

—Published in: *Uskali Mäki, Ionnis Votsis, Stéphanie Ruyhy and Gerhard Schurz (eds.) Recent Developments in the Philosophy of Science. EPSA13. Heidelberg: Springer (2015): 99-109—*

*This is the penultimate version. For the definitive text, please consult the published version:*

*[https://link.springer.com/chapter/10.1007/978-3-319-23015-3\\_8](https://link.springer.com/chapter/10.1007/978-3-319-23015-3_8).*

## Why the Realist-Instrumentalist Debate about Rational Choice Rests on a Mistake

Christine Tiefensee

### Abstract

Within the social sciences, much controversy exists about which status should be ascribed to the rationality assumption that forms the core of rational choice theories. Whilst realists argue that the rationality assumption is an empirical claim which describes real processes that cause individual action, instrumentalists maintain that it amounts to nothing more than an analytically set axiom or ‘as if’ hypothesis which helps in the generation of accurate predictions. In this paper, I argue that this realist-instrumentalist debate about rational choice theory can be overcome once it is realised that the rationality assumption is neither an empirical description nor an ‘as if’ hypothesis, but a normative claim.

### Keywords

Rational Choice Theory • Realism • Instrumentalism • Rationality • Normativity

## 1 Introduction

Within the social sciences, much controversy exists about which status should be ascribed to the rationality assumption that forms the core of rational choice theories. In one corner of the ring, we find realists who argue that the rationality assumption is an empirical claim which describes real processes that cause individual action. In the other corner, we see instrumentalists who maintain that the rationality assumption amounts to nothing more than an analytically set axiom, an ‘as if’ hypothesis or useful fiction, which helps in the generation of accurate predictions. In this paper, I approach the realist-instrumentalist debate from a different angle by bringing a distinctly normative interpretation of rationality to the contest. This understanding submits that, contrary to realist or instrumentalist readings, rationality is a normative concept, with ascriptions of rationality being normative judgements that evaluate agents, their actions and intentional states.<sup>1</sup> I will argue that this interpretation is correct: the rationality concept is normative. My main objective, though, will be to show that once the realist-instrumentalist debate is seen to rest on a mistaken interpretation of the rationality assumption, this debate loses its footing. More generally, then, this paper is driven by the conviction that discussions about rationality within the social sciences,

---

<sup>1</sup> To avoid misunderstanding right from the outset, this normative account submits that the rationality concept is normative, not that it is ethical or moral.

metaethics and the philosophy of science can greatly benefit by taking each other's findings more strongly into account. It thus seeks to contribute to an interdisciplinary approach which aims to bring together different strands of enquiry so as to make best use of their respective insights.

I will start by clarifying what is understood by the rationality assumption together with its realist and instrumentalist interpretations. This will be followed by arguments for the normativity of rationality. I will then turn to my main objective, setting out how the normativity of rationality helps overcome the realist-instrumentalist debate. The paper concludes by examining possible objections to the normative account and its implications for positive uses of normative rationality assumptions in the social sciences.

## 2 The Rationality Assumption

Rational choice theory is concerned with instrumental rationality. As such, its underlying rationality assumption can be stated as follows:

- (RA) A rational agent  $\mathcal{A}$  chooses the best means so as to attain a specific end, given  $\mathcal{A}$ 's beliefs.

More formalised, we find the following definition in the rational choice literature:

- (RA\*) A rational agent  $\mathcal{A}$  maximises the expected value of a utility function defined on  $C$ , relative to a subjective probability distribution defined on  $B$ .<sup>2</sup>

Accordingly, (RA)'s concern is threefold (List/Pettit 2011: 24): It deals with the way in which  $\mathcal{A}$ 's attitudes connect with his environment (attitude-to-fact standards), how  $\mathcal{A}$ 's preferences<sup>3</sup> and beliefs relate to one another (attitude-to-attitude standards) and how these propositional attitudes result in action (attitude-to-action standards). With regard to attitude-to-fact relations, (RA) implies that rationality ascriptions are to be based on an agent's subjective perception of the situation, not its objective description. Concerning attitude-to-attitude relations, beliefs and preferences are required to meet certain coherence criteria, such as consistency and transitivity together with constraints on probability distributions. This makes possible consistent belief sets and orderings of preferences. In accordance with modern utility theory, the content of preferences is left open and not limited to specific considerations. Finally, attitude-to-action standards demand that actions maximise expected value, where any such maximisation is again regarded as relative to  $\mathcal{A}$ 's subjective beliefs. It must be noted that this specification of the rationality assumption is one amongst several possible interpretations. I will return to some such alternatives in the penultimate section of this paper.

