# *Cognitive Science and Liberal Contractualism: A Good Friendship*[1]

Óscar L. González-Castán

Recibido: 1 de abril de 2005
Aceptado: 10 de mayo de 2005

**Resumen**

En este artículo argumento que tanto el cognitivismo como el contractualismo liberal defienden una concepción pre-moral del deseo humano que tiene su origen en la tradición hobbesiana y humeana que ambas teorías comparten. Además, la interpretación computacional y sintáctica de la ciencia cognitiva defiende la idea, que Gauthier evidentemente comparte, de que la mente humana –o, en el caso de Gauthier, la mente del "hombre económico"–, es un mecanismo puramente formal que se caracteriza por realizar operaciones lógicas y matemáticas. Concluyo que una concepción unitaria sobre el comportamiento humano subyace a las diversas teorías psicológicas, morales y políticas que han sido dominantes en la tradición analítica.

*Palabras clave:* Ciencias cognitivas, contractualismo liberal, ética, deseo humano.

**Abstract**

In this paper, I shall argue that both cognitivism and liberal contractualism defend a pre-moral conception of human desire that has its origin in the Hobbesian and Humean tradition that both theories share. Moreover, the computational and syntactic themes in cognitive science support the notion, which Gauthier evidently

[1] A previous version of this paper was written during the tenure of a Research Scholarship of the DGICYT (Spanish Ministry of Education) number PB 93-0683.

I should like to thank Prof. John Hyman for comments on earlier drafts of this paper.

shares, that the human mind – or, in Gauthier's case, the mind of "economic man" – is a purely formal mechanism, characterized by logical and mathematical operations. I shall conclude that a single conception of human behaviour runs through the various dominant psychological, moral and political theories of analytic inspiration.

*Keywords*: Cognitive Science, Liberal Contractualism, Ethics, Human Desire.

## Introduction

In this paper, I shall try to show how cognitive science, as it is conceived by some contemporary theories of mind, can be used to support the main psychological presuppositions that underlie Gauthier's liberal contractualism. I shall also explore the historical transformations in the philosophy of science that made it possible for Gauthier's political theory and these theories of mind to have many important psychological insights in common. I shall conclude that a single conception of human behaviour runs through the various dominant psychological, moral and political theories of analytic inspiration.

I should say from the start that I am not sympathetic with either the view of cognitive science or the view of contractualism about which I talk here. However, my main concern is not to criticize these views, which I have done in other papers [González-Castán (1998), (1997), (1996), (1992)], but to argue that they have provided a quite specific and cogent theory on human behavior that have greatly influenced our self-conception.

## I. A Very Brief Historical Survey

Since the beginning of philosophical inquiry, psychology has played a central role in the overall scheme of reality. In particular, psychological ideas have been important because of their ethical and political implications. Authors with quite different philosophical agendas – such as Plato, Aristotle, Descartes, Hobbes and Hume – have all emphasized these implications. To support this assertion, I shall remember some historical landmarks.

Bernard Williams has pointed out that, for Plato and Aristotle, one of the principal tasks of philosophy is to give an account of the human soul in which their conception of a moral life could find an indispensable place. These Greek philosophers believed that the human soul must be described or analyzed in such a way that "if people properly understood what they were, they would see that a life of justice was a good not external to the self but, rather, an objective that it must be rational to pur-

sue" [Williams (1985), pp. 33-34]. Accordingly, psychology is, in a certain sense, subservient to ethics and politics. Our conception of the soul, of its different parts and their relationships to each other, is constrained by the requirement that it must allow us to deem irrational, and undesirable, a human life that does not embody the proper harmony of these parts. If a person fails to respect the hierarchical order between the parts of his soul, he will be unable to achieve the state of *eudaimonía*.

As it is well-known, this classical strategy contrasts dramatically with contractualist theories of morality; for they just proceed in exactly the opposite direction. Instead of trying to devise a psychological theory which guarantees a preeminent place for the ethical life, their aim, as Williams puts it is, "given an account of the self and its satisfactions, to show how the ethical life (luckily) fits them" [Williams (1985), p. 32]. Hobbes and Gauthier, among others, pursue this strategy in moral theory. They believe that the satisfaction of the interests and preferences of the agent has a certain kind of priority over the morality of his acts. Morality has a place in people's lives because, given a certain conception of the psychology of human beings, and of their interactions in a hypothetical situation before the creation of society, to act morally is to act rationally, that is, in the agent's own interest. Thus, the aim of contractualism is to defend the doctrine that being moral is rational and advantageous for the agent.

