a result of social intercourse and by the use of public language.[9] Now, few persons will deny the decisive role of society and of public language in the formation

[9] "Public language" designates here a social mode of communication transmitted from generation to generation (Ross 2005, 288), which must be distinguished from an internal "language of thought" (Ross 2005, 53) and also from "emotional signalling systems" (Ross 2005, 300).

of human beings, both as a species and as individuals, but Dennett and Ross go much further. Public language is described by Ross, who makes explicit one of Dennett's suggestions, as "*the* scaffold that makes humans so strikingly different in their ecology from other intelligent animals" (Ross 2005, 286-287). Indeed, Dennett even claims that the difference between human beings, the "most prodigious intentional systems on the planet" and the poor intentional systems exemplified by frogs is largely explained by the fact that the former are "bathed in words" that allow them to "assert, deny, request, command and promise" (Dennett 1987, 112). How this bathing in words may transform us through a relatively uncontrolled use of public language is what Dennett discusses with great ingenuity in chapter 8 of *Consciousness explained*, even though one may remain unconvinced that this process can ultimately generate the capacity for human beings to take intentional stances toward themselves.

It is now possible to recapitulate. If Dennett was right in chapter 8 of *The intentional stance* when saying that intentionality is derived from our genes thanks to natural selection, we would be fully equipped to take intentional stances toward other people or even towards ourselves, but this process would be redundant since, thanks to natural selection, people would be already endowed with intentional states such as beliefs and wants. One may admire the ingenious analogies between the working of a computer and the working of the brain that Dennett explores in this chapter, but if natural selection was so generous, it is difficult to understand how one could still maintain that beliefs and desires are *nothing but* the fact that intentional stances must be taken toward those to whom such intentional states are attributed. Why should human beings—whose brains would be endowed with such a well designed program allowing them to interpret other's actions and their own actions by taking intentional stances—wait to be themselves the object of an intentional stance in order to be able to experience beliefs and desires?

In fact, Dennett and Ross do not want to go so far; while they claim, like so many thinkers, that natural selection has designed the brain as a remarkably efficient hardware, it is a *cultural* selection—based on social intercourse and public language—that they invoke to explain the development of the required software, namely intentionality and consciousness. One would like to know more about the mechanism through which such a highly sophisticated software—about which

Dennett even suggests that its survival to the brain cannot in principle be excluded (Dennett 1991a, 430, 368)—has been designed, not by a learned engineer working in artificial intelligence, not even by natural selection, but by the virtue of the progressive use of a public language learned by cultural interaction. In any case, were Dennett and Ross claiming that social intercourse and public language can really endow *people* with the capacity and the autonomy required to take such intentional stances, the "behaviourist" thesis, according to which intentional states *depend* in some way on being the object of an intentional stance, would no longer keep afloat. Indeed, here again, claiming that the joint effect of natural and cultural selection was to provide human beings with all the intentional states required to take intentional stances brings us back either to the neutralisation of this "behaviourist" thesis in favour of an eliminativist one, or to a revamped folk psychology, which could very well work with these gifts of selection.

Clearly, to understand the role reserved to the intentional stance, behaviourist and externalist components of the discussed thesis must be interpreted in a more radical fashion. As underscored by Ross (in a personal communication), taking an intentional stance "is manifested in 'behavior' not, in the first place, in beliefs". This might ensure a role to intentional stances within the very process of cultural selection, but this move raises new questions. Let us admit that, before being able to believe or to want anything, it is possible to behave in a way that corresponds to taking an intentional stance, and let us try to imagine such behaviour. For example, an individual might be afraid, in a strictly behavioural manner, when facing someone else and predict an aggressive behaviour on this basis, but it is far from clear that no intentional states would be involved in such a situation, especially if we remember that Dennett showed how it is complicated to decide whether frogs really have beliefs and desires or not (Dennett 1987, 106-110). Possibly, most relevant examples of behavioural attitudes corresponding to a purely behavioural intentional stance might be found, but the difficulty becomes still more serious when we consider what is required for an intentional stance to be efficient enough to contribute in some way to the fact that desires and beliefs with those who are the objects of such stances can be considered real.

Indeed, after recalling the criterion invoked by Dennett and Ross for not being a simply metaphorical intentional stance, which is that the intentional stance be the only possible way to predict the considered behaviour, let us suppose that a frog has no desire to avoid predators nor beliefs about the efficiency of jumping away to reach such a goal but nonetheless jumps away when facing an aggressive move of a predator. Can we really say that the frog has taken a behavioural intentional stance that may allow us to claim that the aggressive intentions of the predator were *nothing more* than the fact that they provoked this move of the frog? Answering yes might sound a bit preposterous, but if the answer is no, why would a similar move have this consequence when we replace the frog by pre-evolved human beings who are still unable to have beliefs and desires? Clearly, Dennett and Ross's answer would be that in contrast with these human beings, the frog is not "bathed in words".

But what is the exact role of public language? The point is not to simply claim that public language has strongly contributed to develop typical abilities of human beings, a claim that is perfectly acceptable even for "folk psychologists". Ross goes much further and his externalist approach may even imply that propositional attitudes such as beliefs and desires should be considered real "not as descriptions of patterns in brains, but as descriptions in patterns of social communication" (Ross 2005, 61). If we consider that the development of *real* desires and beliefs in social communication is clearly less problematic than their development in a human brain, this construal might be a way to conciliate significant intentional stances (taken from the social world endowed with desires and beliefs) with the idea that pre-evolved human beings have no intentional states. However, how such a purely behavioural intentional stance might be *socially* taken with the help of social beliefs and desires which do not exist yet in individual human beings is far from clear. And such is the idea that intentional states so generated were nothing more than their association with behaviours voluminously predicted through (and only through) this behavioural and social intentional stance.

In any case, this hypothesis, which can hardly be intuitively or empirically grounded, supposes the existence of a social agency that I will briefly discuss later. For now, let us conclude that the problem lies with the fact that the notion of individual stance, that is perfectly clear when it is attributed to fully developed human beings, becomes less and less clear when, in order to make the theory consistent, it is attributed to stance takers that have little in common with them.

INSTRUMENTALISM VERSUS REALISM

If intentional-stance functionalism seems to hesitate between endowing human beings with full capacity to take intentional stances, and saving the "behaviourist" thesis by subordinating their intentional states to a socio-behavioural intentional stance, it is because an obvious tension exists in Dennett's thought between an "as if" *instrumentalism* associated with this reduction of intentional states to simple instruments of interpretation for those who take an intentional stance and a somewhat hesitant *realism* according to which cultural evolution has installed in human beings ontologically consistent powers.

As recalled by Ross, Dennett's position "has regularly been associated with instrumentalism" (Ross 2005, 64, see also 160, 264). John Davis, for example, associates Dennett with Friedman and claims that, like the latter, "Dennett is not interested in the realism of our assumptions about the mind, but only about their predictive value" (Davis 2004, 96). It is true that various passages in Dennett's work may incite readers to think this way, but the anti-instrumentalist and even realist character of Dennett's philosophy is one of Ross's most consistent claims. After all, what makes Dennett's originality among the most radical of materialist thinkers is his anti-eliminativism. Against those who claim that notions like beliefs, desires, and consciousness must be eliminated, Dennett has devoted a large part of his work to show that such entities *really* exist. Ross, who claims that a basic realism is a presupposition of his own book (Ross 2005, 21-22, 57), still accentuates this realism, which for him is capital for economics, and presents Dennett, surely not as a commonsense realist, but as "radical scientific realist" (Ross 2005, 163-164).

But what about the intentional stance, which is so easily perceived as an instrumentalist trick allowing us to interpret people's actions *as if* they were guided by desires and beliefs? To counter this perception, Ross distinguishes two quite different (while complementary) activities pertaining to intentional-stance functionalism: one is purely *methodological* (MISF), and the other is resolutely *ontological* (OISF) (Ross 2005, 336-337). It is only the former that can be said to be instrumental, for example when I attribute beliefs and desires to an object—a thermostat, for example—just because it is *useful* to predict or explain its behaviour without seriously thinking that the intentional states referred to have any ontological status. By contrast, the latter "aims at explaining, still in intentional terms, the dynamics of systems one already has reasons for believing to be irreducibly intentional" (Ross 2005, 336).

To illustrate such a case, Ross, safely enough, takes the example of "a person". This counterargument simply consists in accepting the validity of the accusation of instrumentalism for non-crucial cases, where the intentional stance has no special ontological pretensions, but since the accusation normally bears on the cases where intentional stances are taken toward a person with, according to Ross, an ontological meaning, introducing such a distinction is tantamount to simply rejecting the accusation in question. However, the borderline between both types of cases remains rather vague as illustrated by the case of frogs discussed non-conclusively by Dennett (Dennett 1987, 106-116, see especially 111-112). Thus, the most important question concerns the foundations of the ontological certification that is granted this way to only *some* intentional stances taken.

On what grounds does intentional-stance functionalism attribute *ontological* status to entities such as intentional states? In his remarkable paper of 1991 entitled "Real patterns", Dennett comes back to an idea he had introduced a few years earlier in *The intentional stance* where he drew attention to the fact that entities such as centres of gravity can be said to be fully real without being pieces of "furniture of the physical world" (Dennett 1987, 72). The 1991 paper characterises such entities as *real patterns*, which are *real* because, when it comes to capturing the phenomenon to which they correspond, "*there is* a description of the data [constituting them] that is more efficient than the bit map" (Dennett 1991b, 34). For example, centres of gravity satisfy this criterion since "we think they serve in perspicuous representations of real forces" (Dennett 1991b, 29).

According to Dennett, the notion of existence should not be treated as univocal, a view which allows him to save at a relatively low cost the existence and reality of intentional states: "beliefs are best considered to be abstract objects rather like centers of gravity" (Dennett 1991b, 29). However, if it is true that the notion of "existence" can be understood as a matter of degree, one may wonder whether the totally passive and abstract notion of centre of gravity can really be put on the same (existential) footing as notions such as belief and consciousness, which correspond also to real patterns, but which have hardly any meaning at all if they are emptied from their *active* connotations.

After publishing his own paper on real pattern (Ross 1995), Ross came back to the central question raised by Dennett's paper in his contribution (Ross 2000) to an edited collection of essays on Dennett's philosophy. This contribution proposes an ontological interpretation of Dennett's views that the latter has received with an evident sympathy, which did not however imply full conviction (Dennett 2000, 359-360). Ross unequivocally claims in his chapter that "reality is composed of real patterns all the way down" (Ross 2000, 160). This must be understood as a strong ontological thesis claiming that "to be is to be a real pattern" (Ross 2000, 161), a thesis which requires a criterion for determining when a pattern can be said to be "real".

Ross formulates a technical criterion that adds a few precisions to Dennett's criterion of being "more efficient than the bit map".[10] For Ross, such conditions seem necessary to protect Dennett's philosophy against a "slide into instrumentalism" (Ross 2000, 160) and to definitively put it on the side of a kind of realism that Ross has christened "Rainforest realism". It might be difficult, however, to avoid thinking that such an intricate criterion required for determining whether something is *a real being* sounds a bit ad hoc and that the way to check whether consciousness, for example, satisfies it is not very clear.

SELVES VERSUS AGENTS

Be that as it may, when it comes to account for the reality of the selves in *Consciousness explained*, Dennett describes it both as an *abstraction* (Dennett 1991a, 368, 414)—an aspect which does not fit very well with Ross's views concerning Dennett's realism—and what he calls a "*center of narrative gravity*" (Dennett 1991a, 418, 429), meaning that it is narratives and biographical accounts that "spin" a self, just as a character in a novel is spun. Ross draws from this idea concerning the formation of the self through narratives (Ross 2005, 280, 285, 286), while emphasising the role of education still more than Dennett does (Ross 2005, 282-289). Moreover, he suggests completing this analysis

[10] The whole text of these conditions reads this way: "To be is to be a real pattern, and a pattern is real iff: (i) it is projectible [sic] under at least one physically possible perspective and (ii) it encodes information about at least one structure of events or entities S where that encoding is more efficient, in information-theoretic terms, than the bit-map encoding of S, and where for at least one of the physically possible perspectives under which the pattern is projectible, there exists an aspect of S that cannot be tracked unless the encoding is recovered from the perspective in question" (Ross 2005, 68-69; but also see Ross 2000, 161).

with the help of game theory by distinguishing three levels of games that would be played in the process of the acquisition of the self. However, probably more interesting for the methodology of economics than these levels—whose relations are carefully described, but which are not directly related to concerns in economics, and are not concretely implemented—is the idea that Ross develops about the relation between agency and selfhood.

In *Consciousness explained,* Dennett reported scientific observations of multiple personality disorders, to the effect that "a single human body seems to be shared by several selves" (Dennett 1991a, 419ff.). Ross made some headway on observations of this kind which allowed him to conclude "the Dennettian theory that separates agency from selfhood conceptually undermines microeconomic individualism" (Ross 2005, 311). Now, Dennett is not an economist and he rarely refers to economics and even relatively rarely to the notion of agent (or agency). In the indices of these books, "agent" appears only once in *Consciousness explained*, nowhere in *The intentional stance*, and ten times in *Brainstorms*; but none of these uses is related to economics or to the concept of self, which never appears in the same chapter as the concept of agent. For Ross, who is an economist as well as a philosopher, the notion of agency was a central one, and the idea that it could not be hardwired to the notion of self, or to the notion of human being, became a central point of his analysis of economics.

This idea plays a decisive role in his discussion of Gary Becker (even if it is only a marginal piece in his book), and above all in his discussion of the economics derived from Samuelsonian revealed preference theory (which is a central piece in the economic application of his theory). Since he does not see any problem in dissociating agents from selves, Ross can claim that "the biography [which is related to the self] of a typical person can't be the biography of a single (diachronic) economic agent" (Ross 2005, 156). According to him, only a contestable Aristotelian assumption incites us to think that human lives should be modelled "as single projects aimed at achieving (some) consistent goals" (Ross 2005, 159). Thus, if there are such things as human agents (and utility functions), they do not have to be coextensive with selves (and their biographies).

Ross's position in this discussion was based on his decisive adoption of revealed preference theory (RPT) as the paradigm of neoclassical theory. Many economists would contest such an assessment, and their contestation would probably not be diminished by the fact that this assessment is based on philosophical as much as economical considerations. Indeed, Ross took "Samuelson's generic late-positivist philosophy more literally than he did" (Ross 2005, 156). He illustrates this by admitting that its usual application to typically economic matters was not relevant in his considerations: "I have treated RPT just as a set of axioms, leaving completely open the question of which phenomena, if any, the axioms describe" (Ross 2005, 156). A consequence of this decision is that Ross's conception of an agent is strictly determined by these axioms as illustrated by the discussion of Becker's thesis on stable preferences: "Of course, as a matter of logic, an *economic agent* must have stable preferences; otherwise RPT does not apply to it" (Ross 2005, 157). Indeed, how can transitivity be conciliated with changing tastes?

But this decision has still more radical consequences. Once it is admitted that any straightforward agents must strictly respect the Samuelsonian axioms, which imply consistency, it is clear that human beings are disqualified, given their lack of consistency clearly manifested in experimental economics to which Ross devotes a well-documented section of his book.[11] But where can we find agents if human being are such poor candidates for this title? For Ross, there is a crowd of other candidates, and among them we find various inferior animals and especially insects, whose paradigmatic quality of agent he so frequently refers to: "[a] good example of a prototypical economic agent is an insect".[12] This is hardly surprising, since insects have little propensity to modify their behaviour and, consequently, to fail a consistency test. Moreover, they are good optimisers as, incidentally, Marx had noted in a famous passage of *Das Kapital* according to which bees are able to surpass many architects in the construction of her cells, even though, the worst architect is still superior to the best bee for being able to build a house in his or her head before building it in the real world.[13]

For Ross, however, the criterion is rather "the central locus of control" (Ross 2005, 381), which is present in bugs, but not in humans. It is not clear on what ground one can rely on this central control to generate consistency and maximisation, but it seems that insects are

[11] Ross 2005, 165-190; see especially pages on preference reversals that experimentation has put into light (Ross 2005, 177ff.).
[12] Ross 2005, 251; but see also 95, 241, 252, 253, 256, 290, 331, 377, 381, 393.
[13] *Das Kapital*, Book one, Third section, chapter VII, part I.

successful from this point of view. In any case, there is no doubt that some electronic devices can be modelled with the capacity to satisfy the criterion, but a bit more unexpectedly, the other kind of candidates that are systematically considered by Ross are neurons. Paul Glimcher has developed a research program called neuroeconomics, consisting of the application of economic analysis to neural behaviour. According to Ross, "the economics in question must obviously be Samuelsonian, since no one imagines that parts of organisms have utility functions based on internally represented preferences" (Ross 2005, 325). Thus, brains would be "an ideal site for Samuelsonian microeconomics" where the agents are neurons and perhaps modules (organised groups of neurons) (Ross 2005, 334). About one of these modules, the visuomotor cortex, Ross even concludes that "it is thus a straightforward economic agent, just like an insect" (Ross 2005, 331).

Even if these modules were straightforward sub-personal agents, taking an intentional stance toward them would be purely methodological; it helps to understand their working, as Ross admits, without implying that they really have intentional states. By contrast, according to Ross, "the various long-, medium- or short-term interests" (Ross 2005, 337), which were analysed, equally with economic tools, by the psychiatrist George Ainslie in his *Picoeconomics*, are such that an *ontological* intentional stance can legitimately be taken toward them, which means that their dynamics is considered "to be irreducibly intentional" (Ross 2005, 336). So much so that each interest is considered to be "as clever as a person" in such a way that "their strategic cunning will tend to unravel all equilibria" (Ross 2005, 345), at least in some types of games.

Now, the idea that human beings are constantly divided between trends that draw their decisions in various directions is an old one, which goes back at least to Augustine, but, according to Ross, this tension is explained by the fact that a typical human is constituted by a "colony of agents" which "emerge under analysis as a complex assembly of buglike homunculi" (Ross 2005, 252). This leads us to what Ross designates as "this book's central thesis". It might look odd that this central thesis does not directly concern economics—at least as it is traditionally understood—but rather the very nature of a person, which is defined as "a set of basically compatible long-range interests that have co-opted a sufficient army of short-range interests into their coalition to maintain stable equilibrium" (Ross 2005, 351), but one must recall that the book bears on microexplanations in both economic theory *and cognitive science.*

A consequence of this is that "selves aren't straightforward economic agents. They are more like nations than insects" (Ross 2005, 290). However, since people do take actions, they should be agents in some sense; according to Ross, they are "agents-by-extension" (Ross 2005, 256), in contrast with straightforward agents. Therefore, they are put on the same footing as nations, which can take actions just as well.[14] Consequently, "[t]he application of economics to people will thus have to follow the same methods, and meet the same ontological and epistemological demands, as the application of economics to countries and corporations" (Ross 2005, 257). Since they are no more unified than nations, it is not surprising that, when treated as agents, humans "show ubiquitous preference reversal and time inconsistencies" (Ross 2005, 253). Instead of acknowledging that human agents, who used to be prototypical agents, are far from being consistent, Ross defines agency through consistency and concludes that humans are not prototypical agents.

The idea that human beings could be constituted of many centres of decisions is not new. As underscored by Ross, Davis, who has devoted his own 2003 book to this question, has pointed out that it must go back at least to Hume. More recently, Jon Elster edited a book, entitled *The multiple selves*, gathering papers that gave a fair idea of the state of the question in mid-1980s.[15] It is interesting to note, however, that Ross bases his intrapersonal community thesis on the idea that the person, whether or not identified with a self, is a *community of many agents*, whereas most of his predecessors refer, more or less metaphorically, to a *multiplicity of selves* who are present in a single person, usually identified with an agent. Davis, for example, suggests that individuals of neoclassical economics are constituted as a community of selves; but since he is looking for "an adequate conception of the individual in economics" (Davis 2003, 80), he turns away from orthodoxy and adopts

[14] Ross does not seem to be bothered by the fact that nations are intermingled between themselves, vaguely circumscribed and constantly redefined according to people's sensibilities and ideologies.

[15] Elster 1987. However, in his Introduction to the book, Elster insists on the fact that most of these tensions inside a person do not imply the duplication of selves (Elster 1987, 10, 13, 14, 23, 24, 26-27, and especially 30); however, there are two mentions of a possible exception to this, associated to the phenomenon of self-deception, on pages 28, 31. Davis also, in note 9 of his chapter 4 (Davis 2003, 196) quotes Elster, Steedman and Krause, as well as Kavka, who all deny to be claiming that there are *literally* multiselves.

the "socially embedded individual conception" (Davis 2003, 117) developed in heterodox economics. In contrast, Ross, who does not care for the unity of the individual, turns toward a version of neoclassical economics based on Samuelson's revealed preferences, which Davis denounces for its formalist indifference to the question of individuals (Davis 2003, 93-94).

CAN METHODOLOGICAL INDIVIDUALISM SURVIVE?

These circumstantial remarks are far from doing justice to the careful analyses of Davis; and even Ross's book, whose discussion was the core of the present paper, is so dense that only some of its main theses have been discussed. However, these books and a few others recently published force us to reconsider some of the fundamental pieces of traditional methodology of economics. These books ruthlessly reject methodological individualism and they invite us, at the very least, to reconsider the relation between rationality and agency.

What can still be said in favour of methodological individualism, which was the favourite target of so many theoreticians? Personally, I always had trouble to see what exactly was the point of the long debate between methodological individualism and the opposite thesis, whether holism, collectivism, or whatever. Of course, straw men built up by opponents to define either holism or individualism correspond to strongly opposite perspectives, but it is very difficult to find serious defenders of these extremist theses. Usually, those who consider themselves as champions of one of these theses, in order to make their position tenable, take care to introduce so many nuances in their characterisation of their favourite "ism" that the division between such opposing views becomes more or less blurred.[16]

On the one hand, it is not difficult, for example, to underscore various individualist features in the methodology of a radical collectivist thinker such as Marx and of a *macro*economist such as Keynes.[17] On the other hand, it should be admitted that methodological individualism cannot be separated from the mechanism through which unwanted consequences of human actions—consequences that explain most economic phenomena like an endogenous increase in the price of

[16] Malcolm Rutherford argues for a middle way between extreme versions of holism and individualism, which are "taken to be unacceptable" and "unappealing" (Rutherford 1994, 50; see also 36-37).

[17] For Marx, see Elster 1985; Lagueux 2001, 698-701. For Keynes, see Lagueux 2001, 696-698.

potatoes or a rise in the rate of unemployment—result from the interference of a multitude of anonymous but, in principle, understandable individual actions.[18] Very few methodological individualists, if any, would deny that the consequences of these multiple (rational) actions are constantly modified and deviated by the unforeseeable impact of the natural and, above all, social environment.[19]

It is also important to reject any association of methodological individualism with the imaginary economics of Robinson Crusoe, since methodological individualism is a methodology adapted to essentially *social* sciences such as economics. As aptly underscored by Ross himself (Ross 2005, 216ff.), Robinson Crusoe is just a useful pedagogical device, but methodological individualism has nothing to do with this kind of pedagogy. More importantly, methodological individualism is *not* a reductionism. The point of methodological individualism is not to *reduce* social phenomena to individual ones; given what was said above about unwanted consequences, such a reduction would be doomed to fail. It is to *understand* social phenomena by explaining why rational human actions produce social consequences significantly different from those which would be expected by people who take such actions.

Now, human actions have a lot of social consequences, but it is clear, as abundantly illustrated by Ross, that the causal link goes at least as much the other way around. Society influences individuals possibly still more than individuals influence society. No doubt that cultural evolution of humanity can hardly be understood otherwise than as a complex interaction between societies and individuals. However, most phenomena that are explained by interventions of society have traditionally been excluded from the domain of economics. Ross has shown how far social structures, through public language, education and imitation have been determinant for the very genesis of individuals, but the genesis of individuals concerns traditionally anthropology,

---

[18] Among others, J. W. N. Watkins, who was among the first authors to devote important papers to methodological individualism, underscored the link between this methodology and unwanted consequences (Watkins 1953, 26), a notion so closely associated with the thought of Friedrich Hayek, another defender of methodological individualism.

[19] In a recent paper, Geoffrey Hodgson (Hodgson 2007, 220, 222) claims that methodological individualism understood in this fashion is misnamed since an individualism that makes room for explanations requiring interaction between people is not a pure individualism, and therefore cannot easily be distinguished from approaches that are not considered individualistic. He is surely right, but my point concerns the validity of the methodology not of the label, since it is the former and not the latter that is challenged by Ross's arguments.

psychology, and neurobiology, but not economics which is concerned by the genesis of a subset of *social* phenomena. And methodological individualism is relevant only when it comes to explain social phenomena.

Ross could oppose at least two considerations to this position. Since, according to him, people are assemblies of homunculi (whether, neurones, modules, or interests), the relation between people and societies is mirrored by the relation between such homunculi and individual people; a state of things which suggests that microeconomics should analyse the relation between homunculi and individual people as well as the relation between people and societies. However, a well-grounded analysis of the way these sub-personal agents are more or less coordinated and related to the whole person might be a great triumph for behavioural psychology and for neurobiology, but not for economics as such, even if neurobiologists use RPT or other economic tools. Here it comes, the second and more important of Ross's objections, which radically rejects the traditional distinction between the respective domain of economics and of psychological and cognitive sciences.

Ross even presents his conception as based on what is probably the most respected tenet of the methodology of economics, namely Robbins's famous definition of economics as "the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses" (Robbins 1935, 16; quoted by Ross 2005, 87). Three lines further, Ross adds, however, that he drops from it the word "human", a move which allows him to enlarge the scope of economic analysis in such a way that economics becomes the science (with Robbins's provisos concerning ends and scarce means) of the behaviour of insects (and other animals), robots, neurons and interests as well. Those whom Ross pejoratively refers to as "humanists" are apparently those who would object to dropping this quite significant element of Robbins's definition.[20]

Be that as it may, Ross can defend his position by arguing: (1) that Robbins's definition encourages a generalised and formal conception of economics rather than a conception restricted to the questions related to material wealth; (2) that Samuelson has already developed a theory which is neutral from the point of view of the object to which it is applied; and (3) that game theory and a few other techniques have

[20] Davis and Mirowski count among those who Ross considers humanists (Ross 2005, 46, 70, 118, 257, 258, 270).

broken the frontiers between the various sectors of the larger science concerning any kind of behaviour. For sure, such a general science of behaviour cannot be monopolized by economics, but economics can offer some useful tools to it in such a way that an economist may be tempted to intervene qua economist in this larger domain.

If such was the case, it is clear that methodological individualism would not be a methodology adapted to *this* general science, except in its restricted domain specifically concerning the social consequences of agents' actions. However, if it can still be appropriate to invoke methodological individualism (as described above namely as a strategy for understanding the social consequences of human actions) when dealing with *human* microeconomics, it is not because the latter benefits from an indefensible epistemological privilege, but, because, human beings—the only "systems" who have to understand something—*need* to understand why the consequences of *their individual actions* are typically incorporated in social structures that escape them and seem to impose their laws on them. And this is precisely the type of explanation that a methodologically individualistic economics can offer them.

RATIONALITY AND AGENCY

Another aspect of traditional economics that has to be reconsidered is the fundamental principle of rationality, which is necessarily linked to the notion of agency, since it is decisions and actions that are labelled rational or irrational. Throughout the second half of the 20th century, economic rationality was progressively associated with consistency, an association that characterises what Giocoli (Giocoli 2003) has called a "system of relations" in contrast with a "system of forces" in which rationality was rather associated with the notion of maximisation. As is well known, rationality-consistency is closely related to RPT, which taken together constitute the core of the neoclassical economics that Ross defends and that Davis criticises. RPT and rationality-consistency require a type of agency that leads Ross to declare that insects and neurons—but not humans—are straightforward agents, and that brings Davis to look in heterodox economics for a more satisfactory notion of an individual agent.

I suggest that the source of such opposite reactions to rationality-consistency lies chiefly in the way the notion of rationality has evolved with economic theory. A certain concept of rationality was already playing a central role in classical economics, especially when the first theories of the market were developed. However, the rationality involved was a relatively *minimal* rationality,[21] requiring just what is necessary to induce a buyer to stop paying more when the possibility to pay less for the same service becomes clearly available, or to incite a farmer to produce more wheat when the price obtained for a barrel is largely superior to the cost involved and to lower production when the contrary situation prevails. I call this type of rationality minimal because it just requires that people not be so thoughtless to produce more and more units when it is clear that they lose a lot of money every time they produce a new unit, which is roughly similar to requiring someone be only so mindful to turn off the faucet when the bathtub is full! Happily most people have at least this degree of rationality that is enough to allow a market to work more or less properly.

With the so-called marginalist revolution, economists have introduced a more precise notion of rationality. If people are not that thoughtless, why not strictly maximise whatever valuable they can obtain? This seemed to be the logical way to elaborate, with the help of calculus, a much more precise economic analysis. But treating economic agents as utility maximisers relied on a questionable psychology. As it is well known, the next eight decades or so were largely devoted to de-psychologise economic theory, a process that has culminated with RPT.

There is little doubt that this long process going roughly from Jevons to Samuelson's respective contributions corresponds to a tremendous theoretical, if not empirical, progress. There is little doubt that, thanks to the analyses developed during this process, various economic phenomena were literally discovered and others were understood with much greater precision. But this was realised through a set of systematic idealisations of the notions of rationality and the corresponding notion of agency that required forgetting the specific features of the minimal rationality that guides actual human agents.

Human beings rarely maximise and are far from being consistent; they hesitate, make mistakes, change their mind, regret, suffer from myopia, are sensitive to frames in which questions are raised and are

[21] This notion was introduced by Cherniak (1992), but I use it in a slightly different way and context. More generally, the idea of dissociating rationality from maximisation was explored by different economists, the most influential being Herbert Simon with his notion of "bounded rationality". However, maximisation and consistency remain by far the conceptions of rationality that economists evoke the most spontaneously. In any case, in the present context, I refer to minimal rationality uniquely in order to question the way in which Ross derives so many conclusions from his adoption of RPT and strict consistency as the criterion of rationality.

influenced by superficial similarities, but they are not stupid for all that, and they make decisions and take actions in order to reach their ends. This is exactly the kind of rationality that the theories that *explain* economic phenomena require. When they invoke nothing more than this minimal rationality, these theories are protected from the otherwise devastating objections such as those that are raised by experimental economics, whose conclusions are generally that people rarely maximise and are inconsistent without being thoughtless for all that.[22] Idealisations of economics are very helpful, but if we chose one of them—possibly the most distant from the actual behaviour of human beings—as the prototype of rationality and agency, it is clear that human agents will be disqualified as straightforward agents, a status which will be, by hypothesis, reserved to "systems" who satisfy the assumptions of the theory chosen.

It is great that Samuelson's revealed preference theory can be applied to entomology, artificial intelligence, and neurobiology, in spite of being non applicable where *human* behaviour is concerned; however, for those who are specifically interested in the phenomena which are covered by the "humanist" notion of economics, the Samuelsonian agent remains a theoretical construction useful only to illustrate the working of an agent that is idealised to the point of being able to compete with an insect in matter of perfect consistency (or perfect rigidity)!

It seems reasonable to conclude that human economic agents perform actions after being involved in the more or less complex and erratic reflective activity that their consciousness makes possible and that, while they choose certain means judged to be optimal, they rarely aspire to strictly maximise the way bugs can do at their own level. This is in no way denigrating the program of research promoted by Ross, which aims to shed some light on the way the highly complex resources of the brain might help to account for the behaviour of economic agents, straightforward or not. But the question is whether this valuable program of research should be substituted for the more traditional economic program in such a way that the whole methodology of economics, including the very definition of economics, be radically transformed. This brings us back to the question that opened the present paper.

[22] For an attempt to justify such a claim, see Lagueux 2004.

A REVOLUTION OR NOT?

Are we witnessing a revolution in methodology of economics? Note that I found appropriate to raise the question this way, even though neither Ross nor the other authors that I have mentioned at the beginning of this paper have explicitly pretended to promote a revolutionary way to understand economics. In any case, I think that the publication of their respective books testifies that methodologists of economics can no longer simply ignore the questions raised by the integration of this discipline in the context of the rapid development of cognitive sciences and of artificial intelligence.

It is not clear, however, to what extent these developments may have transformed the way economists should treat the questions which were traditionally their own. It is true that the rate of innovation in the methods used by economics is relatively high in the present times, especially with the increasing place occupied by game theory, but, as far as I know, it was mostly classical game theory that was involved in the spectacular transformation of conventional economics in the last two or three decades.

Ross refers more systematically to evolutionary game theory, which can find various applications in the whole domain corresponding to his larger conception of economics, but the impact of this kind of game theory on the questions traditionally treated by economists is less manifest: at least, this impact was not made clear by Ross in the present book, which however must be completed by a forthcoming second volume, more promising on this ground. Indeed, its proper subject-matter, macroeconomics must be reinterpreted with the help of evolutionary game theory in a way that might propose an original solution to the persistent problem of micro-macro relations.

But, for the time being, the proper question to ask is whether microeconomics has been endowed with the type of categories that Rosenberg hoped to see in this discipline. Folk categories such as beliefs, desires, and consciousness may have been tentatively explained with the help of combined neurological and socio-anthropological analyses, but one may wonder in what sense their alleged dependency on intentional stances constitutes a gain from this point of view. The paradigmatic example proposed by Rosenberg was the folk notion of a fish that was replaced by anatomical concepts that "cut nature at the joints" because they are defined on the basis of a genetic analysis in a way that makes them perfectly precise and apt to make accurate predictions possible.

The concept of a (straightforward) agent has been precisely redefined by Ross along the lines proposed by Samuelson, but its direct applications do not concern traditionally defined economic questions. In contrast, the notions of beliefs, desires, consciousness, and selves, as redefined by Dennett and Ross, can hardly be safely described as cutting nature at the joints. It does not seem much more appropriate to refer to a new explanation of economic phenomena which should replace folk explanation like the Copernican scientific explanation of diurnal movement replaced the one provided by folk astronomy. Irrespective of the appreciation we may have of the validity of Ross's new type of explanation, a notable difference concerns the fact that, in this case, it is the explanative concepts of folk economics (beliefs and desires) which are themselves explained otherwise, and not the traditional explanation of economic phenomena (markets, level of prices, cost and production, and so on) which is replaced by a more scientific one, at least in the present book.

May we say, at least, that a serious attempt has been made to replace with scientific and empirical foundations the philosophical bases of traditional concepts used by economists? At some points in his book,[23] Ross insists on the importance of granting primacy to empirical and scientific considerations over philosophical ones. This seems to me a very sound principle to follow when both of these considerations really point in rival directions. If you are interested in the causes of the 1929 economic crisis, I strongly recommend you to turn toward empirical analyses provided by economists or historians specialised in this question rather than toward philosophical speculations about the origin of crises. But such situations are rather rare.

Happily, most philosophers have learned to avoid attempts to offer answers to questions reserved to specialists of an empirical domain. But when the most fundamental questions are involved, the debate is not between philosophical and empirical considerations, but between differently oriented philosophies. Sometimes, the speculations of one of them are more anchored in recent scientific developments, but that does not warrant them superiority as philosophical speculations. The history of philosophy is full of cases of unfortunate philosophical interpretation of various up-to-date scientific theories. The most famous of them, the

[23] For example, Ross 2005, 124.

positivism of Auguste Comte, was conceived by his author as the first philosophy to be really based on scientific and empirical considerations, but it opens the road to series of divagations leading to the proclamation of the dogmas of the religion of Humanity. Don Ross's philosophy is far from being threatened by this kind of divagations; however, it remains a philosophy that intelligently takes account of an impressive quantity of empirical results, but which, for an important part, is highly speculative and controversial. In spite of the fact that it may open new roads for possible inquiries that can put us on the track to potential revolutionary developments of great interest for economics, one cannot conclude that this controversial philosophical contribution, in its present state, should lead economists to massively redefine their concepts, or even to follow its author along the way out of traditional economics, which has been opened by Samuelsonian revealed preference theory.

REFERENCES

Cherniak, Christopher. 1992 [1986]. *Minimal rationality*. Cambridge (MA): MIT Press.

Churchland, Paul. 1984. *Matter and consciousness: a contemporary introduction to the philosophy of mind*. Cambridge (MA): MIT Press.

Davis, John. 2003. *The theory of the individual in economics*. London: Routledge.

Dennett, Daniel. 1981. *Brainstorms*. Cambridge (MA): MIT Press.

Dennett, Daniel. 1991a. *Consciousness explained*. Boston: Little, Brown, and Co.

Dennett, Daniel. 1991b. Real patterns. *The Journal of Philosophy*, 88 (1): 27-51.

Dennett, Daniel. 1993 [1987]. *The intentional stance*. Cambridge (MA): MIT Press.

Dennett, Daniel. 2000. With a little help from my friends. In *Dennett's philosophy: a comprehensive assessment,* eds. Don Ross, Andrew Brook, and David Thompson. Cambridge (MA): MIT Press, 327-388.

Elster, Jon. 1985. *Making sense of Marx*. Cambridge: Cambridge University Press.

Elster, Jon. 1987 [1985]. Introduction. In *The multiple self*, ed. Jon Elster. Cambridge: Cambridge University Press, 1-34.

Giocoli, Nicola. 2003. *Modeling rational agents*. Cheltenham: Edward Elgar.

Hausman, Daniel. 1992. *The inexact and separate science of economics.* Cambridge: Cambridge University Press.

Hodgson, Geoffrey. 2007. Meanings of methodological individualism. *Journal of Economic Methodology*, 14 (2): 211-226.

Lagueux, Maurice. 2001. Individualisme, subjectivisme et mécanismes économiques. *Dialogue*, 40 (4): 691-722.

Lagueux, Maurice. 2004. The forgotten role of the rationality principle in economics. *Journal of Economic Methodology*, 11 (1): 31-51.

Mirowski, Philip. 2002. *Machine dreams: economics becomes a cyborg science*. Cambridge: Cambridge University Press.

Robbins, Lionel, 1935 [1932]. *An essay on the nature and significance of economic science*. London: Macmillan.

Rosenberg, Alexander. 1988. *Philosophy of social science.* Boulder, Colorado: Westview Press.

Rosenberg, Alexander. 1992. *Economics: mathematical politics or science of diminishing returns?* Chicago: The University of Chicago Press.