How do (RA) and (RA\*) relate? I will assume here that (RA\*) is a formal representation of RA: It offers a mathematical formalisation of our folk psychological approach to intentional, purposeful action (Hausman 2012). Acknowledging that (RA\*) is RA's mathematical representation is crucial for our understanding of the realist-instrumentalist debate. For, as Lehtinen/Kuorikoski (2007: 123) stress, since utility functions are nothing more than mathematical representations of preference

---

<sup>2</sup> This is an adaptation of Binmore's (1998: 360-361) definition.

<sup>3</sup> I will use the terms 'preference', 'desire' and 'end' interchangeably here.

orderings which are, moreover, open to certain mathematical transformations, (RA\*) can only ever be read as an ‘as if’ claim. Consequently, the realist-instrumentalist debate should not be understood as revolving around (RA\*): the question is not whether or not agents really employ the mathematical terms of rational choice theory, have utility functions and calculate expected utilities. Rather, the focal point of the realist-instrumentalist debate concerns the status of (RA) and thus the question of whether or not agents really choose in the way that (RA) envisages. Let us turn to this debate next.

### 3 Realism vs. Instrumentalism

Realists can be distinguished from instrumentalists by their stances on the following three theses:

- (1) (RA) is an empirical claim about unobservables.
- (2) It is an open question whether or not actual agents are rational.
- (3) Actual agents are rational.

Realists endorse all three claims. According to them, the concept of rationality refers to the property of rationality, just as (RA) depicts causal mechanisms that underlie the decision-making process of human beings (1). Realists thus interpret RA empirically: Although RA concerns unobservables—after all, what we can directly observe is agents’ behaviour, not their rationality—ascribing rationality to agents amounts to putting forward an empirical description of their psychologies.<sup>4</sup> This enables (RA) to feature not only in predictions, but also in the explanation of phenomena that interest us. Since (RA) is an empirical claim, tests must show whether or not agents really are rational (2). Yet, realists submit that this is a question that we should answer in the positive: (RA) is, even if maybe not outright true, at least a good approximation of the truth (3).<sup>5</sup>

Instrumentalists disagree. Contrary to (1), they argue that (RA) should be understood as nothing more than an ‘as if’ hypothesis<sup>6</sup> or stipulated axiom which we employ in analytic models so as to develop accurate predictions. Since (RA) makes no claim whatsoever about processes that are operative in actual agents’ psychologies, seeking to test (RA) is beside the point. Yet, if we leave our analytic models behind and enquire whether or not actual agents do indeed choose as this axiom envisages, instrumentalists agree with (2) that this is an open question. Nonetheless, they reject (3): Either, they declare that actual agents are not rational,<sup>7</sup> or they stay non-committal on whether or not they are. Either way, they hold that (RA) retains its usefulness as a model-theoretic axiom as long as it helps produce accurate predictions. Whether or

---

<sup>4</sup> See Popper (1967) for the peculiar position that the rationality assumption is empirical and indeed false, but that we should nevertheless hold on to it, and Lagueux’s (1993) attempt to salvage this position. See Hempel (1962) for the claim that the rationality assumption might be abandoned in tests.

<sup>5</sup> Realists’ endorsement of (3) is compatible with the thesis that (RA) is a deliberately *false*, idealised empirical description (Mäki 2000). Realists’ claim would then be that actual conditions sufficiently approximate these idealised conditions. Since this variation on realism does not affect my argument, I will neglect it here.

<sup>6</sup> As Lehtinen (2013) emphasises, only certain interpretations of ‘as if’ signal an instrumentalist outlook.

