Economics and Philosophy, **16** (2000), 117–138 Copyright © Cambridge University Press

FOUNDATIONS OF PROCEDURAL RATIONALITY: COGNITIVE LIMITS AND DECISION PROCESSES

Frederic Laville
Ecole des Hautes Etudes en Sciences Sociales

1. INTRODUCTION

Many criticisms have been made of optimization theory (Laville, 1999a). These objections may be explained by the fact that human rationality is bounded – that decisions are constrained by cognitive limitations (Simon, 1982). In the present paper, I will show that if rationality is bounded, then we must study the processes of decision. My thesis is that *cognitive limitations lead to procedural rationality*. Although this assertion has already been sustained implicitly by Simon (1959) and explicitly by Mongin (1986), it has not been argued in sufficient depth.

In order to make this point, we shall need to ascertain optimization theory's methodological status. This issue has always been a subject of much dispute, and a number of questions are being debated. Does optimization theory describe decision processes or does it only characterize decision results? What part do cognitive limitations play in the determination of behavior? Can we predict behavior if we know only goals and constraints? Can optimization theory rationalize all types of behavior? Does it assume idealized agents? Must we question the traditional division of labor whereby psychology studies decision processes and economics characterizes the substance of decisions? All of these questions may be answered if, in accordance with my thesis, one regards optimization theory as a formal theory of decision rather than a

This paper incoroporates material from my dissertation at the Ecole des Hautes Etudes en Sciences Sociales. I would like to thank Daniel Andler, Bernard Conein, Olivier Favereau, Philippe Mongin, Bertrand Munier, Laurent Thévenot, Bernard Walliser, and an anonymous referee for detailed and useful comments on earlier drafts of this paper.

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

cognitive theory of deliberation. Indeed, I shall argue that optimization theory disregards both decision processes and cognitive limitations. Far from being independent, these two assumptions are deeply correlated; if we assume bounded rationality, then we must investigate procedural rationality.

This paper is organized as follows. Section 2 outlines the optimization controversy among realists and instrumentalists as to the methodological status that should be granted to optimization theory. Section 3 introduces a cognitive methodology for explaining behavior which acknowledges that behavior is complex and must therefore be examined at more than one level of explanation. This methodology will help us to assign to optimization theory a definite methodological status. I shall argue that optimization theory rests fundamentally upon being abstracted from the issues of both decision processes and cognitive limitations. Section 4 discusses the first assumption, and argues that optimization theory is a computational theory, that is, a theory which describes behavior independently of decision processes. Section 5 discusses the second assumption, and argues that optimization theory is an idealized theory, that is, a theory which assumes idealized agents. My thesis will then have been established: the assumption of cognitive limits (bounded rationality) requires the study of decision procedures (procedural rationality).

2. THE OPTIMIZATION CONTROVERSY

Optimization theory's methodological status is highly controversial. Two interpretations clash, resting upon different epistemological positions and leading to opposite assessments of the theory. The first interpretation, formulated by Friedman (1953), assigns to optimization theory an instrumentalist status; the second interpretation, advocated by Winter (1975), accords it a realist status. Friedman and Winter were not the first to raise this problem of interpretation. In the 1940s, the marginalist controversy had already opposed Lester (1946) to Machlup (1946) concerning the empirical investigations of Hall and Hitch (1939). Nevertheless, Friedman and Winter were the first to state the problems precisely and combine their arguments into systematic positions.

2.1. The instrumentalist defense

According to Friedman, optimization theory has one single aim, namely, the prediction of behavior. Hence, the only way to assess optimization theory is to compare its predictions with observed behavior. Optimization theory has only one purpose: to capture the substance of decisions. One does not require it to reproduce decision processes; it is needed only to predict decision results. Such is the meaning of Friedman's 'as if'

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

argument; people behave as if they maximized their utility. For this reason, optimization theory cannot be criticized on the grounds that it would require excessive cognitive abilities. Such a critique would only mean that we are granting optimization theory too large a validity domain. Optimization theory should not be considered a theory of decision processes. Since optimal behavior does not always result from an optimization procedure, one cannot criticize optimization theory for the excessive burden of computation it imposes. Although Machlup (1946) had already underscored this point, it was only with Friedman that the full consequences of this insight were worked out. To illustrate his position, Friedman compared the economic agent to a billiard player. Let us suppose that we want to predict a professional billiard player's behavior. The best strategy might be to compute his behavior as if – in order to determine a ball's trajectory – he were trying to solve the underlying mathematical problem. Obviously, such an assumption does not require the billiard player to solve mathematical equations; even so, it helps us to characterize the player's behavior and the ball's movements. To predict behavior, optimization theory need not reflect decision processes.

How can optimization theory's predictive success be explained? Even if optimal behavior does not always result from an optimization procedure, we must nonetheless indicate the nature of the underlying processes. A few years after Alchian (1950), Friedman postulated a selection process in answer to this question. Optimization theory is an adequate theory of behavior, because a process similar to natural selection is at work within the economy. Such is the nature of competition; a firm whose behavior deviates from optimization theory's predictions will be eliminated from the market. Even if people do not explicitly solve optimization problems, optimization theory offers a good account of the survival conditions imposed by economic competition. Since prediction is an economic theory's only purpose, and since optimization theory accurately predicts economic behavior, there is no point in conducting a more careful study of the analogy.