Rosenberg, Alexander. 1994. What is the cognitive status of economic theory? In *New directions in economic methodology*, ed. Roger Backhouse. London: Routledge, 216-235.

Ross, Don. 1995. Real patterns and the ontological foundations of microeconomics. *Economics and Philosophy,* 11 (1): 113-136.

Ross, Don. 2000. Rainforest realism: a Dennettian theory of existence. In *Dennett's philosophy: a comprehensive assessment,* eds. Don Ross, Andrew Brook, and David Thompson. Cambridge (MA): MIT Press, 147-168.

Ross, Don. 2005. *Economic theory and cognitive science: microexplanation.* Cambridge (MA): MIT Press.

Rutherford, Malcolm. 1994. *Institutions in economics: the old and the new institutionalism.* Cambridge: Cambridge University Press.

Searle, John. 1980. Mind, brains and programs. *The Behavioral and Brain Sciences*, 3: 417-457 [Including Searle's introductory paper: 417-424), Open peer commentary: 424-450, and Searle's response: 450-456].

Searle, John. 1997. *The mystery of consciousness.* New York: New York Review Books.

Watkins, John William Neville. 1953. Ideal types and historical explanation. *British Journal for the Philosophy of Science,* 3 (9): 22-43.

Maurice Lagueux is currently associate professor in the Department of philosophy at Université de Montréal where he has also taught history of economic thought in the Department of economics. His research interests are economic methodology, philosophy of architecture, and philosophy of history. Contact e-mail: <maurice.lagueux@umontreal.ca>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 56-60.
http://ejpe.org/pdf/1-1-art-2r.pdf

Reply to Lagueux: on a revolution in methodology of economics

DON ROSS
University of Cape Town
University of Alabama at Birmingham

I thank Maurice Lagueux for his thoughtful reflections on my book *Economic theory and cognitive science: microexplanation* (henceforth: *ETCS:M*). Given space restrictions, in this comment I won't say much about his criticisms of my use of the intentional stance, except to observe that they seem to rely on equivocation between thinking of beliefs and desires as internal states of people, and as descriptions of relationships between patterns in their behaviour and external circumstances. I think that one must *completely* reject all traces of the former view if one hopes to avoid the sorts of logical conundrums that Lagueux worries about; and I find various passages in his arguments to suggest that he has not. The literature arguing for radical externalism about propositional attitude states is now vast, and I can do no better than refer readers to it.[1] I am surprised that Lagueux still finds objections based on bafflement about recursion—that to be an intentional agent requires that the intentional stance be already possible—partly persuasive. Intentionality comes in a continuum of degrees of sophistication, and has historically expanded its scope incrementally, just like most biological phenomena including life itself.[2]

As for Lagueux's fear that my criterion for ascribing reality to a pattern is "a bit ad hoc", I refer the interested reader to Ladyman and Ross (2007), where this difficult issue is considered with the care and large freight of relevant evidence that it demands.

[1] See especially Keijzer 2001; Morton 2002; Millar 2004; and Melser 2004.
[2] Consider in this context Dennett's *reductio*: Every mammal must have a mammal for a mother. There was a time before there were any mammals. Therefore, there are no mammals.

Here, I will concentrate instead on Lagueux's closing concern, which, based on inference from his title, is seen by him as his main theme. He puts it as:

> the question of whether this valuable program of research [i.e., explaining the psychological and neural influences on economic behaviour] should be substituted for the more traditional economic programme in such a way that the whole methodology of economics, including the very definition of economics, be radically transformed (Lagueux 2008, 51).

He frames the matter this way despite acknowledging that I do not "explicitly pretend […] to promote a revolutionary way to understand economics" (Lagueux 2008, 52). This is an understatement. The first two pages of *ETCS:M* include the following words: "it is not the aim of this book to try to tell economists they should go about their business in a fundamentally different way than they do […] I don't want to advertise myself as promoting—heaven forbid—yet another 'paradigm shift'".

In contrast to Lagueux, I take my book to give more comfort to methodological conservatives than to methodological revolutionaries. I can summarize my interpretation of its core thesis as follows. Most economists have tied at least one hand behind their collective back with respect to answering critics of their standard analytical and empirical methods as a result of philosophical commitment to methodological individualism (MI). That it impedes their capacity for self-defense provides one motivation for dropping this commitment. A more general motivation is that all scientists ought, in general, to steer clear of philosophical commitments that are other than banal, and MI is not banal.

Lagueux claims that MI is, in fact, either anodyne or a straw man. I do not agree. The following description of MI is compatible with Lagueux's stated reason for thinking it important to economics: according to MI, the basic unit of economic explanation—basic in the sense that explanations averting to properties of this unit are templates for complete explanations so far as economics is concerned—is a (spatially and temporally) whole, normal, human person. Thus, as Lagueux says, when we find communities of these units jointly frustrated in their consumption and production ambitions, the economist sets out to explain this as resulting from the interaction of the incentives and constraints that impinge on them as whole individuals.

Recently, a perspective that seems very close to the one Lagueux invokes has been expressed in an uncompromising way by Faruk Gul and Wolfgang Pesendorfer (2008). They argue that there is a fact of the matter about the proper domain of economics: its subject matter is (constrained) *choice*, by whole humans. (Note that I add 'constrained' because an economist would not be concerned, according to their view, with a person choosing to think about purple flowers instead of blue ones. But an economist might well be interested in a person choosing to *pick* purple rather than blue flowers.) Gul and Pesendorfer do not say what they mean philosophically by choice, though they clearly think that classical decision theory has effectively axiomatized it, and to that extent identified it. It seems also fair to attribute to them the idea that choice is some kind of computation of relative costs and benefits, under guidance of a prior notion of what constitutes a solution (e.g., maximizing a utility function or identifying a Nash equilibrium strategy). The crucial polemical claim of their essay is that the economist is professionally interested in *what* people compute as economic agents but not in *how* they compute. On this basis, Gul and Pesendorfer conclude that neuroeconomics is misbegotten, on grounds that discoveries of neuroscience are in principle irrelevant to economics. The same point applies, for the same reasons, to psychology.

As attested by most of the other papers in the volume their paper leads off, Gul and Pesendorfer's perspective is highly controversial among economists. That alone suffices to show that Lagueux is hasty in depicting me as an isolated would-be prophet of change standing well ahead of the methodological herd. In fact, I am rather *more* sympathetic to Gul and Pesendorfer's position than are *most* of the eminent, and often incontrovertibly mainstream, economists who have objected to it. In *ETCS:M*, I too defend the idea of economics as a science concerned with abstract optimization under scarcity. Ultimately, however, I reject Gul and Pesendorfer's extreme separateness thesis for two reasons. First, the sociology of science is such that its institutions do not tolerate completely isolated disciplines. Total abandonment of interest in unity is treated as a symptom of quackery. Second, economists very frequently cannot achieve their ideal of describing phenomena by means of elegant reduced-form models that uniquely estimate quantitative values of dependent variables. Especially with the fall in the price of computation over the past three decades, large structural models that require a great deal of epistemically risky econometrics have become the norm, not the exception, in economics (Harrison 2008; Humphreys forthcoming). Thus economists *do*—constantly—put forward hypotheses that are partly about the 'how' as well as the 'what' of economic computation. In light of this, why should there be a special ban on independent variables that range over neural or psychological properties, where these help to constrain estimations and improve the fit of models?

Once one gets this far, one encounters a vicious undertow if one attempts to cling to MI. The problem is simply that, at the level of both the brain and the whole person—which, as I argue in *ETCS:M*, are *not* the same thing—the computational processes by which choices over whole-person-scale alternatives are computed turn out, empirically, to be more like games than like parametric optimization exercises. One should therefore drop the assumption that the whole person is the basic unit of explanation. (This does not mean that one should seek some *other* such unit; one should just stop restricting oneself with such extra-empirical metaphysics altogether.) Then, I argue in the book, it turns out that nothing that ever should have mattered to economists is sacrificed anyway.[3] The Samuelsonian method defended by arch-conservatives such as Gul and Pesendorfer applies at least as well to sub-personal agents as it ever did to people. It doesn't apply anywhere without mediation through models; *no* material entity is literally and only an economic agent. But many highly useful sciences are about virtual entities.

Lagueux exactly echoes Gul and Pesendorfer when he says, in criticism of me, that:

> a well-grounded analysis of the way these sub-personal agents are more or less coordinated and related to the whole person might be a great triumph for behavioural psychology and for neurobiology, but not for economics as such, even if neurobiologists use RPT or other economic tools (Lagueux 2008, 48).

[3] A good deal of the forthcoming third volume of *Economic theory and cognitive science* is devoted to showing that included in the set of commitments not lost with, but indeed strengthened without, MI is *normative* individualism. A crucial basis for MI in economics has been a mistaken impression that it is essential for standard welfare analysis

To this I respond that insofar as one uses economic tools one is to that extent doing economics, just as, when one applies the quantum formalism to (say) computers one is to that extent doing physics. I reject the thesis, common among philosophers, that the division of labour among the sciences is derived from a set of deep ontological 'joints' at which we try to carve nature. Aspects of the behaviour of many types of systems—commitment to which *as* systems is provisional and *also* not based on metaphysics—is such that they will do things we can *only* predict and explain if we model them as optimizing utility or production functions or playing games. That is why there is a thriving science of economics that should be continuously open to enrichment, but should indeed be spared the violence and dislocation of revolution.

REFERENCES

Gul, Faruk, and Wolfgang Pesendorfer. 2008. The case for mindless economics. In *The foundations of positive and normative economics: a handbook*, eds. Andrew Caplin, and Andrew Schotter. Oxford: Oxford University Press, 3-40.

Harrison, Glenn W. 2008. Neuroeconomics: a critical reconsideration. *Economics and Philosophy*, 24 (3): 303-344.

Humphreys, Paul. [forthcoming 2009]. Computational economics. In *The Oxford handbook of philosophy of economics*, eds. Harold Kincaid, and Don Ross. Oxford: Oxford University Press.

Keijzer, Fred. 2001. *Representation and behaviour*. Cambridge (MA): MIT Press.

Ladyman, James, and Don Ross. 2007. *Every thing must go: metaphysics naturalized*. Oxford: Oxford University Press.

Lagueux, Maurice. 2008. Are we witnessing a revolution in methodology of economics? About Don Ross's recent book on microexplanation. *Erasmus Journal for Philosophy and Economics*, 1 (1): 24-55. http://ejpe.org/pdf/1-1-art-2.pdf

Melser, Derek. 2004. *The act of thinking*. Cambridge (MA): MIT Press.

Miller, Alan. 2004. *Understanding people: normativity and rationalizing explanation*. Oxford: Oxford University Press.

Morton, Adam. 2002. *The importance of being understood: folk psychology as ethics*. London: Routledge.

Ross, Don. 2005. *Economic theory and cognitive science: microexplanation*. Cambridge (MA): MIT Press.

Don Ross is professor of philosophy, and professor of economics, at the University of Alabama at Birmingham; as well as professor of economics at the University of Cape Town in South Africa. He specializes in game theory; picoeconomics and neuroeconomics of learning and addiction; industrial economics in Africa; philosophy of science; and the historical and philosophical foundations of economics.
Contact e-mail: <dross1@uab.edu>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 61-86.
http://ejpe.org/pdf/1-1-art-3.pdf

The impossibility of finitism: from SSK to ESK?

DAVID TYFIELD
Institute for Advanced Studies, Lancaster University

Abstract: The dramatic and ongoing changes in the funding of science have stimulated interest in an economics of scientific knowledge (ESK), which would investigate the effects of these changes on the scientific enterprise. Hands (1994) has previously explored the lessons for such an ESK from the existing precedent of the sociology of scientific knowledge (SSK). In particular, he examines the philosophical problems of SSK and those that any ESK in its image would face. This paper explores this argument further by contending that more recent literature in SSK exposes even deeper philosophical problems than those identified by Hands. Meaning finitism has emerged as the philosophical core of SSK. An examination of the profound problems with this position is used to show that an underlying extensionalism is the root of SSK's intractable philosophical difficulties, and to illustrate the entirely different approach of a critical philosophy that is advocated in its place. In this way, the project of an ESK is shown to depend upon a critical philosophy.

Keywords: economics of science, sociology of scientific knowledge (SSK), meaning finitism, extensionalism, critical philosophy

JEL Classification: B40, Z13

The dramatic changes currently occurring in the funding and economic imperatives of scientific research have naturally led to an upsurge in interest in an economics of science (e.g., Mirowski and Sent 2008; Mirowski and Sent 2002). Indeed, allied to a number of disciplinary developments in the philosophy of science and economics, recent years have seen the emergence of a plethora of research projects all calling themselves the 'economics of science' but with little in common beyond

AUTHOR'S NOTE: I would like to thank Francesco Guala, Uskali Mäki, the participants of the LSE Conference on Philosophy of Social Sciences 2005, and two anonymous referees, for comments on earlier drafts of this paper. Any errors that remain are, of course, mine.

the name (Sent 1999). Few if any of such projects offer explanations of why these changes have occurred. Nor do they offer theoretical frameworks for rigorous examination of the effects of these changes on the production of scientific knowledge, what may be called an 'economics of scientific knowledge' (ESK). Such a research programme, however, seems to be of exceptional practical importance in an age characterized by the social penetration of, and dependence upon, scientific knowledge. Indeed, there have been widespread expressions of concern regarding the potentially corrosive effects of the deepening presence of economic incentives in scientific research.[1] There is thus a significant gap calling for this ESK to be established.

As Hands (1994) notes, the sociology of scientific knowledge (SSK) seems to be an obvious place to start for such an ESK. Yet he points out that any SSK-based ESK would need to face up to the particular philosophical problems that beset SSK; in 1994, this was revolving primarily around the so-called "problem of reflexivity". However, much of recent debate regarding SSK (and particularly the strong programme of the Edinburgh School) has been concerned with the related issues of meaning "finitism", "interactionist" social ontology, and Kripke's Wittgenstein, marking a definite shift in the philosophical debate from earlier concerns about reflexivity.[2] Indeed, finitism must now be acknowledged as a (or perhaps even *the*) central element of the model of science & technology studies (STS) associated with the Edinburgh School.

Far from this shift in the debate leading to stronger philosophical grounds, it seems that SSK's philosophical problems are as deep as ever. It is argued here that the philosophical problems of SSK are much more profound than the familiar problems of "reflexivity". In particular, finitism is intelligible only if it is false. It follows that SSK is not merely self-refuting, but, insofar as it holds onto finitism, it is unintelligible. If SSK is even to be able to sustain its *own* research programme, let alone act as role model for an ESK, it must therefore forsake finitism.

---

[1] See, e.g., Boyle 1996, Boyle 2003; Brown 2000; Campbell, et al. 2002; Eisenberg 1987; Eisenberg 1996; Geiger 2004; Heller and Eisenberg 1998; Krimsky 2003; Nelson 2001; Nelson 2004; Newfield 2003; Resnik 2007; Washburn 2005; and references in Mirowski and Sent 2002. For more sanguine assessments of the changes see, e.g., Callon 2002; Greenberg 2001; Shapin 2003; Tijssen 2004.

[2] See, e.g., the debates between Kusch (2004), Bloor (2004), and Sharrock (2004); and between Stueber (2005, 2006), King (2006), and Bernasconi-Kohn (2006). For the purpose of brevity, unless otherwise stated I will be using 'SSK' to refer exclusively to the Edinburgh School in this paper.

Analysis of this problem reveals the root of these philosophical woes to be SSK's implicit philosophical commitment to an "extensionalist" theory of meaning, in which (the development of) the meaning of a term is understood in terms of (the growth of) the set of objects incorporated under that label. Repudiation of this extensionalism demands taking a completely different approach to the philosophical examination of the nature (or ontology) of meaning. This novel approach is effectively "transcendental" or "critical" in nature, involving examination of the necessary conditions of possibility of the premise; in this case, the familiar but problematic possibility of intelligible application of meanings and rules. In short, in order to resolve SSK's philosophical problems so that it can fulfil its potential as an insightful examination of the social nature of scientific knowledge production and act as model for an ESK, the entire approach to the philosophical issues that plague SSK must be rethought.

The paper proceeds as follows. In the next section, I introduce SSK in more depth and explore the centrality of meaning finitism in its philosophical vision. In the following sections, I proceed to explore the philosophical problems with SSK, first reviewing the familiar problems already discussed in the ESK literature, and then turning to the deeper problems regarding meaning finitism and its underlying extensionalism. The latter argument is then developed in a discussion of the resolution of these problems offered by the alternative approach of a critical and transcendental philosophy, before concluding in the final section.

WHAT IS SSK? THE CENTRALITY OF FINITISM

In order to appraise SSK, we must first work out what it is. In brief, SSK is the empirical examination of the generation of scientific knowledge as an open-ended and contingent social process, situated in specific socio-historical locations.[3] As is often (always?) the case, one may perhaps understand its project more clearly by considering what it is against; in this case, that is: Parsonian functionalist sociology of norms; Mertonian sociology of science; and, underlying both of these, what Barnes and Bloor dub 'rationalist' *ex ante* philosophy of science.[4]

---

[3] The literature on SSK is now very large. For overviews, see Barnes and Edge 1982; Barnes, Bloor, and Henry 1996; Barnes 1974; Bloor 1991; Collins 1983; Shapin 1995, and references therein. On the 'social turn' in the philosophy of science more generally following Kuhn 1970; see Hands 2001, chapter 5.

[4] For a discussion of Parsons, see Barnes 1995. For the sociology of science, see Merton 1973. Barnes and Bloor use the term 'rationalist' philosophy of science, for example, in

As such SSK is both a sociological enquiry into the actual generation of beliefs in the social world of 'science', and a naturalistic (if not anti-philosophical) philosophy of science upon which the former is based.[5] The key move in the development of SSK is the shift from the investigation of science for the *truth* (or rationality) of scientific knowledge to the question of why belief A rather than B (or C, or...) is accorded *credibility* by the scientific community.[6] The history of science reveals that the development of scientific knowledge is ridden with controversy. The 'facts' can be, and are, interpreted in many different ways. It follows that the 'facts' themselves cannot determine scientific knowledge. SSK instead turns its attention to the causal explanation of how different beliefs come to be believed. Given that all beliefs must come to be believed, this leads to the "symmetry principle", which demands that both 'true' and 'false' scientific beliefs must be treated equally as regards how people came to accept them (Barnes and Bloor 1982, 23).

How is this position reached? Starting from the Kuhnian insights into the social relativity of beliefs and the theory-ladenness of observation, and the broader changes in post-positivist (e.g., Quinean) philosophy towards a non-foundational epistemology, SSK argues that whether our beliefs are true or false is entirely inaccessible to us, for we cannot step outside ourselves and our social world in order to compare our beliefs with the world as it is.[7] It follows that there is no ultimate appraisal of scientific knowledge, only the situating of it in *further* scientific understanding of how 'scientific' knowledge is produced and the status of that 'knowledge'.

Barnes and Bloor 1982. The phrase seems to include not only classical logical positivism of the "Received View" (Suppe 1977; Hands 2001) but also post-positivist developments that seek to uncover the rationality of the development of science. Thus Lakatos 1970; Laudan 1977; and Worrall 1990, are all explicitly cited as examples on various occasions. For an extended debate between the positions see Laudan 1981, 1982; and Bloor 1981.

[5] Classic examples of the former include Collins and Pinch 1993; and Shapin and Schaffer 1985. Note also that 'inductivist' in this context means simply that the logic of this process is ampliative and not only logically determined, as per deductive schemas of reasoning.

[6] There is a possible ambiguity in the term 'credibility', noted by Haddock (2004, 3, 5). As I use the term, it refers to the actual social acceptance given to a belief and not the belief's plausibility.

[7] Barnes and Bloor explicitly refer to Quine (1960, 1980) much less often than to Wittgenstein. Nevertheless he is acknowledged as a major source of their work. See, e.g., Bloor 1998, 632; Barnes 1982; and Barnes 1983.

Another way in to the argument proceeds from (a reading of) Wittgenstein's *Philosophical investigations* (2001).[8] This starts from a position in which social life and interaction, including the development of (scientific) knowledge, is a matter of extending meanings, rules, and classes to new instances; a theoretical position called "extensionalism". Pickering (1992, 4) summarizes the resulting argument well:

> Since the central problematic of SSK is that of knowledge, the first move is to characterize the technical culture of science as a single conceptual network, along the lines suggested by the philosopher of science Mary Hesse (1980).[9] Concepts [...] within the net are said to be linked to one another by generalizations of varying degrees of certainty, and to the natural world by the piling up of instances under the headings of various observable terms. When scientific culture is specified in this way, an image of scientific practice follows: practice is the creative extension of the conceptual net to fit new circumstances.[10]

This process of extending the net to new instances, however, is not logically determined by the meaning (or rule, or class) itself. In the famous example deployed by Wittgenstein (2001, §185), for instance, a child is asked to "add" 1 to a particular number, in order to test their understanding of arithmetic. Instead of counting "1, 2, 3, 4...", however, the child continues "1, 11, 111, 1111..." One may rebuke the child for not understanding, but in fact "plus", or any other term, cannot be exhaustively and unambiguously defined so as to make its application always certain and uniquely logically determined. Hence the chastened child may now simply proceed "1, 2, 3, 4, 11, 12, 13..." instead and may continue to offer unexpected variations that fit the further specified and refined requirements of the rule *ad infinitum*.

[8] The validity of this reading is the subject of much of the recent debate. See, for example, the exchange between Bloor (1992) and Lynch (1992a, 1992b), as well as more recent work by Kusch (2004, 2006), with replies from Bloor (2004) and Sharrock (2004). Perhaps it would be more accurate to use the common neologism of 'Kripkenstein' rather than Wittgenstein when referring to SSK's philosophical influences, following Kripke's (1982) exposition of Wittgenstein, though even this differs in important respects from SSK's argument. More on this below, but also see Bloor 1997.

[9] For 'Hesse nets' see also Hesse 1976.

[10] Note that the two communities party to this debate, philosophers and sociologists, tend to use the term 'extension' in two slightly different ways. For philosophers, the 'extension' is the extent of the particulars covered by that class. For sociologists (e.g., Pickering 1992, 4) 'extension' refers to the act of extending this class to the next instance. I will be using the term in the philosophical sense.

In short, logical determination of the application of rules cannot reside in the rules themselves, but can only be determined for all instances where they are themselves already specified. As such, an infinitely specified definition is impossible, the resulting theory of meaning is "finitism", so-called because at any one time the existing extension of a meaning is finite, and it is precisely because of this that extending it to the *next* instance is not *already* determined. According to Barnes, et al. (1996), finitism may therefore be defined by five criteria, namely:

1) Future applications of terms are open-ended;
2) No act of classification is ever indefeasibly correct;
3) All acts of classification are revisable;
4) Successive applications of a kind term are not independent; and
5) The applications of different kind terms are not independent of each other.

In short, this presents an inductivist account such that:

> a class is its accepted instances at a given point of time: those instances are the existing resources for deciding what else belongs in the class, the available precedents for further acts of classification, the basis for further case-to-case development of classification (Barnes, et al. 1996, 105).[11]

This argument is generally used to argue that, in the absence of determination of future applications by existing meanings, there is no (private, mentalistic) fact of the matter regarding what is meant by a proposition; what may be called "meaning scepticism".

In the case of SSK, however, the particular application of this argument regards the process of science, with such a model taken to represent the development of all scientific knowledge. This leads to the conclusion that 'philosophy', which attempts to explain how the development of science is a rational process determined by the internal logic of scientific knowledge, is entirely wrong-headed, attempting the impossible. Nor does SSK shy away from the radical implications of this thesis. Thus it is argued that logic itself cannot be deductively, i.e., logically, justified; for the meaning of the logical operations themselves

[11] See also Barnes and Bloor 1982, 39.

are classifications whose extension is also open-ended.[12] It follows that, *pace* 'rationalist' philosophy of science, neither logic nor the empirical evidence determines the development of science.

If this is the case, it follows that something else must determine what scientists believe and how these beliefs change. SSK's solution is that social interests are the relevant determining factor and thus social science can explain the development of science more generally (Barnes 1982, 35; Barnes, et al. 1996, 29). A strict dichotomy is thus set up between investigating the process of science philosophically (wrong, according to SSK) and sociologically (right). As Mäki (1992) claims, this is a radically pro-science, even scientistic programme, in which science is to be explained by more science, and there is never deemed to be any need for philosophical justification.[13]

In order to be able to examine the empirical and contingent process of knowledge production as a social process, SSK also needs a social ontology that can make sense of the contact between social factors and the production of science, thus conceived. This takes us to the second element of SSK's argument—set against Parsonian functionalism—namely the social ontology of "interactionism", so named because social 'reality' is argued to be the outcome of the concrete interactions of actual (sociable and mutually-susceptible) individuals.

Interactionism is effectively a social ontology of finitist social rules. It acknowledges the experience of apparently irreducible social facts, particularly as social rules and norms, and so rejects methodological individualism. But these social rules are not accorded ontological status as 'real', and so reified as in Parsonian functionalism, because the apparent intransigence of society is simply the result of taking too narrow a perspective (King 1999a, 1999b, 2006).

Clearly, social rules are meaningful or else they could not be followed (nor transgressed) by human agents. However, given the picture of meaning discussed above, the application of a social rule in any given instance is precisely to extend a rule so as to include a new particular. Finitism shows, though, that the pre-existing meaning of the social rule cannot logically determine this process. It follows that social

[12] See Barnes, et al. 1996, 198, et seq., for consideration of the paradox of the heap as the *reductio ad absurdum* of *modus ponens*; also see Barnes and Bloor 1982, 42.

[13] Barnes (1982, 38) calls SSK a "totally naturalistic approach to semantic problems". Hands (1997, 2001) also argues that SSK is simply a philosophical naturalism, like those deferring to biology or cognitive science.

rules cannot determine—nor therefore explain—any apparent 'following' of the rule nor any other associated social process.

For interactionism, therefore, social rules are merely the finitist precedent produced by past concrete interactions of individuals. The resulting social ontology is 'interactionist' in that it consists of the output of the negotiations and consensus of all the interactions of humanity throughout history regarding the extension, and hence meaning, of 'social rules', i.e., a conception of 'Social life as bootstrapped induction' (Barnes 1983). From the perspective of any one individual, therefore, social reality will seem given and real, but in fact this is simply because the social 'reality' confronting us is the result of the interactions of all the rest of humanity, which are obviously always greatly beyond our individual control.

Taking these two strands of analysis of science and social ontology together, then, what is the effect of this argument as regards SSK's empirical and sociological program for studying the interaction of science and society? If we acknowledge that both, social rules and scientific theoretical propositions, are meanings (part of the conceptual 'Hesse net') and that these are only extended 'inductively', it follows immediately that the very *content* of scientific knowledge will also be responsive (however indirectly) to the social positioning, and hence to the particular understanding associated with given social interests, of the scientists.

Furthermore, given that there is only ever comparison of beliefs *within* the net of meaning and so no discrimination of 'true' and 'false' beliefs by comparing them directly with the world, social factors can be seen to feature in explanations of *all* scientific knowledge and not just lapses or corruptions of the 'pure' logic of scientific discovery through reference to perversion of the specifically scientific social norms. It follows that, as regards the third and final limb of SSK, Mertonian sociology of science is seen to be wrong in the 'rationalist' assumption of a scientific method and its consequent exclusive focus on the social conditions necessary for the emergence of the particular social norms that characterize the institution of this disinterested scientific enquiry.[14] For SSK, such sociology of science does not go far enough in its employment of sociological analysis in science, i.e., right into the heart

[14] That is, the (in-?)famous four norms of 'Disinterestedness', 'Communism', 'Scepticism', and 'Universalism': see Merton 1973; and the discussion in Hands 2001, 180, et seq.

of scientific *knowledge* and not just concerning the institutional norms of 'science'.

In summary, in the context of massive changes to the economics of science, examination of the impact of economic conditions on the production of scientific knowledge—an economics of scientific knowledge (ESK)—would seem to be extremely important. SSK seems to afford the examination of the interaction of social beliefs and (the development of) scientific knowledge itself in just the way we are seeking for such an ESK. However, were we to consider an SSK-based ESK, we must immediately acknowledge the significant problems with SSK, to which we now turn.

PROBLEMS WITH SSK 1: THE FAMILIAR PROBLEM OF REFLEXIVITY

Probably the most high profile of SSK's theoretical problems is its perennial problem of reflexivity, as it has been discussed in earlier examinations of the suitability of SSK for ESK (Hands 1994). Indeed, "all of the authors involved in the recent SSK feel impelled to give *some response* to the question of reflexivity and the relativism (that many suggest) it implies" (Hands 1994, 93, original emphasis). Furthermore, "what tends to happen [in SSK studies] is that the sociological theories and (anti) philosophical arguments upstage" its empirical work (Hess 1997). But SSK's anti-philosophical naturalism is so domineering precisely because of the intractable philosophical and theoretical problems it throws up. If we are to resolve these problems and fulfil SSK's promise as an examination into the interaction of social factors and the production of knowledge, then we must pay some explicit attention to these philosophical problems and their origins.

What, though, is the problem of reflexivity? As Hands summarizes it:

> Many of the advocates of the SSK claim to undermine the hegemony of the natural sciences by showing that what is purported to be objective and 'natural' is neither one of these things, but rather simply a product of the social context in which it is produced. If this is true for all human inquiry, then it must be said for the SSK as well; this makes everything socially/context dependent and thus *relative*. (Hands 1994, 92, original emphasis).

It follows that there would be no grounds, other than social happenstance, for accepting any belief, and this includes SSK *itself*.

Hence the "problem of reflexivity" is that if the SSK argument is correct, we have no grounds to accept SSK itself.

I agree with this point (though it is made rather too quickly here, as we shall see), but I do not draw the same conclusions as Hands. For Hands (1994, 96) concludes that the problems of reflexivity of SSK are "not so great as to deter entry" into an economics of science in SSK's footsteps. Rather, he sees the experience of SSK as informative, offering cautionary tales about the 'wilderness' through which it has walked and for which economics of science must also steel itself (Hands 1994, 97).

But on what grounds can Hands counsel that reflexivity does not present such a problem for SSK so as to rule out economics of science *ab initio*? For when Hands writes that: "Those involved in the SSK have travelled through much of this wilderness [of reflexivity problems and philosophical disorientation] before us, and to neglect their signposts would surely be a folly" (Hands 1994, 96), this can only be read so as to license a recommendation to follow them on the condition that SSK has actually travelled 'through' the wilderness and not merely 'into' it, i.e., it must have come out the other side. SSK's route must take us somewhere worth travelling to.

It is by no means clear to me that SSK is not, philosophically, still wondering adrift. Indeed, to be fair to Hands, his more recent writings on SSK and economics of science (Hands 1997, 2001) do not make such a bold claim as regards the 'role-model' SSK can provide, perhaps precisely because the intervening period has seen merely an exacerbation of this problem as parts of science & technology studies take ever-more outlandish stances in an attempt to deal with it. Indeed, the relative philosophical *conservatism* of SSK is a major reason that I have chosen it in particular as the STS tradition addressed in this paper, with the philosophical critique being offered applying *a fortiori* to other, more radically anti-philosophical STS perspectives.

But it follows that if SSK is still stranded, then surely the best signpost to follow would not be those SSK has posted that lead nowhere, but the one that says 'Danger, Wilderness Ahead, Do Not Enter'. The only other alternative is that SSK is, like democracy according to Churchill, the worst option apart from all the others. Nonetheless, for this to be the case, two points must be established: (1) just how bad it is, for it may be that *anything* would be better, even the status quo; and (2) what the alternatives are, if there are any. In answering these questions, we will also see how the problem of reflexivity arises from the deeper problem of SSK's finitism.

Let us consider each of these points in turn. First, it may be retorted that this argument assumes that the wilderness is a particularly inhospitable place—that reflexivity presents a particularly devastating problem for SSK—and this is not the case. Certainly, this line of argument is perfectly defensible given one reading of the reflexivity problem. This states that the relevant criterion for assessment of scientific knowledge is its credibility, and that this is the case no matter whether the belief is in fact, coincidentally, 'true' or 'false'. SSK itself, therefore, must also be susceptible to this kind of reflexive investigation, which would show how social factors have influenced its acceptance by some groups and rejection by others. But this requires only that the credibility conferred to all beliefs, whether 'true' or 'false', demand social explanation, and this is not the same as claiming that there is no difference between 'true' and 'false' beliefs, which *would* lead to reflexivity being a problem.

Thus stated, it is quite right that the credibility of SSK is a social phenomenon and that this does not entail that accepting beliefs is merely a matter of whim. In this case, the reflexivity is a satisfying, not a negating, one. But then, we have been worrying about nothing! Reflexivity is not a problem at all. There is no wilderness ahead but civilization, science!

Unfortunately, this is clearly not the case, as a more in-depth consideration of SSK shows. To criticise SSK in this way demands particular caution if we are to give it a fair hearing. We have seen SSK's argument is in fact a radical repudiation of mainstream philosophy of science. It is thus no surprise that it has both generated much controversy, and that misinterpretations abound. For instance, it must be appreciated that SSK does not claim, *pace* some vociferous critics, that there are no such things as true or false beliefs; or that there is no way the world is, independent of our knowledge of it. It is only claiming that we cannot *know* (in the traditional sense of having justified belief) whether our beliefs are true or false, and so this cannot feature in any explanation of why a belief is held, hence the symmetry principle. We can have true or false beliefs but this is merely a matter of coincidental correspondence, and this correspondence, or lack thereof, is not accessible to us in any particular case and so cannot count as one of the causes of actual acceptance of that belief.

Nevertheless, even if we are careful about avoiding a straw man, SSK's stance is highly problematic. For instance, consider the argument that there can be no conclusive appraisal of scientific truth, only the shifting allocation of credibility amongst different scientific belief, all *within* the finitist net of meaning and never by direct comparison of meaning and world. We cannot know whether our scientific beliefs are true or false and so we cannot refer to 'truth' or 'falsity' of beliefs: the terms become idle and superfluous. Yet if we cannot take account of truth or falsity, we have no grounds on which to discriminate 'X' from 'not X', so that we can believe both. As such, the 'truth' and 'falsity' of our beliefs is a necessary condition of the possibility of rational judgement, and without judgement we fall prey to an all-consuming relativism that makes all beliefs equally 'defensible'.

In other words, if we cannot refer to 'truth' or 'falsity' (as per symmetry), we must forsake altogether all use of these concepts, and *this includes tacit presupposition* as well as explicit usage. But this rules out rational judgement and so abandons us to relativism. In the case of SSK, this relativism is simply displaced into social terms so that the social context 'decides' what is and what is not 'knowledge', now redefined as merely "that system of beliefs that a community collectively accepts as knowledge" (Bloor 1991, 3). SSK is thus neither more nor less 'sound' than any competing argument. Nor, crucially, can it provide 'reasons' *at all*, thus belying the pleas of Barnes (1974, 156) that, while not presentable in any particular argument, SSK is to be accepted because "this whole volume is crammed with proffered reasons why its main tenets should be accepted; its justification lies within itself." Such talk of 'justification' is simply ruled out for SSK.

It is crucial to recognize that what is being argued here is *not* that SSK is *avowedly* relativist in this way.[15] Indeed, I have stressed above how SSK's view on truth is not to deny that beliefs do in fact have a truth-value, only that we cannot know it either way in any particular case. However, it is a necessary condition of the possibility of rational judgement that we *can* employ the concepts of truth and falsity in the way that the symmetry thesis prohibits. And symmetry follows ineluctably from finitism and the Hesse-net picture of meaning, because these entail that all beliefs are simply a matter of shifting the credibility accorded to definitions in "the creative [and undetermined] extension of

[15] It is in this sense that Hands's reconstruction of the reflexivity argument above could be said to be too quick.

the conceptual net" (Pickering 1992, 4). The present argument, therefore, is rather that, *regardless* of whether SSK is explicitly judgementally relativist or not, its allegiance to finitism and the symmetry thesis *commits* it to this relativism. Hence no amount of express protest about its rejection of this position can prove the contrary, just as the sceptic, conversely, cannot but display his lack of scepticism once outside the classroom by always leaving the building through the front door and never through the second floor window (Bhaskar 1998).

Should evidence for this theory and practice inconsistency in SSK's program be needed, it is available in abundance. For instance, given that use of the concept 'truth' is a necessary condition of the possibility of a rational discourse, and given further that SSK is participating in such a rational discourse while simultaneously proscribing use of the concept 'truth', it follows immediately that there is an insoluble paradox at its very heart that can only play itself out in interminable fractiousness and disagreement. And this, of course, is exactly what has happened to the wider SSK programme, splintering into mutually incompatible sub-programmes at loggerheads in a lethal but never-ending game of 'epistemological chicken', in which protagonists are challenged to take ever greater risks in the explicit affirmation of such a self-refuting judgemental relativism (Collins and Yearley 1992a, 1992b; Woolgar 1992; Callon and Latour 1992). In short, it seems that SSK throws itself out with its own bathwater.[16] If this is the case, not only do these philosophical problems effectively prevent SSK from sustaining its critical challenge to Received-View philosophy and sociology of science, but also any economics of science that would follow SSK's lead would be beset by exactly the same errors.

But—it can be retorted—you cannot blame SSK for this! For as SSK shows, even logic itself cannot be justified in a non-circular way and all SSK is doing is pointing this out; we cannot blame the messenger. In other words, the logical circularity of deductive logic itself shows reflexive inconsistency to be inevitable. This is the typical defence employed by SSK. Barnes (1974, 39) argues, for instance, that such reflexive inconsistency is merely "the appalling, unresolved difficulties of philosophy" which "do not", and by implication should not, "worry

[16] Hands (1994, 95) uses the phrase, but I note that he is not referring directly to SSK when he does so. See also Rosenberg 1985 for statements to the same effect; and Callon and Latour 1992 for the original joke about the 'Bath school', though from a radically different perspective.

the layman".[17] The tactic thus is to claim innocence by way of universal guilt or to point out that, like it or not, the wilderness is the only option, because it is everywhere. The position is not hopeless, however, according to SSK, because its general inductivism leads to a bootstrapping philosophy, where 'truth' is accorded as a mark of *post hoc* success.