<sup>7</sup> Friedman (1953) sometimes appears to take up this stance, but compare Mäki (2000).

not actual agents really do choose as postulated by (RA) is none of instrumentalists' business.<sup>8</sup>

#### 4 The Normativity of Rationality

In contrast to both realist and instrumentalist interpretations, I suggest that (RA) should be understood as a normative claim. Holding rationality to be a normative concept is a widespread, albeit not uncontroversial, position in metaethics.<sup>9</sup> Nor is it unknown in the social sciences, where Harsanyi (1976: 90), for instance, declares that "already at a common-sense level, rationality is a *normative* concept: it points to what we *should* do in order to attain a given end or objective."<sup>10</sup> Two features of the rationality concept support this normative interpretation.

Firstly, rationality ascriptions are in a thin sense normative in that they are based on norms or standards: When determining whether or not an individual behaves rationally, we need to apply the rationality standards mentioned above. What endows these standards with thick normativity, though, is that they do not simply divide choices into the categories 'rational' and 'irrational', but that these categorisations carry inherent positive and negative valence: Rational action makes sense, or can be supported by reasons, or has gone right in an important way, whereas irrational behaviour makes no sense, or cannot be supported by reason, or has gone wrong in an important way. We approve of rational action and criticise irrational choices. Rational and irrational behaviour are, therefore, not normatively on a par. Secondly, normative concepts are characterised by their intimate link to action in that, *prima facie*, we would expect individuals to be motivated to act in accordance with their normative judgements. This is particularly clear within the moral context: If John judges that one ought to help people in need, say, we would expect him to be inclined to help needy people when the situation arises. The rationality concept also possesses this practical import which is paradigmatic of normative concepts. For example, if John intends to travel to Berlin and agrees that it would be most rational for him to catch the train, we would expect him to step on a train, and not take the car, say. Rationality judgements, then, are not only associated with normative pressure to conform to them, but also subject irrational agents to criticism and the demand to correct failures of rationality in future.

Yet, the rationality assumption is not only normative, it is also constitutive of agency. Donald Davidson (1980, 1984, 2004) indefatigably explains why. When ascribing intentional states to others and trying to understand their actions, we are engaged in the business of interpretation: Against the background of our own beliefs, we want to make intelligible each other's actions and the intentional states that lead to these actions. The linchpin of interpretation is, in turn, provided by (RA). In a

---

<sup>8</sup> Rational choice theory can also be understood as being engaged in the explication of the concept of rationality. If understood along these conceptual lines, analytic and normative readings of the rationality assumption need not be mutually exclusive, as the rationality assumption could be understood as the conceptual specification of the normative concept of rationality.

<sup>9</sup> Again, the term 'metaethics' should not mislead readers into thinking that the normative account interprets instrumental rationality as an ethical or moral concept. Strictly speaking, 'metanormativity' would be a more appropriate label than 'metaethics' here.

<sup>10</sup> Compare Gibbard (1990), Southwood (2008), Broome (2007). See also Hands (2011, 2012) and Grüne-Yanoff/Lehtinen (2012) for discussions of normative interpretations of rational choice theory.

nutshell, we impute to agent  $\mathcal{A}$  the desire to drink some water, say, if this desire makes intelligible his drinking the glass of water offered to him, provided that he has the belief that the glass contains water. Similarly, the belief that the glass contains water should be attributed to  $\mathcal{A}$  if this belief makes sense of his drinking the water offered to him, provided that he has the desire to drink water. Finally,  $\mathcal{A}$ 's drinking water will count as an action, provided that his wanting to have a drink of water and believing that the glass contains water are  $\mathcal{A}$ 's reasons for his drinking water. More generally, then, when asking which propositional attitudes should be attributed to agents and how to understand their actions, we are guided by the normative question of which intentional states it would be most rational to have and which action would render the agent most intelligible. To adapt Sellars' (1956/1997) words, interpreting agents thus amounts to embedding their actions within the space of reasons.