2.2. The realist critique

For Winter (1975), optimization theory is more than a theory of the substance of decisions. Optimal behavior is not enough; behavior must also result from an optimization process. The fact that people behave as if they maximized their utility does not warrant optimization theory. Optimization theory would be credible only if people were to explicitly compare their alternatives and systematically determine which one best satisfies their preferences. This constraint may, of course, sometimes be fulfilled. Some decisions do result from explicit optimization procedures;

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

TABLE 1. Two interpretations of optimization theory

Point	*Instrumentalism*	*Realism*
Status	Optimization theory is an instrumentalist theory; its only purpose is to predict behavior.	Optimization theory is a realist theory; it must not only predict behavior, but also describe it.
Goal	Optimization theory is a substance theory; the point is that people behave as if they maximized their utility.	Optimization theory is a process theory; it is not enough that people behave as if they maximized their utility.
Capabilities	There is no point criticizing optimization theory's cognitive burden; optimality does not always result from optimizing.	What counts is that optimization theory describes decision processes; its cognitive burden is crucial.
Grounds	Evolution supplements optimization; optimal behavior can be selected without any optimization procedure.	Optimization theory is auxiliary; an evolutionary process can generate suboptimal behavior.
Adequacy	Optimization theory has proved reliable; there is no use studying the underlying evolutionary process.	Evolution is the fundamental process; being an approximation, optimization theory is not worth focusing upon.

but to impose such a constraint would significantly reduce optimization theory's empirical range of validity.

Winter makes another criticism: even if optimization theory is not a theory of decision processes, one must justify its predictive success. A mere suggestion that optimization theory can be given an evolutionary grounding is not enough; it must be proven that optimal behavior actually arises from an evolutionary process. Such was the research program that Winter (1964) originally proposed. His goal was to determine optimization theory's validity domain within an evolutionary framework. He demonstrated that economic selection generates optimal behavior only under certain assumptions, whose validity could only be empirically demonstrated. Why then should one say that optimization theory is the fundamental principle? If evolution produces optimal behavior, why shouldn't we regard evolution as primary? Nelson and Winter (1982) argued in this way. For them, the optimality of behavior is only a theorem that may be demonstrated under certain assumptions. Table 1 summarizes the whole optimization controversy.

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

2.3. A first synthesis

What lessons may be learned from the optimization controversy? Despite their fundamental disagreement, the two interpretations converge on several points.

Status. Even though Friedman and Winter interpret optimization theory from two diametrically opposed perspectives, they have both set aside epistemological issues in order to consider the question from a theoretical viewpoint. Instead of reiterating the epistemological controversy between instrumentalism and realism, both have tried to focus the discussion on more specific issues, such as optimization theory's purpose, cognitive burden, ground, and validity range.

Goal. It is important that two questions be distinguished. The first question is general: should decision theory aim only at predicting behavior, or should it also describe underlying decision processes? Both goals must be pursued. We shall see that Friedman's position is unacceptable. Winter is right to stress each goal's importance. The second question is more specific: does optimization theory predict only the substance of decisions, or does it also describe cognitive processes? Both Friedman and Winter agree upon distinguishing *optimization* from *optimality*, that is, optimization as a *procedure* from optimization as a *result*; but while Friedman regards optimization theory as a theory of the substance of decisions, Winter considers it to be a decision process theory as well.

Capabilities. Friedman's 'as if' argument is the major strength of optimization theory's instrumentalist interpretation. From this argument it follows that an optimization procedure is not the only possible explanation for optimal behavior. Optimization theory's adequacy as a formal theory does not always result from its adequacy as a procedural theory. Hence, the cognitive critique of optimization theory is incomplete; no conclusion may be drawn from the fact that people have neither the computational nor the informational capacity to execute optimization procedures. Such a critique concerns only the procedural interpretation of optimization theory. Winter is right in criticizing the theory's unrealistic requirements, but Friedman is right in restricting this critique to procedural interpretation.

Grounds. If optimization theory is assigned two different goals, its success at predicting behavior may be explained by its accuracy as a theory of decision procedures; but if it is regarded only as a theory of the substance of decisions, its potential success will remain mysterious unless it can be explained by underlying cognitive or evolutionary processes. Friedman realized this, since he postulated a selection process to complement his methodological arguments. It remains to be determined, however, whether Winter is right in considering evolution to be a

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

theoretical alternative to optimization, or if Friedman is right in regarding evolution as an auxiliary hypothesis to optimization.

Adequacy. Even if we accept Friedman's 'as if' argument, and restrict optimization theory to just the prediction of the substance of behavior, Winter may still be right in arguing that optimization theory requires too much cognitive capacity. The link between optimal behavior and cognitive abilities is certainly more complex – everything depends upon the underlying cognitive processes – but it has not disappeared. As we shall see, optimization theory does assume idealized agents. This paper's main conclusion will be that postulating bounded rationality leads to the rejection of optimization theory in favor of procedural rationality.

3. LEVELS OF EXPLANATION

In this section I shall introduce a cognitive methodology that will transform the epistemological dispute between instrumentalism and realism into a methodological distinction between different levels of explanation. Borrowed from cognitive science, this methodology rests upon the following principle: *to understand cognitive systems, one must study them at different levels*.

3.1. A cognitive methodology

Marr (1977) distinguished two steps in the process of solving an information-processing problem. The first step aims at what he calls a *computational theory*, which is a formal theory of the problem and is independent of the cognitive processes used to solve it. The goal of the theory is to specify what function is computed and to explain why this computation does solve the information-processing problem. Answers are needed to both 'what' questions – what is the computed function? – and 'why' questions – why is it this function that is computed? After the first step has been completed, the second step aims at specifying how the function is implemented by the cognitive system. The goal, then, is to design algorithms capable of computing the previously determined function and to find a physical implementation for them. Now the relevant questions are not 'what' or 'why' questions, but rather 'how' questions – how is the function abstractly computed, and how is the algorithm physically implemented?