While it is not clear that this offers sufficient defence, it does seem that if we concede SSK's critical points—viz. that meaning is not unique and fully determining so that even logical terms do not logically determine—then there is at least a shift of the burden of proof onto those who would like to claim that reflexive inconsistency is a problem for theories because it is a criterion that discriminates (between consistent, and so tenable, beliefs and those that are not), rather than a ubiquitous and insuperable condition of all discourse. Indeed, in the absence of any demonstration to the contrary, SSK has been able to withstand such criticism despite the manifest inconsistency of its position for the last two decades. I believe that the critical arguments SSK makes against its various 'rationalist' opponents are sound, so it seems the challenge is to show that there is a way out of SSK's problems.

This takes us to the second problem of evaluating the alternatives to SSK's wilderness. The rest of the paper seeks to argue not only that there is such an alternative, but also that we can find it by looking at SSK itself, though not at what it would point out to us explicitly. Indeed, in order to see the alternatives what is needed is not some miraculous philosophical *deus ex machina* but a closer examination of the philosophical *problems* that are central to SSK's project, namely those associated with finitism: the theory of meaning that is pivotal in its conception of the interaction of science and society and that lies at the root of these intractable difficulties. What is needed is to conduct a transcendental philosophical examination of SSK's philosophical problematic itself. As we shall see, however, SSK's philosophical naturalism acts to *preclude* any such examination on its own part, and thus serves to prevent SSK from addressing, let alone resolving, its problems of reflexivity.

[17] Nor, it seems, the social scientist. Similarly, Collins, and Yearley (1992a, 308) argue that we all, not just SSK, find ourselves in an epistemic state of "permanent insecurity". See also Barnes and Bloor 1982, 41; and Bloor 1998, 629. The problem with such statements is not that they are *wrong* but, like all such sceptical positions, that they are hugely overstated.

PROBLEMS OF SSK 2: THE NOVEL PROBLEMS OF EXTENSIONALISM

Let us, therefore, take a closer look at finitism. SSK sets itself against the thesis that the development of science is determined by the meanings of the propositions formulating the proto-scientific laws to be tested by empirical observation. Against this thesis, it argues that because scientific laws can be interpreted in numerous ways, they do not have unique meaning regarding their application or testing in any given case. Instead, SSK argues positively, the infinite number of extensions logically compatible with the existing set shows that the development of scientific knowledge is unconstrained by the meaning of the proto-scientific laws, which merely act as 'precedents' facilitating any subsequent inductive determination of its extension.

The argument against 'rationalist' philosophy of science is cogent, but the derivative positive conclusion of finitism simply does not follow. That there is not one unique determinate meaning to any given proposition or rule does not entail that it can mean anything, but that it has many meanings, and 'many' does not equal 'any'.[18] For SSK to be persuasive here, we must overlook this step, or be presupposing something that acts as a minor premise to validate the inference, by justifying the false dilemma of theories, rules, and the like, either having a unique meaning or meaning anything at all.

Similarly, consider the argument regarding social ontology and rules. The argument for finitism establishes that norms and rules cannot be formulated with sufficient precision to obviate the possibility of their systematic 'misunderstanding', giving them a logically consistent but alternative interpretation to that which is commonly socially accepted. Given that Parsonian norms are supposed to be such clearly-formulated and *sui generis* rules, it is plain that these could not possibly determine our social interaction in the way Parsons claims. Once again: so far so

[18] Compare to comments by Mermin (1998, 610) that: SSK's stretching of the related point of the underdetermination of theory by evidence to its radical conclusions overlooks the fact that it is "a trivial logical point [that] almost entirely misses the actual character of scientific practice." For "the problem confronting physicists […] is rarely an overabundance of plausible theories [but…] is to find even a single *reasonable* theoretical structure […]" (original emphasis). Bloor's (1998) response to this is particularly revealing for its characteristic shifting of meaning that conceals disagreement as agreement: "The problem is not that lots of theories fit perfectly: it is that nothing ever works properly […] so why do we prefer this imperfection to that imperfection?" This entirely distorts Mermin's point, however, for he has not said that "nothing" ever works but that it is *difficult, if possible*, to find something that does. See also Laudan 1998; and Hacking 1992, 55.

good. But once again, SSK presents a false dilemma of either rules determining social interaction or social interaction determining rules to derive the latter as its positive conclusion.

Given that these positive claims are finitist conclusions, therefore, we can conclude that finitism is ungrounded, the arguments in its favour resting on false dilemmas. Furthermore, once we dispense with finitism, the problems of reflexivity and relativism do not arise, because in each case these hang on SSK's positive claims, not its critical ones. Thus, for instance, it is the finitist argument regarding 'rationalist' philosophy of science, and the specifically finitist picture of meaning as extending 'Hesse nets' to new instances, that leads to ruinous relativism, as we have seen. But why does SSK consider finitism and the false dilemmas that justify it to be compelling? The answer to this question lies in SSK's extensionalist theory of meaning, on which its entire problematic is built.[19]

We can readily accept that extending rules/theories to the next instance is not logically determined by their existing extensions. But this only licenses finitism if the development of rules/theories is identified with this process of extending extensive sets, i.e., given extensionalism. Nevertheless, from this extensional theory of meaning and its development, it follows that the only conceivable form of determination is *logical* determination, which is binary: in a given case either there is logical determination or there is not, as when an argument is deductively valid or not.

Thus it is clearly extensionalism that licenses the false dilemma of uniquely determining or wholly indeterminate meaning; of meaning either one thing or anything/nothing. For given that the determination of the development of meaning is *logical* determination, the existing meaning (i.e., extensive set) determines how it is developed either *uniquely* or *not at all*. With 'meaning', 'development', and 'determination' thus defined (as 'extension', 'extending', and 'logical

[19] Ironically, Barnes (1982) explicitly contrasts finitism and extensionalism, and affirms that "an alternative to extensional semantics is essential" (Barnes 1982, 24). However, the question he then goes on to address is "what determines whether or not a concept properly applies to its next instance?" The terms of his analysis of meaning are thus explicitly extensionalist, for it is extensionalism alone that sets this as the relevant question regarding semantic issues. Barnes, et al. (1996, 105) also explicitly state that "a class is its accepted instances at a given point of time." It follows that Barnes's whole investigation is conducted within an extensionalist paradigm, even if his conclusions are not aligned with the main protagonists in the 'extensionalist' philosophical debate (e.g., Putnam as 'realist', and Searle as 'description theorist' in this case).

determination' respectively), it follows also that meaning does not determine its own development at all, i.e., finitism. But notice that SSK has not *concluded* that there is no such thing as intrinsic meaning, for its non-existence is assumed in the extensionalism of the very formulation of the question it addresses.

Therefore, it is SSK's prior commitment to extensionalism, what underlies the false dilemmas that would justify their positive conclusions. Extensionalism, however, is a common philosophical position that, in SSK's useful terms, is accorded much credibility by the social community called 'philosophy'. The question is thus: why must it be discarded?

TOWARDS AN ALTERNATIVE PHILOSOPHICAL APPROACH

The first point to note in response to this question is that extensionalism is responsible for the positive claims of SSK and, as we have seen, these are blighted with an inimical relativism that renders them insupportable and self-refuting.[20] Conversely, were we to dispense with extensionalism, we can retain the valuable critical arguments of SSK without being forced by the false dilemma to take that next step; one that then leads to the extinction of their critical challenge. This surely provides at least *prima facie* evidence, as a negative argument, to challenge extensionalism.

But positive arguments in favour of taking intensionality seriously are also easily marshalled. For, in each case, the intelligibility of SSK's claims rests on an unchallenged ambiguity that allows the tacit presupposition of what it is expressly denying to go unnoticed. In other words, it is not merely the case that SSK refutes itself but also that its positive claims, if true, would be unintelligible. It follows that if we understand the claims, they must be wrong.

Consider the Wittgensteinian (2001, §185) example of the '+1' rule, discussed above. SSK argues that extending the rule, and its extension at any given point in time, is not logically determined; no formulation of the rule can ever be sufficiently precise that it rules out all interpretations but one (the one intended). Thus '+1', and every other rule, is actually indeterminate; anything goes (e.g., Barnes, et al. 1996).

[20] There are legion other examples of direct inconsistency in SSK's pronouncements, e.g., regarding finitism and consistency see Barnes 1982, 38, where finitism is argued to be the result of consistency; and the various Barnes and Bloor quotations noted above at footnote 17 regarding the impossibility of consistency given finitism. Laudan (1998, 321) also notes the blithe inconsistency of SSK.

But if such a rule can mean anything at all—if it can mean 'punch the teacher' or 'make porridge' or *anything else at all*—then it is utterly without content and so is totally unintelligible. Alternatively, we can see that we can only understand the point being made by the '+1' example because we *do* understand the rule in a determinate way and are struck by the possibility of understanding in a *different* way. But this is miles away from saying that the rule has no determinate meaning at all, in which case it would be totally unintelligible, as would *any statement about it* including *SSK's argument itself*.

Thus consider, for instance, trying to make SSK's finitist point using the nonsense example 'trung tring' instead of '+1'. I say 'trung tring' to the difficult child and he proceeds to stand on his chair, or cry, or leave the room, or stare at me blankly, or else. Clearly, the finitist philosophical point cannot be demonstrated by this example because we do not already have some idea of what the rule means against which to compare the supposedly unusual interpretation of the child. Nor, therefore, can we say anything about this rule unless and until it does have some meaning for us. In other words, if '+1' means anything at all, then SSK cannot intelligibly make any argument about it or making use of it. The only possible conclusion is that the extensionalist picture on which this argument is based is *not* what is happening in meaning use. As such, we can show that it is a necessary condition of the possibility of rule or meaning use that these are *not*, at any given time, totally unlimited in application.

Exactly the same criticism may be made of SSK's argument that the rules or theories underdetermine the development of scientific beliefs, which leads to the conclusion that social factors, such as interests, are the determining factor. But what are these social factors? According to the social ontology of SSK itself they can be no more than other social rules (e.g., the rule 'make money' or 'find a partner') 'known' by the individual, but these cannot, *ex hypothesi*, determine the social interaction that extends the other social rule. It follows that on SSK's model, there is no determinate connection between social interaction and social rules, in which case it seems misleading to argue that there are any social rules in the first place. Conversely, their conclusion is only plausible if social rules are intelligible, in which case they do have intrinsic, determinate content, i.e., they are intensional and not just extensional.

A perfect illustration of this (which is also particularly relevant to issues of an ESK) is presented in Barnes et al. (1996) in discussion of a court case concerning a patent for aniline dyes.[21] It is argued that what decides the outcome of the patent case is not the result of balancing of arguments, the "weighing" of "weightless quantities", but is "a consequence of the balance of power", i.e., as a matter of which decision will best serve the interests of the constitutional order and/or the legal establishment itself (it is not specified which). This sounds terribly scandalous, and of course it is welcome to remind us that a totally 'disinterested' judiciary is just nonsense, but on SSK's account, when there is *nothing* in the arguments made before the court that determines the outcome, we cannot even *distort* the decision by seeing how it would interact with 'our' or other power-political interests. Nor is there even any point in a judicial decision, because how the decision itself is then interpreted, whether in its implementation or in its future use as jurisprudential precedent, is entirely undetermined. Note that this also immediately makes a mockery of criteria 4 and 5 of finitism (mentioned above) that finitist meaning functions like a system of precedent, because finitism actually deprives such a system of any material that could ever act as such: precedents must *constrain* as well as enable; yet finitism systematically denies the former.

Similarly, unable to see any implications of 'knowing' that we have particular 'interests'—for there are none—then all decisions become impossible, as, ironically, Barnes et al. (1996, 124) themselves notice. In short, SSK cannot argue simultaneously that rules underdetermine and that the deficit is made up by other rules, yet its plausibility trades on this systematic ambiguity. SSK therefore argues that it can explain how scientists, judges, and the rest of us choose between belief A and belief B, but overlooks the fact that it cannot explain how we can choose *at all*, because, in its repudiation of philosophy and ontology, it denies the *material cause* of a relatively autonomous intensional meaning upon which all such agency depends.

Similar arguments can be presented for all of SSK's positive finitist claims because in each case close reading shows that SSK is arguing as follows:

[21] Note also that, as presented, the argument suggests that the court is deciding on the actual chemistry and not merely on the patent dispute. It seems that Mermin (1998) also notices this.

(1) *not A* (critique of 'rationalism', compatible with extensionalism but also with intensionality)
(2) *either A or B* (the disjunction of the false dilemma, from extensionalism)
Therefore
(3) *B* (unintelligible if true, intelligible only if false)

Thus if we are to be able to understand SSK's claims at all, we must admit from the outset that rules do have determinate content that *constrains* (as well as enables), but does not itself fully *determine*, our subsequent action, including development of the rule itself. This yields the distinction between determin*ate* (i.e., a *material* cause that is constraining in the instant, but transformable in the future) and (fully, uniquely) determin*ing* meaning.[22] But this is just to admit an intensional theory of meaning that acknowledges internal relations of necessity between different meanings, hence rendering meaning relatively resistant to our use of it so that we cannot simply do as we please—*even collectively*—with meaning, *pace* SSK. And this, in turn, is to refuse to identify meaning with extensive classes and extending them to new instances.

We have, therefore, repudiated the extensionalist theory of meaning that is the root of SSK's philosophical problems. In order to take this step, though, we have had to employ a transcendental, i.e., a specifically *philosophical*, argument, examining the necessary conditions of intelligibility of the philosophical problem itself. Such a *sui generis* philosophical analysis is exactly what SSK repudiates in its philosophical naturalism. However, this step is necessary for SSK if it is not to be forced to choose sides on the false dilemmas that arise from its philosophical neglect, in each case forcing SSK into a self-refuting position as we have seen. Thus SSK sees only that if extensionalism is right, then it is wrong, but sees it as inevitable, because it does not see that intensionality is presupposed anyway so that extensionalism is simply wrong: the initial premise is false and the argument collapses. Conversely, if we employ a transcendental approach as opposed to one

[22] Note also that, because meanings do not themselves fully determine, they also do not have agency, just as SSK correctly claims in its criticism of its various opponents. Instead, agency is reserved to human meaning-users in whom meaning resides, but here agency amounts to *changing* and not *creating* meanings (Bloor 1997, 70). There must be meanings in the first place or there is no possibility of an agent having *any* understanding, which in turn rules out agency, including the agency to manipulate and create new meanings.

starting from the presumption of an extensionalist model of meaning, we can readily admit the intensionality of meaning and thus secure each of SSK's critical points, while (in the schema above) the step to B or conclusion (3), via the false dilemmas and ambiguities of premise (2), and the insoluble paradoxes that come with that step are avoided.

The refutation of premise (2) by transcendental reasoning, however, demands first that we admit the problems of finitism, for this alone can provide the motivation actually to *examine* the philosophical problems and not merely rest on an anti-philosophical complacency characteristic of naturalist perspectives. That is to say, so long as finitism is accepted one cannot even consent to the problems of reflexivity to be *problems*. For the 'net of meaning' picture makes it evidently absurd to attempt to step outside or beyond the net. Thus it is finitism, and the interactionist social ontology it sponsors, that licences the conflation of knowledge and social structure to the single un-tethered level of the net of meaning, and it is on this basis that problems of reflexivity are simply *accepted* as *irresolvable*. In the context of an underlying philosophical naturalism that precludes the analysis offered here and the admission of the importance of tackling philosophical problems, it follows that in order to begin even to address these problems, finitism must first be rejected.

The fundamental false dilemma underlying SSK is, therefore, that between first or *ex ante* 'rationalist' philosophy and naturalist or sceptical anti-philosophy. Taking its stand against the former, SSK immediately sides with the latter, but thereby finds itself cast into the fogs of relativism, which it then fully embraces in its defiant dismissal of the problems of reflexivity. It is SSK's (anti-) philosophical naturalism that explains the *ex ante* and unquestioned acceptance of extensionalism responsible for its intractable theoretical difficulties; thereby also explaining why SSK, despite its anti-philosophism, paradoxically finds itself dominated by its philosophical dimension, as noted by Hess above. But SSK need not take this option of philosophical naturalism because there exists the third option of a *critical* philosophical perspective, which asks [regarding rules/meaning]: *given* that [meaning use] is possible (at the price of ruling oneself into dumb silence), what is ontologically presupposed by this? With the clear alternative of a critical philosophy, then, we see that this element of SSK's program can, and indeed must, be relinquished.

CONCLUSION

We have investigated SSK and concluded that, as it stands, it does not afford a profitable basis for the development of an ESK, but rather forecloses such a project. Furthermore, we have found this to be caused by the inadequate philosophical understanding at its heart: its extensionalist theory of meaning, which manifests in the problems of reflexivity and meaning scepticism. Confronting this problem forces us to take an entirely different approach, examining the ontological presuppositions of this impossible but apparently ineluctable challenge. This thereby repudiates the fundamental philosophical problematic from which extensionalism itself arises, namely the attempt to provide a watertight philosophical 'solution' to the problems of meaning use that arise when it is treated in terms of logical determination of the application of labels of extensive sets.

Such explicit transcendental philosophical examination, however, also significantly reorganizes the project of a social study of science, without thereby sacrificing its significant critical advantages over alternative research projects; which leave scientific knowledge sealed off, pristine, and inviolable for a wholly separate philosophy of science. Indeed, the exact opposite is the case: it is SSK itself that cannot sustain these critical points because, deprived of any possibility of knowing whether beliefs are true and even of any determinate meaning, it must rule itself into silence, taking even its own arguments with it. Time and again SSK points to important truths but on each occasion it then goes on to snatch them away, denying there are any truths at all.

The implications of such a reorientation of SSK are of central importance for any project of an economics of scientific knowledge. This is not only in the sense of offering a model that is itself not riddled with problems, but also because an ESK, if it is to do anything at all, must be able to offer a critique of how and where the imposition of economic imperatives on scientific research has a detrimental effect on the "scientific knowledge" thereby produced. In the age of the ubiquitous penetration of such economic issues into research, the failure, or rather refusal, of any SSK-inspired ESK to make such judgements would be a grievous loss to social criticism.

REFERENCES

Barnes, S. Barry. 1974. *Scientific knowledge and social theory.* London: Routledge and Kegan Paul.

Barnes, S. Barry. 1982. On the extensions of concepts and the growth of knowledge. *Sociological Review*, 30 (1): 23-44.

Barnes, S. Barry. 1983. Social life as bootstrapped induction. *Sociology*, 17 (4): 524-545.

Barnes, S. Barry. 1995. *The elements of social theory*. London: UCL Press.

Barnes, S. Barry, and David Bloor. 1982. Relativism, rationalism and the sociology of knowledge. In *Rationality and relativism*, eds. Martin Hollis, and Steven Lukes. Oxford: Blackwell, 21-47.

Barnes, S. Barry, David Bloor, and John Henry. 1996. *Sociology of scientific knowledge*. London: Athlone Press; Chicago: University of Chicago Press.

Barnes, S. Barry, and David Edge (eds.). 1982. *Science in context: readings in the sociology of science*. Milton Keynes: Open University Press.

Bernasconi-Kohn, Lorenzo. 2006. How not to think about rules and rule-following: a response to Stueber. *Philosophy of the Social Sciences*, 36 (1): 86-94.

Bhaskar, Roy. 1986. *Scientific realism and human emancipation*. London: Verso.

Bhaskar, Roy. 1998 [1979]. *The possibility of naturalism*. London: Routledge.

Bloor, David. 1981. The strengths of the strong programme. *Philosophy of the Social Sciences*, 11 (2): 199-213.

Bloor, David. 1991 [1976]. *Knowledge and social imagery*. London and Boston: Routledge and Kegan Paul.

Bloor, David. 1992. Left and right Wittgensteinians. In *Science as practice and culture*, ed. Andrew Pickering. Chicago: University of Chicago Press, 266-282.

Bloor, David. 1997. *Wittgenstein, rules and institutions*. London and New York: Routledge.

Bloor, David. 1998. Changing axes: response to Mermin. *Social Studies of Science*, 28 (4): 624-635.

Bloor, David. 2004. Institutions and rule-scepticism: a reply to Martin Kusch. *Social Studies of Science*, 34 (4): 593-601.

Boyle, James. 1996. *Shamans, software and spleens: law and the construction of the information society*. Cambridge (MA) and London: Harvard University Press.

Boyle, James. 2003. The second enclosure movement and the construction of the public domain. *Law and Contemporary Problems*, 66 (1-2): 33-73.

Callon, Michel. 2002. Technology, politics and the market: an interview with Michel Callon, conducted by Andrew Barry and Donald Slater. *Economy and Society*, 31 (2): 285-306.

Callon, Michel, and Bruno Latour. 1992. Don't throw the baby out with the bath school: a reply to Collins and Yearley. In *Science as practice and culture*, ed. Andrew Pickering. Chicago: University of Chicago Press, 343-368.

Campbell, Eric, Brian Clarridge, Manjusha Gokhale, Lauren Birenbaum, Stephen Hilgartner, Neil Holtzman, David Blumenthal. 2002. Data withholding in academic genetics: evidence from a national survey. *Journal of the American Medical Association*, 287 (4): 473-480.

Collins, Harry M. 1983. The sociology of scientific knowledge: studies of contemporary science. *Annual Review of Sociology*, 9: 265-285.

Collins, Harry M., and Trevor Pinch. 1993. *The Golem: what everyone should know about science*. Cambridge: Cambridge University Press.

Collins, Harry M., and Steven Yearley. 1992a. Epistemological chicken. In *Science as practice and culture*, ed. Andrew Pickering. Chicago: University of Chicago Press, 301-326.

Collins, Harry M., and Steven Yearley. 1992b. Journey into space. In *Science as practice and culture*, ed. Andrew Pickering. Chicago: University of Chicago Press: 369-389.

Eisenberg, Rebecca. 1987. Proprietary rights and the norms of science in biotechnology research. *Yale Law Journal*, 97 (2): 177-231.

Eisenberg, Rebecca. 1996. Public research and private development: patents and technology transfer in government-sponsored research. *Virginia Law Review*, 82 (8): 1663-1727.

Geiger, Roger. 2004. *Knowledge and money*. Stanford: Stanford University Press.

Greenberg, Daniel. 2001. *Science, money and politics*. Chicago: University of Chicago Press.

Hacking, Ian. 1992. Putting agency back into experiment. In *Science as practice and culture*, ed. Andrew Pickering. Chicago: University of Chicago Press, 29-64.

Haddock, Adrian. 2004. Re-thinking the 'strong programme' in the sociology of knowledge. *Studies in the History of the Philosophy of Science Part A*, 35 (1): 19-40.

Hands, D. Wade. 1994. The sociology of scientific knowledge: some thoughts on the possibilities. In *New directions in economic methodology*, ed. Roger Backhouse. London: Routledge, 75-106.

Hands, D. Wade. 1997. Conjectures and reputations: the sociology of scientific knowledge and the history of economic thought. *History of Political Economy*, 29 (4): 695-739.

Hands, D. Wade. 2001. *Reflection without rules: economic methodology and contemporary science theory*. Cambridge: Cambridge University Press.

Heller, Michael, and Rebecca Eisenberg. 1998. Can patents deter innovation? The anti-commons in biomedical research. *Science*, 280 (5364): 698-701.

Hess, David. 1997. *Science studies: an advanced introduction*. New York and London: New York University Press.

Hesse, Mary. 1976. Truth and the growth of scientific knowledge. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association: Symposia and Invited Papers*, 2: 261-280.

Hesse, Mary. 1980. *Revolutions and reconstructions in the philosophy of science*. Bloomington (IN): Indiana University Press.

Hollis, Martin. 1982. The social destruction of reality. In *Rationality and relativism*, eds. Martin Hollis, and Steven Lukes. Oxford: Blackwell, 67-86.

King, Anthony. 1999a. The impossibility of naturalism: the antinomies of Bhaskar's realism. *Journal for the Theory of Social Behaviour*, 29 (3): 267-288.

King, Anthony. 1999b. Against structure: a critique of morphogenetic social theory. *Sociological Review*, 47 (2): 199-227.

King, Anthony. 2006. How not to structure a social theory: a reply to a critical response. *Philosophy of the Social Sciences*, 36 (4): 464-479.

Krimsky, Sheldon. 2003. *Science in the private interest: has the lure of profits corrupted biomedical research?* Lanham: Rowman and Littlefield.

Kripke, Saul. 1982. *Wittgenstein: on rules and private language*. Oxford: Blackwell.

Kuhn, Thomas. 1970 [1962]. *The structure of scientific revolutions*. Chicago: Chicago University Press.

Kusch, Martin. 2004. Rule-scepticism and the sociology of scientific knowledge: the Bloor-Lynch debate revisited. *Social Studies of Science*, 34 (4): 571-591.

Kusch, Martin. 2006. *A sceptical guide to meaning and rules: defending Kripke's Wittgenstein*. Chesham: Acumen.

Lakatos, Imre. 1970. History of science and its rational reconstruction. Reprinted in *The methodology of scientific research programmes: philosophical papers, vol. 1*, eds. John Worrall, and Gregory Currie. Cambridge: Cambridge University Press, 102-138.

Laudan, Larry. 1977. *Progress and its problems: towards a theory of scientific growth*. London: Routledge and Kegan Paul.

Laudan, Larry. 1981. The pseudo-science of science? *Philosophy of the Social Sciences*, 11 (2): 173-198.

Laudan, Larry. 1982. More on Bloor. *Philosophy of the Social Science*, 12 (1): 71-74.

Laudan, Larry. 1998. De-mystifying underdetermination. In *Philosophy of science: the central issues*, eds. Martin Curd, and Jan A. Cover. New York and London: W. W. Norton & Company, 320-353. Previously printed in [1990] *Scientific theories, Minnesota Studies in the Philosophy of Science, vol. 14*, ed. C. Wade Savage. Minneapolis (MN): University of Minnesota Press, 267-297.

Lynch, Michael. 1992a. Extending Wittgenstein: the pivotal move from epistemology to the sociology of science. In *Science as practice and culture*, ed. A. Pickering. Chicago: University of Chicago Press, 215-265.

Lynch, Michael. 1992b. From the "will to theory" to the discursive collage: a reply to Bloor's "Left and right Wittgensteinians". In *Science as practice and culture*, ed. Andrew Pickering. Chicago: University of Chicago Press, 283-299.

MacKenzie, Donald. 1981. Notes on the science and social relations debate. *Capital and Class*, 14: 47-60.

Mäki, Uskali. 1992. Social conditioning in economics. In *Post-Popperian methodology of economics: recovering practice*, ed. N. de Marchi. Boston: Kluwer, 65-104.

Mermin, N. David. 1998. The science of science: a physicist reads Barnes, Bloor and Henry. *Social Studies of Science*, 28 (4): 603-623.

Merton, Robert K. 1973. *The sociology of science: theoretical and empirical investigations*. Chicago: Chicago University Press.

Mirowski, Philip, and Esther-Mirjam Sent (eds.). 2002. *Science bought and sold*. Chicago: University of Chicago Press.

Mirowski, Philip, and Esther-Mirjam Sent. 2008. The commercialization of science and the response of STS. In *New handbook of STS*, eds. Edward Hackett, Judy Wacjman, Olga Amsterdamska, and Michael Lynch.Cambridge (MA): MIT Press, 635-689.

Nelson, Richard. 2001. Observations on the post-Bayh-Dole rise of patenting at American universities. *Journal of Technology Transfer*, 26 (1-2): 13-19.

Nelson, Richard. 2004. The market economy, and the scientific commons. *Research Policy*, 33 (3): 455-471.

Newfield, Christopher. 2003. *Ivy and industry: business and the making of the American university, 1880-1980*. Durham (NC) and London: Duke University Press.

Pickering, Andrew. 1992. From science as knowledge to science as practice. In *Science as practice and culture*, ed. Andrew Pickering. Chicago: University of Chicago Press, 1-28.

Quine, Willard V. O. 1960. *Word and object*. Cambridge (MA): Harvard University Press.

Quine, Willard V. O. 1980 [1961]. *From a logical point of view*. Cambridge (MA) and London: Harvard University Press.

Resnik, David. 2007. *The price of truth: how money affects the norms of science*. Oxford: Oxford University Press.

Rosenberg, Alan. 1985. Methodology, theory and the philosophy of science. *Pacific Philosophical Quarterly*, 66 (4): 377-93.

Sent, Esther-Mirjam. 1999. Economics of science: survey and suggestions. *Journal of Economic Methodology*, 6 (1): 95-124.

Shapin, Steven. 1995. Here and everywhere: sociology of scientific knowledge. *Annual Review of Sociology*, 21: 289-321.

Shapin, Steven. 2003. Ivory tower. *London Review of Books*, 25 (17): 15-19.

Shapin, Steven, and Simon Shaffer. 1985. *Leviathan and the air pump: Hobbes, Boyle, and the experimental life*. Princeton: Princeton University Press.

Sharrock, Wesley. 2004. No case to answer: a response to Martin Kusch's 'Rule-scepticism and the sociology of scientific knowledge'. *Social Studies of Science*, 34 (4): 603-614.

Stueber, Karsten. 2005. How to think about rules and rule-following. *Philosophy of the Social Sciences*, 35 (3): 307-323.

Stueber, Karsten. 2006. How to structure a social theory? A critical response to Anthony King's The structure of social theory. *Philosophy of the Social Sciences*, 36 (1): 95-104.

Suppe, Frederick. 1977 [1972]. *The structure of scientific theories*. Urbana (IL): University of Illinois Press.

Tijssen, Robert. 2004. Is the commercialisation of scientific research affecting the production of public knowledge? Global trends in the output of corporate research articles. *Research Policy*, 33 (5): 709-733.

Washburn, Jennifer. 2005. *University Inc.: the corporate corruption of American higher education*. New York: Basic Books.

Wittgenstein, Ludwig. 2001 [1953]. *Philosophical investigations*. Oxford: Blackwell.

Woolgar, Steve. 1992. Some remarks about positionism: a reply to Collins and Yearley. In *Science as practice and culture*, ed. Andrew Pickering. Chicago: University of Chicago Press, 327-342.

Worrall, John. 1990. Rationality, sociology and the symmetry thesis. *International Studies in the Philosophy of Science*, 4 (3): 305-319.

David Tyfield is a research associate at the Institute for Advanced Studies, Lancaster University. His research is on the (global) political economy and philosophy of science and innovation, with a particular focus on climate change and the life sciences.
Contact e-mail: <d.tyfield@lancaster.ac.uk>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 87-106.
http://ejpe.org/pdf/1-1-art-4.pdf

Bernard Mandeville and the 'economy' of the Dutch

ALEXANDER BICK
Princeton University

Abstract: Studies of Bernard Mandeville by economists and historians of economic thought have focused overwhelmingly on the problem of situating his work within the development of the theory of *laissez-faire* and evaluating his influence on major figures in the Scottish Enlightenment, especially Adam Smith. This paper explores Mandeville's economic thought through the lens of a very different transition: England's rapid growth following the Glorious Revolution and its gradual eclipse of Dutch economic hegemony. By situating Mandeville within an Anglo-Dutch context and carefully examining his comments on the Dutch in Remark Q of *The fable of the bees*, the paper shows the manner in which Mandeville's ideas both appropriated lessons from Dutch history and sought to revise ideas about the Dutch current among his English contemporaries. The paper thus sheds new light on core concepts in Mandeville's economic thought and permits exploration of an important moment in the development of political economy.

Keywords: Mandeville, Netherlands, Late Development, William Temple

JEL Classification: A11, B11, B31, N10, N13

The means by which poorer countries appropriate and adapt the technologies and ideas of their richer neighbors is one that has fascinated economists and economic historians for generations. In general, this problem has been addressed in what are somewhat pejoratively called "developing economies", but on rare occasions the tools and questions of development economics have been deployed in

AUTHOR'S NOTE: An earlier version of this paper was submitted to Mary Morgan as a Masters Dissertation in Economic History at the London School of Economics. I would like to thank her for her inspiration and good judgment, along with Tim Hochstrasser, who first introduced me to Mandeville and provided essential guidance. Mario Bick, Diana Brown, Bill Bulman, Robert Darnton, Anthony Grafton, Jamie Kreiner, Peter Lake, Megan Lindsay, Harro Maas, Nealin Parker, Nick Popper, Eric Schliesser, Aaron Tugendhaft, and two anonymous referees offered challenging and helpful criticism. Responsibility for remaining errors of interpretation or fact is my own.

the analysis of early modern Europe itself. In the late 1970s, the economic historian D. C. Coleman suggested, in a parenthetical remark, that, "a late start hypothesis for England might be worth investigating" (Coleman 1977, 197). This idea drew inspiration from the earlier work of F. J. Fisher, who argued that England's economy in the Tudor and Stuart periods was in important respects "underdeveloped" (Fisher 1958). In the paper that follows, I will pursue this line of inquiry by examining the model of commercial development that the Netherlands presented to English political economists in the late 17th and early 18th centuries, and the manner in which this model changed as England began to surpass Dutch economic hegemony.

This territory is in some ways familiar thanks to the detailed and sophisticated studies of Charles Wilson (1984), Joyce Appleby (1978), David Ormrod (2003), and others. My paper will build on their important work, but will focus on a figure that has rarely been examined in this context: the Dutch-born doctor and London émigré, Bernard Mandeville. In particular, I would like to look closely at Remark Q, a lengthy footnote in the manner of Pierre Bayle that Mandeville included in the first edition of his famous *Fable of the bees: private vices, publick benefits*, published in 1714 (Mandeville 1929).[1] There, Mandeville offered a wide-ranging discussion of the differences between the English and Dutch economies and suggested the lessons that the Dutch could, and—more importantly—should not, provide to England. As I will show, his treatment of the interplay between consumption, natural resources, and commercial policy revised important ideas current in England concerning the lessons of Dutch history. Mandeville's text, along with the criticism it received, helps us to understand the complex problems involved in the longer-term appropriation and translation of models between these two commercial rivals.

My approach thus places Mandeville firmly within what, borrowing and extending a concept from Jonathan Israel (1991), might be called the longer *durée* of the "Anglo-Dutch Moment". As a Dutchman and foreigner in London, Mandeville was an observer and minor actor in the great drama of England's political, commercial, and financial

[1] All citations to this text will be to the authoritative F. B. Kaye edition, published by Oxford University Press in 1929 and subsequently reprinted by the Liberty Fund. The text that appears in the Kaye edition as "Remark Q" originally appeared as "Remark P" in the first edition of the *Fable of the bees*, printed for J. Roberts in 1714. For a publishing history of the *Fable*, see Kaye in Mandeville (1929, I, xxxiii-xxxvii; and II, 386-400).

transformation following the Glorious Revolution of 1688-1689. His analysis of core problems in political economy reflects his mixed identity to a degree that has received remarkably little attention among economists and historians of economic thought. Indeed, for much of the past century, in-depth studies by these scholars primarily focused on the rather abstract problem of situating Mandeville's work within the development of the theory of *laissez-faire* and evaluating his influence on major figures in the Scottish Enlightenment, especially Adam Smith.[2] While this debate helped to clarify a number of important aspects of Mandeville's ideas on the balance of trade and the role of government, as well as his contributions to the theory of unintended consequences, it confined discussion to Mandeville's place within a formal theoretical transition of which he by definition could have been only dimly aware.

Here I will look instead at a problem of which Mandeville was keenly aware: England's rapid commercial growth and her changing relationship to the Netherlands in the early 18th century. My paper begins by reviewing Mandeville's ideas on luxury consumption in order to provide a sense of his approach to the London economy and to establish some of the ways that he employed his unusual comparative perspective.[3] I then turn to his analysis of the differences between England and the Netherlands in Remark Q. What, according to Mandeville, were the proper lessons for England to draw from Dutch commercial success? How did these lessons challenge existing interpretations of Dutch history, and what more general principles did Mandeville offer to explain the flourishing of commercial societies?

MANDEVILLE AND THE LONDON ECONOMY

Although Mandeville's later biography remains frustratingly incomplete, a number of recent studies by historians in England and the Netherlands have helped to establish the details of his early life with some precision (Dekker 1992; Goldsmith 1992; Cook 1999, 2007). Born in Rotterdam in

[2] The entire course of this debate to the mid-1970s is reviewed in Landreth (1975). For two particularly illuminating and contrasting viewpoints, see Viner (1958); and Hayek (1978).

[3] A great deal has been written on the "luxury debates" that raged across Europe in the mid-18th century. My paper overlaps with this literature in important ways, but focuses more closely than is often done on the problem of saving and its implications for the ways that English writers understood Dutch economic development in the Netherlands. On the luxury debate, see especially Hont (2006). For analysis of England's "Consumer Revolution", including discussion of Mandeville's ideas on consumption, see the introductions to the edited volumes by Brewer and Porter (1993); and Berg and Clifford (1999).

1670, Mandeville came from a family of physician-magistrates well connected in city politics. He studied philosophy, perhaps even meeting the skeptic and editor Pierre Bayle, and then medicine, completing his doctorate at the University of Leiden in 1691. His dissertation, which examined disorders of the stomach, bore the intellectual imprint of the French-born philosopher René Descartes, who lived in the city in the 1630s and 1640s and whose ideas contributed to a vigorous climate of radical philosophical materialism that lasted well into the 18th century. In addition to his studies, Mandeville was something of an agitator: in 1690, he and his father were implicated in the Costerman affair, a political conflict that swept Rotterdam and ultimately led to his father's expulsion from the city. Sometime in the early 1690s Mandeville left the Netherlands for good, arriving in England no later than November 1693, when he was cited by the College of Physicians for practicing medicine without their permission.

Mandeville's arrival in London came fresh on the heels of the Glorious Revolution, when a Dutch army invaded England, deposing the government of James II and ultimately installing the Dutchman William III and his English wife, Mary Stuart, as joint sovereigns. This event ushered in a long period of Anglo-Dutch political and military cooperation against France that drew England into Continental affairs and dramatically increased the financial burden on English taxpayers. It also led, in the summer of 1694, to the establishment of the Bank of England as a vehicle with which to raise additional money and buttress government credit (Dickson 1967).

As the pamphlet literature of the period amply demonstrates, the explosion of government debt, the increasing importance of "city" financiers in national politics, and the profits in land and money that accrued to some of William's chief Dutch advisors exposed William to charges of corruption and criticism that his government was siphoning English resources for an alliance that was more in the interest of the Netherlands than England itself (Rose 1999). As the Tory MP Charles Davenant put in 1701, "everyone is on the scrape for himself, without any regard to his country, each cheating, raking, and plundering what he can, and in a more profligate degree than ever was known" (Davenant 1771, II, 301).