The ethical significance of psychological theories is also evident in some other philosophies, such as Descartes' and Hume's. Thus, Descartes´ attempt to prove the immateriality of the rational soul – not its immortality – was partly motivated by the desire to prove that certain Christian doctrines are reasonable. Hume, for his part, criticized Cartesian dualism on the grounds that it is incapable of providing a non-religious foundation for morality.

All of these direct relationships between psychology and practical thinking disappeared gradually as psychology tried to establish an independent and respectable place among the natural sciences, and it became common to suppose that the evident and deliberate connections between psychology and ethics which I have mentioned, were explained by the fact that psychology had formed a part of philosophy, which was considered to be a speculative and non-scientific discipline. With this assertion I do not mean to imply that professional psychologists took no interest in the practical consequences of their theories. Many did. Skinner, for example, in his book *Beyond Freedom and Dignity*, went to considerable pains to explore the ethical consequences of behaviorism. His technology of conduct was designed to solve practical and moral problems by providing us with a new explanation of human behaviour that would, ideally, have the same predictive power as biological or physical theories. Freud's psychological theories also have immediate practical consequences. For if an agent is not always responsible for his behaviour, because there are unconscious determinants of his acts, then there will be cases in which we

should treat people clinically, rather than punish them, when they fail to act as they should.

Nevertheless, a psychologist who regards psychology as one of the natural sciences may feel inclined to object as follows. Certainly, Skinner made inferences from behaviorism, to support a rational conception of what human moral life should look like. However, his conclusions should not be regarded as part of psychology itself; but rather, as a speculative entertainment which draws on psychology; something a psychologist may engage in as a sort of intellectual hobby. And if psychoanalysis appears to have important consequences for human practice, this is due to the fact that psychoanalysis is not a theory that meets the standards of scientific inquiry – as Popper, among others, has argued. Popper argued that since psychoanalytical theory is consistent with any logically admissible description of human conduct, it is not falsifiable, and hence it is unscientific [Popper (1982), p. 46]. And one can also argue that psychoanalysis can make its incursions into the practical realm because it uses a vocabulary which is not strictly scientific, and which therefore leaves more room for imaginative speculation, unconstrained by the rigour of scientific reasoning. Psychoanalysis would not be, and has never been, as some psychoanalytical schools of Lacanian inspiration have pointed out, a scientific theory but rather a philosophical theory about the self.

Now, a basic intuition about the nature of science, many times attributed to Hume, underlies this sort of detachment of psychology from ethics. The intuition is that the propositions of natural sciences have no practical or normative content, and carry no implications concerning matters of value. Wittgenstein, for example, persuasively defended the radical separation of science and ethics, particularly in the *Tractatus*. According to Wittgenstein, the description and explanation of states of affairs, which is the business of science, has no ethical implications whatsoever. The description of a murder is no more or less significant than the description of a falling stone, for both alike are descriptions of natural facts. The first differs from the second, inasmuch as it involves psychological facts. But neither description has any ethical implications. Accordingly, Wittgenstein considers that ethics belongs outside the limits of our language, and our world, because sentences which have a sense can only refer to natural states of affairs, and our world is made exclusively out of these states of affairs. Thus, if psychology aspires to be a natural science, none of its propositions can have any ethical significance. And both cognitive science and contemporary analytic philosophy of mind have inherited this self-imposed limitation.

## II. Cognitive Science and Ethics

As is generally acknowledged, cognitive psychology developed, at least partly, as a reaction to behaviourism. One of the key aims of this reaction was to accord a respectable scientific status to the intermediate variables that had been postulated by psychological theorists – such as Hull and Tolman – before the ascendancy of Skinner's behaviorism. Cognitive science, therefore, interpreted these intermediate variables between inputs and outputs as inner states of the organism, and regarded these inner states in turn as mental states. In this way, mental states of various sorts regained the explanatory role in psychology that they had lost during the reign of behaviorism. Conduct, the subject matter of psychology, was therefore conceived as the outcome of causal interactions between mental states. The classical formula 'belief + desire *cause* behaviour' is a brief way of expressing this fundamental doctrine.