According to Marr (1982), three different *levels of explanation* underlie the preceding two steps: the computational level, the algorithmic level, and the physical level (Figure 1). They may be arranged according to the following hierarchy:

Computational level. The problem must first be studied. This involves three main tasks: the computed function must be characterized, the

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

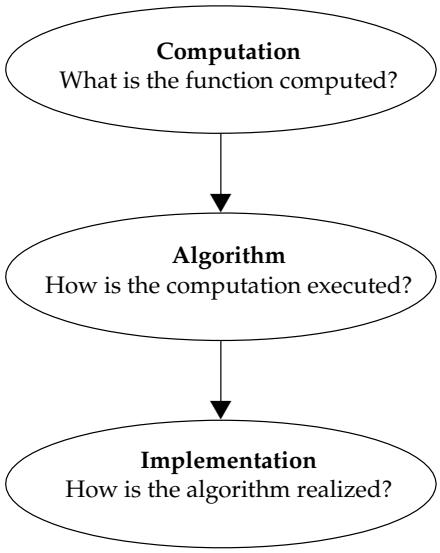


FIGURE 1. The three levels of explanation for a cognitive system (from Marr, 1982, p. 25). To study a cognitive system, three different levels are necessary: the computational level, the algorithmic level, and the physical level. Though each level is relatively independent from the others, the computational theory has logical priority. One must specify the computed function before treating the problems of algorithmic and physical implementation.

properties of the function must be determined, and the function must be shown to solve the problem. A computational analysis disregards the underlying algorithm. The task is to study the problem formally by specifying the system's goal and the environmental constraints involved.

Algorithmic level. After the computed function has been specified, the underlying algorithm must be determined. The task is to specify the representations the system uses and the transformations they go through. What counts are the formal mechanisms which execute the computation.

Physical level. Once the algorithm is specified, its material implementation must be determined. The task is to specify the physical mechanism that instantiates the abstract algorithm. What counts is not the formal representation, but the material device that implements it.

What are the relations between the different levels of explanation? Three points must be emphasized. First, the levels are relatively independent; the same function can be computed by various formal

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

algorithms, and a given algorithm can be implemented through various physical devices. Secondly, computational analysis possesses a logical priority. On the one hand, the computed function must be specified before the underlying algorithm is (how could a function not yet specified be implemented?). On the other hand, the computed function does not depend upon the cognitive system, but upon the problem task (a computational analysis may be seen as a formal description of the problem task). Thirdly, although the computational level possesses a logical priority, the different levels are complementary. For a cognitive theory to be complete, we must know what function is computed, how the computation is formally implemented, and how the algorithm is physically realized.

3.2. The example of addition

To illustrate this methodology, I shall reproduce Marr's example of the cash register (Marr, 1982, pp. 22–4). Since a cash register may be regarded as a cognitive system, it must be studied at three different levels.

At the computational level, two questions must be answered: what function does a cash register compute and why does it compute this function? The first question is easy; a cash register is designed to do arithmetic – in particular, addition. The computational theory is nothing but the theory of addition. Addition is a function, designated by '+', which associates to each pair of numbers a single number, their sum. What are the properties of this function? It is commutative, associative, has a unique element with no effect, and attributes to each number a unique inverse. Note that the computational theory of a cash register starts with the theory of addition, and has no relation to the representations that people use or to the algorithm by which addition is carried out. Consider now the second question: why does a cash register add the prices of the various purchases? This time, the problem is not to identify the computed function, but to explain why it is precisely this function which is appropriate. For a cash register, everybody knows the answer: the properties of addition correspond to the logical constraints of exchange. Table 2 compares the rules of addition to the principles of exchange.

Once we know the computed function, we must find a formal algorithm and a physical device that can implement it. Two problems must be solved: what algorithm can implement the function of addition, and how this algorithm can be physically realized. To obtain an algorithmic theory of addition, one must choose a representation for the arguments of addition and a procedure that can carry out the corresponding transformation. In the case of a cash register, the classical algorithm resembles the one applied when performing addition: the

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

Table 2. Why a cash register adds (from Marr, 1982, pp. 22–3)

Properties of addition	*Rules of exchange*
(1) There is a unique element, zero, the adding of which has no effect: $(4 + 0) = 0$.	(1) Buying something and nothing will cost you as much as buying only something.
(2) Addition is commutative: $(3 + 4) = (4 + 3)$.	(2) The presentation order of your purchases to the cashier will not affect the total.
(3) Addition is associative: $3 + (4 + 5) = (3 + 4) + 5$.	(3) Making two piles with the goods and paying each separately will not affect the total.
(4) Each number has an inverse that gives zero when added: $4 + (-4) = 0$.	(4) If you buy something and then return it for a refund, your total expenditure will be zero.

algorithm works on Arabic numerals, starts from the right, and carries if a partial sum exceeds nine. The problem of physical realization can be solved in different ways. Just as various algorithms may implement the same function, various technologies may implement the same algorithm. A child who carefully performs an addition uses an algorithm similar to the one used by a cash register, be it mechanical or electronic.

3.3. Competence and performance

A computational theory is similar to a theory of competence. Recall that Chomsky (1965) distinguished between two kinds of linguistic theories – theories of competence and theories of performance. While a theory of *competence* describes the abstract knowledge speakers have of their language, a theory of *peformance* describes the situated use speakers make of this knowledge. Competence theories differ from performance theories for they characterize ideal knowledge, that is, knowledge defined in abstraction from the contingencies that pervade linguistic practice.

Generative grammars are the paradigm of what Chomsky is aiming at when he speaks of competence theory. These grammars share the two main properties of competence theories: they have logical priority, and they are independent from cognitive processes. Consider the first property: Chomsky has always urged the priority of competence theories. No theory of linguistic performance is conceivable so long as linguistic competence has not been specified. In the same manner, it is impossible to study the process of language acquisition if one ignores its final state. From a methodological perspective, language training and language use are subordinate to language knowledge. What comes first is competence theory.