To these cries of corruption and moral decay Mandeville offered both a defense of the Revolution settlement and an analysis of the London economy that made moral depravity a central component of commercial growth. If, as H. T. Dickson has argued, Mandeville kept party politics at arm's length, his general orientation is fairly clear (Dickson 1974; Kramnick 1992; Goldsmith 1999). In a 1703 pamphlet, *The pamphleteers: a satyr*, Mandeville defended William's legacy against detractors, arguing that the recently deceased King had successfully protected England from popery and that the "gaudy crown he wore" was not worth "one-tenth the indignities that he bore" (Mandeville 1703, 6).[4] Two years later, in 1705, Mandeville published another work, *The grumbling hive, or knaves turn'd honest*, a satirical pamphlet in doggerel verse that would later serve as the core of *The fable of the bees*. Here he famously compared England to a thriving beehive, "well stockt with Bees/That liv'd in Luxury and Ease" (Mandeville 1929, 24). Greed, corruption, and decadence could be found at all levels of the social hierarchy, but these same characteristics made the hive rich. Mandeville's poem thus played on the image of bees as busy and industrious (Johnson 1966; Hundert 1994, 24-29), while inverting their traditional association with social order: "Every Part was full of Vice", Mandeville wrote, "Yet the whole Mass a Paradise" (Mandeville 1929, 24).

Some scholars believe that this pamphlet was intended primarily to deflect charges of corruption and fraud that had been leveled against England's military commander on the Continent, John Churchill, the first Duke of Marlborough (Kramnick 1992, 201). But the message of *The grumbling hive* was more profound: with its new constitutional monarchy, densely populated capital, and industrious working poor, Mandeville believed that England was poised to become the envy of the world. The vices that he so relentlessly exposed might be unfortunate, but they were inescapable in a flourishing commercial society, and could not be eliminated without undermining growth itself. Here Mandeville drew inspiration from a number of scholars in England and the Netherlands who argued that the human passions could be managed in a way that would maintain or even promote social order.[5] Thus Mandeville argued that:

---

[4] *The pamphleteers: a satyr* (London 1703): To my knowledge, Mandeville has not been definitively identified as the author of this work, but the internal evidence supports attribution and most scholars continue to treat the work as his. The British library lists the pamphlet's author as Mandeville.

[5] Of particular importance were Thomas Hobbes, Pierre Bayle, the Leiden cloth merchant Pieter de la Court, and Baruch Spinoza. For their respective influences on Mandeville's thought, see F. B. Kaye's introduction to the *Fable* (1929, I, xxxix-cxiii); Hundert (1994); and Cook (1999). For the broader outlines of Dutch Republican theory, and the role of the passions in de la Court and Spinoza, see Kossman (1960).

The Root of Evil, Avarice,
That damn'd ill-natur'd baneful vice,
Was Slave to Prodigality,
That Noble Sin; whilst Luxury
Employ'd a Million of the Poor,
And odious Pride a Million more:
Envy it self, and Vanity,
Were Ministers of Industry (Mandeville 1929, I, 25).

If the terms were pitched to offend, the mechanism here was fairly simple: avarice, which led individuals to work hard and to accumulate resources, was balanced by prodigality, or spending beyond one's means, which would cause these same resources to be put back into circulation. The spending of the rich would thus create work for England's poor. It was a formulation that put urban, luxury consumption at the center of the economic system.[6] In the early 1690s, the successful London merchant and real estate developer Nicholas Barbon had made a similar argument, writing that the "chief causes promoting trade are the industry of the poor and the liberality of the rich". Especially important was consumption of fashions and other goods that, as Barbon put it, "serve the pomp of life". Prodigality might be bad for the individual, but it was good for trade (Barbon 1690, 36; Letwin 1963; Finkelstein 2000).

Mandeville appropriated these controversial ideas and presented them to his adopted countrymen in a language that was rich in meaning and would have been familiar to many English readers: prodigality was, after all, a vice to which the English were thought to be particularly prone.[7] Avarice, by contrast, was chiefly associated with the Dutch—perhaps most memorably in Daniel Defoe's catalogue of national stereotypes in *The true-born Englishman*, published just five years before *The grumbling hive*, in 1700 (Defoe 1700, 9). The productive

[6] Unlike his contemporary Daniel Defoe, who travelled extensively throughout England and reported on the diversity of the country's industries, I am aware of no evidence that Mandeville ever strayed more than a few miles from London. As the city was, in F. J. Fisher's memorable phrase, a "center of conspicuous consumption", this may help to account for the disproportionate role Mandeville devoted to the demand-side of the English economy. See Fisher (1990).

[7] See, for instance, Thomas Mun: "this great plenty which we enjoy, makes us a people not only vicious and excessive [but also] wasteful of the means we have" (1664, 178). Although these words were likely written in the early 1620s, Mun's book continued to be read later, and such lamentations concerning the lack of restraint practiced by English consumers were commonplace throughout the 17th century.

alliance between the two passions thus mirrored the productive alliance between England and the Netherlands themselves. In this sense, Mandeville's was a very specific portrait of London; one that wove together allegiance to the government and tolerance for the ethical contradictions that he believed were implicit in its burgeoning consumer economy. Unless the English would prefer to return to poverty and simplicity, they must set aside their incessant "grumbling" and accustom themselves to the pace and character of rapid commercial development—a process that Mandeville had presumably encountered first-hand as a young man living in Rotterdam and Leiden, two of the Netherlands' most important centers of trade and industry.

REMARK Q AND THE DUTCH 'PATTERN'

The lessons of the Dutch experience were not so simple, however. As F. B. Kaye, editor of Mandeville's *Fable of the bees*, pointed out long ago, Mandeville's defense of luxury consumption in *The grumbling hive* presented something of a paradox, in that it squarely contradicted the view—often repeated by English writers—that the Dutch economic miracle of the 17th century had been built, at least in part, on the Dutch people's frugal spending habits and their careful, even obsessive attention to saving (Mandeville 1929, I, 188n). Mandeville's response to this paradox provides a fascinating window into his ideas about economic growth and the manner in which he differentiated processes of commercial development in each country.

If, as Mandeville later quipped, the English knew less about the Netherlands than might be expected of so close a neighbor and ally (Mandeville 1709, 137), many among them were extremely keen to understand the sources of Dutch wealth and the ways in which Dutch examples might inform policy in England. As Joyce Appleby (1978) has demonstrated, the Dutch economy exercised a profound influence on English political economists and served as an important "source of evidence" for their analysis of the English economy. To cite but one of many examples, Josiah Child, perennial Director of the East India Company, enumerated no less than fifteen specific lessons that could be drawn from Dutch commercial practice in his widely distributed *Brief observations concerning trade and interest of money*, published in 1668 and then reprinted and enlarged in 1689, 1690, 1693, and 1694 (Child 1668, 3-6; Ormrod 2003, 313). These lessons—which included the promotion of shipping, the participation of merchants on Dutch councils of state, and the advantages of a low rate of interest—were, in Child's words, "sufficiently obvious, and in a great measure imitable by most other Nations, but more easily by us of this Kingdom of *England*" (Child 1668, 3).

Within the formal ranks of government, interest in the Dutch was equally intense. Senior English officials such as George Downing and Joseph Williamson, each of whom served Charles II in the 1660s and 70s, eagerly collected information on Dutch trade and industry, translating Dutch pamphlets and, on at least one occasion, sending spies to steal and copy secret government documents (Scott 2004; Wilson 1984, 166-175). While these two individuals' work remained largely outside the public view, Sir William Temple's popular *Observations upon the united provinces of the Netherlands*, first published in 1672, reached a much broader audience (Temple 1972; Haley 1986). As ambassador to The Hague from 1668-1670, Temple was friendly with Holland's Grand Pensionary, Johan de Witt, and opposed to the third Anglo-Dutch war that resulted in his death. Temple's book offered an elegant, authoritative, and highly sympathetic survey of Dutch culture, history, religion, politics, and trade. It quickly became a standard reference on the Dutch, and by 1700 was already in its eighth printing.

Important to this literature, and to Temple's account in particular, was the image of the Dutch as a frugal, parsimonious people who had built a thriving economy by shunning luxuries and profiting from the prodigality of their neighbors. Child, for instance, placed "parsimonious and thrifty living" sixth on his list, noting that, "a merchant of one hundred thousand pounds estate with them, will scarce expend so much *per annum*, as one of fifteen hundred pounds estate in *London*" (Child 1668, 4). Temple was equally impressed, arguing that even if trade was their main activity, the "true ground" of Dutch success lay with their "industry and parsimony":

> For never a Countrey traded so much, and consumed so little: They buy infinitely, but 'tis to sell again […] They are the great masters of the *Indian* Spices, and of the *Persian* Silks; but wear plain Woollen, and feed upon their own Fish and Roots. Nay, they sell the finest of their own Cloath to *France*, and buy coarse out of *England* […] They send abroad the best of their own Butter into all parts, and buy the cheapest out of *Ireland*, or the north of *England*, for their own use. In short, they furnish infinite luxury, which they never practice, and traffique in Pleasures which they never taste (Temple 1972, 119).

If Mandeville was to establish the credibility of the argument he advanced in *The grumbling hive*, he needed to address this supposed Dutch virtue and the author who had so conspicuously celebrated it. In 1714, when Mandeville added a preface, explanatory notes, and other new material to the original poem and reissued it as *The fable of the bees*, he did just this, devoting Remark Q to frugality and the historical lessons of the Dutch.

Mandeville made two main arguments: the first sought to undermine the idea that the Dutch were in fact as frugal as many Englishman supposed. The second attempted to show that what restraint was practiced in the Netherlands owed less to virtue than it did to necessity, and thus that the absence of this necessity in England created a novel situation in which consumption could play an entirely different role than it had in his own country.

It was true, Mandeville conceded, that the Dutch were more modest than some of their neighbors on the continent. As a commonwealth, with greater equality of income, one could not expect to find in the Netherlands princely palaces or the sorts of display associated with the court. Instead money was spent elsewhere. "In Pictures and Marble they are profuse", Mandeville reported, and "in their Buildings and Gardens they are extravagant to Folly" (Mandeville 1929, I, 187).

As Simon Schama's analysis of Dutch consumption and culture has shown, this assessment was broadly accurate. Home furnishings and other luxury goods were available and eagerly bought up in every major Dutch city, especially Amsterdam, where Melchior Fokkens, a Dutch gazetteer, in 1665 noted "houses full of priceless ornaments […] splendid alabaster columns, floors inlaid with gold, and the rooms hung with valuable tapestries or gold or silver-stamped leather worth many thousand of guilders" (Schama 1987, 303). Consumption was more discreet, perhaps, kept indoors or partially concealed, but it was not thereby any less prodigious. Temple could praise the Dutch for keeping the same fashions longer than others, and for wearing apparently simple black clothing, but Schama points out that "the black was very often satin or velvet, sometimes discretely trimmed with fur" (Schama 1987, 310).

If the Dutch were in general still more frugal than the English, the second component of Mandeville's argument was designed to show how this was a necessary response to the natural, economic, and political constraints the Dutch faced. Each of these, Mandeville believed, made luxury consumption more difficult and a higher degree of saving imperative. Indeed, Mandeville was at pains to demonstrate to his English audience the terrible burdens under which the Dutch labored. In a parallel passage in a separate work, published in 1709, Mandeville sought to dispel criticisms that the English were paying a disproportionate amount in taxes to support the war:

> If, I say, some of our [English] People should know how they [the Dutch] are oblig'd to pay certain Sums, at which they are rated for using Salt, and Soap, whether they consume little or much; how every Family, that will drink Tea, Coffee, or Chocolate, must pay a great Tax for it, tho' they had but one Dish of any of the three in the whole Year: Should they consider all this, and that the very Cows pay for having Horns, they would think our Burden much lighter than theirs, and cry out, *Blessed* England! (Mandeville 1709, 139).

Indeed, modern research in economic history has estimated that the Dutch paid two and a half times as much in taxes as their counterparts in England, a factor that placed dramatic constraints on the possibilities for growth (Ormrod 2003, 307).

But taxes were only one part of a larger story. In a passage that may have influenced Mandeville directly, Charles Davenant suggested in 1698 that the Dutch might have been forced into thrift by the constant threat of invasion by land and their precarious natural environment. The Dutch, he wrote "are continually forced, in a manner, to pump for life, and nothing can support them but the strictest œconomy imaginable, both in private and in public" (Davenant 1771, 390). Although he viewed luxury spending with considerable skepticism, Davenant recognized that England was an island with abundant natural produce, and thus that its people could probably afford to indulge in greater luxury than their neighbors:

> But our case is far from being the same [as the Dutch]; we are not easily invaded; the expense of our government in time of peace, is much less than theirs; we have a large and fertile country, and a great native product; so that the whole public of this kingdom may grow rich, though the people [...] are more luxurious than in other nations (Davenant 1771, 390).

And perhaps, Davenant continued, "it is not impossible, but that our industry would be less active, if it were not awakened and incited by some irregular appetites, which are more easily found fault with than cured" (Davenant 1771, 390-391).[8]

In Remark Q, Mandeville developed this observation into a complex portrait of the relationship between the passions, natural resource endowments, policy, and growth. Men may initially differ in temperament, and thus be disposed to covetousness, prodigality, or saving, he argued, but "if anything ever draws 'em from what they are naturally propense to, it must be a Change in their Circumstances or their Fortunes". The most important of these circumstances were the "Fruitfulness and Product of the Country, the Number of Inhabitants, and the Taxes they are to bear" (Mandeville 1929, I, 184). If the first is great, but the number of people and the taxes low, nothing will progress beyond a happy and slothful ease:

> Man never exerts himself but when he is rous'd by his Desires: While they lie dormant, and there is nothing to raise them, his Excellence and Abilities will be forever undiscover'd, and the lumpish Machine, without the Influence of his Passions, may justly be compar'd to a huge Wind-mill without a breath of Air (Mandeville 1929, I, 184).

But shift things around through laws and "good management", and then people's dispositions will change. It is worth quoting Mandeville at length on this point:

> Would you render a Society of Men strong and powerful, you must touch their Passions. Divide the Land, tho' there be never so much to spare, and their Possessions will make them Covetous: Rouse them, tho' but in Jest, from Idleness with Praises, and Pride will set them to work in earnest: Teach them Trades and Handicrafts, and you'll bring Envy and Emulation among them: To increase their Numbers, set up a Variety of Manufactures, and leave no Ground Uncultivated; Let Property be inviolably secured, and Privileges equal to all Men; Suffer no body to act but what is lawful, and every body to think what he pleases; for a Country where every body be maintained that will be employ'd, and the other Maxims are observ'd must always be throng'd and can never want People [...] But would you moreover render them an opulent, knowing and polite nation, teach 'em Commerce with Foreign Countries, and if possible get into the Sea, which to compass spare no Labour nor Industry, and let no Difficulty deter you from it: Then promote Navigation, cherish the Merchant, and encourage Trade in every Branch of it; this will bring

[8] Mandeville's debt to this passage was first proposed by F. B. Kaye (Mandeville 1929, I, 187n)

> Riches, and where they are, Arts and Sciences will soon follow (Mandeville 1929, I, 184).

Thus, the Dutch had achieved their remarkable success, through effectively managing the passions, attracting large numbers of people, and encouraging industry and trade. If poor land and high taxes had made the Dutch frugal, this was no virtue. In fact, Mandeville argued, these constraints had if anything only encouraged the Dutch to adopt aggressive tactics to improve their situation. To admit that frugality was necessary under the circumstances was a very different thing than to claim that it had made the Dutch rich in the first place. Frugality was thus only an effect, a response to circumstance rather than an economic virtue.

But in England, circumstances were more favorable and frugality unnecessary. In his classic work of republican and commercial theory, *The true interest and political maxims of Holland and West-Friesland*, published in 1662, Pieter de la Court had written enviously of England's situation. Asking, "why the great inconveniences of Taxes and Wars that we have laboured under, have not occasioned the Fishing, Manufactury, Traffick, and Navigation, to settle and fix in other Countries", he cited England,

> where if all be well considered they have had far greater Advantages of Situation, Harbours, a clean and bold Coast, favorable Winds, and Opportunity of transporting many unwrought Commodities, a lasting Peace, and a great freedom from Taxes than we have (de la Court, 1702, 45).

De la Court's answer to this question was historical, citing the higher taxes and other restrictions placed on foreigners in London at the time of the fall of Antwerp, in 1585. But the implication was that circumstances might change at any time. With government support for trade, taxes lower than in the Netherlands, and abundant resources, England was, by the last decade of the 17th century, in the superior position.

As Mandeville saw it, the English could now afford to rouse their passions through luxury if it was economic growth they desired. Whereas the Dutch were running out of land and had to import much of their food, the English were nearly agriculturally self-sufficient and there was still plenty of uncultivated land available. Here his argument stressed the advantages of England's underdevelopment, or what we might call her "relative backwardness". Elsewhere in the *Fable of the bees*, in fact, Mandeville suggested that work still remained for 100,000 poor for "300-400 years" for England to bring all of its territory into productive use (Mandeville 1929, I, 318). These figures probably would not bear contemporary scrutiny, but the basic analysis was sound: despite considerable progress during the last third of the 17th century, England had not yet achieved the level of intensive resource usage—in terms of land-reclamation, the construction of canals, and the use of wind-power—for which the Netherlands was famous. And the fact that wages were lower in England than in the Netherlands suggests that labor was more abundantly available, as well. What was missing was circulating capital.

Mandeville's argument was not simply that the English could *afford* to consume, but that consumption would actually *contribute* to national wealth. In England, the relative paucity of capital resulted in interest rates higher than those in the Netherlands, which had abundant capital and already had nearly achieved what Adam Smith would later describe as a "full complement of riches" (Smith 1963, I, 76). The Bank of England was beginning to address this problem, but Mandeville evidently believed that the additional reserves of money cautiously squirreled away in chests could be put to better use. It is perhaps only in this context that we can begin to see the logic in Mandeville's otherwise outrageous claim that stealing from the miser would benefit the public good: without the general circulation of capital—by theft or mere profligacy—England's full range of human and natural resources could not be brought into productive employment. Luxury consumption was an important part of this vision, but it was only one part among many.

WILLIAM TEMPLE'S LEGACY

This interpretation of England's potential and the lessons of the Dutch model were not universally accepted among Mandeville's contemporaries. Responses to Mandeville's theory of luxury consumption have received considerable attention from scholars, but one response in particular helps us to understand the ways in which consumption fit within a broader disagreement concerning the historical lessons of the Dutch. In an anonymous pamphlet published in 1725, in direct response to the second edition of Mandeville's *Fable of the bees*, which had been released in the previous year, the lawyer George Bluett took direct issue with Mandeville, in part by attacking his use of William

Temple (Bluett 1725).[9] In his discussion of Dutch spending habits, Mandeville had attempted to show that Temple's *Observations* had been written during a time of particular distress:

> The Nation I speak of was never in greater Straits, nor their Affairs in a more dismal Posture since they were a Republick, than in the Year 1671, and the beginning of 1672. What we know of their Œconomy and Constitution with any Certainty has been chiefly owing to Sir *William Temple*, whose Observations upon their Manners and Government, it is evident from several Passages in his Memoirs, were made about that time. The Dutch indeed were then very frugal; but since those Days and that their Calamities have not been so pressing […] a great Alteration has been made among the better sort of People in their Equipages, Entertainments, and whole manner of living (Mandeville 1929, I, 189).

This interpretation allowed Mandeville to argue that, once free from the burdens of war, the Dutch took to consuming much as people in England were doing at present. Mandeville thus tipped his hat to Temple as the chief source of information on the Dutch at the same moment that he undermined Temple's credibility by suggesting that his observations had been distorted by the unusual events taking place during Temple's tenure in the Netherlands.

Bluett was wholly unconvinced by this reading. "Was there ever a more injudicious Remark?" Bluett asked, continuing:

> In what a perverse manner must he have read the Author he quotes. In the very same Paragraph in which Sir *William Temple* tells him, that his observations were made about that Time, he ascribes the Decay of their Wealth to the Luxury he had for several Years observed to be growing among them (Bluett 1725, 45).

Moreover, Bluett argued that the necessity under which the Dutch labored did not make frugality any less of a virtue:

> If the Dutch in their present Condition are oblig'd to be more frugal than their Neighbours, from the vast Expense they are at in Repairing their Dykes, the Weight of other taxes, and the Scantiness of their Dominions; would not the same Frugality in their Neighbours, who have a greater Extent of Land, and no such

[9] F. B. Kaye attributes this pamphlet to the London lawyer George Bluett. See his comments excerpted in (Stafford 1997, 229-230).

> Demands of Expense, keep them in a Condition still proportionately above them, and continue them still proportionately richer?

How was it, in other words, Bluett asked, that frugality could be such a "whimsical Virtue, that it always makes a poor Country [the Netherlands] rich, and a rich Country [England] poor?" (Bluett 1725, 50).

On point after point Bluett provided a meticulous, logical critique of Mandeville's thought, based on a close reading not only of the *Fable of the bees*, but also several other of Mandeville's books and even some of his favorite sources (Stafford 1997). Bluett's orientation was traditional. He likened the national economy to the economy of the household, as generations of English mercantilists had done before him, and thereby concluded that consumption could only serve to drain the national coffers. Spain's descent during the 17th century and Temple's account of both the Dutch rise and the early stages of their decline provided his evidence. Bluett's reading of Temple was faithful, in so far as Temple held a neo-Polybian view of the cyclical cresting and falling of empires as frugality gave way to luxury and decadence.[10] Temple's *Observations* thus helped Bluett to put Mandeville's interpretation of London's consumer society into historical perspective: "If Vice in general, and luxury in Particular, be the Road to Wealth", Bluett concluded, with a perhaps overly generous sense of irony, "we [in England] bid fair for growing prodigiously Rich" (Bluett 1725, 52).

That Mandeville's assessment seems to have more closely captured the spirit of commercial development in early 18th century England says nothing about the deceptive manner in which he tried to get around the lessons in Temple's *Observations*. But Mandeville read Temple in another way, as well; one that helps us to better understand the overall thrust of his ideas. Later in Remark Q, Mandeville drew on Temple's ideas concerning the relationship between a people's disposition and their passions:

> All Men, as Sir *William Temple* observes very well, are more prone to Ease and Pleasure than they are to Labour, when they are not prompted to it by Pride or Avarice, and those that get their Living by daily Labour, are seldom powerfully influenc'd by either: So that they

[10] As Istvan Hont has argued, Temple balanced this classically-inspired interpretation of the Netherlands' economic troubles with an analysis of the ethically-neutral effects of increasing competition from other trading nations, including England. See Hont (1990, 55).

> have nothing to stir them up to be serviceable but their Wants, which it is Prudence to relieve, but Folly to cure (Mandeville 1929, I, 194).

Here Mandeville had read Temple carefully and was not attempting to deceive. Instead, he used Temple's ideas to show that people were not inclined to work without some sort of stimulus, or anticipated reward. If this reward could in the Netherlands be only very limited, due to the difficult conditions there, in England it could be that much greater.

Where Bluett had tried to analyze the economy from the perspective of virtue and vice, and thus gave the same prescription for England as that which had worked for the Netherlands, Mandeville followed the strand in Temple's *Observations* that explored the ways that virtues and vices developed in intimate relation with the natural environment and institutional context. Since this environment was different in England, a different sort of disposition could be expected to assist its people. Man's nature might be "everywhere the same", as Temple put it, but his manners and customs differed from place to place (Temple 1972, 80, 88-94). This was an observation that presumably came less easily to Bluett than to Mandeville, who had made the same journey between England and the Netherlands as Temple, though twenty-five years later and in the reverse direction.

CONCLUSION

In the space of Mandeville's lifetime, England was transformed into Europe's largest economy, and the Netherlands, though still growing, had lost its position of pre-eminence (Omrod 2003, 307-309). By 1740, seven years after Mandeville's death, an English writer could say that now "England could only borrow money where once she had sought inspiration" (Feingold 1996, 259). The analysis in this paper suggests that this inspiration was itself multi-faceted, and changed over time. In the context of England's rapid commercialization in the first decades of the 18th century, older lessons gleaned from authors like Child and Temple needed to be re-evaluated and new ones given firm footing. Mandeville's Remark Q might be seen in this light, and further taken as an indication that the historical lessons of the Dutch economy continued to animate public discussion in England well into the 1720s, if not later.

Mandeville's analysis of the differences between English and Dutch economies led him to articulate a theory of commercial development that rested heavily on commercial policy and the role of government in managing men's passions. The Dutch had become rich through careful attention to merchant interests, investment in shipping and navigation, intensive exploitation of land, religious toleration, and protection of private property. History showed the importance of these measures and thus buttressed the initiatives of those in England who aimed to institute or protect similar policies at home. Luxury consumption could be an important ingredient not because it was universally beneficial, but because England's circumstances facilitated a positive relationship between indulgence and employment that circumstances in the Netherlands simply could not support. It is this combination of probing psychological analysis, observation, and comparative history that made Mandeville's ideas so powerful, and which led him to the conclusion that England had the capacity to exceed, rather than simply approximate, the model offered by the Dutch.[11]

One last issue requires attention, and may be taken as a sort of postscript. In their classic studies of "late development", Walter Rostow and Alexander Gerschenkron generally assumed a stable industrial model, towards which countries in Eastern and Central Europe, Latin America, Asia, or Africa would strive (Rostow 1960; Gerschenkron 1962). It was Gerschenkron's great innovation to suggest that the form taken by development in a relatively backward country would differ in important respects from the processes of growth in the most advanced economies, directly in proportion to their level of backwardness. In particular, he emphasized the variability of speed, industrial scale, and the role of banks. But he also suggested that late development would require an ideology; a set of principles that could break routine and provide ordinary citizens with faith that improvement lay ahead (Gerschenkron 1960, 24). Of course, Mandeville knew nothing of the Industrial Revolution and even less of the ideological programs that would animate industrialization in the 19th and 20th centuries. But his work might be interpreted as having counseled a similar kind of faith—in the improving qualities of commercial development and the ability of the English state to harness men's private vices, as his subtitle famously argued, for public benefit.

[11] It is clear that Mandeville saw England's potential as vast; it would be interesting to know also whether he saw this potential as posing a direct threat to his native country and its economic and political future. If he did, I have found no indication to this effect in his writings, though it is of course possible that Mandeville would not have addressed such concerns to his English audience.

REFERENCES

Appleby, Joyce. 1978. *Economic thought and ideology in seventeenth-century England*. Princeton: Princeton University Press.

Barbon, Nicholas. 1690. *A discourse of trade*. London: Printed by Tho. Milbourn.

Berg, Maxine, and Helen Clifford (eds.). 1999. *Consumers and luxury: consumer culture in Europe 1650-1850*. Manchester: Manchester University Press.

Bluett, George. 1725. *An inquiry whether a general practice of virtue tends to the wealth or poverty, benefit or disadvantage of a people*. London: Printed for R. Wilkin.

Child, Josiah. 1668. *Brief observations concerning trade and interest of money*. London: Printed by Elizabeth Calvert, and Henry Mortlock.

Coleman, D. C. 1977. *The economy of England, 1450-1750*. Oxford: Oxford University Press.

Cook, Harold. 1999. Bernard Mandeville and the therapy of 'the clever politician.' *Journal of the History of Ideas*, 60 (1): 101-124.

Cook, Harold. 2007. *Matters of exchange: commerce, medicine, and science in the Dutch golden age*. New Haven: Yale University Press.

Court, Pieter de la. 1702. *The true interest and political maxims of Holland and West-Friesland*. London.

Davenant, Charles. 1771. *The political and commercial works*, ed. Sir Charles Whitworth, 5 Vols. London: Printed for R. Horsfield, T. Becket, P. A. de Hondt, T. Cadell, and T. Evans.

Defoe, Daniel. 1700. *The true-born Englishman*. London.

Dekker, Rudolf. 1992. 'Private vices, public virtues' revisited: the Dutch background of Bernard Mandeville. *History of European Ideas*, 14 (4): 481-498.

Dickinson, H. T. 1975. The politics of Bernard Mandeville. In *Mandeville studies: new explorations in the art and thought of Dr. Bernard Mandeville*, ed. Irwin Primer. The Hague: Martinus Nijhoff, 80-119.

Dickson, P. G. M. 1967. *The financial revolution in England: a study in the development of public credit, 1688-1756*. London: Macmillan.

Feingold, Mordechai. 1996. Reversal of fortunes: the displacement of cultural hegemony from the Netherlands to England in the seventeenth and early eighteenth centuries. In *The world of William and Mary: Anglo-Dutch perspectives on the revolution of 1688-89*, eds. Dale Hoak, and Mordechai Feingold. Stanford: Stanford University Press, 234-261.

Finkelstein, Andrea. 2000. *Harmony and the balance: an intellectual history of seventeenth-century economic thought*. Ann Arbor: University of Michigan Press.

Fisher, F. J. 1958. The sixteenth and seventeenth centuries: the dark ages in English economic history? *Economica*, 24 (93): 2-18.

Fisher, F. J. 1990. The development of London as a centre of conspicuous consumption in the sixteenth and seventeenth centuries. In *London and the English economy 1500-1700*, eds. P. J. Corfield, and N. B. Hart. London: Hambledon Press, 105-118.

Gerschenkron, Alexander. 1962. *Economic backwardness in historical perspective*. Cambridge: Harvard University Press.

Goldsmith, M. M. 1992. Bernard Mandeville and the virtues of the Dutch. *Dutch Crossing*, 48: 20-38.

Goldsmith, M. M. (ed.). 1999. *By a society of ladies: essays in the female tatler, by Bernard Mandeville*. London: Thoemmes.

Haley, K. H. D. 1986. *An English diplomat in the Low Countries: Sir William Temple and John de Witt, 1665-1672*. Oxford: Oxford University Press.

Hayek, Friedrich August von. 1978. *New studies in philosophy, politics, economics, and the history of ideas*. Chicago: University of Chicago Press.

Hont, Istvan. 1990. Free trade and the economic limits to national politics: neo-Machiavellian political economy reconsidered. In *The economic limits to modern politics*, ed. John Dunn. Cambridge: Cambridge University Press, 41-120.

Hont, Istvan. 2006. The early Enlightenment debate on commerce and luxury. In *The Cambridge history of eighteenth-century political thought,* Vol. I, eds. Mark Goldie, and Robert Wokler. Cambridge: Cambridge University Press, 379-418.

Hundert, E. J. 1994. *The Enlightenment's fable: Bernard Mandeville and the discovery of society*. Cambridge: Cambridge University Press.

Israel, Jonathan (ed.). 1991. *The Anglo-Dutch moment: essays on the glorious revolution and its world impact*. Cambridge: Cambridge University Press.

Johnson, James W. 1961. That neo-classical bee. *Journal of the History of Ideas*, 22 (2): 262-266.

Kossman, E. H. 1960. The development of Dutch political theory in the seventeenth century. In *Britain and the Netherlands,* Vol. 1, eds. J. S. Bromley, and E. H. Kossman. London: Chatto & Windus, 91-110.

Kramnick, Isaac. 1992. *Bolingbroke and his circle: the politics of nostalgia in the age of Walpole*. Ithaca: Cornell University Press.

Landreth, Harry. 1975. The economic thought of Bernard Mandeville. *History of Political Economy*, 7 (2): 193-208.

Letwin, William. 1963. *The origins of scientific economics: English economic thought 1660-1776*. London: Methuen.

Mandeville, Bernard. 1703. *The pamphleteers: a satyr*. London.

Mandeville, Bernard. 1709. *The virgin unmask'd*. London: F. Morphew, and J. Woodward.

Mandeville, Bernard. 1929 [1714-32]. *The fable of the bees, or private vices, publick benefits*, 2 Vols., ed. F. B. Kaye. Oxford: Oxford University Press.

McKendrick, Neil, John Brewer, and J. H. Plumb (eds.). 1992. *The birth of a consumer society*. London: Europa.

Mun, Thomas. 1664. *England's treasure by forraigne trade*. London: Printed by J. G. for Thomas Clark.

Ormrod, David. 2003. *The rise of commercial empires: England and the Netherlands in the age of mercantilism, 1650-1770*. Cambridge: Cambridge University Press.

Rose, Craig. 1999. *England in the 1690s: revolution, religion, and war*. Oxford: Blackwell.

Rostow, W. W. 1960. *The stages of economic growth*. Cambridge: Cambridge University Press.

Schama, Simon. 1987. *The embarrassment of riches: an interpretation of Dutch culture in the golden age*. New York: Knopf.

Scott, Jonathan. 2004. Downing, Sir George, first baronet (1623–1684). Oxford Dictionary of National Biography online edition. http://www.oxforddnb.com/view/article/7981 (accessed 21 July 2008).

Smith, Adam. 1963 [1776]. *An inquiry into the nature and causes of the wealth of nations*, 2 Vols. Illinois: Richard Irwin.

Stafford, J. Martin. 1997. *Private vices, publick benefits? The contemporary reception of Bernard Mandeville*. Solihull: Ismeron.

Temple, Sir William. 1972 [1673]. *Observations upon the united provinces of the Netherlands*, ed. G. N. Clark. Oxford: Oxford University Press.

Viner, Jacob. 1958. *The long view and the short: studies in economic theory and policy*. Illinois: Free Press.

Wilson, Charles. 1984. *England's apprenticeship, 1603–1763*. New York: Longman.

Alexander Bick is a PhD candidate in history at Princeton University. His current research is on the merchant-scholar Johannes de Laet and the politics of the Dutch West Indies Company in the 1640s.
Contact e-mail: <abick@princeton.edu>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 107-123.
http://ejpe.org/pdf/1-1-art-5.pdf

Is history of economic thought a "serious" subject?

MARIA CRISTINA MARCUZZO
Sapienza, Università di Roma

Abstract: The purpose of this paper is to clarify the nature of research methods in the history of economic thought. In reviewing the "techniques" which are involved in the discipline, four broader categories are identified: a) textual exegesis; b) "rational reconstructions"; c) "contextual analysis"; and d) "historical narrative". After examining these different styles of doing history of economic thought, the paper addresses the question of its appraisal, namely what is good history of economic thought. Moreover, it is argued that there is a distinction to be made between doing economics and doing history of economic thought. The latter requires the greatest possible respect for contexts and texts, both published and unpublished; the former entails constructing a theoretical framework that is in some respects freer, not bound by derivation, from the authors. Finally, the paper draws upon *Econlit* records to assess what has been done in the subject in the last two decades in order to frame some considerations on how the past may impinge on the future.

Keywords: history of economic thought, textual exegesis, rational reconstructions, *Econlit* records

JEL Classification: A20, B00, B40

In her 1932 booklet dedicated to Piero Sraffa, Joan Robinson addressed the question of whether economics is a serious subject. She did so in the form of an apologia of the economist "to the mathematician, the scientist and the plain man". A serious subject in the academic sense, she claimed, "is neither more nor less than its own technique" (Robinson 1932, 3). The point she wanted to drive home was that in economics any

AUTHOR'S NOTE: This is a revised version of my presidential address to the XXII Conference of the European Society for the History of Economic Thought (ESHET), held in Prague, May 15-17, 2008. I wish to thank, without implicating them, Duncan Foley, Wade Hands, Nerio Naldi, Alessandro Roncaglia, Annalisa Rosselli, and Fernando Vianello, for their comments to an earlier version.

attempt at more ambitious endeavours (for instance making realistic "assumptions", and giving up abstract and simplified models) is doomed to failure, since the right techniques to tackle the complexity of the real world are often unavailable.

For my address as the first woman president of the European Society for the History of Economic Thought (ESHET) I decided to look to Joan Robinson, for two reasons: firstly as a homage to a great woman economist, and secondly to clarify the nature of research in the history of economic thought (HET). By reviewing which "techniques" are involved in our discipline, my aim is to find out whether for the historian of economic thought, too, there is any need for an apologia to the economist, the historian and the general audience at large.

For this purpose, I identify four broader categories in which HET can be classified: a) textual exegesis; b) "rational reconstructions"; c) "contextual analysis"; and, with a sort of catch-all definition, d) "historical narrative".

Although I refrain from endorsing a ranking of these techniques, my preferences—or, better, my favourite way of doing HET—deriving from my own personal experience and practice, will become apparent. Finally, I will draw upon *Econlit* records to review what has been done in our subject in the last two decades in order to frame some considerations on how our past may impinge on the future.

TEXTUAL EXEGESIS

It could be argued that textual exegesis (TE) is the technique *par excellence* for doing HET.[1] In return for the toil and trouble of the scholarship—namely the laborious and punctilious skill required in this type of exercise—it accords its practitioners the right to establish the "true" meaning of given texts. This technique defines the scope and method of the professional activity of a historian of economic thought, within the accepted hermeneutic codes.

Stigler (1965) has given us his recipe for good textual exegesis in HET, which encapsulates the demarcation criterion for deciding its scientific character, and which consists of reviewing texts in the light of the interpreter's contemporary economic knowledge. What is required is the ability to reconstruct the general position of "the theoretical core of

[1] It may be objected that textual exegesis is a tool that can be applied to any type of HET research. For my purpose, however, it is used to indicate that specific approach centred on "making sense" of a text or a set of propositions.

an author's work […] in a manner compatible with contemporary economic theory" (Emmett 2003, 525). Since for Stigler "the meaning of the text is determined not by the individual interpreter or even the original author, but by the scientific community of economists" (Emmett 2003, 525), it follows that doing HET is in all relevant respects no different from doing economics, and that any economist will, at least in principle, be endowed with the necessary skills.

Outstanding examples of the Stigler type of TE approach to HET are Hollander's large scale enterprise in interpreting classical political economy and Patinkin's investigation into Keynes's major works.[2] Far from being uncontroversial, in both cases the interpretations were challenged precisely on the grounds of their readings of the texts being framed by inappropriate theoretical contexts[3]

Hollander's reconstruction of Ricardo's corn ratio theory of profit and Patinkin' s identification of the principle of effective demand in Keynes's early 1930s writings are good examples of trying to make sense of the relevant passages or sentences by employing the logic of the particular neoclassical theory which was standard at the time the interpreter was writing.

RATIONAL RECONSTRUCTIONS

Rational reconstructions (RR) were the favoured and, indeed, the most popular technique for doing HET in the 1980s and 1990s. The ideas and insights of Smith, Ricardo, Marshall, Keynes, and Schumpeter were reconstructed in the light of either contemporary problems or modern economic analysis.[4] In just a few praiseworthy cases, investigations were undertaken with philological zest, contextualization, and excursus into unpublished materials, but in most cases the "reconstructions" were made to stand on the shoulders of the contemporary understanding of the issues addressed by past authors.