This implies that judgements of rationality do not evaluate as rational or irrational independently ascertainable propositional attitudes of belief and desire—judgements of rationality are not normative add-ons to otherwise non-normative statements about agents' intentional states. Instead, our conception of propositional attitudes is itself imbued with standards of rationality: It is they which determine which desires and beliefs to impute to an agent by subsuming an agent's behaviour under a consistent, rational set of propositional attitudes. As such, they determine not only what counts as an action and who counts as an agent, but also how to individuate and ascribe beliefs and desires. Standards of rationality, then, are constitutive of intentional states and agency in that an individual who does not satisfy demands of rationality at some minimal level cannot count as an agent in the first place. In Føllesdal's (1982: 312) words: The "assumption that man is rational is ... inseparable from other hypotheses that we make about man: that he has beliefs and values, that he acts, etc. We may in a given case be forced to give it up, but then we have to give up these other hypotheses, too."

## 5 The Mistake of the Realist-Instrumentalist Debate

It is only a small step to see how the insight that (RA) is a normative, necessary precondition for thought and agency affects the realist-instrumentalist debate. Let us remind ourselves that this debate is sparked off by the following theses:

- (1) (RA) is an empirical claim about unobservables.
- (2) It is an open question whether or not actual agents are rational.
- (3) Actual agents are rational.

Realising that ascriptions of rationality are normative shows that both instrumentalists and realists are misguided, either in their respective take on these theses or in their reasons for adopting them. Contrary to (1), (RA) is a necessary, normative precondition for understanding ourselves as intelligible agents who act on the basis of their intentional states. As such, realists are wrong to conceive of (RA) as an empirical claim, just as instrumentalists are wrong to regard it as nothing more than an 'as if' hypothesis. At the same time, instrumentalists are indeed right to point out that attempts to falsify (RA) miss the mark. However, the reason why questions of empirical confirmation and falsification do not apply to (RA) is not grounded in (RA)'s allegedly analytic status, but in its normative nature: Since (RA) is located in the normative and not in the empirical realm, it cannot be empirically confirmed or disconfirmed. Contrary to (2), both realists and instrumentalists are mistaken in

thinking that it is a live question whether or not actual agents are indeed rational. Since rationality is constitutive of thought and agency, being an agent amounts to being a rational agent—without rationality, there is no agency. On the normative reading, then, the question of whether or not agents are rational is closed. Finally, with regard to (3), we can see that instrumentalists cannot withdraw to the evasive stance on (RA)’s truth which they intend to take up. In light of the special status of rationality norms as framework principles for agency discourse, there is no logical space for holding that we merely assume that agents act as if they were rational, without declaring that they are rational. Realists, in turn, are right to endorse (3), yet are wrong to support (3) on empirical grounds. The rationality of agents is not an empirical finding, but a precondition of the interpretive process.

To sum up, the realist-instrumentalist debate is fuelled by the assumption that (RA) is a non-normative claim. Arguments to the effect that rationality is a normative concept which is constitutive of thought and agency show that this is a mistake. (RA) is neither open to empirical confirmation and disconfirmation—as realists would maintain—nor allows instrumentalists to take up a non-committal stance on agents’ rationality. The question ‘Are actual agents really rational or do we only speak as if they were rational?’ is, therefore, beside the point.

## 6 Objections

This argument may be too quick and simple to be fully convincing, so let us look at some possible objections to it. Here, I will limit my discussion to objections that accept the claim that rationality ascriptions are normative whilst questioning the impact that this normativity supposedly has on the realist-instrumentalist debate. I will focus on three worries that I deem the most pertinent. Many of these objections raise further, often thorny lines of enquiry. Pursuing these enquiries in full would go beyond the scope of this paper, so my responses will have to remain schematic at times.

### 6.1 First Objection

Any plausible theory of rationality must be able to account for cases of irrationality. After all, not only do cases of irrationality abound, we can even identify the processes that cause irrational decisions—just think of framing effects, weakness of will, cognitive biases or wishful thinking, to name but a few. Bearing in mind that the normative account renders questions about agents’ rationality closed, it might be concluded that it clearly fails this requirement and should, therefore, be rejected.