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

Consider now the second property: competence theories may be conceived independently of the cognitive processes underlying linguistic performance. For Chomsky, generative grammars do not describe how speakers generate sentences: they characterize the abstract linguistic knowledge speakers have of their language. The way in which such knowledge is formally represented – in terms of rules – says nothing about the cognitive mechanisms according to which people represent and manipulate this knowledge. Distinguishing between competence and performance prevents us from being led into harmful confusions. However, the formulation of generative grammars in terms of rules has led many people astray, and the discussion became very confused when Chomsky (1980) changed his position. Indeed, it now seems that Chomsky conceives of generative grammars not only as competence theories but also as performance theories. Grammatical rules, although inaccessible because they are unconscious, would on this account be explicitly represented in the minds of speakers.

This is the first divergence between Chomsky and Marr; but their methodologies differ on a second point as well. Whereas Marr presents his distinction between computation and algorithm in terms of implementation, Chomsky's distinction involves more than a mere hierarchy of levels. Contrary to performance theories, competence theories assume that people possess *idealized capabilities*. Competence theories disregard all factors that could cause a disparity between the competence of a speaker and his actual performance. Chomsky enumerates a series of assumptions corresponding to such an idealization: a competence theory involves an ideal speaker who belongs to a perfectly homogeneous linguistic community, knows his language perfectly, and is free from all the contingencies which are relative to linguistic practice but independent from grammar.

4. FIRST ASSUMPTION: THE PROCESSES OF DECISION

My thesis is that optimization theory should be conceived as a theory of decision competence. I shall therefore argue that optimization theory may be abstracted from questions of both decision processes and cognitive limitations. In this section I shall discuss the first assumption; in the next section I shall consider the second.

4.1. A computational theory

According to the first assumption, optimization theory must not be interpreted as a procedural theory of decision. Indeed, optimization theory does not require that people solve optimization problems each time they make decisions. This does not, however, prevent optimization

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

Table 3. Why people maximize expected utility (from Friedman and Savage 1952, pp. 468–9)

Postulates	*Constraints*
(P_1) *Completeness* $f \leq g$ or $g \leq f$.*	Facing two alternatives, we must be able to say if we prefer the first to the second, the second to the first, or consider them as equivalent.
(P_2) *Transitivity* If $f \leq g$ and $g \leq h$, then $f \leq h$.	If we prefer an alternative g to an alternative f, and an alternative h to this same alternative g, then we must prefer h to f.
(P_3) *Continuity* If $af + (1 - a)h \leq g$ for all a such as $0 \leq a < 1$, then $f \leq g$.	If we prefer an alternative g to a combination of f with a probability a and h with a probability $(1-a)$ whatever the value of a, then we must prefer g to f.
(P_4) *Independence* For all a such as $0 < a < 1$, $af + (1 - a)h \leq ag + (1 - a)h$ if and only if $f \leq g$.	If we prefer an alternative g to an alternative f if a contingency happens, and still prefer g to f if the contingency does not happen, we must prefer g to f (and reciprocally).

*The following axioms constrain probability distributions defined on the same opportunity set. The following notations are used: (1) X is the opportunity set; (2) F is the probability distribution on X; (3) $\leq$ is the relation 'is not preferred to' defined on F. Axioms are formulated for all f, g, h belonging to F.

theory from having a definite purpose. Optimization theory should be regarded as a computational theory – a theory characterizing optimal behavior independently of decision-making processes. Such is the actual significance of the 'as if' argument, first proposed by Friedman and Savage (1948). Urging that optimization theory is a competence theory means that, if one describes behavior as the solution of an optimization problem ($\underset{x \in X}{\text{Max}}\ u(x)$ under constraint $f(x) \leq a$), optimization theory will predict the corresponding behavioral function ($\underset{x}{\arg}$ [Max $u(x)$ under constraint $f(x) \leq a$]).

Why would people maximize their expected utility? Optimization theory answers the 'what' question, but there remains the question of why maximizing expected utility is the appropriate computation. Here is the alleged role of axiomatization: the expected utility criterion can be inferred from a set of intuitively valid principles. By showing that intuitive axioms guarantee optimal behavior, axiomatizations answer the 'why' question. Table 3 illustrates this point for the axiomatic treatment of von Neumann's and Morgenstern's (1944) notion of utility. Such a treatment is powerful because it legitimates the use of mathematical optimization:

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

If preferences verify (P_1), (P_2), (P_3), and (P_4), there exists a utility function u defined on X such that $f \leq g$ if and only if $\sum_{x_i \in X} f(x_i)u(x_i) \leq \sum_{x_i \in X} g(x_i)u(x_i)$.

By their conjunction, the mathematics of optimization and the axiomatization of utility together form a computational theory of decision: on the one hand, the criterion of expected utility answers the 'what' question by describing what function is computed; on the other hand, the axiomatization of expected utility answers the 'why' question by showing what logical constraints account for the computed function.

4.2. Towards unification?

As soon as we begin to consider optimization theory a computational theory, deep similarities become apparent between decision theory and related disciplines, such as game theory, evolutionary biology, philosophy of mind, artificial intelligence and cognitive psychology. What is remarkable is that all of these disciplines distinguish computational theories from algorithmic theories. Such a movement toward the theoretical unity of the behavioral sciences lends support to the methodological status I have assigned to optimization theory.

Game theory. During the past decade, game theory has been undergoing a deep transformation. Until recently, it was based upon two computational theories: optimization theory answered the question of what the rational behavior for a player is; and Nash equilibrium answered the question of what the rational outcomes of a game are. Lately, these 'what' questions have been supplemented by 'how' questions: how to formalize the players' decision processes; and how to formalize the game's equilibration dynamics. The positions of Binmore (1986) and Selten (1991) both illustrate this transformation. They each emphasize the need to supplement optimization theory and Nash equilibrium by theories of decision process and equilibration dynamics. In game theory, optimization theory and Nash equilibrium are regarded as computational theories in need of algorithmic theories.