RR differs from TE in significant respects, the most important of which consists in the reformulation of the arguments of past authors into a modern theoretical framework, rather than "the construction of a theoretical position from the past author's work that can be contrasted with current knowledge" (Emmett 2003, 525).

---

[2] For my purpose, it will suffice to refer here to Hollander 1979; and Patinkin 1982.
[3] For Hollander, see Garegnani 1982; and Peach 1993; for Patinkin, see Kahn 1984.
[4] See Blaug 1990.

But what exactly is the technique identifying the RR mode of doing HET? Formalisation, model building, or other translations into more rigorous economic language are called for. Unlike TE, which is a search into the *meaning* of a text, RR is a *translation* of the ideas of past authors into concepts recognisable to modern eyes by dressing them up with modern tools. Since the two exercises are somewhat different, so are the techniques involved.

A related point is whether RR is used to support existing economic knowledge or to challenge current theory. In the former case, the past is sifted for the predecessors of modern theory and present ideas—I call this a "quest for ascendancy"; while in the latter case the past is searched for what has been lost and can no longer be found in modern theory—I call this a "quest for an alternative".[5]

RR is thus not just a variety of "Whig" history, whereby present-day theory is appointed the judge *of the past*, but can also be practised as a search *into the past* for alternatives. Pasinetti's (1974) early work on Ricardo and Keynes can be seen as an outstanding example of the RR approach; alongside Hicks' or Samuelson's incursions into HET, with their reconstruction of Hume-Ricardo-monetary trade theory or Hicks's work on the contributions to monetary theory by classical political economists and Keynes.[6]

However, the pursuit of precursors of contemporary concepts and theories sometimes gives way to what I dub "HET seasoning". This is the technique of identifying ascendancy, or seasoning current economic analysis with references to authors of the past: Keynesian, Schumpeterian, Wicksellian are adjectives intended to add "flavour" to models, such as fixed-price or short-period AS-AD, or endogenous growth, or to interest rate determination in dynamic disequilibrium.

This "HET seasoning" can be performed with more or less concern for the source from which it is derived. For example, when the terms Marshallian or Keynesian are used in analysis of current economic facts or problems, they are interpreted *as if* the theory explaining them were the same as that formulated by Marshall or Keynes. Thus short period unemployment (i.e., Keynesian equilibrium) or external economies (i.e., dynamic competition) identify special cases of what is assumed to be a more general framework of analysis. Thus, within the RR perspective,

[5] See Marcuzzo and Rosselli 2002.
[6] To give only two examples, see Samuelson 1971; and Hicks 1967.

historians of economic thought are seen to give their "serious" contribution to the advancement of knowledge whenever they *adapt* original concepts to fit contemporary analysis.

The RR mode ceased to be seen as a legitimate and respectable mode of doing HET with the rise of an alternative competing technique—historical reconstruction—which endowed HET with a more distinct character and autonomy as a discipline, but was also a factor in alienating the community of historians of economic thought from that of the economists. Historical reconstructions meant mastering a new technique involving, besides the published work, perusal of manuscripts and letters, and in general a familiarity with archival research methods, thus situating HET more firmly in the past rather than the present.

In fact, by making the past its present, HET's scope and concerns shift from textual exegesis and translation into modern economic language, to a more complex, puzzle-solving type of investigation; it requires acquaintance with facts, circumstantial and presumptive evidence, which have legitimacy per se, and not as a subsidiary to economic analysis. There are, however, two varieties of historical reconstruction, which share a straightforward endorsement of the historical method and a common suspicion of rational reconstructions, but differ in some significant respects. The first, I call "contextual analysis" and the other, for want of a better name, "historical narrative".

CONTEXTUAL ANALYSIS

By context I mean the set of questions and answers which framed theories and concepts, the intellectual interlocutors to whom they were addressed and 'the state of the art' at the time of their conception. The framework consists of facts regarding time, place and circumstances, about which knowledge and information have first to be dug out and then used to make sense of what is being interpreted or, as far as possible, illuminated.

The first conundrum is a matter of source evaluation; in historical investigations we are always confronted with this problem, but as far as HET is concerned there are two aspects to consider: firstly, how related materials, for instance correspondence, stand vis-à-vis purely scientific work (whether published or unpublished); and secondly, in each individual case, the importance of exploring archives rather than relying on published material alone.

Correspondence is the material upon which biographers build their narrative, providing clues to facts, circumstances and above all motives. Moreover, correspondence helps us place ideas in time and context, and thus leads us to ask questions which would probably not have been asked in those terms.

The second conundrum is theoretical legitimacy: the role of archives in filling the gaps in our knowledge of the personal and intellectual lives of the economists concerned is also unquestionably valid, but what is their value in increasing our grasp of their *theories*? My answer, drawing on my own work—mostly done jointly with Annalisa Rosselli—is that "papers and correspondence" afford insight into the *motivations* behind the choices of a particular set of questions, assumptions or tools. These are not always explicitly stated in the published version, where the solutions discarded and definitions abandoned are left out. Archives therefore allow us to travel the road towards a theory rather than, as it were, visit the final destination.

Sraffa's contextual analysis of Ricardo's theory of value and distribution, conducted using unpublished drafts, parliamentary speeches, and correspondence, besides published works, remains the unsurpassed model of scholarship and mastery of this technique (Sraffa 1951–1973). It is the model which has inspired much of my work on the Cambridge approach to economics, attempting to capture the lives and works of Keynes, Kahn, Joan Robinson, and Sraffa in *their* contexts, drawing on the evidence of the intertwined relationships they formed and the conditions of their times; my aim has been more to grasp their individual and shared concerns, rather than to fit their contributions into a unified core of doctrine.

HISTORICAL NARRATIVE

More recently, rational reconstructions and philological inquiries have been directed towards "minor" or lesser-known figures, in order to answer to the need to survey the broader picture. It is a sort of "stepping down from the shoulders of giants", searching for less theory-laden investigations, connecting intellectual circles, linking characters and events, mapping "tribes and territories". HET thus appears to be progressively diverted from economic ideas and concepts *strictu sensu*, bringing in new perspectives and evidence from hitherto unexplored sources, crossing the boundaries of a single discipline to leap into the broader realm of intellectual history.

Browsing the programmes of the most recent ESHET and HES conferences, or the works awarded recognition as best articles or books by HET societies, this trend can easily be detected. Sessions (or research) on the major classical or neoclassical authors (with the possible exception of Keynes) have shrunk in number and more and more intellectual energy is being devoted to penetrating the less explored territories.

This is a trend which holds particular appeal for young scholars, and it may indeed be a source of misapprehension for those who think that relations with the economic profession should be reinforced, and not loosened. However, some very interesting insights and reconstructions are coming in from these investigations, enlarging and broadening the scope of HET. One should welcome these new openings and encourage the upcoming generation to lead the way.

HET APPRAISAL

After this very brief overview of the techniques and styles of doing HET, I will now turn to the question of its appraisal. What is good HET? How can or should we appraise it?

This issue was raised by Roy Weintraub in his editorial piece in the HES list (Weintraub 1996). HET, he argued, requires a style of scholarship that is standard among historians (use of primary sources, circumstantial evidence, background knowledge, and so forth), but not among economists. His conclusion was that a "good" economist is not necessarily a "good" historian of economic thought and vice versa.

This line of argument has become increasingly popular since the 1990s,[7] being applied to criticise both mainstream and non-mainstream economists as prone to writing mainly "internalist" or even "whig" histories, drawing on their economics rather than mastering the historians' skills.

I personally believe that good HET shows the capacity to be versed in *both* economic theory and historical methods. It is the combination of these two skills and not simply being knowledgeable in one or the other that makes the difference in the quality of our scientific output. However, I would claim that there is a relevant distinction to be made between doing economics and doing HET.

[7] See, for instance, the works by Schabas (1992), Hands (1994), and Henderson (1996).

The task of the "historian" entails close adherence to the reconstruction of theories, personal contributions and the relevant circumstances, requiring the greatest possible respect for contexts and texts, both published and unpublished. The task of the "economist", working on texts and documents but, when expedient, also taking a certain distance from them, entails constructing a theoretical framework that is in some respects freer, and not bound by derivation from the authors. In this respect the theoretician is licensed to compose differences, connecting levels associated with different designs and conceptualisations in diverse authors or in the work of one and the same author.

An example of this distinction between historical and theoretical work which, significantly, has divided and continues to divide the economists working in the Cambridge tradition, has to do with compatibility between the approach of Keynes and the Keynesians, on the one hand, and that of Sraffa and the neo-Ricardians on the other. I feel that in our work as historians we must not be afraid of letting the differences emerge: rather, we must go on bringing to light all those elements of knowledge that offer an understanding of our authors in their historical backgrounds. In our theoretical work we are, in a sense, "freer" to interpret, integrate, and combine concepts and propositions that were quite distinct when formulated. This does not mean taking a cavalier approach to the historian's task but, rather, being clearer about the fact that a different aim is being pursued.

WHOSE HET?

For an understanding of how HET has actually been practised as opposed to how it should be practised according to some declared norms, it may help to review what has been done in our field in the last 20 years or so. This I did by browsing *Econlit* on the basis of descriptors B000-B590, which were introduced in 1991; although the classification and the chronology are far from being accurate and acceptable, this affords us a comprehensive overview of the extant HET literature.

Of course it has to be borne in mind that among the top ranked journals in our field only *History of Political Economy* has full *Econlit* coverage, while for *European Journal for the History of Economic Thought*, *Journal of the History of Economic Thought*, *History of Economic Ideas*, *History of Economic Review* the records start from the early 1990s.[8] However, these gaps are partly offset by the presence, up to the late 80s, of some HET articles in generalist journals such as *Journal of Political Economy*, *Economic Journal*, and even *American Economic Review*.

FIGURE 1

HET *ECONLIT* DESCRIPTORS POST 1991			
B000	Schools of Economic Thought and Methodology: General	**B250**	History of Economic Thought since 1925:Historical; Institutional; Evolutionary
B100	History of Economic Thought through 1925: General	**B290**	History of Economic Thought since 1925: Other
B110	History of Economic Thought: Preclassical (Ancient, Medieval, Mercantilist, Physiocratic)	**B300**	History of Thought: Individuals: General
B120	History of Economic Thought: Classical (includes Adam Smith)	**B310**	History of Thought: Individuals
B130	History of Economic Thought: Neoclassical through 1925 (includes Austrian, Marshallian, Walrasian)	**B400**	Economic Methodology: General
B140	History of Economic Thought through 1925: Socialist; Marxist	**B410**	Economic Methodology
B150	History of Economic Though through 1925: Historical; Institutional	**B490**	Economic Methodology: Other
B190	History of Economic Thought through 1925: Other	**B500**	Current Heterodox Approaches: General
B200	History of Economic Thought since 1925: General	**B510**	Current Heterodox Approaches: Socialist; Marxian; Sraffian
B210	History of Economic Thought: Microeconomics	**B520**	Current Heterodox Approaches: Institutional; Evolutionary
B220	History of Economic Thought: Macroeconomics	**B530**	Current Heterodox Approaches: Austrian
B230	History of Economic Thought: Econometrics; Quantitative Studies	**B590**	Current Heterodox Approaches: Other
B240	History of Economic Thought since 1925: Socialist; Marxist		

Since more than one descriptor can be assigned to each individual record in *Econlit*, we can sum up all the B descriptor records only by making sure we have eliminated multiple attributions. This I have not

[8] HOPE (1969–present); EJHET (Autumn 1993–present); JHET (Spring 1990–present); HEI (1993–present); HER (Winter 1994–present).

attempted to do. However, within each descriptor we can compare the numbers and identify the time profile of each of them.

In the great majority of cases, *Econlit* records only articles from journals and edited volumes, and books which have been reviewed in JEL, in the latter case with a strong bias for those written in English. Notwithstanding all these limitations, we can get a reasonably reliable picture of HET output over the last 20 years. The following figures offer a bird's-eye view.

FIGURE 2

CODE	HET *ECONLIT* DESCRIPTORS	N. PUBBL.	N. DISS.	CODE	HET *ECONLIT* DESCRIPTORS	N. PUBBL.	N. DISS.
B310	History of Thought: Individuals	10941	32	**B200**	History of Economic Thought since 1925: General	639	3
B410	Economic Methodology	3868	20	**B100**	History of Economic Thought through 1925: General	505	3
B220	History of Economic Thought: macroeconomics	2292	8	**B000**	History of Economic Thought and Methodology: General	450	
B120	History of Economic Thought: Classical	1760	9	**B190**	History of Economic Thought through 1925: Other	419	2
B250	History of Economic Thought since 1925: Socialist; Marxist	1714	11	**B530**	Current Heterodox Approach: Austrian	409	2
B130	History of Economic Thought Neoclassical through 1925	1539	1	**B230**	History of Economic Thought: Econometrics; Quantitative and Mathematical Studies	409	3
B520	Current Heterodox Approach: Institutional; Evolutionary	1389	8	**B290**	History of Economic Thought since 1925: Other	391	4
B210	History of Economic Thought: Microeconomics	908	2	**B300**	History of Thought: Individuals	254	1
B110	History of Economic Thought: Preclassical	856	3	**B500**	Current Heterodox Approach: General	149	2
B240	History of Economic Thought since 1925: Socialist; Marxian	820	4	**B320**	Obituaries	134	0
B510	Current Heterodox Approach: Socialist; Marxian; Sraffian	794	8	**B540**	Feminist Economics	104	0

B150	History of Economic Thought through 1925: Historical; Institutional	743	1	**B590**	Current Heterodox Approaches: Other	45	0
B400	Economic Methodology: General	706	1	**B490**	Economic Methodology: Other	39	0
B140	History of Economic Thought through 1925: Socialist; Marxist	705	3	**B160**	History of Economic Thought: Quantitative and Mathematical	27	0

As to the *topics*, the highest score is for individual authors, B310–HT: Individuals (10941); followed by B410–Economic Methodology (3868); B220–HET: Macroeconomics (2292); B120–HET: Classical (including Adam Smith) (1760); B250–HET since 1925: Historical, Institutional, Evolutionary, Austrian (1714); and B520–Current Heterodox Approaches: Institutional, Evolutionary (1389).

FIGURE 3

NUMBER OF *ECONLIT* RECORDS ON AUTHORS

Smith	1366	**Samuelson**	170	**Edgeworth**	87	**Cantillon**	42
Keynes	1207	**Pareto**	148	**Kaldor**	78	**Kahn**	42
Hayek	599	**Malthus**	147	**Harrod**	74	**Turgot**	42
Marx	535	**Friedman**	138	**Thornton**	73	**Einaudi**	34
Veblen	452	**Hicks**	131	**Bentham**	70	**Galiani**	33
Sraffa	381	**Say**	122	**Quesnay**	69	**Modigliani**	33
Schumpeter	376	**Wicksell**	121	**Cournot**	63	**Neumann**	29
Marshall	360	**Kalecki**	119	**Arrow**	62	**Nash**	27
Ricardo	281	**Knight**	107	**Robertson**	62	**Mandeville**	26
Walras	230	**Jevons**	106	**Bohm-Bawerk**	58	**Wieser**	23
Mill	206	**Menger**	104	**Robbins**	54	**Petty**	22
Fisher	196	**Hume**	98	**Georgescu-Rogen**	43		
J. Robinson	193	**Pigou**	93	**Sismondi**	43		

As to *individuals,* about 40% (4540) of the records (10941) are for the top 5 of 50 most studied authors: Smith: 1366; Keynes: 1207; Hayek: 599; Marx: 535; Veblen 452; Sraffa: 381. We may note in passing that the "10 great economists from Marx to Keynes", according to Schumpeter's list (namely, Marx, Walras, Menger, Marshall, Pareto, Bohm-Bawerk, Taussig, Fisher, Mitchell, and Keynes) are very poorly matched in the *Econlit* ranking of records.

FIGURE 4

B310: INDIVIDUALS

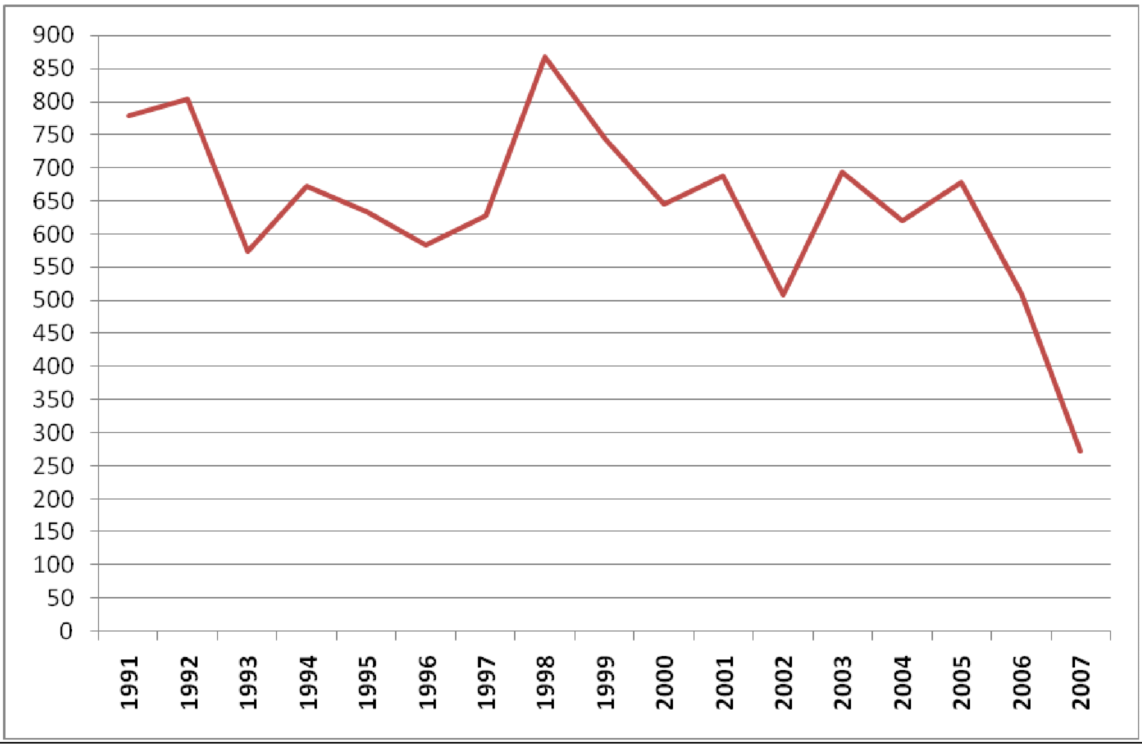
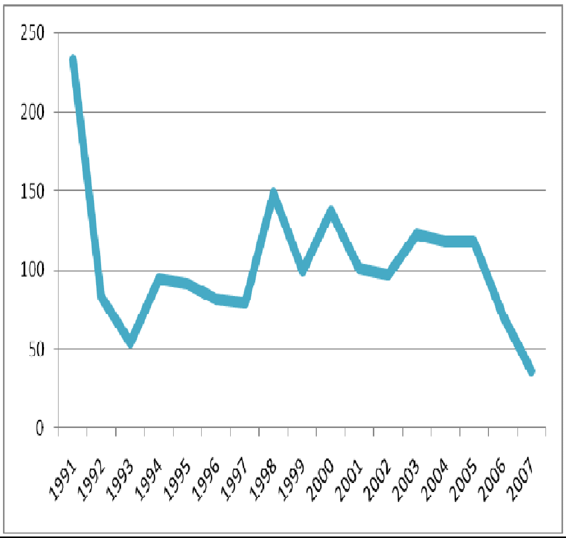


FIGURE 5

B410 ECONOMIC METHODOLOGY **B220 MACROECONOMICS**

FIGURE 6

B120: CLASSICAL (INCL. ADAM SMITH) **B250: SOCIALIST, MARXIST SINCE 1925**



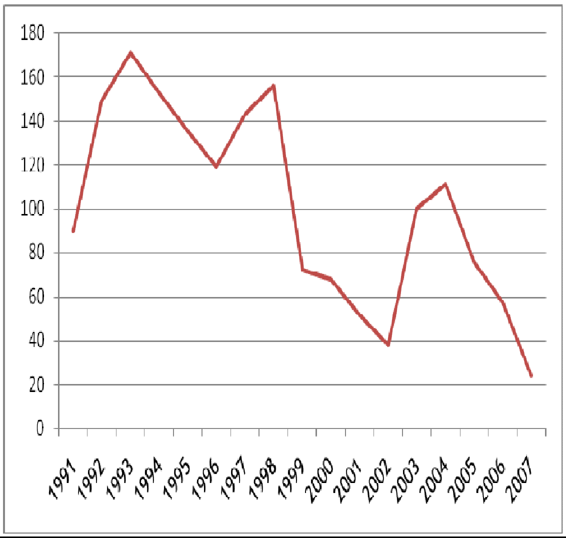
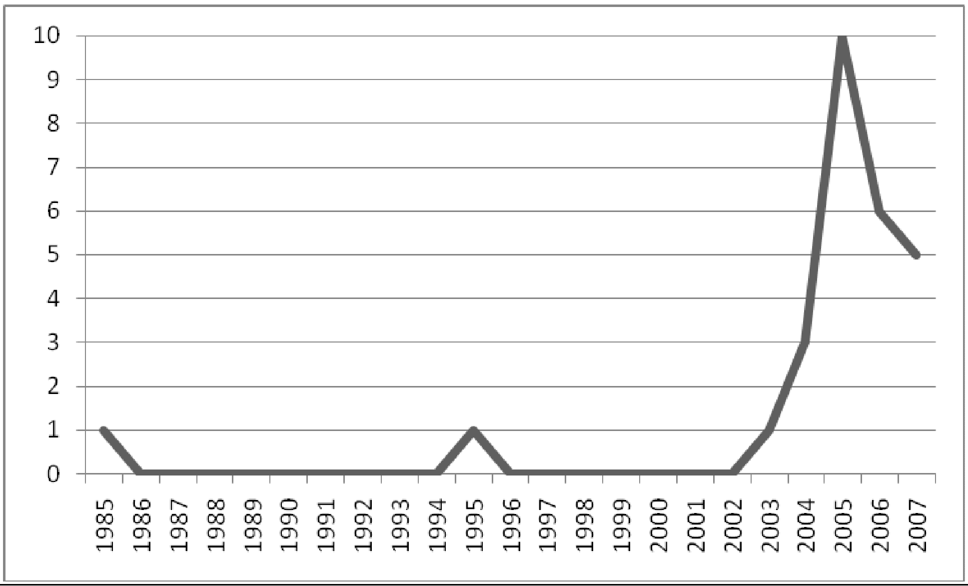


FIGURE 7

B160: QUANTITATIVE AND MATHEMATICAL



As to the *time profile*, there is a general tendency for a fall in the numerosity of records, as can be seen by taking the top 5 ranked descriptors, i.e., those with the highest number of records: B310–History of Thought: Individuals (10941); B410–Economic Methodology (3868); B220–HET: Macroeconomics (2292); B120–HET: Classical (1760); with the notable exception of Econometrics and Quantitative Studies, which has the lowest number of records.

As to the *dissertations*, although their number is small—totalling only 131 over the relevant period—they give us an interesting picture of the trend in the topics of research in the field (See Fig. 2). To be noticed is that the ranking slightly differs from that of the articles; while the top 2 ranked descriptors in both cases are B310 (History of Thought: Individuals) and B410 (Economic Methodology), the lowest numbers of dissertations are found in topics, such as B130 (History of Thought: Neoclassical through 1925), B210 (HET: Microeconomics) which are ranked among the first ten in the case of articles.

In order to be more significant, the exercise should be extended to the pre-1991 period, converting the old descriptors set to the new one; this would sometimes involve the arbitrary matching of topics and a large computation of records, which I have not attempted to do. Hopefully, however, what I have collected is sufficient to provide some evidence in presenting the outlook of our field in the last 20 years.

RESEARCH AGENDA

Can the past activity in HET offer us some insights into its trend for the future? Are the gaps likely to be filled and how? I will venture to point to two areas in which I think we may expect some rise in interest and activity.

The first is feminist economics. HET could be an important tool in the work of exposing an impossible neutrality and pervasive gender-blindness.[9] A gender-sensitive reading of past works and theories could open our eyes to the gradual shifts in meaning of the terms, the slow movement of the boundaries of the discipline, the progressive exclusion from it of whole areas of economic activity (housework for example) and of concepts which, though meaningful, lack a quantitative dimension.

The second is the broadening of geographical areas in which we may see interest and country-related research activity blossom in HET. My recent travels to Mexico, India, Japan, and China have proved to me that HET is a vital key to connecting ideas, as well as preserving identity and individual intellectual histories, and this in turn shows that there are indeed ways and means to establish a multicultural rather than monocultural discipline.

HET may help to enhance the ability to speak the same economic language, a necessary requisite of any scientific communication, with

[9] For a recent attempt in this direction see Marcuzzo and Rosselli (2008).

awareness of the variety and diversification in approaches to economic ideas and problems.[10] National societies for the history of economic thought are already there for us to connect and they could play a prominent role in bridging economic cultures and national backgrounds. Personally, I am looking forward to increasingly globalised research activity in HET.

CONCLUSIONS

I began my paper by asking whether there is any need for an apologia of the historian of economic thought, addressing this question to the economist, the historian, and the general audience at large.

To the economist, I would urge the importance of maintaining in the economics departments keepers of ideas and concepts born in the past, to preserve them from oblivion and the risks of being misused when uprooted from their context. Engaging in conversation and confrontation with contemporary economic discourse is an intellectual duty and while we should be wary of the consequences of alienating ourselves from it, we should make our case as boldly and fearlessly as possible. This is particularly true today, there no longer being a unified, mainstream core in economic analysis.

To the historian, I would plead the importance of ranking our priorities in fact finding and digging into unexplored sources; treasure hunts in economic archives do not have the same pay-off as in much intellectual history and should be done sparingly. Historical investigation should be a benchmark of scholarship for HET, not just another role-model, as mathematics or the natural sciences are for economists.

Finally, to the general audience at large I will make a plea for pluralism in economic analysis, in terms not so much of tools as of ideas. The critical awareness which HET cultivates of how economics is rooted in a given context of interests, ideology, and culture is the road to intellectual freedom and the recipe to advance our knowledge. HET is well placed to cater for the needs of the general public to have this clearly spelt out and understood.

Let me quote from Joan Robinson once again in conclusion: “Economics limps along with one foot in untested hypotheses and the

[10] For a recent illustration of the same point, see Foley 2008.

other in untestable slogans. Here our task is to sort out as best we may this mixture of ideology and science" (Robinson 1964, 28).

This is the challenge, the duty, but also the pleasure of doing history of economic thought.

REFERENCES

Blaug, Mark. 1990. On the historiography of economics. *Journal of the History of Economic Thought*, 12 (1): 27-37.

Emmett, Ross B. 2003. Exegesis, hermeneutics and interpretation. In *The Blackwell companion to the history of economic thought*, eds. J. Biddle, J. Davis, and W. Samuels. Oxford: Blackwell, 523-537.

Foley, Duncan K. 2008. The history of economic thought and the political economic education of Duncan Foley. Distinguished Guest Lecture, History of Economics Society Annual Conference, Toronto, June 27–30, 2008.

Garegnani, Pierangelo. 1982. On Hollander's interpretation of Ricardo's early theory of profits. *Cambridge Journal of Economics*, 6 (1): 65-77.

Hands, D. Wade. 1994. The sociology of scientific knowledge and economics: some thoughts on the possibilities. In *New directions in economic methodology*, ed. Roger Backhouse. London: Routledge, 75-106.

Henderson, James P. 1996. Whig history is dead: now what? The History of Economics Society Website. HES Archives. http://eh.net/pipermail/hes/1996-November/004822.html (accessed November 2008).

Hicks, John. R. 1967. *Critical essays in monetary theory*. Oxford: Clarendon Press.

Hollander, Samuel. 1979. *The economics of David Ricardo*. Toronto: University of Toronto Press.

Kahn, Richard F. 1984. *The making of Keynes's General Theory*. Cambridge: Cambridge University Press.

Marcuzzo, M. Cristina, and Annalisa Rosselli. 2002. Economics as history of economics: the Italian case in retrospect. *History of Political Economy*, 34 (Annual Supplement): 98-109.

Marcuzzo, M. Cristina, and Annalisa Rosselli. 2008. The history of economic thought through gender lenses. In *Frontiers in the economics of gender*, eds. F. Bettio, and A. Verashchagina. London: Routledge, 3-20.

Pasinetti, Luigi. 1974. *Growth and income distribution. Essays in economic theory*. Cambridge: Cambridge University Press.

Patinkin, Don. 1982. *Anticipations of the General Theory? And other essays on Keynes*. Oxford: Blackwell.

Peach, Terry. 1993. *Interpreting Ricardo*. Cambridge: Cambridge University Press.

Robinson, Joan. 1932. *Economics is a serious subject: the apologia of an economist to the mathematician, the scientist and the plain man*. Cambridge: Heffer and Son.

Robinson, Joan. 1964. *Economic philosophy*. Harmondsworth: Penguin, Pelican Books.

Samuelson, Paul A. 1971. An exact Hume-Ricardo-Marshall model of international trade. *Journal of International Economics*, 1 (1): 1-18.

Schabas, Margaret. 1992. Breaking away: history of economics as history of science. *History of Political Economy*, 24 (1): 187-203.

Sraffa, Piero. (ed., with the collaboration of M. Dobb). 1951–1973. *D. Ricardo. Works and correspondence*, 11 vols. Cambridge: Cambridge University Press.

Stigler, George J. 1965. Textual exegesis as a scientific problem. *Economica*, 32 (128): 447-450.

Weintraub, Roy. 1996. Legitimate contribution to H of E. The History of Economics Society Website. HES Archives. http://eh.net/pipermail/hes/1996-September/005765.html (accessed November 2008).

Maria Cristina Marcuzzo is professor of economics at the University of Rome, "La Sapienza". She is co-author of *Ricardo and the gold standard* (1991), co-editor of *The economics of Joan Robinson* (1996), and co-author of *Economists in Cambridge* (2005).
Contact e-mail: <cristina.marcuzzo@uniroma1.it>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 124-146.
http://ejpe.org/pdf/1-1-int.pdf

Realism from the 'lands of Kaleva': an interview with Uskali Mäki

USKALI MÄKI (Helsinki, 1951) is a philosopher of science and a social scientist, and one of the forerunners of the strong wave of research on the philosophy and methodology of economics that has been expanding during the last three decades. His research interests and academic contributions cover many topics in the philosophy of economics, such as realism and realisticness, idealisation, scientific modelling, causation, explanation, rhetoric, the sociology and economics of economics, and the foundations of new institutional and Austrian economics. He is a co-editor of *The handbook of economic methodology* (1998); *Economics and methodology: crossing boundaries* (1998); *Rationality, institutions and economic methodology* (1993). And the editor of two compilations of essays that have become highly influential to the shaping of the field: *The economic world view: studies in the ontology of economics* (2001), and *Fact and fiction in economics: realism, models, and social construction* (2002).

Currently, Uskali Mäki is academy professor at the Academy of Finland. He is also director of the project Trends and Tensions of Intellectual Integration (TINT), based at the department of social and moral philosophy, University of Helsinki. Before settling in Helsinki, he was professor of philosophy of science at Erasmus University Rotterdam from 1995 to 2006, where he was academic director of the Erasmus Institute for Philosophy and Economics (EIPE) since its foundation in 1997. He was co-editor of the *Journal of Economic Methodology* from 1996 to 2005; founding member, executive board member, and from 2007 to 2008, Chair of the International Network for Economic Method (INEM); and has been a research area coordinator for the European Association of Evolutionary Political Economy since 1992.

EJPE is pleased to present this interview with Professor Mäki, in which he offers some reflections on the aims, current situation, and prospects of the field, as well as on the development of his own thought about the philosophy and methodology of economics.

EJPE'S NOTE: This interview was conducted by Luis Mireles-Flores, PhD candidate at the Erasmus Institute for Philosophy and Economics (EIPE), Erasmus University Rotterdam, and co-editor of the *Erasmus Journal for Philosophy and Economics.*

EJPE: ***Professor Mäki, you have had formal training in both areas: philosophy and economics. When and why did you decide to specialize in philosophy of economics?***

USKALI MÄKI: How nice to be asked about those early years. That sweet nostalgia! It all happened in the early 1970s, at a time when our field—as an institutionalized research field—did not yet exist. So I was crazy enough to devote myself to a field that would come into existence only many years later. I never thought of it as a risky investment, but in a sense that is what it was, and indeed it turned out to be one that was to yield lovely returns later on. I don't think I actually anticipated the eventual emergence of our research field, but I did have a very strong opinion that it should. This normative obsession put me on that track.

The intellectual commitment and normative obsession derived from my early experience as an economics student. Having previously studied statistics, math, sociology, and philosophy, I started studying economics during my second undergraduate year. I recall I made the choice since I wanted a subject that would be both intellectually rigorous and socially relevant. But the early experience was somewhat shocking.

Based on everyday experience, I knew that I am not an expected utility maximizer, and I knew that the economy out there was far from perfectly competitive, and I thought I knew many other facts about society and human behaviour that those models that were taught to us appeared to distort so shamelessly. So I wondered what to make of economics, whether this is good science after all, and how on earth I could judge whether it is.

Another feature of the situation in the early 1970s that prompted similar questions was the popular claim that economics is in a crisis, and the related proliferation of rival schools such as versions of Keynesian and Monetarist, Austrian and Marxian, Institutionalist and early Behaviouralist approaches (no wonder Kuhn's notion of scientific crisis was frequently cited at that time). I wondered how to rationally judge the relative merits of these approaches. Where should my intellectual sympathies go, and how could I possibly justify my choice, whatever it might be?

So there was a challenge that could not be escaped. And it was a philosophical challenge, a challenge that could only be treated by exploiting philosophical concepts and theories about science. The step I took was to combine my studies in economics with my studies in philosophy: to look at economics from the point of view of the philosophy of science. Otherwise I could not possibly have survived my further studies in economics. Indeed, I did survive them, and this combination of the two disciplines itself not only survived, but was destined to flourish collectively in the later years.

I would like to mention that the world then was very different from the one that the present generation of aspiring philosophers of economics lives in. The literature was far more limited than today, there were no educational programmes or even competent individual guidance (I envy those of you with this privilege at EIPE and elsewhere today!), no collectively held research agenda was in place, authority structures characteristic of a research field were missing. The future was open, and the adventure could begin. It would become a wonderful adventure.

Were there any particular readings or authors that you recall as having an important influence on your interest in philosophy of economics?

Oh boy, those were years when I must have used most of my time for reading! Richard Lipsey's *An introduction to positive economics* was the textbook used in the introductory course I had taken. This book had an unusually long opening chapter that dealt with methodological issues, and the whole book was designed so as to bring theory and empirical evidence in some contact with one another. Lipsey had been a member of the M2T (for 'methodology, measurement, and testing') group that was influenced by Popper's ideas. Lipsey's introduction led me to read Friedman's 1953 essay on 'The methodology of positive economics' that I considered, on my first reading, a scandal, indeed an intellectually irresponsible apology for dubious economic theories. (As you know, my perception of Friedman's F53 has gone through major changes since that early exposure.)

I then started reading everything that I got hold of. I dug into the history of methodological and philosophical statements about economics and read all the classics and examined the debates that there had been, from Senior and Mill, Marx and Menger, Cairnes and Marshall, through Mises and Robbins, Hutchison and Hayek to Machlup and Samuelson, including the German and British *Methodenstreiten*, Keynes versus Tinbergen, the measurement without theory debate, and of course the F53 debates in the 1950s and 1960s.

I also read virtually everything that had been written in the course of the history of Finnish economics on methodological and philosophical issues. Later, I published a lengthy essay outlining the history of methodological thinking in Finnish economics around issues such as history versus theory, role of math, nature of models, and role of values in economics. The currents I was able to identify were similar to those in many other countries.

There was little new published during those early years in the 1970s, so it was easy to read everything that was. They included some of Larry Boland's articles as well as Spiro Latsis's papers and his edited 1976 volume on *Method and appraisal*. There was a peculiar book entitled *Rational economic man* (1975) by Hollis and Nell that I studied with great care. Later, much more was to appear by people like Alex Rosenberg, Dan Hausman, Mark Blaug, Neil De Marchi, Bruce Caldwell, Wade Hands, and an increasing number of others. But that then meant there was a research field in the making.

Naturally, readings in the philosophy of science were very important. On this I was not on my own, but rather received top rate guidance from professors in Helsinki: Raimo Tuomela, Ilkka Niiniluoto, and others. I studied basic texts such as Nagel's *The structure of science* (1961) and Carl Hempel's *Aspects of scientific explanation* (1965) plus many other authors popular at that time, such as Stegmüller, Popper, Lakatos, Feyerabend, and Laudan. My particular philosophical outlook, scientific realism, was influenced by authors such as Sellars, Cornman, Hooker, Smart, Boyd, Bunge, and Putnam. Among other very important readings were Ilyenkov on the abstract and the concrete, Nowak on idealizations, and Vaihinger on the *Als-Ob* (in 1979 I taught a course on idealizations and fictions in economics). I also read quite a lot of Rom Harré's works and found them inspiring. And I must confess I was influenced by Roy Bhaskar's first two books (1975 and 1979). I even used some of their vocabulary when I started teaching undergraduate courses in the philosophy of the social sciences in the last years of the 1970s. But as I tried to apply Bhaskar's ideas in my emerging realist philosophy of economics, within a few years I abandoned them as too simplistic for the purpose (as you know, some years later Bhaskar's ideas were discovered by Tony Lawson and used in arguments that I think distort facts about economics).

This early disappointment with Bhaskar helped me realize there was nothing available in the philosophical literature that would be directly applicable to such a complex and peculiar subject as economics (partly for these reasons, I also never got very excited about the project of applying Popper and Lakatos to economics). I had to start creating my own framework. This is in no way surprising. An up-to-date philosophy *of economics* did not exist. It had to be created.

From your personal perspective, what are the principal aims of a discipline like philosophy of economics?