This objection rightfully challenges the normative account to show how cases of irrationality can be reconciled with rationality’s constitutive role for thought and agency. To conclude that this challenge remains unmet, though, would be premature. For, whilst the normative account does indeed rule out *global* irrationality, it can allow for cases of *localised* irrationality. To elaborate, it might not always be possible to subsume an agent’s behaviour under a set of propositional attitudes which fully satisfies the rationality constraint. If so, we are indeed forced to admit that an agent is, in a certain respect, irrational. As long as these lapses of rationality remain sufficiently limited, though, they need not undermine agency status. How pervasive failures of rationality may be without threatening agency is, in turn, a further, normative question—it enquires how much irrationality we can absorb whilst still being able to conceive of an individual as an intelligible agent. Accordingly, the normative account

does acknowledge cases of irrationality, but stresses that this irrationality can only ever be local. Global irrationality entails loss of agency status.

## 6.2 Second Objection

I have mentioned above that the normative account employs a specific interpretation of the rationality assumption, whilst different specifications are also available. One such alternative could be the following:

(RA\*\*) A rational agent  $\mathcal{A}$  maximises his material self-interest relative to objective probability distributions.

Contrary to (RA), (RA\*\*) strictly constrains the content of preferences by focussing on material self-interest only and requires agents to have objectively true beliefs about their situation so as to count as rational. Here, critics may insist that it clearly is a live question whether or not individuals act in line with (RA\*\*): Whether or not they pursue their self-interest and have correct beliefs about probability distributions is, after all, far from being settled. It might be objected, then, that the conclusions of the normative account either miss their target by wrongly focussing on (RA), or must be limited in scope so as to apply to (RA) only.

Bearing in mind (RA)'s prominence and widespread employment in the social sciences, criticising the normative account's focus on (RA) is not convincing. For the same reason, even if its conclusions applied to (RA) only, this would be a significant result. Still, questions about the scope of its conclusions are pressing, so let us look more closely at the status of alternative rationality specifications such as (RA\*\*). Initially, we must note that just like (RA), (RA\*\*) is neither an empirical claim nor an 'as if' hypothesis, but a normative statement: It puts forward a specific interpretation of the value of rationality. Yet, in contrast to (RA), this suggested understanding is substantive in that it evaluates agents' ends by identifying material self-interest as a rational desire and imposes stricter demands on rational belief. As such, (RA\*\*) does not merely demand, like (RA), that an agent's preferences, beliefs and actions stand in certain structural relations to one another, but also requires rational agents to have propositional attitudes with a specific content. Accordingly, if focus on (RA\*\*) were to revive the realist-instrumentalist debate, this would have to be because of these additional demands on the content of agents' propositional attitudes: The question 'Are actual agents really rational?' would now have to be interpreted as the question 'Do actual agents really have these particular desires/beliefs?'. Here, critics are right: The normative account does not close questions about the specific content of agents' desires and beliefs. For, although it entails that if an individual is an agent, his beliefs and desires adhere to (RA), it does not settle which beliefs and desires these are. As such, it does make sense to ask 'Are actual agents really self-interested, or do we merely assume that they are?'.<sup>11</sup> However, against the background of the normative account, this question now takes on a different hue. For, since the normative account has shown propositional attitudes and instrumental rationality to be indispensable parts of agency, no fundamental doubts about instrumental rationality, unobservables, testability or propositional attitudes can drive this realist-instrumentalist question. Far from doubting the interpretive process, this query must rather be seen as being firmly located within it by enquiring about its outcomes in the form of specific attitude attributions. Consequently, focus on (RA\*\*) does not reinstate the original realist-

---

<sup>11</sup> Unlike (RA), (RA\*\*) is thus not constitutive of agency: Not pursuing one's material self-interest, say, does not undermine agency status.

instrumentalist contest, but reduces it to a truncated debate about the outcomes of the interpretive process, not its principles. As a result, it has lost much of its bite.