Evolutionary biology. Compared to economics, biology embraced optimization theory only recently. But, from the start, biology has interpreted optimization as a computational theory. What matters is that optimization theory permits evolutionary theory to characterize the results of natural selection in complete abstraction from the underlying neural structures. Let us take an example: when Maynard Smith (1982) uses dynamic programming to formalize animal behavior, he does not attribute to animals the ability to solve optimization problems; he postulates only that some underlying mechanisms will support the most effective behavior within a given population. With optimization theory

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

and game theory, the theory of evolution gained the computational theory that was lacking (Maynard Smith 1978).

Philosophy of mind. In this field too, a deep controversy has opposed realists and instrumentalists as to the nature of rational explanation. While Fodor (1975) defended a realist interpretation of decision theory, Dennett (1987) supported a position close to Friedman's instrumentalism. For Dennett, intentional psychology – the underlying conceptual structure of optimization theory – is to be considered a pure theory of competence, for it disregards the underlying cognitive processes. This hypothesis explains both the predictive strength and the empirical weaknesses of intentional psychology. For this reason, Dennett proposed supplementing what he calls 'instrumental psychology' with a 'subpersonal psychology'. Here is the corresponding division of tasks: whereas instrumental psychology would have to give a formal explanation of behavior (at the computational level), subpersonal psychology would have to describe the underlying cognitive processes (at the algorithmic level).

Artificial intelligence. In artificial intelligence, logic is used mainly as a computational theory in order to characterize knowledge independently of its underlying representation. People regard logic as describing the 'task environment' (Newell and Simon, 1972), or as a theory of 'the knowledge level' (Newell, 1982). But this computational interpretation has not always prevailed. In the same way that optimization theory in economics was considered at first to be a true description of decision processes, logic has long been considered in artificial intelligence to be an algorithmic theory. It was supposed to describe human reasoning on the model of theorem-proving – conclusions resulting from a series of deductions in accordance with logical rules (Genesereth and Nilsson, 1987). As the fields are related, however, the methodological status of logic has gradually changed. Whereas logical formalism is still widely used to investigate abstract knowledge, few researchers follow Nilsson (1991) in interpreting logic to be a true description of human reasoning processes.

Cognitive psychology. A similar transformation took place in this field as well. Logical formalism was at first considered an algorithmic theory of human reasoning. People were supposed to be equipped with a 'mental logic', that is, a system of rules and representations describing their inference processes (Brain, 1978; Rips, 1983). Under the pressure of experience, however, the mental logic assumption has gradually lost its plausibility, and many psychologists nowadays treat logic as a partial theory of deductive competence (Johnson-Laird, 1983; Johnson-Laird and Byrne, 1991). The psychology of reasoning has passed, therefore, from a period where logic was considered both an algorithmic theory

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

and a computational theory to a period where logic is supposed only to describe deductive competence.

4.3. Reassessing criticisms?

Optimization theory has been subjected to many criticisms. From a computational perspective, such criticisms assume a completely different characterization. The following dilemma must be faced: either we restrict optimization theory to the computational level, or we abandon it in the face of so much criticism.

Let us first consider the *empirical critique*. Optimization theory has been sharply criticized for its falsified predictions. Kahneman and Tversky (1986) have described many laboratory experiments where a violation of optimization theory's most basic principles produced incoherent choices. For many people, such experiments discredited expected utility models. This first objection is fundamental, for it concerns not only the algorithmic interpretation of optimization theory, but also its computational interpretation. Yet, it is important to distinguish two ways of interpreting the empirical critique. The first is that, if the observed violations do not contradict the hypothetically computed function, then the empirical biases concern only the algorithmic level. By determining the limits of performance, empirical work provides the psychologist with valuable information about the underlying algorithm. Biases must then be conceived as empirical constraints that algorithmic theories must respect, but they do not falsify optimization theory so long as it is conceived to be a computational theory. But there is a second interpretation: if the observed violations are large enough to call into question the – hypothetically – computed function, then optimization theory has been falsified at the computational level itself. Depending upon their amplitudes, empirical biases involve divergent consequences. This is why the empirical violations of the expected utility criterion, which Kahneman and Tversky described in their works, do not always require economists to reject optimization theory.

Next, we turn to the *cognitive critique*, which complains that maximizing utility requires unrealistic capabilities. Optimization theory is a poor theory of human behavior, for it imposes upon the agent an excessive computational burden. In retrospect, it appears that Friedman and Savage (1948) had already answered this critique. Optimization theory must be regarded as a theory of competence rather than as a theory of performance. When economists use optimization theory, they do not assume that people solve the corresponding optimization problems. Using a computational theory, one cannot directly assess the cognitive abilities required by the underlying algorithm. Hence, one cannot criticize optimization theory for the computational burden that it

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

Table 4. Critiques of optimization and levels of explanation

Critique	*Description*	*Evaluation*
Empirical	Experimentally observed behavior falsifies the theoretical predictions derived from optimization theory.	Optimization theory describes decision competence. Empirical biases may only concern performance.
Cognitive	People have neither the informational nor the computational capacities required by optimization theory.	Optimization theory does not describe decision performance; it describes only decision competence.
Logical	Introducing computational costs into optimization programs leads to an infinite regress of decisions.	If optimization theory is regarded as a competence theory, it should not include computational costs.

seems to involve; it could only be what I shall call a *level error*. In a sense, the cognitive critique is akin to the preceding one, for it has often been asserted that the empirical heuristics postulated by Tversky and Kahneman (1974) contradicted the course of an optimization procedure, whereas this reproach could only be addressed to an algorithmic theory of behavior.