There are many goals. Descriptive analysis of theories, methods and practices; diagnosis and explanation of epistemic performance; normative assessment and institutional design; and of course, not a fully separate task, the clarification of tricky concepts and implicit presuppositions. Economics is a very complex subject matter, and any given account of it will only highlight some of its limited aspects, serving only limited purposes. Overgeneralized and oversimplified accounts abound, and they are just that: overgeneralized and oversimplified. One cannot do all at once, both accounting and appraising, and perhaps suggesting revising the core features of economics in terms of one simplistic formula—even though these sorts of endeavour appear to have a lot of rhetorical appeal.

On the other hand, given that economics is such an immensely powerful epistemic institution in contemporary society, philosophy of economics should not remain an insulated puzzle-solving activity exercised in tall academic ivory towers. It should take on societal responsibilities in the collective and interdisciplinary monitoring of the epistemic and political performance of economics. Economics is all too important to be left to economists alone.

I recall the time when we had just created EIPE in Rotterdam in 1997. I envisioned a possible vocation for our future graduates in the philosophy of economics, one that would give the field a high profile as socially responsible and influential activity. The idea was simple, and I still think the world should welcome it. Given that economic theories and research results play such a powerful role in shaping policies and worldviews all over the place, and given that decision makers (with or without an education in economics) must rely on the expertise of professional economists, decision makers should consult experts *on* economics on top of experts *in* economics in order to be in a better position to judge the quality of information and advice provided by economists. This quality has to do with things such as reliability and various hidden background presuppositions. So there should be demand for expertise *on* economics, and this demand should be met by producing a supply of such expertise by way of educating specialists in the philosophy and methodology of economics. I optimistically envisaged EIPE would do just that.

Well, that vision has still some way to go to be fully implemented, but let me mention a small example that gives a hint as to what such a dream world could be like. In 2002, the Central Bank of Austria in Vienna organized a one-day workshop on 'truth in economics', and invited me to play a major role in it. I understood Austrian economists had been challenged with some sort of epistemic legitimacy issues, and they felt like needing some philosophical guidelines for making their case. The deliberations of the workshop were recorded and later broadcast on the Austrian radio. I found it a fascinating experience. Central Bank economists interested in issues of truth!

One gets an entirely different idea of the goals of the philosophy of economics when looking at it from the point of view of the philosophy of science. I can see two kinds of services. First, philosophy of economics is just one of the many philosophies of *X* (where *X* can denote physics, chemistry, biology, cognitive science, archaeology, etc.). At the highest level of abstraction, general philosophy of science produces accounts of science in general. Philosophies of economics, geology, psychology, and the like, produce accounts of their target disciplines at lower levels of generality. But naturally there is interaction between these levels of generality in both directions. Thus philosophy of economics produces accounts of its target discipline that may be used for purposes such as testing and developing more general accounts of science. In other words, philosophy of economics may have the goal of providing evidential and productive services to the rest of philosophy of science. A philosopher of science examining economics is welcome to inform other philosophers of science (who examine science in general or some other special discipline such as biology) about his or her discoveries concerning economics.

The second sort of service amounts to contributing to the "naturalization" (or rather "socialization" or generally "scientifization") of the philosophy of science. Consider science as one institutionalized form of the use of the human brain. In order to understand science, one has to understand its institutions as well as the human brain. Together with cognitive and other sciences, economics can be used as a scientific resource for this purpose. The supposition is that science has an economic aspect; it can be viewed as an economy. Now if economics is utilized in such a project as a resource, then it becomes necessary to analyze and assess that resource for its credibility and reliability. This is where philosophy of economics becomes indispensable.

You asked about my personal perspective. Let me take this to allow for tracing a development in my own orientation in producing and publishing my research. Even though I was trained in philosophy (like Alex Rosenberg and Dan Hausman), I was first employed by an economics department and regularly taught 'economic methodology' to economics students for more than a decade, from the late 1970s to the early 1990s. During that period, most of the other activists (like Neil De Marchi, Mark Blaug, Larry Boland, Bruce Caldwell, and Wade Hands) had a background in economics and worked for economics departments, and also were closely connected to the rising wave in the study of the history of economics.

Recall this was also the period when there was a lot of talk about the "crisis" of economics that I mentioned earlier. All this shaped much of the agenda of the field. It became largely a project of historically spirited normative appraisal, with some participants having the hope of somehow helping make economics better as an empirically controlled science. But the appraisal was based on rather limited concepts and questions, shaped by Popperian and Lakatosian frameworks. I never shared these frameworks, but I did address pretty much the same audiences as these fellow workers: namely other economic methodologists, historians of economics, practicing economists—rather than philosophers of science. Yet I think I largely acted like a philosopher, annoyed by conceptual confusions and obsessed with conceptual clarification. This must have had an impact on my style of writing, too. But I now feel I may have been too optimistic about making an impact. I fear some of my nuanced analyses may have been too much even for my fellow economic methodologists.

I have only later engaged myself more closely in the debates in the general philosophy of science. I have made the pleasant (and also in some other ways unpleasant) discovery that many of the ideas I developed while addressing non-philosophical audiences are now fresh and topical stuff for philosophers. This means I need to make another effort in presenting and reframing those ideas to a new audience (while, happily, getting a chance to refine them further). There are gaps between intra-field conversations that need to be bridged. Intra-field inquiries have not proceeded in step with one another, so there is a challenge for inter-field coordination. This task is helped by the fact that philosophy of economics has gradually established itself as a serious partner in the philosophy of science.

Let me mention an example of this last observation. Elsevier is presently busy with a giant publication project: a series of 16 volumes of handbooks in the philosophy of science. Next to volumes devoted to the philosophy of mathematics, physics, chemistry, biology, and so on, there will be one volume on the philosophy of the social sciences and another volume on the philosophy of economics. Not bad at all?

In 1992, you wrote that the method of isolation was ubiquitous in economics. What are your ideas today about the method of isolation in economic science? Has there been any significant change or expansion on your ideas about this topic?

This was indeed an important insight. Among other things, it helped me see the point of many of those disturbingly unrealistic assumptions that so badly annoyed me when I started my economics studies. The idea emerged gradually through my readings of Marshall, von Thünen, and Nowak and the rest of the Poznan school in the course of the 1980s. (Let me say here that it is a shame that many Anglo-American philosophers of science tend to ignore the fact that the philosophers of the Poznan school were the pioneers of the study of idealization in science. This is another example of harmful and unfair metropolitan provincialism, as I would call it.)

One important idea was to connect idealization and isolation. Idealizations are performed by false assumptions that suggest that a variable takes on values such as zero or infinity or other distorting value (zero transaction costs, instant adjustment, complete and transitive preferences, *ceteris paribus*). Such assumptions have a function, and one cannot judge those assumptions without understanding their function. I argued their function is often to help effect an isolation. This is what I've called the experimental moment in theoretical modelling. The key notion is that of controlling for other things in order to isolate one thing. The economist neutralizes those other things in order to let the isolated thing act on its own, as it were. In laboratory experiments, this takes place through causal manipulation, while in theoretical modelling it is accomplished by way of assumption. They both isolate.

I recall it was a relief to reach this insight. It would transform the terms of debate, I believed. From now on we could ignore simple criticisms of falsehood in assumptions. We should not focus just on individual assumptions and their realisticness without a good grasp of the function they serve in a larger context. This would also change the terms of philosophical labelling: just accepting and using false assumptions would not make anyone an instrumentalist. Falsehood is a tool for a realist, too. Among other things, in the exegesis of Friedman's 1953 essay, this helped me turn against the mainstream reading of him as an instrumentalist. I have argued for reading the essay as a realist statement instead.

There is a special challenge that a realist account of idealization and isolation must meet. This is the fact that I mentioned in my 1992 paper: many idealizing assumptions are motivated by mathematical rather than metaphysical reasons. They enhance the tractability of modelling and facilitate mathematical derivations. This is a concern that has been discussed by people like Frank Hindriks, Nancy Cartwright, and Anna Alexandrova. As I see it, the challenge is to develop criteria for assessing such tractability assumptions from a realist point of view so as to tell those that distort facts that shouldn't be distorted from those that don't.

Change or expansion? Oh yes, the framework keeps evolving. And it appears to apply widely. In my recent interventions into the debates over models and modelling, I have employed the idea of isolation in my MISS account of models (models as isolations and surrogate systems). In my work on explanatory progress, I have expanded the framework by incorporating the idea of isolations and de-isolations (as well as re-isolations) among both potential *explananda* and potential *explanantia* of a model or theory. These operations take place as responses to challenges in a "dynamics of debate" that drive explanatory progress. The roles of causal mechanism and explanatory unification can also be highlighted in the isolation framework. In 1992, I drew a distinction between intra- and inter-disciplinary isolation, but have only recently started applying it as part of the present project on interdisciplinarity. I am presently working on incorporating ideas about contrastivity and difference-making in the overall framework of theoretical isolation. I think this will further extend its applicability and fruitfulness.

Some of your initial work in the field has been focused on the analysis of two topics: Austrian economics and the Rhetoric of economics. Can you elaborate on how these subjects have played a part in the development of your thinking? Perhaps you can also briefly sketch your current opinions on both themes.

Indeed, that's right. The two stories are somewhat different. At least three reasons lie behind my early interest in Austrian economics. It was one of those traditions that experienced a mass scale revival in the 1970s. It offered what seemed to be the strongest case for free market thinking. Perhaps most importantly, maybe next to Marxian economics, it has been the most philosophically self-reflective tradition in economics. Thanks to this last feature, there was a lot to read and analyze in the philosophical and methodological writings of Menger, Mises, Hayek, Lachmann, as well as in the secondary philosophical literature. I also read what these and other authors, most importantly Israel Kirzner, wrote in their economic work.

The papers I then published on Austrian economics dealt with its methodology and metaphysics. They were intended to serve two purposes: to provide novel interpretations of some Austrian ideas and to develop ideas for more general use. For this latter purpose, Austrian economics served as a source of inspiration and as a test ground for philosophical inquiry. I offered a new reading of Menger's idea of economics as an exact science in terms of recent philosophical work on laws as second-order universals (by David Armstrong and others). I analyzed notions such as money (as a collection of causal powers), the market process (as a causal process; here I modified Wesley Salmon's account of causal process), entrepreneurship (as a causal power), the relationship between realism and subjectivism (as a combination of ontic subjectivism and ontological objectivism), and the invisible-hand mechanism and invisible-hand explanation (as essentialist and how-possibly explanation). There was also a contribution to the literature on hermeneutics and Austrian economics that made an interesting start in the late 1980s but seems to have discontinued (which is a big pity, in my view). I believe all these ideas are still relevant to contemporary concerns not only in regard to Austrian economics, but more widely. I should perhaps now emphasize their possible broader relevance given that few Austrian economists seem to have paid much attention to these papers. Partly due to the unresponsiveness of Austrian economists to my work, I haven't done anything in this area for many years (yet one day I hope to pull together these contributions in the form of a book). My current work looks more into areas such as new institutional economics, behavioural economics, and geographical economics.

The story behind my interest in the rhetoric of economics is different. In the beginning of the 1980s, Ronald Coase and Willie Henderson published papers on rhetoric and metaphor in economics, then in 1983 there appeared D. McCloskey's famous piece in the *Journal of Economic Literature*, and Arjo Klamer's *Conversations with economists*. McCloskey and Klamer launched a campaign in support of the rhetorical perspective, combining it with some very radical philosophical claims. What happened was that they proposed joining the recognition of rhetoric in economics with antirealist philosophy in one package, as if they belonged together: if you choose rhetoric, you also must choose antirealism. Many readers were misled to consider the recognition of rhetoric as part of such a package. Some bought the package, some others didn't. Some bought antirealism because they believed they had to, otherwise they wouldn't get the valuable idea of rhetoric. Some others did not buy rhetoric because they believed they would then have to buy antirealism as well, and this turned them away.

What I saw was conceptual confusion and ungrounded antirealism, and this triggered my pedantic obsessions and realist instincts. I set out to demolish the package. One result was an ongoing debate with McCloskey and Klamer; it has now lasted more than twenty years. Another result was an account of rhetorical realism, or realist rhetoric.

So I have tried to show that rhetoric and antirealism do not necessarily belong together, and that a much better option would be to combine rhetoric with realism. Rhetoric is real and powerful in scientific practice, so it must be recognized and examined. But rhetoric is neutral with respect to the realism versus antirealism issue. So one is free to link rhetoric with realism.

This project, and the controversy with McCloskey, has been a lot of fun and also very useful. I have been forced to develop an account of rhetorical realism as an alternative to rhetorical antirealism. Another nice thing is that the study of the rhetoric of economics has highlighted one way in which economics is a socially shaped activity. It is unfortunate that the rhetoric of economics project does not seem to have made progress for many years now. There is much more to be done here by serious students of the rhetoric of scientific inquiry.

Not long ago, in 2005, you published an article explaining and arguing for what you have labelled 'local scientific realism'. Can you elaborate on how this conception differs from traditional realist positions towards science?

This idea is related to the differences of levels of generality in philosophical accounts of science that I mentioned earlier. The arguments for local scientific realism provide an instructive case against the popular practice of borrowing ideas from general philosophical literature and applying them directly to economics or any other specific discipline.

The dominant conceptions of scientific realism in the philosophy of science are supposed to offer general accounts of science. But I think they largely fail as such accounts. And I think they do not fail because some disciplines are not real sciences or because some disciplines had better be interpreted in antirealist terms. I think they fail because they are too thick and specific. And I think they are too thick because they are designed so as to fit with some of the most successful parts of physics. So in fact they are local realisms, but they are typically presented as global or general views of realism about science.

Among the typical ingredients in these supposedly general conceptions one can find the ideas that scientific theories postulate unobservables (the electron serving as the paradigm example); that those entities exist mind-independently; that current theories about them are mostly at least approximately true about them; that thanks to these achievements, scientific theories are predictively and technologically successful.

These conceptions of what scientific realism entails about science have then prompted criticisms and debates such as those related to the 'no-miracle' argument and pessimistic induction. They all take place within the framework of those principles without questioning them.

In my view, this is fine and nice, but only within limits. Beyond those limits, the consequences are unpleasant for those scientific disciplines and research fields that do not conform to such principles. Either they do not qualify as science at all or they are expelled into the arms of antirealist philosophies of science. Like many other disciplines, economics would immediately go to one of these dustbins.

The way to avoid such consequences is to do two things. First, if we want to have a global or general scientific realism, it must be made very thin and abstract. I have called such a global version minimal scientific realism. I have suggested its principles include that the objects of scientific theories *may* exist (rather than exist); that they exist (if they do) *science*-independently (rather than mind-independently); that current scientific theories are *possibly* true (rather than true). And nothing is required about unobservability or technological success.

The second thing is to go local when considering any particular scientific discipline in realist terms. Those global minimal principles are then specified and amended depending on what is the case at any particular local level. This means we are likely to have a number of local scientific realisms tailored for specific disciplines and fields, and perhaps theories: scientific realism about chemistry, about geology, about quantum mechanics, about evolutionary biology, about microeconomics, and so forth. Naturally, local and global realisms should be in harmony with one another. Minimal global realism should be implied by all local realisms.

I hope this vision will help to modify the terms of the realism-antirealism debate and also to rehabilitate the importance of local philosophies of science, such as the philosophy of economics.

And what would a local realist approach to economics look like then?

Well, this is exactly the big ongoing project, so no final formulations can be given yet. But surely many ideas can be outlined at this point. That economics is largely a non-experimental social science has major ramifications for any idea of realism about it.

Economic theories do not seem to postulate unobservable entities akin to electrons. Economics is about *commonsensibles* as I've called the various objects that are familiar from everyday experience: firms and households, preferences and expectations, money and prices, wages and taxes, etc. These things do not exist mind-independently, but they do have a fair chance of existing science-independently provided we take this in a constitutive rather than in a causal sense. The causal sense of science-dependence can be permitted to take care of situations in which ideas produced by academic economics are adopted by social actors with consequences for their behaviour, as in the so-called self-fulfilling and self-defeating prophesies.

Scientific realism about economics is an apt position also because the *explananda* of economic theories are so often products of various invisible-hand mechanisms. The causation of what happens is often not transparent; therefore scientific models of these non-apparent invisible-hand mechanisms are needed. And what science identifies as causally responsible exists independently of that science in the sense of not being conceptually constituted by it.

However, predictive and technological success cannot be required of economics in order for it to be compatible with scientific realism. And given the massive epistemic uncertainty when dealing with an immensely complex and effectively uncontrollable subject matter like society, we cannot require that economic theories and models be established as true as a condition for realism to apply. What we can include in a realism about economics is a normative dictum that truths about the real world should be pursued.

As a special realist principle of epistemic justification let me mention the idea of ontological unification. The capacity of a theory to unify a variety of different kinds of phenomena can be taken as speaking in favour of the theory's truth. The intuition is that it would be strange if a false theory had this capacity. But the realist should add that not just any sort of unification will do. A theory that can be used only for logically deriving descriptions of various classes of phenomena may also be false. So one should require that the theory unifies the phenomena ontologically by showing that they are of the same kind after all: they are made of the same stuff or are produced by the same causes, and so on. Now this is very relevant to the analysis of economics given that economics is obsessed with taking unification as far as possible, also beyond its traditional disciplinary boundaries by aspiring to explain not only phenomena of money and trade but also those of marriage and crime in terms of rational choice in a market. For a realist to regard this favourably, ontological unification is required.

In addition to your endorsement of scientific realism, and perhaps as a consequence of it, your work is also full of references to the notion of truth. What is the role of truth in your philosophy of economics? And, furthermore, is economics a science that aims at truth?

You are right, I've been rather unashamed in talking about truth. As you know, we are living the age of "bullshit" (as Harry Frankfurt puts it), characterized by an irresponsible lack of interest in truth and what is true. I don't share this cultural inclination at all, but rather follow my own strong and perhaps naïve intuitions. If a representation suggests that *F* is the case and if as a matter of fact *F* is the case, then the representation is true. That's about the simplest way of putting the intuition.

Economists—and it seems most philosophers of economics—have an uneasy relationship with the notion of truth. At least they largely try to avoid using terms such as 'true' and 'truth'. At the same time, there seems to be no similar difficulty with 'false' and 'falsity'. Numerous surrogate terms are actively used, such as 'right', 'correct', 'valid', and so on. But I don't think I've ever seen the meanings of such "escape terms" explained. So using them offers no improvement compared to using the terminology of veracity.

Then there are those who are happy to use the terminology of truth, but do not intend it to be taken literally. When they talk about 'truth', they turn out to mean something different. They reduce truth to something else, such as predictive success, persuasiveness, coherence, or socially constructed agreement. Such views of truth enjoy some popularity, and I have resisted them by trying to reveal their counter-intuitive implications. For example, long ago there was a socially constructed agreement that the earth lies at the centre of the universe, yet this collectively held belief was false; not because people have changed their minds, but because of the structure of the universe. On my rhetorical realism, one does not produce the world and truths about it by persuading audiences. Truth is not a matter of persuasiveness as on the antirealist view of rhetoric, but persuasion in appropriate institutional conditions may promote the discovery, communication, and acceptance of truths about the world.

I can be pretty precise about the role that the notion of truth has played in my arguments about economics. There is the normative role: economics should pursue truths about the economy. The descriptive role is more nuanced. I don't claim economists pursue truths (while I believe some do, some others don't). I don't claim that (most, many, or any) economic models and explanations are true. Many of my arguments are even-if arguments: Even if so-and-so, this model or explanation may be true. Even if its assumptions are false, a model may be true. Even if the model radically simplifies an immensely complex real-world phenomenon, it may be true. Even if a model predicts poorly, it may be true. Even if only few economists (or none at all!) are persuaded to accept a model, it may be true. This I take to be sufficient for a realism about truth in economics.

So I have put forth possibility arguments. These arguments do not imply that any given theory or model is actually true. They just suggest that a theory or model may be true even though it has some further properties that might appear to speak against its truth. Now, it would be nice to get from such possibility judgements to claims about actually achieved truth. I have tried to take steps towards this direction by outlining some further constraints the meeting of which makes a difference for the likelihood of actual truth acquisition. These include the idea that persuasion among economists had better take place in dominance-free institutional conditions and the idea that economic theories and models had better be in line with our general views about the way the world works (this is what I've called the ontological www constraint). But I think there are limits beyond which one cannot get in one's capacity as a philosopher of economics. It is ultimately up to (perhaps philosophically informed) practicing economists to judge whether any given theory or model actually is or is not true.

Of course, the difficult issue remains: what is it for a model to be true? Philosophers of science have neglected this issue or have adopted the straightforward position that models are the sorts of entities that cannot be true. In a forthcoming article, I argue that perhaps they can, and I support this with some novel arguments that re-examine both notions, those of 'model' and 'truth'. I am very curious to see how this initiative will be received.

Indeed, some of your most recent work has been on the role of theoretical models in economic science. Can you give a more detailed account of your position on this topic and of how it connects to your previous research?

I must say I am very excited about this theme. Given that economics is very much a modelling discipline, this helps me understand many characteristics of the subject. Models and modelling are also among the most popular topics in the philosophy of science today. In the philosophy of economics, valuable contributions have been produced by Mary Morgan, Robert Sugden, Marcel Boumans, Nancy Cartwright, Tarja Knuuttila, Julian Reiss, and others. I am developing an account of theoretical models that is slightly different from the others that have been proposed.

In my account, models are imagined small worlds that can be described variously, such as verbally, visually and mathematically. These imagined toy worlds serve as surrogate systems that can be used as representatives of some real world systems (or some other target systems, such as theories). The epistemic point of such surrogate systems is the wish that by directly examining their properties and behaviour ("let's see what happens in this model"), the modeller will indirectly learn about the target system. For this to be possible, the model must resemble the target system in certain important respects. In order for the model to qualify as a representation of a target, I do not require that it actually does resemble, but only that issues of resemblance can be reasonably raised. Moreover, in order for a model to resemble its target, it is sufficient that there be just very limited similarities between the two; no detailed and comprehensive correspondence is needed. The desired and required respects of resemblance in any particular context are determined by ontological constraints (the properties of the target) and pragmatic constraints (the purposes and audiences of modelling). All these various elements of models as representations are identified and coordinated by what I call a 'model commentary'.

Isolation is in the picture as a major part of this MISS account. Models as imagined toy worlds have the function of isolating limited aspects of their targets. Models do not characterize their targets in all their rich detail, but rather pick out features that are viewed as relevant for some purpose of model use. Idealizing assumptions are elements in model descriptions that serve as vehicles for making these isolations explicit. Models often isolate causal mechanisms in very skeletal form, and it is hoped that the mechanism in the imagined model world is also in operation in the real target world.

While this account should help to swallow a lot of unrealisticness in models as entirely reasonable and well-taken, it should also help distinguish good and bad modelling exercises from one another. I have suggested a distinction between *surrogate* models and *substitute* models for this purpose. Surrogate models can be intended as bridges to their targets thanks to the fact that issues of resemblance are taken seriously, so one hopes to learn about the target by examining the model. Substitute models, on the other hand, literally substitute for their target, and no issues of resemblance arise: examination of the model only informs about the model but provides no information about the properties of any target system.

As you can see, this is a way of reconceptualising some of the age-old issues in, and about, economics. Naturally, the notions of surrogate model and substitute model, and how one should go about recognizing and distinguishing them in practice, require a lot of further scrutiny.

You are currently heading a seemingly very ambitious research project called Trends and Tensions in Intellectual Integration (TINT), sponsored by the Academy of Finland. Can you please explain what are the main characteristics and goals of this project?

This is indeed an ambitious endeavour. It is motivated by current developments in the social sciences in their interdisciplinary relations. Looking at these developments from the point of view of economics, there are two trends, both involving interesting tensions. Economic ideas are increasingly used in the study of phenomena traditionally examined by other social sciences such as sociology, political science, law, human geography, and science studies. Economics itself, in particular its depiction of human behaviour, is increasingly put under the pressure of progress in experimental psychology and neurobiology. It is various aspects of this complex web of trends and tensions that we examine. The perspective is mainly philosophical. This means that the core concepts include those of model, mechanism, explanation, unification, reduction, emergence, level, domain, progress, and of course, those of discipline and field.

We expect to learn many sorts of thing. One is the variety and mechanisms of interdisciplinary interactions, including how special disciplines may resist engaging in such interactions. The other is the nature of participant disciplines. Interdisciplinary interactions provide a particularly revealing source of information in this respect. So we learn about political science and sociology, psychology and neurobiology, and given that economics is a major focus of attention, we learn new things about economics by looking at how it relates itself to other disciplines.

The undertaking is not only interdisciplinary itself, but also very international. We have been able to recruit some of the best post-docs in the field to our ranks from abroad, such as EIPE graduates Emrah Aydinonat and Caterina Marchionni, as well as Till Grüne-Yanoff who graduated from the LSE. We also have an active visitors programme through which TINT has hosted many PhD students and more advanced players from abroad. The Finnish team members are very competent, creative and productive, including Petri Ylikoski, Aki Lehtinen, Jaakko

Kuorikoski, Tarja Knuuttila, and others. It is a fantastic team. Regular seminars, workshops and conferences are among our working tools.

And to conclude, could you offer a succinct diagnosis of the present state of philosophy of economics and perhaps some of your expectations about its future development?

Well, that's much to ask, but it is a nice challenge. First of all, through all these years I have witnessed tremendous progress in the field. The field is now far larger than ever. Many more people are active in it. And many of the new activists are well educated for the task. Educational programmes—most notably that of EIPE—are making a difference. Institutes and greater concentrations of experts (e.g., Rotterdam, Helsinki, London, Amsterdam, Madrid, Duke, Alabama, Buenos Aires) are proliferating. New topics and issues have been addressed; many previously dark aspects of economics have been illuminated. Standards of quality are improving. All this makes me very happy.

From a social point of view, five trends strike me as important. A numerous and capable younger generation is entering the field and is making new important initiatives. A growing proportion of the activity now takes place in Europe relative to North America. A growing share of the activity now takes place in philosophy departments relative to economics departments. The division of intellectual labour is growing in the field: practitioners increasingly specialize in limited topic areas. This trend is an indication of the maturity of the field, but it has the unfortunate consequence that sound synoptic overall visions will become harder to create. Finally, partly in reaction to that, there will be more collaboration, both among specialists in the philosophy of economics and between them and others, such as practicing economists.

As to the topics of our inquiries, some presently popular ones will stay and others will emerge. I expect issues related to assumptions, models, and their realisticness to remain at the core of the field. Further inquiries into economic causation will be made. Economic explanation is one of the under-researched topics, and much more will be done on it. Prediction and forecasting are badly neglected, so there is a call for more attention. Traditional issues of testing will stay on the agenda and will be addressed by looking at the nature and roles of a variety of kinds of evidence. I expect the issue of scientific progress to make a comeback onto the agenda, but this time framed in updated philosophical terms.

The trend has been towards analyses informed by case studies, and I expect this to continue. I also expect this to be balanced by sophisticated conceptual work on some of the meta-theoretical notions that are not sufficiently well understood. There will also be further analyses of some core concepts of economics, such as those of rationality, wellbeing, market, money, firm, and others.

I expect more focus on interdisciplinary relations since they increasingly shape and reshape theoretical and explanatory activity in economics, both in relation to other social sciences and in relation to cognitive and life sciences. The coming years will see more analyses of fields such as behavioural economics, neuroeconomics, institutional economics, evolutionary economics, and geographical economics, but also of the prevailing ambitious trends towards integrating the social sciences with the cognitive and life sciences. The contributions by folks like Don Ross, Jack Vromen, John Davis, Harold Kincaid, Erik Angner, Caterina Marchionni and myself will be followed up by many others.

I expect there to be more attention to interdisciplinary relations also in practical policy contexts, such as those of climate change and health care. Related to this, I expect the philosophy and methodology of applied economics outside of academia to become an honourable area of inquiry. This is also very much needed, given its scale and societal importance (maybe some of our young experts who will not stay at universities, but will find jobs at the various economic research institutes, are able and willing to take on this task as part of their new job description).

I expect methods used or usable by economics to attract more attention. The scrutiny of econometric and experimental methods will continue, building on the work of people like Aris Spanos and Francesco Guala, and this will be supplemented by analyses of simulations and surveys. I expect there to be debate over qualitative methods as in other social sciences. In this connection, hermeneutics will make a comeback. And I expect there to be more work on the philosophy of macroeconomics, not just by Kevin Hoover and Roger Backhouse.

I expect the trend towards a more social image of economics to continue. Not only are we going to see more detailed analyses of the social structure and dynamics of academic (and hopefully non-academic) economics, but we are going to be more informed about the external social contexts in which economics has evolved and is being done. Others will follow Phil Mirowski's footsteps. It is a special challenge to draw philosophical conclusions from these studies.

I expect the trend to continue towards economics itself, together with cognitive science, playing a growing role as a scientific resource in science studies as in Jesus Zamora-Bonilla's work. Economics of economics will be an exciting special case that offers many new opportunities for self-referential investigations. I also expect and hope that philosophers of economics will contribute, in the spirit of social epistemology, to the redesign of academic institutions that are likely to enhance the capabilities of economic inquiry to produce true and (societally and humanly) relevant information about the real world.

May I conclude with some observations about audiences? There has been a chronic complaint (sometimes associated with a sort of self-pity) that practicing economists are not interested in what philosophers and methodologists say about economics. Some think this is an outright failure of our field. The premise behind this is that making a difference for research practices in economics is a major goal of our meta-scientific activity.

I would look at this issue somewhat differently. I do strongly believe that a close contact with economic research practice is important for epistemological reasons so to speak: to be adequate, philosophical accounts of economics must be well informed about what they are about. I also think that reaching the audience of practicing economists would be nice, and we should work for it, but still I would not think of a failure in this task as manifesting a fatal failure of the field (it is perhaps as much a failure of economics: it takes two to tango after all). There are other important audiences—such as philosophers of science, other social and natural scientists, policy makers, lay public—that have or should have an interest in being enlightened by our philosophical analyses of economics. At least the first two of these are increasingly receptive to what we say. Philosophers of science have welcomed philosophers of economics to their ranks and are eager to learn from us. Other social scientists are puzzled by the increasing intrusion of economic ideas into their disciplines and are also eager to learn from us in deciding what to make of this.

As I said earlier, economics is too important to be left to economists alone. By extension, the fate of the philosophy and methodology of economics should not be left at the mercy of economists alone.

SOME RECOMMENDED TEXTS WRITTEN OR EDITED BY USKALI MÄKI:

[Forthcoming] Models and the locus of their truth. *Synthese*. http://www.helsinki.fi/filosofia/tint/maki/materials/ModelsSyntheseC.pdf

[Forthcoming] MISSing the world: models as isolations and credible surrogate systems. *Erkenntnis*. http://www.helsinki.fi/filosofia/tint/maki/materials/ModelsErkenntnisE.pdf

2009. *The methodology of positive economics: reflections on the Milton Friedman legacy*, ed. Uskali Mäki. Cambridge: Cambridge University Press.

2005. Reglobalising realism by going local, or (how) should our formulations of scientific realism be informed about the sciences. *Erkenntnis*, 63: 231-251.

2005. Economic epistemology: hopes and horrors. *Episteme. A Journal of Social Epistemology*, 1 (3): 211-220.

2004. Theoretical isolation and explanatory progress: transaction cost economics and the dynamics of dispute. *Cambridge Journal of Economics*, 28 (3): 319-346.

2004. Realism and the nature of theory: a lesson from J. H. von Thünen for economists and geographers. *Environment and Planning A*, 36 (10): 1719-1736.

2003. 'The methodology of positive economics' (1953) does not give us *the* methodology of positive economics. *Journal of Economic Methodology*, 10 (4): 495-505.

2002. *Fact and fiction in economics: realism, models, and social construction*, ed. Uskali Mäki. Cambridge: Cambridge University Press.

2001. *The economic world view: studies in the ontology of economics*, ed. Uskali Mäki. Cambridge: Cambridge University Press.

2001. Explanatory unification: double and doubtful. *Philosophy of the Social Sciences*, 31 (4): 488-506.

2000. Kinds of assumptions and their truth: shaking an untwisted F-twist. *Kyklos*, 53 (3): 303-322.

2000. Performance against dialogue, or answering and really answering: a participant observer's reflections on the McCloskey conversation. *Journal of Economic Issues*, 34 (1): 43-59.

1999. Science as a free market: a reflexivity test in an economics of economics. *Perspectives on Science*, 7 (4): 486-509.

1998. *The handbook of economic methodology*, eds. John B. Davis, D. Wade Hands, and Uskali Mäki. Cheltenham: Edward Elgar.

1998. *Economics and methodology: crossing boundaries*, eds. Roger Backhouse, Daniel Hausman, Uskali Mäki, and Andrea Salanti. Houndmills: Macmillan.

1998. Against Posner against Coase against theory. *Cambridge Journal of Economics*, 22 (5): 587-595

1995. Diagnosing McCloskey. *Journal of Economic Literature*, 33 (3): 1300-1318.

1994. Isolation, idealization and truth in economics. In *Idealization in economics*, eds. Bert Hamminga, and Neil De Marchi, special issue of *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 38: 147-168.

1993. *Rationality, institutions and economic methodology*, eds. Uskali Mäki, Bo Gustafsson, and Christian Knudsen. London: Routledge.

1992. On the method of isolation in economics. In *Idealization IV: Intelligibility in Science*, ed. Craig Dilworth, special issue of *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 26: 319-354.

1992. The market as an isolated causal process: a metaphysical ground for realism. In *Austrian economics: tensions and new developments*, eds. Bruce Caldwell, and Stephan Boehm. Boston: Kluwer Academic Publishers, 35-59.
1990. Scientific realism and Austrian explanation. *Review of Political Economy*, 2 (3): 310-344.
1990. Mengerian economics in realist perspective. *History of Political Economy*, annual supplement, 22 (5): 289-310.
1988. How to combine rhetoric and realism in the methodology of economics. *Economics and Philosophy*, 4 (1): 89-109.

Uskali Mäki's Website: <www.helsinki.fi/filosofia/tint/maki/index.htm>

TINT Website: <www.helsinki.fi/filosofia/tint/>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 147-153.
http://ejpe.org/pdf/1-1-br-1.pdf

Review of Donald MacKenzie's *An engine, not a camera: how financial models shape the markets.* Cambridge (MA): MIT Press, 2006, 377 pp.

JOB DAEMEN
EIPE, Erasmus University Rotterdam

This is an important book which has received a lot of attention from various corners. That attention is well deserved. *An engine, not a camera*, written by eminent sociologist Donald MacKenzie, is a compelling and accessible story about the links between science and reality, theory and practice, in the setting of financial markets. He focuses on the performativity of finance theory: how the theory has profoundly influenced, shaped, and constructed the practice of financial markets. The book is the culmination of a number of papers authored and co-authored by MacKenzie on this subject. The performativity of economics, an idea coined by Michel Callon, entails that economics performs, shapes, and formats the economy rather than merely observes how the economy or a particular economic process functions. Given the number of citations and references, *An engine, not a camera* appears to present a strong case for the performativity thesis.

In taking financial economics and financial markets as his case in point MacKenzie has chosen well. The insight that academic theory has deeply influenced the practice of financial markets is not new by any means. Peter Bernstein makes that claim in his 1992 book *Capital ideas: the improbable origins of modern Wall Street*. Prominent finance scholars such as Stephen Ross and Merton Miller have done so as well. But MacKenzie's approach is interesting in at least two regards. First, using sociological concepts, methods, and tools, he provides a plausible explanation of how exactly the performative effect has come about. Second, and this is where his self-professed main interest lies, he explores how far up the performative effect reaches, to the point where the question arises whether the theory has created its own reality. That would present an important challenge to traditional ideas about the relation of theory and practice.

In the first chapter we are introduced to MacKenzie's methodological approach: an extensive number of interviews, narratives, and other existing literature. Chapter 1 also provides an introduction of the performativity concept, grounding it specifically in actor network theory (ANT) and the work of Callon, and Bruno Latour, and the Edinburgh 'strong programme' in the sociology of scientific knowledge (SSK). MacKenzie distinguishes various types of performativity going from weak to strong. Generic performativity implies that an aspect (model, theory, data) of economics is used in an economic process: theory is used as a tool or instrument. Effective performativity involves practical, difference-making use of such an aspect: theory acts as an engine of change. The strongest variety is Barnesian performativity: the practice is shaped along the lines of the theory/model. Connected to Barnesian performativity is the concept of counterperformativity. Here the practice develops contrary to what the theory or model posits. In the latter two, theory operates as a constitutive mechanism. Many have considered Barnesian performativity as a form of self-fulfilling prophecy. That is not the case. Unlike self-fulfilling prophecies which imply falsehood, performativity is a priori neutral with regard to truth attribution.

Chapters 2 to 5 set the table. MacKenzie presents a nice sketch of the rise to prominence of financial economics through the main theoretical breakthroughs of the 1950s and 1960s: the Modigliani-Miller propositions, Markowitz's portfolio theory, the capital asset pricing model of Sharpe and others, and the efficient market idea, chiefly inspired by Samuelson and Fama. The social setting is described: the initial controversy within the finance community and the resistance of practitioners to buy into the new theory. And later on, the questions about the empirical validity of the theories and the realisticness of the assumptions are dealt with. In particular, the treatment of assumptions regarding distributions of returns, in chapter 4, is compelling. MacKenzie's background in applied mathematics shows here. The development of option pricing theory, the crown jewel of finance, is described in chapter 5; again in a fine and eminently readable way.

However well these first five chapters are written, there is not really much news in them. That changes in chapter 6. The story turns to how in the early 1970s option pricing theory, in particular the Black-Scholes-Merton model, made its way into the practice of the derivatives markets and the subsequent thriving of those markets enabled by the model. But that journey from theory to practice has by no means been straightforward and impersonal. On a sociological note, MacKenzie emphasizes the substantial involvement of individuals ("bodies") in this process. For instance, many members of the University of Chicago faculty were involved in efforts to legitimize derivatives trading and the setting up of the Chicago derivatives exchanges in particular. And Fischer Black himself sold sheets with option prices to traders. This entangled process of theory and people led to a situation where the practice started resembling the theory more and more. Initially, the Black-Scholes prices did not match that well with the actual prices in the market, but the fit improved over time. Moreover, stringent assumptions in the model, such as the absence of transaction costs and the possibility of unlimited short selling, became less unrealistic as the derivatives markets flourished.