### 6.3 Third Objection

The final objection consists of three interrelated components. Firstly, the realist-instrumentalist debate seems to be as much about the status of certain assumptions as it is about the aim of science. According to realists, science aims at true descriptions and explanations of the phenomena that interest us, which cannot be based on mere 'as if' hypotheses. For instrumentalists, science aims at accurate predictions on grounds of parsimonious, generalisable assumptions, of which the 'as if' hypotheses of rational choice theory form one example. The normative account, it might thus be objected, does nothing to address this aspect of the realist-instrumentalist debate. Secondly, critics may object that the normative account assumes the social sciences to be interested in the explanation of individual action, whereas their research clearly focuses on the explanation of macro-phenomena. Whilst the rationality of agents may feature heavily when explaining individual action, these critics explain further, it plays no role within explanations of macro-phenomena, where the main explanatory burden is carried by situational and structural characteristics of the agent's environment (cf. Lehtinen/Kuorikoski 2007). This important distinction and its implications are simply neglected by the normative account. Finally, as these previous objections indicate, social science widely employs theories of rationality in positive explanations and predictions. However, by interpreting (RA) as a normative claim, the normative account seems to preclude such positive uses of rational choice theory.

The first component of this objection is partly right and partly wrong. It is wrong in that the normative account does intervene in the controversy about explanation and prediction insofar as it rules out arguments to the effect that rational choice explanations are impossible because (RA) can only ever be an 'as if' hypothesis. Since fundamental doubts about rationality and propositional attitudes are rejected by the normative account, such worries can no longer provide the grounds for rejecting rational choice explanation. It is right in that considerations about the normativity of rationality neither decide whether science should generally aim at explanation or prediction, nor settle whether there might be legitimate uses of 'as if' assumptions in the social sciences. However, since we should not expect normativity to be a magic wand that we can wave so as to solve all questions, and bearing in mind the multifaceted objectives of science, this is just as well.

Secondly, there is indeed a difference between explanations of individual action and those of macro-phenomena. However, this difference is not categorical. To elaborate, take the example of high cost situations which are often quoted in order to illustrate the relevance of structural constraints. In cases such as these, situational restrictions are taken to be so severe that they determine behaviour and thus render detailed enquiries into individuals' preferences superfluous. As a consequence, the explanatory burden is carried by structural features, not preferences. But even if so, we must realise that whilst high cost restrictions make it easier to ascribe preferences, they do not make preference ascriptions redundant. For, what allows us to narrow down the content of preferences so considerably in high cost situations is that it is easier to determine which preferences it would be rational for agents to have, given their severe



situational restrictions. (RA) is, therefore, not only involved in the explanation of individual actions, but is also implicitly at work in structural explanations.<sup>12</sup>

Finally, it is true that it is not obvious how normative considerations of rationality can feature in positive social science, so we must ask how the normative and positive functions of rationality can be reconciled. Luckily, though, much work has already been done on the connection between reason and cause. Since a full account of this connection would go far beyond the scope of this paper, let me merely indicate here what would be required in order to achieve a reconciliation of rationality's positive and normative functions. With regard to explanation, we would have to explain how normative accounts of action that appeal to agents' rationality and reasons relate to causal explanations of action. Any such explanation must include, amongst others, considerations about the connection between reason and cause, the non-normative supervenience base of normative rationality judgements and its link with mental causation, the possibility of psycho-physical laws and questions about normative explanations.<sup>13</sup> If prediction is understood as the symmetric counterpart to explanation, the same thoughts apply. With regard to predictions that do not harbour any explanatory ambitions, though, the normativity of rationality does not appear to pose much of an obstacle. Since in this case, we are not concerned with questions as to *why* predictions are correct, but only *that* they are correct, we can assume anything we like. As long as the generated predictions prove to be accurate, neither the status of these assumptions nor their truth matter. Of course, we may find such non-explanatory predictions highly unsatisfying. The reason for this, though, is not found in considerations about the normativity of rationality, but in general questions about the relation between prediction and explanation.

## 7 Conclusion

In this paper, I have endorsed and argued for the position that the rationality assumption is normative. I have further suggested that the normativity of rationality helps to overcome the realist-instrumentalist debate about rational choice. The question 'Are actual agents really rational or do we only speak as if they were rational?' is, therefore, beside the point. Still, many questions remain unanswered. Most pressing, we need to understand better the role of normative considerations of rationality in positive social science. It is this question which should attract our attention, not debates about realism and instrumentalism.