The last objection I shall consider is a *logical critique*. Maximizing utility requires time, effort, a search for information, and the use of sophisticated tools for computation. Let us assume that all of these factors may be summarized by a cost that will be incorporated into optimization programs. Stigler (1962) initiated this strategy when he postulated costs for informational search. A decision will be optimal only if it reflects this new cost. But the argument is not yet complete: solving this optimization program will itself generate new computational costs. Thus, there is no way of avoiding an *infinite regress*. This complaint, first mentioned by Winter (1964), was put forward in terms of infinite regression by Mongin (1986), and then formalized by Mongin and Walliser (1988). For them, such an infinite regression is a decisive argument against optimization theory. The problem with their objection, however, is that the regression argument assumes an algorithmic interpretation of optimization theory. Insofar as optimization theory is regarded as a computational theory, the logical critique loses its force. Indeed, cost models rest upon an ambiguity; they are computational models into which a cost variable, assumed to capture the computational difficulty of the decision program, is grafted, but the matter of computational difficulty appears only at the algorithmic level. Table 4 summarizes the whole discussion.

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

5. SECOND ASSUMPTION: THE LIMITS OF COGNITION

Optimization theory rests upon two fundamental assumptions. I have already suggested that it is independent from *decision processes*. I shall now argue that it is likewise independent from *cognitive limitations*.

5.1. An idealized theory

As a theory of competence, optimization theory assumes *idealized agents*. On this assumption, the computed function coincides with the problem's solution. Underscoring this second assumption does not mean falling into a level error. Optimization theory is not criticized because it requires people to solve difficult optimization problems. We know already that optimal behavior does not always result from explicit optimization. Nevertheless, economic agents must possess sufficient capacity to execute an algorithm leading to optimal behavior. Given their cognitive abilities, it is perfectly possible that agents could not handle *any* algorithm which would lead to optimal behavior. Therefore, optimization theory predicts optimal behavior only under an assumption of ideal agents – that is, under an assumption of perfect abilities. When Savage (1972) recognized this new assumption – in the second edition of *The Foundations of Statistics* – he sharply deviated from the position of Friedman (1953). The 'as if' argument no longer suffices to bridge the gap between real behavior and theoretical predictions: optimization theory formally describes behavior; but not any kind of behavior, only the behavior of an idealized agent. On this question, the evolution of Savage's thought is quite significant; after he encountered the Allais paradox (Allais, 1953), Savage qualified his theory of decision by a new assumption – that of the ideal agent.

I can now state my thesis in full. Like any competence theory, optimization theory rests upon two fundamental assumptions:

Decision processes. Optimization theory disregards decision processes. It provides a formal analysis of the decision problem and specifies the computed function independently of the underlying algorithm. As Friedman (1953) puts it, agents behave as if they maximized their utility.

Cognitive limitations. Optimization theory disregards cognitive limitations. In order to analyze a decision problem, it assumes perfect cognitive capabilities. As Savage (1972) puts it, only idealized agents can be sure to behave as if they maximized their utility.

These two assumptions rest upon different 'as if' arguments: in the first assumption, 'as if' means that people behave as if their decision processes were identical to algorithmic optimization; in the second

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

Table 5. The four possible goals of optimization theory

Goal	*Description*
Descriptive	Optimization theory describes the cognitive processes underlying behavior. Decisions result from a systematic comparison of alternatives.
Predictive	Optimization theory predicts only the substance of behavior. It remains valid as long as observed behavior is optimal.
Normative	Optimization theory states the best solution for a given decision problem. It summarizes the constraints that one must respect in order to be rational.
Postdictive	Optimization theory can rationalize almost any behavior, insofar as the theorist is prepared to modify costs, objectives, and constraints whenever necessary.

assumption, 'as if' means that people behave as if cognitive limitations never constrained behavior.

The preceding interpretation does not empty optimization theory of its content. Let me clarify this point. Three different goals are commonly assigned to optimization theory: descriptive, predictive, and normative. But there is a fourth possible goal: optimization theory may also be given a *postdictive* interpretation (Schoemaker, 1982). Table 5 summarizes optimization theory's four possible goals.

The postdictive interpretation trivializes optimization theory. Given sufficient degrees of freedom, virtually any behavior can be described as optimal (Schoemaker, 1991). This interpretation underlies the economic approach to human behavior advocated by Becker (1976). The idea is that any behavior may be *rationalized* after the fact by formulating anew the postulated decision problem. The postdictive and the competence interpretations, although frequently confused, are in fact radically opposed to one another. The confusion seems to result from the fact that the same criticism has been made of both. Both have indeed been criticized because they – allegedly – protect optimization theory from empirical falsification. But make no mistake: the competence interpretation has different consequences from the postdictive interpretation. In the first case, optimization theory describes a definite behavior; in the second case, no behavior can escape the net of optimization theory. While the competence interpretation objectifies optimization theory, the postdictive interpretation empties it of all its content.

5.2. The limits of optimization theory

The force of optimization theory resides in the two preceding assumptions. If it describes behavior with such remarkable parsimony, this is

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

only because decision processes and cognitive limitations have been abstracted away. To predict behavior, we need only know the agent's goal and environment. It is this economy which explains the unique status of optimization theory (Popper, 1967; Latsis, 1972). By setting aside cognitive limitations, it guarantees the theoretical independence of economics from psychology. This issue was the real focus of the optimization controversy; whereas Friedman and Savage supported the conceptual autonomy of economics from psychology, Winter and Simon advocated the need for psychological foundations of economics.

Weber (1949) was the first to understand why the rationalistic method – which used to be in his time the counterpart of optimization theory for sociology – is such a powerful theory of action. Indeed, he proposed an interpretation for the concept of instrumentally rational action that shares many characteristics with the competence interpretation I have been arguing for. According to Weber, optimization theory is what he called an *ideal type*. The concept of ideal type contains, more or less explicitly, the two assumptions I have been focusing upon: rational explanation is independent from psychic mechanisms (psychological independence), and rational explanation constitutes a logical standard which helps the social scientist to reveal potentially irrational behavior (idealized agents). Besides, Weber was quite aware of the fundamental problem involved in the optimization controversy; conceiving of rational explanation as an ideal type helps the social scientist to guarantee the independence of sociology from psychology.