But then, enter chapter 7, something happened which would defy much of the established theory on financial markets: the 1987 crash. MacKenzie discusses the challenges that event posed to the paradigm of efficient markets, in particular with regard to rational expectations and information processing. But that debate has been conducted extensively by others. Where it gets interesting is with the observation that, after the crash, the empirical fit of derivatives prices with the original Black-Scholes-Merton model deteriorated. "While it could reasonably be said of this technosystem [i.e., derivatives markets before the 1987 crash] that it performed theory [...] what is now [after the crash] performed is no longer classic option pricing theory" (MacKenzie 2006, 201-202).

What happened was that empirical option prices started to exhibit a volatility skew. Contrary to the assumption made by Black, Scholes, and Merton, the expected standard deviation of asset returns implied by the option prices was no longer constant between various options on the same asset. MacKenzie traces the emergence of the volatility skew to an awareness of a form of systemic risk in financial markets. Contrary to the theoretical assumptions, there do exist various limits and constraints in the market, for instance imposed by regulation, or for risk management purposes, or caused by liquidity issues. Skew, or "smile", then is the mechanism built-in by the market to deal with the shortcomings of the theory employed. The emergence of skew constitutes a counterperformative move: practice develops in a direction diverging from the model that is supposed to describe that practice accurately.

In chapter 8, MacKenzie presents his take on the failure of long term capital management (LTCM) as an illustration of the points he made before. LTCM was a high profile and initially spectacularly successful money management operation, headed by some of the biggest names from Wall Street but also with the prominent involvement of eminent academics such as Nobel laureates Robert C. Merton and Myron Scholes. Its blow-out in 1998 presented a major shock to the financial system, prompting government action and a concerted cleanup effort. The use of theory in practice and the subsequent frictions between model and reality, the entanglement of theories and models with actual individuals, the importance of culture and social setting, they all come along. MacKenzie appears to have dual purposes here. On the one hand, he wants to show the relevance of social and sociological aspects, visible for instance in the run-on-the-bank situation that arose. On the other hand, he provides a striking example of a changing relation between model and reality, which fits nicely with the performativity concept.

MacKenzie also offers a contribution of his own with regard to the demise of LTCM. Besides greed, blind faith in models, overleverage, and the systemic chain reactions that unfolded, imitation played a role in his opinion. This entails that market players copy each other's strategies resulting in what MacKenzie labels a "superportfolio". That is not one of his most convincing suggestions, I think. It sounds a bit like saying that there are more buyers than sellers as an "explanation" for when the market goes up. Because, by definition, for every buyer there has to be a seller implying that there are never more buyers than sellers and vice versa. Likewise, for the positions taken by LTCM and its fellow funds there had to be counterparties.[1]

The true philosophical beef of this book resides in its last chapter: 'Models and markets'. Finance has never been a uniquely academic endeavour, but there is a tension between the theoretical and the practical. This ambivalence ultimately boils down to the particular goals of modelling: modelling to obtain tools to use in practice or modelling as an academic activity with the goal of improving knowledge. Or from a different point of view: plausible, analytically tractable models resulting in good abstractions versus realisticness. But the reality is that it is not either/or; rather "the boundary that separated academic financial economics and practical activity was very porous" (MacKenzie 2006, 249). That applies to data, concepts, tools, and people.

[1] This also becomes evident in the current debt crisis. In 2006 Goldman Sachs decided to take the opposite side of many of its competitors in the CDO market. That netted them some handsome profits but it did not isolate Goldman from the systemic fallout of the crisis.

That brings us back to the performativity issue. MacKenzie argues convincingly that finance theory has become incorporated in the infrastructure of the financial markets in three ways. First technically, as evidenced by the use of models in trading software not only in trading and investing but also in regulation, market organization, and risk management. Second linguistically, as can be seen in the use of originally theoretical terminology, such as “beta” and “volatility”, becoming standard jargon. Third legitimatizing: (financial) economists actively helped in the advent and development of certain markets, when options and securities trading was still very much seen as shady speculation and not unlike pure gambling. Thus, “finance theory’s incorporation into market infrastructure was consequential” (MacKenzie 2006, 252).

So far then the case has been made for generic and effective performativity. What about the Barnesian variety and counter-performativity? MacKenzie is cautioning that these two may be very hard to prove, but he does attempt to make the case with regard to the Black-Scholes-Merton model of option pricing. And, despite his caution, both notions do apply, he asserts. Originally, empirical option prices and market conditions started resembling those postulated by the model. That could be regarded as “simply a consequence of the discovery of the right way to price options” (MacKenzie 2006, 258). But there is a reason to doubt that statement and rather consider this as a case of Barnesian performativity. That reason is the appearance of skew after the 1987 crash leading to a counterperformative move. The emergence of skew tells us one of two things, according to MacKenzie: “if Black-Scholes is the “right” way to price options, then the market has been wrong since 1987; on the other hand, if a pronounced volatility skew in options is “correct”, then the market was wrong before 1987”. The latter is more plausible, as a case of rational learning, and that makes the Black-Scholes model not “a “true” discovery of what was already there” (MacKenzie 2006, 258-259).

This argument of a “false” or “inaccurate” model having such a profound effect on practice has been widely picked up by philosophers of economics and sociologists, in particular those active in the social studies of finance (SSF) program. But some caution is in order here. Is the model indeed false because the volatility pattern implied by empirical option prices differs from the assumption in the model? I don’t think so. Option pricing theory states that the price of a derivative depends on the variability, or volatility, of the underlying object. It is not a theory which claims to predict, explain, or understand that variability. Constant volatility is an assumption in the model; an unrealistic assumption that is. Black, Scholes, and Merton were well aware of this, even in their seminal papers: "the valuation formula assumes that the variance rate of the return on the optioned asset is constant. But the variance of return on an option is certainly not constant" (Black and Scholes 1973), and "the expected return is not directly observable" (Merton 1973), and therefore volatility is also not observable. So, rather than the model being false, it is one of the assumptions that fails.[2]

In connection to this, there also appears to be a misconception about the nature of the phenomenon of volatility. Being the standard deviation of returns over a certain period, the one true correct volatility number can only be determined ex post. That means that the unambiguously correct price of a derivative can only be ascertained after the contract has expired. What is entered when calculating the price of a current option is an estimate of volatility and what can be inferred from empirical option prices is the market consensus about those estimates. If one adheres to some form of randomness in the returns on assets, it should become evident how difficult making those estimates is. And there is no reason why the process of estimating cannot change, like it did after the 1987 crash. In fact, Fischer Black refused to postulate or accept any model of volatility because he considered this impossible (Mehrling 2005).

Financial economics is characterized by a positivist methodology; a rigorously quantitative and formalistic approach dominates. But it is only a positivist veil because the phenomena that it deals with are still very much social, contingent, and contextual. The relation between theory and practice is not like it is in physics or chemistry, even when the style of modelling may look alike. Many in economics talk about inexact laws, tendencies, stylized facts, good abstractions and the like. Labelling theories right and wrong then becomes a complicated matter.[3] Yet, it appears that MacKenzie's Barnesian and counterperformativity

[2] That of course brings this issue close to one of the most debated items in (philosophy of) economics: Milton Friedman's (1953) article: The methodology of positive economics.

[3] Emanuel Derman, co-author with Fischer Black of a number of papers and a former colleague of Black at Goldman Sachs, has made a comparable point with regard to option pricing theory.

hinges exactly on dishing out such tags. That is even more surprising given the supposed neutrality of the performativity thesis with regard to truth; a neutrality which appears to be one of the main attractions of the concept.

MacKenzie's stepping in the positivist trap that finance presents, highlights another important aspect of this book. On the one hand, it is a supreme effort to cross the boundaries of disciplines, in this case finance and sociology. But, on the other hand, it also shows how difficult, perhaps impossible, it is to get it exactly right. Be that as it may, MacKenzie has still been able to drive home the relevance of sociology to financial markets. And while he ultimately cannot deliver a knock-out punch with his case for Barnesian performativity—his self-professed main interest—he does deliver a persuasive, and to some extent novel, account of how knowledge travels from theory to practice (and back again) and the consequences thereof.

REFERENCES

Bernstein, Peter L. 1992. *Capital ideas: the improbable origins of modern Wall Street*. New York: John Wiley & Sons Inc.

Black, Fisher, and Myron Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy,* 81 (3): 637-654.

Friedman, Milton. 1953. The methodology of positive economics. In *Essays in positive economics*, Milton Friedman. Chicago: University of Chicago Press, 3-43.

MacKenzie, Donald. 2006. *An engine, not a camera: how financial models shape markets.* Cambridge (MA): MIT Press.

Mehrling, Perry. 2005. *Fisher Black and the revolutionary idea of finance*. New York: John Wiley & Sons Inc.

Merton, Robert C. 1973. Theory of rational option pricing. *Bell Journal of Economics and Management Science,* 4 (1): 141-183.

Job Daemen is a PhD candidate at the Erasmus Institute for Philosophy and Economics, Erasmus University Rotterdam. His research focuses on the philosophy and methodology of finance.
Contact e-mail: <job.daemen@planet.nl>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 154-164.
http://ejpe.org/pdf/1-1-br-2.pdf

Review of Stephen T. Ziliak and Deirdre N. McCloskey's *The cult of statistical significance: how the standard error costs us jobs, justice, and lives.* Ann Arbor (MI): The University of Michigan Press, 2008, xxiii+322 pp.

ARIS SPANOS*
Virginia Tech

The stated objective of this book is to bring out the widespread abuse of significance testing in economics with a view to motivate the proposed solution to the long-standing problem of *statistical vs. substantive significance* based on re-introducing 'costs and benefits' into statistical testing. The authors strongly recommend returning to the decision-theoretic approach to inference based on a 'loss function' with Bayesian underpinnings, intending to ascertain substantive significance in terms of "oomph, a measure of possible or expected loss or gain" (Ziliak and McCloskey 2008, 43).

The idea of a 'loss function' was introduced by Wald (1939), but rejected later by Fisher (1955) who argued that when one is interested in the truth/falsity of a scientific hypothesis, the cost of any actions associated with the inference is irrelevant; this does not deny that such costs might be relevant for other purposes, including establishing a range of substantive discrepancies of interest. This is still the prevailing view in frequentist statistics, which, to use one of the authors' examples (Ziliak and McCloskey 2008, 48), rejects the argument that to evaluate the substantive discrepancy from the Newtonian prediction concerning the deflection of light by the sun, one needs a loss function which reflects the relevant 'costs and benefits'.

How do the authors justify wedging the notion of a loss function back into econometrics? They interpret it in terms of 'economic cost' and trace the idea back to Gosset (1904); described pointedly as "a lifelong Bayesian" (pp. 152, 158, 300). How do they make their case? Curiously enough, not by demonstrating the effectiveness of their recommended procedure in addressing the statistical vs. substantive significance problem using particular examples where other 'solutions'

* AUTHOR'S NOTE: I am grateful to Kevin D. Hoover for many valuable comments and suggestions.

have failed. Indeed, in 320 pages of discussion, there is not a single credible illustration of how one can apply their proposed 'solution' to this problem. Instead, they attempt to make their case using a variety of well-known *rhetorical strategies and devices*, including *themes* like battles between good vs. evil, and conceit vs. humility, frequent *repetition* of words and phrases like 'oomph', 'testimation', 'sizeless stare' and 'size matters', picturesque language, metaphor and symbolism, flashback, allusion, parody, sarcasm, and irony. Their discourse in persuasion also includes some 'novel' devices like cannibalizing quotations by inserting their own 'explanatory' comments to accommodate their preferred interpretation, 'shaming' notable academics who 'should have known better', and recalling private conversations as well as public events where notable adversaries demonstrated the depth of their ignorance.

Their main plot revolves around a narrative with several ostensibly corroborating dimensions:

A. Evidence for the chronic abuse of statistical significance in economics.
B. Tracing the problem in statistics and the social sciences.
C. A 'selective' history of modern statistical thought as it pertains to the problem.
D. Discussion of various philosophical/methodological issues pertaining to the problem.
E. A 'what to do' list of recommendations to address the problem.

I will comment briefly on A-C and then focus my discussion on the last two dimensions.

A. The authors' accumulated evidence (chapters 6-7) for the widespread confusion between statistical and substantive significance in the abuse of significance testing takes the form of updating their 1996 scrutiny of applied papers published in the *American Economic Review* in the 1980s, which was based on grading these papers on 19 questions they devised for diagnosing the various facets of the problem. Although most of these questions are highly problematic in themselves, for the purposes of this review I will (reluctantly) take their evidence at face value and assume that most researchers sidestep the problem because they are unaware of a credible way to address it. Indeed, the researchers who scored very high on the M-Z scale only demonstrated *awareness* of the problem, but none of them, as far as I can see, had a credible procedure to ascertain the substantive significance warranted by the data in question.

B. The literature on the problem of statistical vs. substantive significance is almost as old as modern statistics itself, and the authors do make an effort to trace its history all the way back to Edgeworth (1885) by stretching the truth somewhat to fit their narrative (see Hoover and Siegler 2008). Since the dominating objective for the authors is persuasion, this historical retracing is spread into several chapters (4, 10, 11, and 12) for impact, and as a result, it becomes rather diffused and less informative. The gist of the discussion is that, despite its long history, this problem has been raised in economics rather belatedly, and the authors do deserve some of the credit for making an issue of it, even though their discussion obfuscates the issues involved.

C. The narrative concerning the historical development of modern frequentist statistics which 'accommodates' their preferred interpretation of the problem is summarized as follows:

> We want to persuade you of one claim: that William Sealy Gosset (1876–1937)—aka "Student" of Student's t-test—was right and that his difficult friend, Ronald A. Fisher was wrong. […] Gosset, we claim, was a great scientist. He took an economic approach to the logic of uncertainty. For over two decades he quietly tried to educate Fisher. But Fisher, our flawed villain, erased from Gosset's inventions the consciously economic element. We want to bring it back (Ziliak and McCloskey 2008, xv).

Throughout this book, Fisher is painted as the villain of the story and Gosset as the patron saint of modern statistics whose contributions have been overlooked as a result of concerted efforts by Fisher and his disciples. Gosset (an employee of the Guinness brewery) is presented as the source of numerous great ideas in statistics which Fisher (a famed professor) was systematically embezzling while peeling off their 'economic element' (Ziliak and McCloskey 2008, xv). One such idea, as their story goes, was the evaluation of inferences in terms of their 'economic costs', and not the relevant error probabilities as such. Unfortunately for science, Fisher's conception of statistics prevailed, and Gosset's vision was forgotten by both statisticians and economists. One of the book's main objectives is to redress that.

It does not take much effort to discredit their narrative concerning Fisher and his role in the development of modern statistics because its inaccuracies and distortions are legion. The narrative reads like a regurgitated but disconnected fable with Bayesian undertones; its heroes are primarily Bayesian 'at heart' and its villains are mainly Fisherian in perspective. However, even a glance through Savage (1976), one of the heroes, undermines the credibility of their narrative:

> *Just what did Fisher do in statistics?* It will be more economical to list the few statistical topics in which he displayed no interest than those in which he did. […] Fisher is the undisputed creator […] of the modern field that statisticians call the design of experiments, both in the broad sense of keeping statistical considerations in mind in planning of experiments and in the narrow sense of exploiting combinatorial patterns in the layout of experiments (Savage 1976, 449-450).

Acknowledging Fisher's epoch-making contributions to modern statistics does not, in any way, devalue Gosset's pioneering role in founding the frequentist approach in finite sampling theory, and influencing the work of both Fisher and Egon Pearson with insightful ideas and questions (see Plackett and Barnard 1990).

To illustrate the inaccuracy of the authors' narrative, let me simply oppugn one overhasty claim, that Arthur Bowley was a messianic disciple of Fisher who contributed significantly to spreading his statistical 'gospel' to economics (Ziliak and McCloskey 2008, 235, 293). Fisher revolutionized statistical thinking in the early 1920s while he was a non-academic statistician at Rothamsted Experimental Station; his first academic job, as professor of 'eugenics' at University College (London), was in 1933. Indeed, the academic establishment, led by Bowley (second only to Karl Pearson in academic status), fought with ferocity against Fisher's ideas, averted his appointment to several academic positions, and precluded him from most statistical forums, including the Royal Statistical Society (RSS). When this establishment could no longer ignore Fisher, Bowley and his cronies invited him to address the RSS for the first time in 1934, but their real intention was to expose him as a charlatan (see the discussion in Fisher 1935; and Box 1978).

D-E. The formal apparatus of the Fisher-Neyman-Pearson approach to frequentist inference was largely in place by the late 1930s, but its philosophical foundations left a lot to be desired. Several foundational problems, including: (a) the fallacies of acceptance and rejection, (b) the notion of statistical adequacy, (c) the role of substantive information in statistical modeling, and (d) the role of pre-data vs. post-data error probabilities (Hacking 1965), were left largely unanswered (Mayo 1996;

Spanos 1999). In particular, neither Fisher's p-value, nor Neyman-Pearson's 'accept/reject' rules, provided a satisfactory answer the basic question: 'When do data $\mathbf{x}_0$ provide evidence for or against a (substantive) hypothesis or claim?'

Indeed, both approaches are highly susceptible to:

(I). *the fallacy of acceptance*: (mis)-interpreting accept H_0 [*no* evidence *against* H_0] as evidence *for* H_0,

(II). *the fallacy of rejection*: (mis)-interpreting reject H_0 [evidence *against* H_0] as evidence *for* H_1; the best example of this is conflating statistical with substantive significance.

This created a lot of confusion in the minds of practitioners concerning the appropriate use and interpretation of frequentist methods. In the absence of any guidance from the statistics literature, practitioners in different applied fields invented their own favored ways to deal with these issues which often amounted to misusing and/or misinterpreting the original frequentist procedures (see Gigerenzer 2004). Such misuses/misinterpretations include, not only the well-known ones relating to the p-value, but also: (i) the observed confidence interval, (ii) the p-value curves, (iii) the effect sizes, (iv) the fallacy of the transposed conditional, (v) Rossi's real type I error, (vi) Zellner's random prior odds, and (vii) Leamer's extreme bounds analysis.

It can be argued that the authors' high-pitched recommendation of (i)-(vii), in their 'what to do' list to address the problem of statistical vs. substantive significance (Ziliak and McCloskey 2008, chapter 24), constitutes a perpetuation of the same foundational confusions, colored by the authors' Bayesian leanings, which have bedeviled frequentist inference since the 1950s. Space limitations prevent me from repudiating (i)-(vii) in any detail. Very briefly, the primary confusion underlying (i)-(ii) stems from the fact that, although observed confidence intervals do "draw attention to the magnitudes" (p. 73), they are no more informative on substantive significance than p-values; actually, there is a one-to-one mapping between the two, and they are equally vulnerable to the 'large *n* [sample size] problem'. Moreover, the relevant post-data error probabilities in estimation are either zero or one—the observed confidence interval either includes or excludes the true value of the unknown parameter θ—because the underlying reasoning is *factual* (under the true state of nature), as opposed to *hypothetical* (under different hypothetical scenarios) in testing.

The lack of proper post-data error probabilities in estimation explains why the different values of θ within an observed confidence interval are treated on a par, and the various 'effects sizes' proposed in the literature cannot possibly provide a reliable measure of substantive significance. Hence, the use of p-value curves to discriminate among the different values of θ within an observed confidence interval, giving the impression of attaching probabilities to these values (Ziliak and McCloskey 2008, 185), represents a mix-up of two different types of reasoning resulting in obfuscation (Spanos 2004). The charge that error probabilistic reasoning suffers from the fallacy of the transposed conditional stems from a false premise that error probabilities are conditional; there is *nothing* conditional about the evaluation of tail areas under different hypothetical scenarios, unless one conflates that with Bayesian reasoning which is conditional (Spanos 1999).

Among the various 'unsuccessful' attempts to address the problem of statistical vs. substantive significance that the authors dismiss, as yet another 'sizeless stare', is Mayo's (1996) post-data *severity evaluation* of the Neyman-Pearson 'accept/reject' decisions:

> If one returns to Mayo's discussion of what constitutes a "severe test" of an experiment, one finds only sizeless propositions, with loss or error expressed in no currency beyond a scale-free probability. […] A notion of a severe test without a notion of a loss function is a diversion from the main job of science, and the cause, we have shown, of error" (Ziliak and McCloskey 2008, 147).

It is clear from this quotation that the authors did not understand the use of this post-data evaluation in addressing the problem. *First*, contrary to their charge, there is no such thing as a 'severe test of an experiment', there are only severe tests of hypotheses or claims based on a particular test T_α and data $\mathbf{x}_0=(x_1,\ldots,x_n)$. *Second*, the severity evaluation, far from being another 'sizeless proposition', is actually framed in terms of a *discrepancy* parameter $\gamma \geq 0$ from the null, say: H_0: $\theta=\theta_0$ vs. H_1: $\theta > \theta_0$. The relevant post-data error probabilities—which remain firmly attached to the inference procedure itself and not to the hypotheses—evaluate the extent to which a substantive claim, such as $\theta \leq \theta_0 + \gamma$ or $\theta > \theta_0 + \gamma$ (associated with accept or reject), is warranted on the basis of a particular test T_α and data $\mathbf{x}_0$.

Depending on whether the Neyman-Pearson test has accepted (rejected) H_0, the severity evaluation is framed in terms of the smallest (largest) warranted discrepancy $\gamma \geq 0$, measured on the same scale as θ, with its magnitude easily assessable on *substantive* grounds. Hence, contrary to the authors' charge, the post-data severity evaluation of an accept/reject decision, gives rise to warranted discrepancies γ, which, in conjunction with substantive information, can help to address the fallacies of acceptance/rejection (see Mayo and Spanos 2006). Let me illustrate this.

Example 1. Consider the case where data $\mathbf{x}_0$ constitute a realization from the simple Normal model where $X_k \sim \text{NIID}(\mu,\sigma^2)$, k=1,2,…,n. The t-test based on $\tau(\mathbf{X}) = \sqrt{n}\left(\overline{X}_n - \mu_0\right)/s$ is a UMP test for the hypotheses: $H_0: \mu = \mu_0$ *vs.* $H_1: \mu > \mu_0$ (see Cox and Hinkley 1974).

Assuming that $\overline{x}_n = .02,\ n = 10000,\ s = 1.1$, yields $\tau(\mathbf{x}_0)$ =1.82, which leads to rejecting the null μ_0=0 at significance level α=.05, since c_α=1.645. Does this provide evidence for a substantive discrepancy from the null? The post-data evaluation of the relevant claim $\mu > \gamma$ for different discrepancies $\gamma \geq 0$, based on SEV($\tau(\mathbf{x}_0)$; $\mu > \gamma$) = P($\tau(\mathbf{X}) \leq \tau(\mathbf{x}_0)$; $\mu \leq \gamma$) [table 1], indicates that for a high enough severity threshold, say .9, the maximum warranted discrepancy is $\gamma < .006$.

Table 1: Severity evaluation of the claim: $\mu > \gamma$							
γ	.001	.005	.006	.01	.02	.05	.07
SEV($\tau(\mathbf{x}_0)$; $\mu > \gamma$)	.958	.914	.898	.818	.500	.003	.000
POW($\tau(\mathbf{X})$; c_α; $\mu = \gamma$)	.060	.117	.136	.231	.569	.998	1.00

One then needs to consider this in light of substantive information to assess whether the warranted discrepancy $\gamma < .006$ is substantively significant or not. In addition, the severity reasoning can be used to elucidate certain *fallacious claims* repeated by the authors throughout this book, pertaining to the very problem that occupies center stage: "A good and sensible rejection of the null is, among other things, a rejection *with high power*" (Ziliak and McCloskey 2008, 133). And "refutations of the null are easy to achieve if power is low or the sample is large enough" (p. 152).

No! No! You have it backwards. Rejection with high power is actually the main source of the problem of statistical vs. substantive significance, and 'large enough sample sizes' n go hand in hand with high power, not low. For instance, the power of the above t-test increases with the non-centrality parameter $\delta = \sqrt{n}(\gamma)/\sigma$, which is a monotonically increasing function of n. When a test has very high power for tiny discrepancies from the null, as in the large n case, rejection of the null provides less (not more) evidence for the presence of a substantive discrepancy. This is illustrated in table 1, where the power of the test, based on POW($\tau(\mathbf{X})$; c_α; $\mu = \gamma$) = P($\tau(\mathbf{X}) > c_\alpha$; $\mu = \gamma$), is very high for small discrepancies from the null; it is almost 1 at γ =.05. What is even more misleading is that the power increases with the discrepancy $\gamma \geq 0$, in contrast to the severity evaluation.

Analogously, when a test with very low power for sizeable discrepancies of interest rejects the null, it provides more (not less) evidence for the presence of a substantive discrepancy.

Example 2. Let us consider the case where data $\mathbf{x}_0$ in example 1 yielded instead $\bar{x}_n = .633$, s=1.1, for n=10; small sample case. In this case $\tau(\mathbf{x}_0)$= 1.82 leads to *accepting* the null μ_0=0 at α=.05 since the critical value now is c_α=1.833. Does this provide evidence for *no* substantive discrepancy from the null? The post-data evaluation of the relevant claim $\mu \leq \gamma$ for different discrepancies γ, based on SEV($\tau(\mathbf{x}_0)$; $\mu \leq \gamma$) = $\mathbb{P}$($\tau(\mathbf{X}) > \tau(\mathbf{x}_0)$; $\mu > \gamma$), indicates that for a high enough threshold, say .9, the minimum discrepancy warranted by data $\mathbf{x}_0$ is γ >1.1.

Table 2: Severity evaluation of the claim: $\mu \leq \gamma$

γ	.1	.25	.5	1.0	1.1	1.2	1.5
SEV($\tau(\mathbf{x}_0)$; $\mu \leq \gamma$)	.080	.150	.356	.841	.894	.931	.983
POW($\tau(\mathbf{X})$; c_α; $\mu = \gamma$)	.078	.147	.351	.838	.892	.930	.982

Again, substantive information should be used to assess if such a discrepancy is substantively significant or not. These two examples demonstrate how the same test result $\tau(\mathbf{x}_0)$=1.82, arising from two different sample sizes, n=10000 and n=10, can give rise to widely different 'severely passed' claims concerning the warranted substantive discrepancy: γ < .006 and γ > 1.1, respectively. Note that in the case of 'accept H_0' shown in table 2, the power moves in the same direction as severity and the two are close because $\tau(\mathbf{x}_0)$=1.82 is very near the critical value c_α=1.833.

Statistical adequacy. Another inveterate foundational problem associated with the Fisher-Neyman-Pearson frequentist approach has to do with the absence of a reasoned criterion for deciding when an estimated model is adequate on statistical grounds. Goodness-of-fit criteria have been discredited because of their vulnerability to spurious inference results. Gosset, as the authors rightly observe (Ziliak and McCloskey 2008, 59-60), is credited with raising the issue of invalid probabilistic assumptions, such as 'normality', giving rise to spurious results as early as 1923 (see Lehmann 1999). His questions were explored by Egon Pearson in the early 1930s, but largely ignored by Fisher and the subsequent statistics literature for a variety of reasons beyond the scope of this review.

As argued in Spanos (1986), addressing the problem of *statistical adequacy* (the validation of the model assumptions vis-à-vis data $\mathbf{x}_0$) requires, *ab initio*, a purely probabilistic construal of a statistical model, specified in terms of a complete list of (internally consistent) probabilistic assumptions, in a form that is testable with data $\mathbf{x}_0$. That often requires unveiling implicit assumptions as well as recasting assumptions about unobservable errors terms. It also requires distinguishing between *statistical* and *substantive adequacy*, contrary to the current conventional wisdom in economics which conflates the two under the banner of 'specification error'. This is because securing the former is a necessary condition for assessing the latter (Spanos 2006b). Statistical adequacy renders the relevant error probabilities ascertainable by ensuring that the *nominal* error probabilities for assessing substantive claims are very close to the *actual* ones. The surest way to draw invalid inferences is to apply a 5% significance level test when its actual type I error probability is close to 100% due to misspecification (Spanos and McGuirk 2001).

Using statistical adequacy—not 'oomph' (Ziliak and McCloskey 2008, 48)—to select the best model in the sense that it 'accounts for the regularities in the data', can explain why the t-test, the R^2 and other statistics vilified by the authors, are often statistically vacuous when any of the probabilistic assumptions constituting the statistical model in question are invalid for data $\mathbf{x}_0$. Indeed, statistical adequacy helps to place the problem of statistical vs. substantive significance in a proper perspective. Despite the importance of the latter problem, any attempt to address it becomes hopeless unless one deals with the *statistical misspecification* issue first. The very notion of statistical significance becomes ambiguous without statistical adequacy since it is unknown whether the apparent significance is genuine or simply an artifact, i.e., the result of a sizeable discrepancy between the relevant nominal and actual error probabilities; talk about 'baseless size'!

In light of this dubiousness, the researchers accused of 'sizeless stare' and outright ignorance are guilty only of sidestepping a problem which nobody knows how to address adequately, least of all the two authors; paying lip service is far from dealing with it. Continuing this line of reasoning, do the authors expect credit for mentioning a blurred form of the 'specification problem' (Ziliak and McCloskey 2008, xvii) and some vague references to 'other errors', even though they have done nothing about them in their published work? Or do the authors point a finger at the failings of others to distract from the more serious problems that they themselves ignore in their published work?

The problem of statistical misspecification is not only more fundamental, but researchers have known, for some time now, how to handle it using thorough *misspecification testing* and *respecification*. Moreover, Fisher-type significance testing plays a crucial role in model validation (see Spanos 1986, 1999; Mayo and Spanos 2004). Indeed, one wonders how many applied papers published in the *American Economic Review* over the last 30 years are likely to pass the statistical adequacy test; I hazard a guess of less than 1% for the reasons I discuss in Spanos (2006a).

Where does this leave the authors' concern with the problem of statistical vs. substantive significance? Shouldn't they have known that, even if one had a credible procedure to address the problem, one couldn't make any progress on the basis of statistically misspecified models?

In conclusion, do the authors genuinely believe that their 'what to do list', based primarily on (i)-(viii), and some wispy references to "Jeffrey's d, Wald's 'loss function', Savage's 'admissibility' […] and above all Gosset's 'net pecuniary advantage'" (Ziliak and McCloskey 2008, 250), constitute a credible solution to this important problem? If so, they delude themselves far more than those economists at whom they wag their fingers throughout this book.

REFERENCES

Box, Joan. F. 1978. *R. A. Fisher: the life of a scientist*. New York: Wiley & Sons.

Cox, D. R., and D. V. Hinkley. 1974. *Theoretical statistics*. London: Chapman and Hall.

Fisher, Ronald A. 1935. The logic of inductive inference. *Journal of the Royal Statistical Society*, 98 (1): 39-54 [with discussion: 55-82].

Fisher, Ronald A. 1955. Statistical methods and scientific induction. *Journal of the Royal Statistical Society*, B, 17 (1): 69-78.

Hacking, Ian. 1965. *Logic of statistical inference*. Cambridge: Cambridge University Press.

Hoover, Kevin D., and Mark V. Siegler. 2008. Sound and fury: McCloskey and significance testing in economics. *Journal of Economic Methodology*, 15 (1): 1–37.

Lehmann, E. L. 1999. Student and small-sample theory. *Statistical Science*, 14 (4): 418-426.

Mayo, Deborah G. 1996. *Error and the growth of experimental knowledge*. Chicago: The University of Chicago Press.

Mayo, Deborah G., and Aris Spanos. 2004. Methodology in practice: statistical misspecification testing. *Philosophy of Science*, 71 (5): 1007-1025.

Mayo, Deborah G., and Aris Spanos. 2006. Severe testing as a basic concept in a Neyman-Pearson philosophy of induction. *British Journal for the Philosophy of Science*, 57 (2): 323-357.

Plackett, R. L., and G. A. Barnard (eds.). 1990. *Student: a statistical biography of William Sealy Gosset, based on writings by E. S. Pearson*. Oxford: Clarendon Press.

Savage, Leonard J. 1976. On re-reading R. A. Fisher. *Annals of Statistics*, 4 (3): 441-500.

Spanos, Aris. 1986. *Statistical foundations of econometric modelling*. Cambridge: Cambridge University Press.

Spanos, Aris. 1999. *Probability theory and statistical inference: econometric modeling with observational data*. Cambridge: Cambridge University Press.

Spanos, Aris. 2004. Confidence intervals, consonance intervals, p-value functions and severity evaluations. *Virginia Tech Working paper*, Blacksburg.

Spanos, Aris. 2006a. Econometrics in retrospect and prospect. In *New Palgrave handbook of econometrics, vol. 1*, eds. T. C. Mills, and K. Patterson. London: MacMillan, 3-58.

Spanos, Aris. 2006b. Revisiting the omitted variables argument: substantive vs. statistical adequacy. *Journal of Economic Methodology*, 13 (2): 179-218.

Spanos, Aris, and Anya McGuirk. 2001. The model specification problem from a probabilistic reduction perspective. *Journal of the American Agricultural Association*, 83 (5): 1168-1176.

Wald, Abraham. 1939. Contributions to the theory of statistical estimation and testing hypotheses. *Annals of Mathematical Statistics*, 10 (4): 299-326.

Ziliak, Stephen T., and Deirdre N. McCloskey. 2008. *The cult of statistical significance: how the standard error costs us jobs, justice, and lives*. Ann Arbor (MI): The University of Michigan Press.

Aris Spanos is Wilson Schmidt professor at the Department of economics, Virginia Tech. His research focuses on econometrics, modelling speculative prices, and the philosophy and methodology of empirical modelling. Contact e-mail: <aris@vt.edu>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 165-170.
http://ejpe.org/pdf/1-1-br-2r.pdf

Science is judgment, not only calculation: a reply to Aris Spanos's review of *The cult of statistical significance.*

STEPHEN T. ZILIAK
Roosevelt University

DEIRDRE N. MCCLOSKEY
University of Illinois at Chicago

Over the past century the usual (and the conveniently mechanical) procedure devised by the great statistician, geneticist, and racial eugenicist R. A. Fisher has been shown to be scientifically silly again and again and again. Rarely has anyone actually defended NHST (null hypothesis significance testing). That is because it is logically indefensible. Statistical significance is neither necessary nor sufficient for substantive scientific significance. Everyone knows this, once they stop regressing for a minute and actually think.

We have noticed two peculiar features of the rare defenses, exhibited in Aris Spanos's (2008) review. For one thing, when mounted by people sophisticated in statistics, such as Spanos, or his allies Kevin Hoover and Mark Siegler (2008), the defenses are never defenses. They begin on the first page by admitting that NHST does *not* give mechanical assurances that its alleged findings are scientifically important. Spanos acknowledges the salience of this "long-standing problem of statistical vs. substantive significance". It is certainly "long-standing"—the error of mixing one with the other, as we show, dates to the foundation of the journal *Biometrika*, in 1901. Unhappily, though, and every time, the defenders promptly lose sight of their concession. On the second page they re-assert, as for example both Spanos and Hoover/Siegler do, that NHST offers the scientist a way of making a scientific judgment without regard to what is persuasive to other scientists.

For another thing, the defenders are always angry. Ignorant sneering, personal insult, and irrelevant indignation are judged acceptable when defending NHST. We think the anger comes from a psychological tension. The defenders realize uneasily that it is strange to depend for scientific judgment on a sampling statistic without a persuasive context—failing to ask how big is big, which is the only scientific context relevant to a real scientific test. But they have been thoroughly indoctrinated in NHST, and belong to a professional club in which $t > 2.0$ or $p < .05$ or whatever is substituted for scientific judgment. The mechanical procedure of their profession is under attack. So they get angry. They have no reply. So they shout and bluster.

Spanos throws up a lot of technical smoke that has the effect of obscuring the plain fact that he agrees with us. (The mathematics in his piece is irrelevant to anything of importance. The reader may omit it.) His technical smoke billows. For example, he calls NHST "the Fisher-Neyman-Pearson approach". The terminology is conventional, but expresses a revealing historical error. Jerzy Neyman and Egon S. Pearson were in fact enemies of Fisher (true, *anyone* who disagreed with Fisher became instantly his enemy for life, especially if he or she was not academically powerful). The young men, Neyman and Pearson, with the encouragement of William Gosset (aka "Student"), were to be precise *criticizing* Fisher's one-criterion test of significance, from 1928 on. Although they did not then introduce the loss functions that later became routine in statistical and econometric theorizing (despite Fisher's fierce and irrational opposition), they did for example in 1933 emphasize that "how the balance should be struck" between Type I and Type II errors (false positive and false negative errors) "must be left to the investigator" (Neyman and Pearson 1933, 296).

That is a big improvement over elevating Type I error to the only criterion, $t > 2.0$, and pretending that judgment and persuasion therefore do not need to be the crucial last step in any scientific test. Statistical significance according-to-Fisher translated every quantitative question into a probability about the data assuming the truth of the singular hypothesis. It collapsed the scientific world into a Borel space, p (0, 1.0)—a procedure, by the way, that the mathematical statistician Émile Borel himself emphatically rejected. Borel (1871–1956), though a master of abstract imagination, was deeply interested in the substantive side of testing, and in Paris in the 1920s helped convert a young Jerzy Neyman to a life of substantive significance (Reid 1982, 68-70).

But of course that is the *sole* problem we are concerned with in *The cult*, the Fisherian mistake of supposing that *statistical* significance is just the same thing as *substantive*, *scientific*, *economic* significance. Spanos ends by claiming that we have ignored specification errors (which is false: we speak of them, and of twenty-something other errors of statistical and scientific experiments. But in the book we did not want to be distracted from observing the main and elementary problem of lack of scientific substance). That specification errors, and sample-selection bias, and biases of the auspices, and the rest, are *also* problems with the usual mechanism of NHST does not (of course) somehow repair the simpler problem that we and hundreds of other critics since the 1920s have drawn attention to.

The problem is always ignored in econometrics. Arthur Goldberger gives the topic of "statistical vs. economic significance" one page of his *A course in econometrics* (1991), quoting a little article by McCloskey in 1985. Goldberger's lone page was flagged as unusual by someone in a position to know. Clive Granger reviewed four econometrics books in the March 1994 issue of the *Journal of Economic Literature* and wrote: "when the link is made [in Goldberger between economic science and the technical statistics] some important insights arise, as for example the section [well… the page] discussing 'statistical and economic significance', *a topic not mentioned in the other books*" [by R. Davidson and J. G. MacKinnon, W. H. Greene, and W. E. Griffiths, R. C. Hill, and G. G. Judge] (Granger 1994, 118, italics supplied).