## References

- Binmore, K. (1998). *Game theory and the social contract*, vol. 2. London: MIT Press.  
Broome, J. (2007). Is Rationality normative? *Disputatio* II (special issue), 161-178.  
Davidson, D. (1980). Actions, reasons and causes. In his *Essays on actions and events* (pp. 3-20). Oxford: Clarendon Press.  
Davidson, D. (1984). Radical interpretation. In his *Inquiries into truth and interpretation* (pp. 125-139). Oxford: Clarendon Press.

---

<sup>12</sup> It may be argued that there are cases in which (RA) is totally irrelevant because a specific model is 'robust' regarding its behavioural assumptions, in the sense that its results remain the same no matter whether we assume rational, habitual or random behaviour, say (Lehtinen/Kuorikoski 2007: 127). Why this should still count as a rational choice explanation, though, escapes me.

<sup>13</sup> See Davidson (1980), Leach (1977), Spohn (2002).

- Davidson, D. (2004). *Problems of Rationality*, vol. 4. Oxford: Oxford University Press.
- Føllesdal, D. (1982). The Status of rationality assumptions in interpretation and in the explanation of action. *Dialectica*, 36, 301-316.
- Friedman, M. (1953). The methodology of positive economics. In his *Essays in positive economics* (pp. 1-43). Chicago: University of Chicago Press.
- Gibbard, A. (1990). *Wise choices, apt feelings*. Oxford: Clarendon Press.
- Grüne-Yanoff, T. & Lehtinen, A. (2012). Philosophy of game theory. In U. Mäki (Ed.) *Philosophy of economics* (pp. 531-576). Oxford: North Holland.
- Hands, D. W. (2011). Normative Rational Choice Theory. <http://papers.ssrn.com/sol3/papers.cfm?abstractid=1738671>. Accessed 23 Apr 2013.
- Hands, D. W. (2012). The positive/normative dichotomy and economics. In U. Mäki (Ed.) *Philosophy of economics* (pp. 219-240). Oxford: North Holland.
- Harsanyi, J. C. (1976). Advances in Understanding Rational Behavior. In his *Essays on ethics, social behavior, and scientific explanation* (pp. 89-117). Dordrecht: Reidel.
- Hausman, D. (2012). *Preference, value, choice, and welfare*. Cambridge: Cambridge University Press.
- Hempel, C.G. (1962). Rational action. *Proceedings and Addresses of the American Philosophical Association*, 35, 5-23.
- Lagueux, M. (1993). Popper and the rationality principle. *Philosophy of the Social Sciences*, 34, 468-480.
- Leach, J.J. (1977). The dual function of rationality. In: R.E. Butts & Hintikka, J. (Eds.) *Foundational problems in the special sciences* (pp. 393-421). Dordrecht: D. Reidel Company.
- Lehtinen, A. (2013). Three kinds of 'as if' claims. *Journal of Economic Methodology*, 20, 184-205.
- Lehtinen, A. & Kuorikoski, J. (2007). Unrealistic assumptions in rational choice theory. *Philosophy of the Social Sciences*, 37, 155-138.
- List, C. & Pettit, P. (2011). *Group agency*. Oxford: Oxford University Press.
- Mäki, U. (2000). Kinds of assumptions and their truth: shaking an untwisted F-twist. *KYKLOS*, 53, 317-336.
- Popper, K. (1967). The rationality principle. Reprinted in D. Miller (ed.,1985) *Popper selections* (pp. 357-365). Princeton: Princeton University Press.
- Sellars, W. (1956/1997). *Empiricism and the philosophy of mind* (ed. by R. Brandom). Cambridge, Mass.: Harvard University Press.
- Southwood, N. (2008). Vindicating the normativity of rationality. *Ethics*, 119, 9-30.
- Spohn, W. (2002). The many facets of the theory of rationality. *Croatian Journal of Philosophy*, 6, 249-264.