I have shown why optimization theory is such a powerful theory of behavior. But the preceding discussion could be misleading, inasmuch as the two assumptions underlying optimization theory explain not only its strengths, but also its weaknesses. Both are linked, for if one gives up the ideal agent assumption, one can no longer be sure that behavior will always solve the corresponding decision problem. While, under the assumption of perfect rationality, it was enough to know just the goal and the environment of an agent to predict his behavior, under an assumption of bounded rationality, the behavior of an agent will be jointly determined by his goal, his environment, *and his capabilities*. Such is the whole significance of the bounded rationality hypothesis: rationality is bounded insofar as cognitive limits shape decisions and constrain behavior.

5.3. The concept of procedural rationality

Though I have distinguished between the two assumptions, one must still appreciate that they are linked to one another. If the second assumption – as to cognitive limitations – is removed, the first – as to decision processes – no longer holds. It is impossible to assume that agents have cognitive limitations without specifying their decision processes.

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

Insofar as the perfect rationality assumption applies, optimization theory suffices to predict decisions. Algorithmic theories describe only the computations according to which people generate optimal behavior. The construction of algorithmic theories falls outside the range of economic theory, and a division of labor prevails, whereby economists predict behavior and psychologists describe cognitive processes. But the assumption of perfect rationality is inadequate; this is the main conclusion that can be drawn from an examination of the empirical, methodological, and theoretical criticisms addressed to optimization theory (see Simon, 1982; Mongin, 1984; Conlisk, 1996; Laville, 1999a). Since the cognitive capacities of human beings are bounded, optimization theory is unable to accurately predict human behavior. Algorithmic theories not only specify the underlying decision processes, they also contribute to the prediction of behavior. Under the assumption of bounded rationality, the decision processes are manifested in behavior. Such is this paper's main thesis: *cognitive limitations demand procedural rationality*.

In economics, it was Simon (1976) who introduced the concept of procedural rationality. Although this concept underlies all his work, only lately has Simon explicitly formulated the two underlying conceptions of rationality he had assumed from the beginning. Here are the two definitions he proposed:

Substantive rationality. The property of rationality characterizes the substance of decision. Rational behavior is optimal behavior, that is, behavior 'appropriate to the achievement of given goals' (Simon, 1976, p. 130).

Procedural rationality. The property of rationality characterizes the processes of decision. Rational behavior is deliberative behavior, that is, 'the outcome of appropriate deliberation' (Simon, 1976, p. 131).

Simon's definitions are quite general. He seems to assimilate substantive rationality to teleology, and procedural rationality to deliberation. But in the remainder of his paper the definitions are tightened: he assimilates substantive rationality to the maximization of expected utility as conceived in economics, and procedural rationality to information processing theories as developed in psychology. For Simon, then, substantive rationality reduces to optimization theory, and the study of procedural rationality reduces to the study of human cognitive processes.

Contrary to Marr, Simon does not treat the distinction between substantive and procedural rationality as a distinction between levels of explanation. Moreover, he rejects the idea that a theory of the problem task might accurately describe people's behavior at the computational level. This divergence between Marr and Simon may be explained by the

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

kinds of tasks that Simon has studied; his interest in problem solving – in chess, logic, and cryptarithmetic puzzles – explains why he cannot assume that people always conform to the demands of the task environment. In chess, for example, a computational theory of the problem will hardly help to specify the players' decision processes. Simon (1978) agreed with Marr on the following point: the concepts of substantive and procedural rationality, like the levels of computation and algorithm, answer different questions. By predicting decisions, theories of substantive rationality answer 'what' questions, while theories of procedural rationality, by describing cognitive processes, answer 'how' questions. The distinction between substantive and procedural rationality is closer to the distinction between competence and performance theories: each distinction underscores the behavioral impact of cognitive limitations.

6. CONCLUSION

This paper has aimed at clarifying the relations between optimization theory, cognitive limitations and procedural rationality. Two conclusions have been reached. The first is that *optimization theory accounts for decision competence*. Optimization theory rests upon two fundamental assumptions: it is abstracted from both decision processes and cognitive limitations. This interpretation relies upon Friedman's 'as if' argument and Savage's ideal agent assumption and also draws heavily upon Marr's and Chomsky's methodologies, without which the methodological status of optimization theory could not have been clarified. The second conclusion is that *cognitive limitations lead to procedural rationality*. The ideal agent assumption is a prerequisite to the use of optimization theory to accurately predict human behavior. Under this assumption, theories of procedural rationality are superfluous. Insofar as rationality is bounded, however, and cognitive limitations become apparent, we need theories of procedural rationality.

What, then, should be our next task? Since cognitive limitations lead to procedural rationality, we should investigate procedural theories of decision. If we define procedural rationality as appropriate deliberation, two perspectives may be adopted: (1) a *cognitive perspective*, assimilating deliberation to the execution of requisite computation, and (2) an *argumentative perspective*, assimilating deliberation to the elaboration of adequate argumentation (Laville, 1999b). In this paper I have tacitly privileged the first perspective – urging a cognitive methodology for explaining behavior. If we accept the cognitive perspective, then an inquiry concerning procedural rationality may be conducted from either of two different standpoints: (a) that of *bounded rationality*, whereby procedural rationality may be studied by stressing the cognitive limita-

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

tions that prevent people from adopting optimal behavior (Laville 1998a), or (b) that of *situated cognition*, from which procedural rationality may be studied by focusing upon the cognitive resources – human or artificial – that shape the environment (Laville 1998b).