Not mentioned in the other books. *That* is the standard for educating young people on the statistical/substantive distinction in econometrics and statistics at the advanced level. We wonder if Professor Spanos does better for his own students. The three stout volumes of the *Handbook of econometrics* contain a lone mention of the point, unsurprisingly by Edward Leamer (Griliches and Intriligator 1983, I, 325). In the 732 pages of the *Handbook of statistics* (Maddala, Rao, and Vinod 1993) there is one sentence (by Florens and Mouchart on p. 321). In his own impressive *Probability theory and statistical inference* (1999) Spanos himself tried to crack the Fisherian monopoly on advanced econometrics. But even Spanos looks at the world with a sizeless stare (Spanos 1999, 681-728).

The main point of Spanos's piece is that Ziliak and McCloskey do not offer guidance on how to address substantive scientific significance. Yet even if we had not, it would not be a fault. NHST is intellectually bankrupt, as Spanos agrees it is, and it should be abandoned. If you earn your living robbing banks, you should stop, right now, at once. You should not complain, "But how am I now to earn my living?" Go get honest work. And the honest work in the present case is the exercise of scientific judgment, quantified by relevant magnitudes that the best scientists find persuasive. It is quite false that Ziliak and McCloskey offer no such guidance. On the contrary, in scores of places in the book, especially on the economic matters, we offer ideas about what constitutes an oomph-ful, scientifically relevant judgment, on, say, an experiment in paying companies to hire the unemployed. Of course, we have more intelligent suggestions about economics than about psychology or medicine. We are economists, after all. But that is the main point. *There is no discipline-independent criterion for importance, calculable from the numbers alone*. Read that again. *There is no discipline-independent criterion for importance, calculable from the numbers alone*. Scientific judgment is scientific judgment, a human matter of the community of scientists. As vital as the statistical calculations are as an input into the judgment, the judgment cannot be made entirely by the statistical machinery.

That is really what Spanos craves: a machine for making scientific judgments. He is scornful of Bayesians (on the usual illogical and Fisherian grounds that judgment cannot be exercised in scientific decisions, or on the anti-economic and Fisherian grounds that cost and benefit in persuasion are irrelevant). We are rather fond of Bayesians. If Thomas Kuhn and his numerous children and grandchildren in the history, sociology, and philosophy of science have taught us anything it is that science is a community of mutual—preferably honest and logical—persuasion. That is what Bayesians say, and it seems a sensible reminder that science must always entail judgment, not merely calculation.

In the end we are reminded of what the American philosopher William James said about the three stages of a theory's reception: "First, you know, a new theory is attacked as absurd; then it is admitted as true, but obvious and insignificant; finally it is seen to be so important that its adversaries claim that they themselves discovered it" (James 1907, 198). Spanos has examined no archives on the history of statistics, but claims (stage 1) that our theory of how NHST arose from Fisher's disputes is absurd, and that we are silly to reject NHST for model validation in econometrics. Anyway (stage 2), everyone knows that "significance" is not the same thing as scientific importance. The point, he says, is obvious and insignificant: misspecification is what matters. Yet, by-passing our large-scale empirical work on the *American Economic Review*, Spanos offers his own claim to have discovered what we discovered (stage 3): "One wonders how many applied papers published in the *American Economic Review* over the past thirty years are likely to pass the statistical adequacy test; I hazard a guess of less than 1%" (Spanos 2008, 163).

Here is our challenge. If you think, like Spanos, that you have a valid defense of NHST, offer it. Spanos, like Hoover/Siegler, and Anthony O'Brien (2004), have tried. They have failed. But at least they are serious about their intellectual commitments, and *believe* (given their Bayesian priors) that NHST is defensible. It is not.

REFERENCES

Florens, Jean-Pierre, and Michel Mouchart. 1993. Bayesian testing and testing Bayesians. In *Handbook of statistics, Vol. 11*, eds. G. S. Maddala, et al. Amsterdam: North Holland, 303-391.

Goldberger, Arthur. 1991. *A course in econometrics*. Cambridge: Harvard University Press.

Granger, Clive W. J. 1994. A review of some recent textbooks of econometrics. *Journal of Economic Literature*, 32 (1): 115-122.

Griliches, Zvi, and Michael D. Intriligator (eds.) 1983, 1984, 1986. *Handbook of econometrics.* Vols. I, II, and III. Amsterdam: North-Holland.

Hoover, Kevin D., and Mark Siegler. 2008. Sound and fury: McCloskey and significance testing in economics. *Journal of Economic Methodology*, 15 (1): 1-37.

James, William. 1907. *Pragmatism: a new name for some old ways of thinking.* New York: Longmans, Green.

McCloskey, Deirdre N., and Stephen T. Ziliak. 1996. The standard error of regressions. *Journal of Economic Literature*, 34 (1): 97-114.

McCloskey, Deirdre N., and Stephen T. Ziliak. 2008. Signifying nothing: a reply to Hoover and Siegler. *Journal of Economic Methodology*, 15 (1): 57-68.

McCloskey, Deirdre N. 1985. The loss function has been mislaid: the rhetoric of significance tests. *American Economic Review*, Supplement 75 (2): 201-205.

Neyman, Jerzy, and E. S. Pearson. 1928. On the use and interpretation of certain test criteria for purposes of statistical inference, Part I and Part II. *Biometrika*, 20A (1-2): 175-240, 263-294.

Neyman, Jerzy, and E. S. Pearson. 1933. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London, A*, 231: 289-337.

O'Brien, Anthony P. 2004. Why is the standard error of regression so low using historical data? *Journal of Socio-Economics*, 35 (5): 565-570.

Spanos, Aris. 1999. *Probability theory and statistical inference*. Cambridge: Cambridge University Press.

Spanos, Aris. 2008. Review of S. T. Ziliak and D. N. McCloskey's *The Cult of Statistical Significance. Erasmus Journal for Philosophy and Economics*, 1 (1): 154-164. http://ejpe.org/pdf/1-1-br-2.pdf

Ziliak, Stephen T., and Deirdre N. McCloskey. 2004a. Size matters: the standard error of regressions in the *American Economic Review*. *Journal of Socio-Economics*, 33 (5): 527-46.

Ziliak, Stephen T., and Deirdre N. McCloskey. 2008. *The cult of statistical significance: how the standard error costs us jobs, justice, and lives.* Ann Arbor (MI): The University of Michigan Press.

Stephen T. Ziliak is professor of economics at Roosevelt University and his research fields are: welfare and poverty; economic history, rhetoric, and philosophy; and history and philosophy of science and statistics. Contact e-mail: <sziliak@roosevelt.edu>

Deirdre N. McCloskey is Distinguished professor of economics, history, English, and communication, at the University of Illinois at Chicago. Contact e-mail: <deirdre2@uic.edu>
Website: <www.deirdremccloskey.com/>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 171-174.
http://ejpe.org/pdf/1-1-br-3.pdf

Review of David C. Colander's *The making of an economist, redux*. Princeton: Princeton University Press, 2007, 268 pp.

RENÉ L. P. MAHIEU
EIPE, Erasmus University Rotterdam

With *The making of an economist, redux* (henceforth: *Redux*), David Colander is building on his successful earlier work with Arjo Klamer: *The making of an economist* (1990). *Redux* has basically the exact same format as the earlier work. It is meant to be interesting for a very broad public: economists, (prospective) students of economics, and lay persons (Colander 2007, vii). In this review, I will give a short overview of the content of the book. Moreover, I will question how much extra value this new edition has over the old one, especially with an eye to people interested in the methodology and philosophy of economics.

The main body is the presentation of a survey done by Colander. Students of economics at top American graduate schools were asked to fill in questionnaires and some of them were interviewed so as to get a view of what graduate training is like from the perspective of the students. The main motivation behind this effort is that looking at graduate economics education would give insight into the profession as it stands, and into the changes that might be forthcoming. The content of the book is divided into three parts.

In the first part, the results from the survey are presented along with an example of the questionnaire that was used. Colander makes an effort to give us a profile of the students: what they are interested in, what their political orientation is, and what they think about their training. In all cases an effort is made to find interesting differences between the different schools. First, the quantitative data are analysed. This includes data like the age of the students and the number of foreign students, but also how students assess the relevance of neo-classical economics and the importance of a broad knowledge of the economics literature. To compensate for the limiting nature of questions needed to produce quantitative data, room for written comments on answers and open questions were also part of the questionnaire; these are discussed in the following chapter. Questions that were presented include "Do you see economics as relevant?" with answers like "Yes, economists are the only careful, structured, empirical thinkers on most economic, political and social topics" or "Uncertain. While not true to everyone, many economists undervalue the contributions of other disciplines and don't effectively engage in the policy-making process" or "No, normal people solve crosswords; economists write papers (of which 80 percent are never read)". And "What makes a successful economist?" with answers like "Well-published, often talked about, cited, invited to conferences; thinking of new things, happy" or "Curious, rigorous analytically, resourceful, relevant, creative" or "One who gets tenure (I do not interpret 'successful' necessarily to mean 'good')". Finally the results of a study of some of the respondents from the earlier survey are presented.

Colander's new general conclusion is that, contrary to his conclusion in the previous book, there is not much wrong with the education system; students are quite happy with the education that they receive. Overall this first part of *Redux* is a fairly interesting read, since it lives up to the goal of giving the reader some insight into the making of economists. But, it also suffers from problems inherent to the method employed. Colander cannot escape from ending up with many sentences in the following style: "Chicago had the greatest interest in… (x percent); MIT had the least (y percent). MIT had the most interest in… (x percent); Yale had the least (y percent), etcetera, etcetera". As a scientific endeavour, the use of questionnaires might also raise some eyebrows, especially among economists. This is acknowledged by Colander but he makes a strong case for granting the results some value nonetheless.

In fact, in the second part a transcription is given from the conversations Colander held with students, partly as a remedy against the problems inherent in the method. These conversations are presented without any further explicit comment. Although there are many interesting and entertaining things to be found in these dialogues, it does get quite repetitive.

In the third part reflections are given on the overall results, not only by David Colander, but also by Arjo Klamer, Colander's co-author in the previous 1990 book, and Robert M. Solow. Colander's main conclusion is that current graduate training has its problems, but overall is doing quite well. According to him, it has improved since the 80s by becoming more focused on empirical research, and although there is still a strong focus on mathematics, it has become much more focused on application and less on purely mathematical exercises. The main remaining problem is that although economics has supposedly become a more diverse field, the core of graduate programmes has not changed accordingly.

Solow's reflection is mainly a defence of macroeconomics in reaction to a proposal, that Colander makes in part one, that the theoretic macroeconomics course should be taken out of the core of graduate programmes.

Klamer's reflection entitled 'Does this have to be our future?' is the most critical. For him the results show nothing remarkable; indeed a bit more attention is given to empirical work, but this has not solved the main problems within economics education. If anything, things have gotten worse. Differences between schools have gotten smaller, heterodox economics seems to have completely disappeared from the curriculum and so has the history of economics. This, together with an ever decreasing focus on the classical economics literature (Smith, Hayek, Keynes, Marx, and so forth) and on philosophy of science makes economics a less intellectual field than it could and should be.

It is in this final part that things should have been getting really interesting. But, sadly enough the book never achieves the apotheosis I was hoping for. Colander's own reflections are of some interest but they do not convey any sense of need or urgency, and more problematically they do not seem to be built on a clear (methodological) perspective of what graduate training should look like. This could be surprising for those readers who are familiar with Klamer and Colander's 1990 book in which Colander argued for what he called a "sociological approach to methodology" (Klamer and Colander 1990, 191). From that perspective graduate training was found to be defective in several ways, and these deficiencies ultimately came from problems within the science of economics in general. The main problem identified then was that the mainstream positivist methodology was focused on the formal empirical testing of hypotheses, while in economics at that time no generally accepted way of formal testing or other non-formal process for the selection of hypotheses could be found (Klamer and Colander 1990, 189-190). Therefore—Colander argued in 1990—new non-formalist methodological conventions should be established to guide economists and their students in their scientific enquiry.

Starting from this point of view, one would expect that, since Colander finds—in *Redux*—that existing problems have largely been solved, these new methodological conventions have indeed been found. Education must have become more focused on different formal and non-formal ways of discovering truth. However, in *Redux*, Colander is not at all concerned with making such a point. On the contrary, teaching the judgment and wisdom that should guide empirical testing are still not basic elements of graduate training. Nevertheless, Colander argues in *Redux* that new ways of doing formal empirical testing, that are appropriate in all or most cases, have been found. Furthermore Colander suggests that there has been a move to "a more inductive approach in which empirical evidence rather than theory guides research" (Colander 2007, 244). While something of course can be said for the positive developments of empirical methods in economics over the last twenty or so years, no constructive elaboration is made in that direction in *Redux*. Given the proliferation of books and journals such as EJPE, it is clear that questions concerning the methodology of economics have not been solved over the past 20 years. Especially from someone who in the past has shown interest in the more delicate questions of philosophy and methodology of economics, I was expecting a more thorough argumentation for a claim announcing that (almost) all is now well in economic science.

Taking these criticisms into account, this book is still a must-read especially for two groups of people. First, those who are contemplating going into graduate training in economics would be helped to get a clearer picture of what such graduate training will actually be like. This is all the more important since in the book it is argued convincingly that there is a big gap between what students expect to learn and what they actually get to learn in graduate school. Second, the book is valuable for all those who have influence on the content and design of graduate programmes in economics. For all others this can still be a book for the coffee-table, since my own experience has shown that it does function quite well as a conversation starter.

REFERENCES

Colander, David C. 2007. *The making of an economist, redux*. Princeton and Oxford: Princeton University Press.

Klamer, Arjo, and David C. Colander. 1990. *The making of an economist*. Boulder and London: Westview Press.

René L. P. Mahieu is currently a research master student at the Erasmus Institute for Philosophy and Economics, Erasmus University Rotterdam. Contact e-mail: <renemahieu@gmail.com>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 175-180.
http://ejpe.org/pdf/1-1-br-4.pdf

Review of Arjo Klamer's *Speaking of economics: how to get into the conversation.* Abingdon: Routledge, 2007, 199 pp.

ERWIN DEKKER
Erasmus University Rotterdam

Arjo Klamer wants to change the way we think about economics. He argues that economics is not a body of accumulated knowledge, a mirror of the economic world out there, or rhetoric (the art of persuasion), but rather a bunch of conversations. In his recent book, *Speaking of economics* (2007), he introduces the term conversation in order to show that this perspective helps us understand the practice of economics better.

Klamer has previously done a lot of work in the rhetorical approach to economics together with Deirdre McCloskey, for example in their joint book *The consequences of rhetoric* (1988), and their article The rhetoric of disagreement (1989). Klamer's metaphor of the conversation, although it includes rhetoric, is more encompassing and places emphasis on the social and cultural as well as the rhetorical aspects of the practice of economics.

According to Klamer what economists do can be best compared to being in a conversation: it is all about the company they keep and thus economists are those who are in the economic conversation. This conversation has its own social structure, culture, and way of evaluating arguments. To explain the way in which arguments are evaluated, Klamer relies heavily on his earlier work on rhetoric, the social structure and cultural aspects are what is new here.

For Klamer the most important aspect of the social structure of the economic conversation is attention. Every economist is looking for attention and reciprocates that attention to other economists whom he believes are interesting. Attention is not distributed evenly between economists: most of the attention is directed to a few people in the field, the superstars. Viewed as such, the people in the conversation compete for attention; however, Klamer is reluctant to stress the metaphor of competition too much where attention is concerned. Economists are not just in academia for the attention; they often have an intrinsic motivation to be part of the conversation and a personal passion for their subject (Klamer 2007, 60). But more importantly the product of scientific activity is a joint product. The product is the conversation itself with its theories, models, and stories (Klamer 2007, 61). It is not the ideas or best ideas that matter the most, but the amount of attention that these ideas receive within the conversation.

The other important insight that the conversation metaphor provides is the idea that economics has its own culture, which is however embedded within the general culture. A conversation has its own history and language; these can be so specific that economists from one field cannot talk to other economists with a different specialisation. These specialists have different ways (cultures) of approaching the economic world out there. Such differences also exist historically. Take for example John Hicks's IS-LM model and Keynes's own verbal description. Klamer argues that a model such as the one Hicks formulated was more in tune with the then rising cultural values of modernism outside economics, which particularly after World War II (WWII) were also very influential within economics. Therefore Hicks's formulation proved to be much more influential than Keynes's own original verbal description (Klamer 2007, 139-142). We could easily extend the analysis to say that cultural values in a conversation influence which ideas are successful, and which model is elegant and precise. In many ways Keynes's description was more precise than Hicks's model, but Keynes's description was only so in the verbal descriptive culture of the nineteenth century, not in the mathematical abstract culture of the second half of the twentieth century. Klamer's analysis of scientific culture and economic modernism is definitely the most original and valuable part of his book.

However the most important weak point of the book is an unresolved tension between the style of Klamer's argument and the content of his argument. The style of the argument is that of an accessible conversation in which everyone is invited to join in; the content of the argument however is largely negative and depicts a conversation that is almost impossible to join. Klamer compares the conversation of economists with that of a group of Italian men arguing vehemently in small closed clusters on a square:

> I wanted to join in, argue politics, offer my opinion on the Bologna soccer team. But, even apart from my bad Italian, I knew I couldn't. Each group had a history I was not privy to, referenced past conversations, called upon anecdotes that would have been lost on

> me. Even if I had managed to worm my way into one of the groups, I would have been immediately found out. I can't talk with my hands […] No matter what, I was not part of any conversations taking place in the square. I had a similar feeling when I went to my first economics conference (Klamer 2007, 16).

The tension is evident here between the pessimistic substance of the argument, the difficulty of joining the economic conversation, and the open and personal tone, or to put it in Klamer's words, the conversational tone, with which it is expressed. Klamer's style suggests that the conversation of economists does not have to be closed off and highly abstract, but that this results from conscious choices made by the participants of that conversation. It is not so much that conversations are necessarily hard to enter, rather that they are very hard to enter when the participants are quite unwilling to draw outsiders in. Clearly it is very hard for a Dutch college professor to join a conversation in an Italian square; similarly it is very hard to join an academic conversation which is constantly referring only to itself and is full of jargon. To join such a conversation you do indeed need graduate studies as Klamer claims (Klamer 2007, 158). However things might be very different if we try to start a conversation with an Italian passer-by, whom we ask to explain some of his culture and perhaps introduce us to some of his friends. It might actually turn out that this person is eager to explain to us all about the riches of his culture and how much it can teach us. Similarly it is conceivable that the economic conversation, although specialised, could become as open and willing to interact with its surroundings, as say the conversations of Adam Smith, Marshall, or Tinbergen.

The idea that economists should write for other economists is a relatively recent idea. In fact Klamer points out that the turn inwards—the idea that the significant audience comprises the initiated, that is, colleagues and knowledgeable critics—is a modernistic idea (Klamer 2007, 147). So rather than claim that it is a universal characteristic of the academic or economic conversation as Klamer does, I would like to claim that it is a characteristic typical of the (late) modernistic economic conversation. It is the idea of economics for economics sake that is so typical of post-WWII writing.

Economic writing after WWII was no longer about the relevance of economic models for the real world, but about the theoretical possibilities of the general-equilibrium models or capital theory. During the inter-bellum period there was a strong belief that economics could make the world a better place and therefore discussing policy and addressing a wider audience was an essential part of economics. The turn inwards led to the disappearance of policy advice from the economics literature or to its becoming part of the specialised discipline of public finance. One could say that the economic conversation changed from a group of eager Italians trying to tell us about the richness of their culture to the sealed-off group unwilling to talk to strangers.

This change is also reflected in the motivations that economists have for doing economics, as Klamer shows. In the late modernistic phase, economists became sceptical about the applicability of their own theories and models to the real world. Reasons for doing economics changed from overcoming business-cycles and stabilizing economic growth to solving theoretical puzzles and finding firm mathematical foundations. More recently economists' self-justifications seem outright cynical, like: having fun, doing it because it is interesting, or even 'to keep ourselves busy'. The nature of their articles has reflected this attitude: the conversation was not aimed at the world out there, but at other economists, or not even that (Klamer 2007, 146). A vice that Klamer seems to suffer himself when in his introduction he claims that: "Even if no one pays any attention to them [my thoughts], the book has satisfied my hunger to make sense of the world I am part of" (Klamer 2007, xvii). While not as cynical as an article about the dead-weight loss of giving Christmas gifts, such statements reflect a general feeling that what economists do is largely irrelevant to others.

It would be wrong however to believe that the idea of a conversation entails a sealed-off conversation turned inwards. In fact, as I have said, the style of Klamer's book can be taken as a strong argument that an open conversation can exist. He is desperately trying to explain to outsiders why economics is the way it is. He wants to be held accountable by the general public for what he and his colleagues are doing inside the ivory tower, and I would say rightly so. If we accept Klamer's claim that the relevance of an economic argument is evaluated within the conversation of economists, the relevance of the overall economic conversation should be evaluated within a broader conversation, the academic, for example, or the political. This wider conversation should not be a one-way street in which others judge the economic conversation, but it should be a way to show the relevance of economic arguments to others. A conversation that is turned inward is too easy to ignore, something that politicians are very prone to do.

It is truly unfortunate that Klamer does not make this argument explicit. By introducing a concept like accountability within his theory about the conversation, he could have strengthened it greatly. True enough, economics is what economists do. However, what can it tell us about the world? And why is it worth spending taxpayers' money on it? Being accountable means that separated conversations have to explain their relevance to and for each other. In many ways economics is too important to be left just to economists. The accountability that Klamer seeks in his own style should have been developed into an argument that a conversation is not completely autonomous and should at the very least also be judged by its relevance to other conversations.

By not introducing accountability, Klamer grants too much autonomy to the economic conversation. We have already seen that for Klamer one of the defining characteristics of modernism is the turn inwards: the idea that the significant audience comprises the initiated. Part of this idea is the belief that a discipline can provide its own justification, a project that was most explicitly present within mathematics. This idea however has lost most of its force, not in the least because it proved to be impossible even in mathematics. The claim that economics is what economists do, which is so important for Klamer, however, reflects this modernistic attitude that a discipline can justify itself: "Judging economists from the ground floor up is pointless. To judge that conversation high up you need to enter it and that takes a while, a few years at least and preferably graduate study" (Klamer 2007, 158). With this claim he accepts the modernistic belief that only insiders can have relevant opinions about the conversation. Ironically his style of writing provides a perfect example that this is not at all true.

Overall, I think that while Klamer's characterisation might be quite appropriate for the late modernistic economic conversation of the seventies and eighties, it does not do justice to the changes in that conversation since. Most importantly, however, by accepting the idea that economics is what economists do, he unwittingly buys into the modernistic idea that a discipline (conversation) can justify itself. I have suggested here that if we accept his idea that arguments are evaluated within a conversation it is very unsatisfactory not to have a way to evaluate the different conversations. Holding conversations accountable to each other would be a way to evaluate different conversations, even if they operate relatively independently. And by doing so Klamer would give us at least the beginning of an answer to the question that keeps pressing itself to the forefront in this book: what is the justification for the economic conversation and why is that conversation relevant?

REFERENCES

Klamer, Arjo. 2007. *Speaking of economics: how to get into the conversation*. Abingdon: Routledge.

Klamer, Arjo, and Donald N. McCloskey. 1989. The rhetoric of disagreement. *Rethinking Marxism*, 2 (3): 140-161.

Klamer, Arjo, Donald N. McCloskey, and Robert M. Solow (eds.). 1988. *The consequences of economic rhetoric*. Cambridge: Cambridge University Press.

Erwin Dekker is a PhD candidate at the Faculty of history and arts, Erasmus University Rotterdam. His thesis concerns the representation of the economy in the twentieth century, and his main research interests are the history and philosophy of economic thought.
Contact e-mail: <dekker.erwin@gmail.com>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 181-184.
http://ejpe.org/pdf/1-1-ts-1.pdf

PHD THESIS SUMMARY: The imprisoner's dilemma: the political economy of proportionate punishment.

DANIEL J. D'AMICO
PhD in economics, May 2008
George Mason University

The most common heuristic used in economics is not coincidentally entitled the 'prisoner's dilemma'. Several classical economists saw in the social phenomena of crime and punishment an obvious demonstration of the principles of economics. Through changes in the criminal justice system society could deter criminals, thus ridding itself of the baneful costs of crime. Adam Smith (1762) explained the most basic choice inherent to all social action. Men are confronted with the perpetual choice to either truck, barter and exchange—or rape, pillage and plunder. The prisoner's dilemma is a modern and formal presentation of Smith's profoundly subtle insight. The game's namesake scenario describes two criminals so attracted by the personal rewards of defection that they forgo the higher social rewards of cooperation. Not only are the inmates in the narrative strategically pitted against one another, but so are all individuals constantly offered the short term rewards of taking to get ahead instead of trading.

As James Madison explained in the Federalist Paper, No. 51 (1788), institutional design attempts to promote mutual exchange while suppressing coercion. If punishments are levied so the short term rewards of crime are not attractive compared to the long term rewards of production and exchange, then individuals will choose the latter. In the prisoner's dilemma of social interaction, the rule of law acts at the meta-level.

As my title implies, I intend to treat the imprisoner—the central planner, the government itself—as no different from the traditional agents that economists so often analyze. I entertain the possibility that state authorities are no better informed or incentivized than the ordinary individuals within society, and that are also tempted by the rewards of defection.

Men are governed by other men—imperfect and fallible. Real punishments are imperfect because those who create, interpret, and enforce the rules are not God. Yet it is a bizarre irony that today's most popular criminal justice theories—those theories which supposedly justify, legitimize, guide, organize, and motivate the criminal justice system—begin from the premises that the state authority is the necessary and sufficient purveyor of criminal punishments.

The most recent trend in punishment theory calls for adherence to the principle of proportionality: a punishment should be well-fitted to the crime, and like crimes should be treated alike. I accept the proportionality principle as intuitively appealing and theoretically sound, but the question remains, what institutional framework best produces proportionality?

In chapter 2, I begin this dissertation by drawing attention to the dominant trends in sociology and economics that are responsible for preserving the assumed role of the state in providing criminal punishment. Economics and sociology have long sat in opposition to one another. Sociologists sought to understand the essence and operations of punishment institutions, while economists tried to optimize the allocation and production of punishments. Naturally these perspectives were hostile to one another, but today each field has taken significant strides towards a common ground. Sociology and economics have both begun to look at the important role that institutions have on influencing the outcomes of social and economic processes. I present a framework of reciprocally embedded institutional influence to help explain significant historical changes in punishment paradigms over the last half century.

I go on in separate chapters to argue that a centrally-planned criminal justice system often produces dis-proportionate punishments because of its inability to deal with problems associated with dispersed knowledge and non-benevolent incentives. In chapter 3, I argue that knowledge problems inhibit a central-authority, even one guided by benevolent intentions, from knowing how to provide punishments in the quality and quantities that best produce proportionate outcomes. A central-planner can never possess the full scope of knowledge required to achieve proportionate punishments because such knowledge is often dispersed throughout society in the minds of several different people. Even if a central-planner fully embraced the insights of the proportionality principle, he would still lack the knowledge of how to produce real proportionate punishments in practice.

In chapter 4, I argue that central-planners lack the incentives to avoid bureaucratic inefficiencies. If a central authority knew what decision making process—rules or discretion—could achieve proportionality it would still lack the incentives to follow such processes. In theory, the long and variable lags associated with punishment policy give good cause for rule-based sentencing rather than discretion-based sentencing. Rules should provide stable long-run expectations about the future level of crime and therefore promote investments in long-run production processes for security and deterrence technologies. But the government-monopoly over the criminal justice system lacks a credible commitment to obey rule-based criminal sentencing guidelines once they are in place. In the face of incomplete and non-credible rules, various agents in the criminal justice hierarchy wield de facto discretion over the outcomes of criminal sentences.

I take the ends of proportionality as given and ask the critical question whether centrally-planned institutions are ever capable of knowing or discovering the preferred techniques of proportionality or if they ever confront the incentives to produce proportionality. These theoretical exercises of assuming imperfect knowledge and imperfect incentives are not incompatible but are instead complimentary to each other. The incentive problems explained in chapter 4 are not a crucial critique against the knowledge problems explained in chapter 3. The essential knowledge regarding the ins and outs of the criminal justice system is suppressed when market-based decision-making processes are replaced by politics. Decisions must be made according to some criteria; the political process then introduces incentive incompatibilities. The logical inconsistency between the ends of proportionality and the means of central-planning is the fundamental problem that upholds the emergence of socially preferable outcomes in criminal justice. Low crime, low costs, proportionality, and equality before the law are replaced by high crime, high-costs, dis-proportionality, and disparity.

This dissertation is a step to constructing and implementing criminal-justice reform from a logical and philosophically consistent approach. Though my intentions are not necessarily the elimination or deterrence of crime, it seems reasonable to assume that the establishment of a criminal justice system that is more internally consistent, responsive to social preferences, and informed by realistic assumptions will in turn also be more technologically efficient at responding to crime.

Daniel J. D'Amico obtained his PhD in economics from George Mason University on May 2008, under the supervision of Professor Peter J. Boettke (George Mason University). The author is currently assistant professor of economics at the Joseph A. Butt S. J. College of Business, Loyola University New Orleans.
Contact e-mail: <danieljdamico@gmail.com>
Website: <www.danieljdamico.com>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 185-187.
http://ejpe.org/pdf/1-1-ts-2.pdf

PHD THESIS SUMMARY: Rationality and institutions: an inquiry into the normative implications of rational choice theory.

BART ENGELEN
PhD in philosophy, September 2007
Katholieke Universiteit Leuven

In my dissertation, I aim to analyze what a desirable basic institutional structure looks like from the perspective of rational choice theory. While the main topic is thus normative in nature, I start by clarifying in the first part what the notion of rationality exactly entails. I do so by focusing explicitly on the economic conception of rationality, according to which a rational individual is motivated to serve his self-interest on the basis of cost-benefit calculations. Such a *homo economicus* is characterized by intentional and instrumental actions, perfectly informed beliefs and exogenously given and egoistic preferences. In my view, however, this model is inadequate if one aims to understand what it means to be rational. All of the above mentioned requirements turn out to be overly demanding in this respect.

That is why I suggest dropping these assumptions in order to construct what I label the minimal conception of rationality. Since the latter turns out to be very formal indeed, I propose two further alternatives, which focus not so much on the choice of means to achieve certain goals but rather on the choice of those goals themselves. According to the first, broad conception, actions are rational if they are based on good reasons, which are further qualified as well-informed beliefs and autonomous preferences. According to the second, expressive conception, actions, beliefs and preferences are rational if they express the things people care about. The latter requires that people can reflect upon and identify with their reasons, which implies a capacity to reflect upon and distance themselves from their own bundle of preferences.

In the second part of this dissertation, I try to show the value and limitations of these conceptions by applying them to decisions in the context of large-scaled elections. In this respect, it becomes immediately clear that the economic conception fails to explain why quite a lot of people go out and vote. After all, a single vote has only an infinitesimal impact on the electoral result and thus does not enable people to serve their interests or realize their goals. This leads to the so-called voting paradox, according to which no rational individual would decide to vote. The standard solution is to assume that individuals vote because they derive satisfaction from the very act itself. However, this strategy is rather ad hoc and does not explain how people vote once they find themselves inside the voting booth. The expressive conception of rationality does better in this respect. It suggests that people vote because they care about democracy in general or about a particular political candidate or ideology. Since they conceive of themselves as being a good citizen (or, say, a good socialist), they express this aspect of their identity by going out to vote (for the socialist party).

In the third part of this dissertation, I analyze more fully the normative implications of the different conceptions of rationality. More specifically, I try to answer the question of which basic institutional structure is desirable if one assumes that people are by and large rational. This immediately shows that both the normative issue (what should institutions look like?) and the explanatory issue (how do rational individuals act?) are closely connected. In my view, proposals regarding institutional design and reform should be based on empirically adequate models of individual actions and motivations. This search for a realistic utopia goes against the conventional strategy of most economists who rely on the *homo economicus* model even when it fails to explain individual behavior.

To explain more fully what the normative implications are of the counterfactual assumption that all people are economically rational, I focus on the work of James Buchanan. In his theory of constitutional choice, he argues in favor of a minimal state whose only task is to make sure that the market functions properly. Buchanan thus favors a strict constitutional limitation of governments, which tend to expand beyond legitimate borders as soon as politicians and public servants are allowed to serve their own interests. In my view, however, the abovementioned criticisms of the *homo economicus* model have theoretical as well as normative implications. The empirically supported fact that a majority of individuals does not act in economically rational ways creates more room for legitimate government intervention. Expressively rational citizens will, for example, more easily agree on the desirability of a collective provision of certain public goods. In addition, expressively rational politicians and public servants can be more easily trusted to serve the public interest rather than their narrowly defined self-interest.

As an alternative to Buchanan's one-sided focus on economic rationality (at the individual level) and the market (at the institutional level), I focus on the work of Samuel Bowles and Herbert Gintis. More specifically, I explore their work on the phenomenon of strong reciprocity, which refers to the widespread tendency of people to reward prosocial behavior and punish antisocial behavior, even if this is costly for themselves. Since this is clearly economically irrational, Bowles and Gintis propose to complement the *homo economicus* model with the *homo reciprocans* model. This model, which comes close to the expressive conception of rationality, is able to incorporate the insight that social norms surrounding reciprocity, cooperation, and fairness are crucial in regulating interactions.

At the normative level, Bowles and Gintis stress that such norms often lead to socially desirable outcomes, since they enable people to cooperate without relying on coercive and costly government intervention. This suggests that the debate between proponents of the market and of the state neglects the importance of communities in which people spontaneously interact on the basis of informal and generally prosocial norms. As such, the insights of Bowles and Gintis lead me to defend a basic institutional structure in which markets, states and communities mutually complement and reinforce each other. They also justify my general optimism as regards the capacity and motivation of people to try and improve the rules and institutions that govern their everyday lives.

Bart Engelen obtained his PhD in philosophy from the Katholieke Universiteit Leuven, Institute of Philosophy, on September 2007. He has been supervised by Professor Antoon Vandevelde (Dean of the Institute of Philosophy, Katholieke Universiteit Leuven). Currently, the author is a postdoctoral researcher at the Institute of Philosophy, Katholieke Universiteit Leuven. Contact e-mail: <bart.engelen@econ.kuleuven.be>

Erasmus Journal for Philosophy and Economics,
Volume 1, Issue 1,
Autumn 2008, pp. 188-189.
http://ejpe.org/pdf/1-1-ts-3.pdf

PHD THESIS SUMMARY:
The moral trial: on ethics and economics.

ALESSANDRO LANTERI
PhD in philosophy and economics, April 2008
EIPE, Erasmus University Rotterdam

This dissertation investigates the experimental evidence exposing how economists' behaviour differs from that of non-economists, in that economists often display more self-interested conduct. A veritable 'moral trial' has stemmed from that evidence, in which it is argued that economists are selfish, thus immoral, and it is recommended that they change the teaching of economics.

I disassemble the moral trial (Section I), examine the psychological and logical soundness of both: evidence and charges (Section II) and find them lacking in several respects. Finally, I suggest (Section III) a novel and altogether different interpretation of the evidence.

The first section proposes as a starting point a sketch of Economics (chapter 1) and how its main focus is on markets (chapter 2). It also introduces the main theme of this work: the moral trial that has befallen the discipline and its practitioners (chapter 3). I assess the charges—which are unclear—and the evidence—which I find inconclusive. In the moral trial, any guilt of economists ought to depend on their being selfish, in a fashion reminiscent of 'economic man'.

In the second section, therefore, I explore the ways in which we make sense of other people's motivations and behaviour, and outcomes, and expose numerous fallacies people incur when attributing motivation to observed actions (chapter 4). Further, to even constitute a charge, selfishness ought to be described as a moral violation, but such a case may be very hard to make (chapter 5).

Although I show that it does not represent sound evidence of selfishness or immorality, the behavioural gap between economists and non-economists requires an explanation beyond criticism. In the last section, I therefore attend to the task of making sense of economists' peculiar conduct. I start by emphasising the importance of emotions in decision-making and the ways in which the perception of the choice context affects behaviour (chapter 6) and I then proceed to argue that economists, because of their training and their specialised knowledge, frame situations differently from non-economists and more specifically that they frame most decision contexts as market-like. One could thus explain economists' behaviour in the experiments by looking at the way in which different sub-samples of subjects frame situations (chapter 7). That is only part of the explanation because economics students already behave differently from other students at the beginning of their training and this may happen because of an 'economist's stereotype', to which freshmen adjust upon enrolment (chapter 8). To conclude, if the teaching of economics matters at all, is it fair that it stand judgement for its effect on our students? Is it ruining them? (chapter 9).

The conclusion of this dissertation is that economists frame situations in a way that makes them believe self-interested conduct is fine and therefore behave self-interestedly on certain occasions. This peculiar behaviour is probably responsible for the unflattering economic stereotype, which in turn represents a benchmark for young economics students. These explanations of economics students' behaviour seem sounder than the one prevailing in the literature—namely self-selection: the claim that selfish people voluntarily enrol in economics for which no persuasive rationale has yet been proposed. The explanations advanced here, moreover, reject any deep difference between economists and non-economists, which would be difficult to square with the observations that, on many occasions, economists behave no more selfishly than non-economists. Finally, since the behavioural gap narrows after graduation, it seems that economics teaching does have some consequences for its students, but that these consequences wear off with time. The moral trial should therefore not be a cause for too much concern about the ethics of economists.

Alessandro Lanteri obtained his PhD in philosophy and economics on April 2008, from the Erasmus Institute for Philosophy and Economics, Faculty of philosophy, Erasmus University Rotterdam (The Netherlands). He developed his doctoral research under the supervision of Deirdre McCloskey (University of Illinois at Chicago) and Jack Vromen (EIPE, Erasmus University Rotterdam). The author is currently a post-doctoral fellow at the Department of public policy and public choice (POLIS), Faculty of political science, at the University of Piemonte Orientale (Alessandria, Italy). Contact e-mail: <alessandro.lanteri@sp.unipmn.it>
Website: <alelanteri.googlepages.com>