REFERENCES

Alchian, A. 1950. Uncertainty, evolution, and economic theory. *Journal of Political Economy*, 58:211–21

Allais, M. 1953. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*, 21:503–46

Becker, G. S. 1976. *The Economic Approach to Human Behavior*. University of Chicago Press

Binmore, K. G. 1988. Modeling rational players: Part II. *Economics and Philosophy*, 4:9–55

Braine, M. D. S. 1978. On the relation between the natural logic of reasoning and standard logic. *Psychological Review*, 85:1–21

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. MIT Press

Chomsky, N. 1980. *Rules and Representations*. Columbia University Press

Conlisk, J. 1996. Why bounded rationality? *Journal of Economic Literature*, 34:669–700

Dennett, D. C. 1987. *The Intentional Stance*. MIT Press

Fodor, J. A. 1975. *The Language of Thought*. Crowell

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

Friedman, M. and L. J. Savage. 1948. The utility analysis of choice involving risk. *Journal of Political Economy*, 56:279–304

Friedman, M. and L. J. Savage. 1952. The expected-utility hypothesis and the measurability of utility. *Journal of Political Economy*, 60:463–74

Genesereth, M. R. and N. J. Nilsson. 1987. *Logical Foundations of Artificial Intelligence*. Morgan Kaufmann

Hall, R. L. and C. J. Hitch. 1939. Price theory and business behavior. *Oxford Economic Papers*, 2:12–45

Johnson-Laird, P. N. 1983. *Mental Model*. Cambridge University Press

Johnson-Laird, P. N. and R. M. J. Byrne 1991. *Deduction*. Erlbaum

Latsis, S. J. 1972. Situational determinism in economics. *British Journal for the Philosophy of Science*, 23:207–45

Laville, F. 1998a. Modélisations de la rationalité limitée: De quels outils dispose-t-on? *Revue économique*, 49:335–65

Laville, F. 1998b. La cognition située: une nouvelle approche de la rationalité limitée? Forthcoming in *Revue économique*

Laville, F. 1999a. Should we abandon optimization theory? The need for bounded rationality. Working Paper: Ecole des Hautes Etudes en Sciences Sociales

Laville, F. 1999b. Les voies de la rationalité procédurale. Working Paper: Ecole des Hautes Etudes en Sciences Sociales

Lester, A. 1946. Shortcomings of marginal analysis for wage-employment problems. *American Economic Review*, 36:63–82

Machlup, F. 1946. Marginal analysis and empirical research. *American Economic Review*, 36:519–54

Marr, D. 1977. Artificial intelligence: a personal view. *Artificial Intelligence*, 9:37–48

Marr, D. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Freeman

Maynard Smith, J. 1978. Optimization theory in evolution. *Annual Review of Ecology and Systematics*, 9:31–56

Maynard Smith, J. 1982. *Evolution and the Theory of Games*. Cambridge University Press

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press

Mongin, P. 1984. Modèle rationnel ou modèle économique de la rationalité? *Revue économique*, 35:9–64
Mongin, P. 1986. Simon, Stigler et les théories de la rationalité limitée. *Information sur les Sciences Sociales*, 25:555–606
Mongin, P. and B. Walliser. 1988. Infinite regression in the optimizing theory of decision. In *Risk, Decision and Rationality*. B. Munier (ed.). pp. 435–57. Reidel
Nelson, R. R. and S. G. Winter 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press
Newell, A. 1982. The knowledge level. *Artificial Intelligence*, 18:81–132
Newell, A. and H. A. Simon. 1972. *Human Problem Solving*. Prentice-Hall
Nilsson, N. J. 1991. Logic and artificial intelligence. *Artificial Intelligence*, 47:31–56
Popper, K. R. 1967. La rationalité et le statut du principe de rationalité. In *Les fondements philosophiques des systèmes économiques*. E. M. Claassen (ed.). pp. 142–50. Payot
Rips, L. J. 1983. Cognitive processes in propositional reasoning. *Psychological Review*, 90:38–71
Savage, L. J. 1954. *The Foundations of Statistics*. Wiley
Schoemaker, P. J. H. 1982. The expected utility model: its variants, purposes, evidence and limitations. *Journal of Economic Literature*, 20:529–63
Schoemaker, P. J. H. 1991. The quest for optimality: a positive heuristic of science? *Behavioral and Brain Sciences*, 14:205–45
Selten, R. 1991. Evolution, learning, and economic behavior. *Games and Economic Behavior*, 3:3–24
Simon, H. A. 1959. Theories of decision-making in economics and behavioral science. *American Economic Review*, 49:249–83
Simon, H. A. 1976. From substantive to procedural rationality. In *Method and Appraisal in Economics*. S. J. Latsis (ed.). pp. 129–48. Cambridge University Press
Simon, H. A. 1978. On how to decide what to do. *Bell Journal of Economics*, 9:494–507
Simon, H. A. 1982. *Models of Bounded Rationality*. MIT Press
Stigler, G. J. 1961. The economics of information. *Journal of Political Economy*, 69:213–25
Tversky, A. and D. Kahneman. 1974. Judgment under uncertainty: heuristics and biases. *Science*, 185:1124–31
Tversky, A. and D. Kahneman. 1986. Rational choice and the framing of decisions. *Journal of Business*, 54:251–78
Von Neumann, J. and O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton University Press
Weber, M. 1949. *The Methodology of the Social Sciences*. Trans. E. A. Shils and H. A. Finch. Free Press
Winter, S. G. 1964. Economic natural selection and the theory of the firm. *Yale Economic Essays*, 4:225–72
Winter, S. G. 1975. Optimization and evolution in the theory of the firm. In *Adaptive Economic Models*. R. H. Day and T. Groves (eds.). pp. 73–118. Academic Press

https://doi.org/10.1017/S026626710000016X Published online by Cambridge University Press