The immediate appeal of this explanatory scheme lies in the fact that it enables us to connect the most general strategy of scientific psychology, as it is summarized by many leading philosophers of mind, with our commonsense psychology; for our folk psychology is also characterized by the use of terms such as "belief" and "desire". And we explain our own and other people's behaviour in terms of this intentional vocabulary. Smith opened the refrigerator because he desired a glass of milk and believed that the milk was in the fridge, *ceteris paribus*. In this way, it is possible to preserve a categorial continuity between the psychological discourse that we use in our ordinary conversations and the discourse that is characteristic of scientific psychology; and the difference between folk psychology and scientific psychology does not lie in their basic vocabulary, but in certain restrictions that the scientific method imposes on our common way of thinking. Thus, the overt behaviour that we try to explain causally by appealing to mental states of all kinds has such a complex etiology that it forces us, when we apply the standards of scientific method, to postulate entities – sentences in the language of thought and the like – whose properties have little in common with the properties that folk psychology usually attributes to our mental states. But despite this fact, these hypothetical entities preserve the causal powers that we intuitively attribute to mental states. Furthermore, these hypothetical entities are regarded as the true referents of our everyday mental terms, and hence as mental states and processes with causal powers. The advantage of these postulated entities by comparison with mental states as they are conceived of in folk psychological explanations of behaviour, is that they can be easily identified with states of the brain. So there is nothing odd about them from a materialist point of view. These theses are characteristic, for example, of Fodor's and Garfield's psychological realism [Fodor (1987), Garfield, (1988)].

All this can be summarized by saying that folk psychology, as part of the man-

ifest image of man-in-the-world, in Sellars's sense, is the starting point of scientific psychology and the foundation that makes the task of psychology intelligible. Of course, not all philosophers of mind accept this approach to the philosophy of psychology. Eliminative materialists obviously do not accept it. But, for the purposes of this paper, I do not need to discuss the problems involved in these different points of view.

Unfortunately, the attempt to preserve the intentional vocabulary and to introduce inner mental states into the scientific explanation of behaviour did not signal a renewed interest in the practical consequences of psychology. To a great extent, cognitive psychology arose as an internal revolution within scientific psychology, as an attempt to overcome the explanatory schemes, epistemological presuppositions, and ontological limitations of behaviorism. Thus, the main motivation of cognitivism was not the idea that behaviorism had undesirable practical consequences so far as human liberty and dignity are concerned, and that we should try to avoid these consequences, if possible, by reforming the foundations of our psychological theories. In fact, these ethical considerations barely appeared in the philosophical literature that discussed behaviourism and related issues in the philosophy of mind.

One important outcome of the reaction against behaviorism was this: cognitive science rehabilitated the intentional explanation of behaviour – after effecting a major change in our conception of mental states – and with it the prospect of smoothly accommodating this form of explanation within a mechanistic model of the mind, elaborated against a materialistic and scientific background. Nevertheless, for many philosophers, these two achievements are not compatible, appearances notwithstanding. For eliminativists, although intentional explanations have their origin in folk psychology, and are therefore deeply rooted in our culture, we can dispense with them, because we can devise a more powerful way of explaining behaviour by using the non-intentional vocabulary of neurobiology, than by using the intentional idiom of mental states. For other philosophers, the very idea that explanations of behaviour can be both mechanistic and intentional is contradictory, because intentional explanations go hand in hand with the possibility of attributing responsibility to an agent. For example, Dennett has claimed that "explanations that serve to ground verdicts of responsibility are couched at least partly in terms of the beliefs, intentions, desires, and reasons of the person or agent held responsible" [Dennett (1981), p. 234]. The problem is that mechanisms are not responsible for their behaviour or its effects.

If this is true, then there is an inconsistency at the heart of cognitive science: intentionality and mechanism are inconsistent. Nevertheless, Dennett has argued persuasively that intentional explanations of behaviour are compatible with mechanistic explanations, that is, with explanations that are related to the design and the physical stances [Dennett (1981a), p. 237]. If we adopt the intentional stance, we

thereby commit ourselves to attributing rationality to the system whose behaviour we are explaining, but this commitment does not preclude mechanistic explanations of the same behaviour.

> One reason we are tempted to suppose that mechanistic explanations preclude intentional explanations is no doubt that since mechanistic explanations (in particular, physical explanations) are for the most part attempted, or effective, only in cases of malfunctioning or breakdown, where the rationality of the system is obviously impaired, we associate the physical explanation with a failure of intentional explanation, and ignore the possibility that a physical explanation will go through (however superfluous, cumbersome, unfathomable) in cases where intentional explanation is proceeding smoothly. [Dennett (1981a), p. 243]

Mechanistic and intentional explanations are therefore fully compatible. However, according to Dennett, the moral evaluation of behaviour and the attribution of responsibility do not come into the picture with the intentional explanation of behaviour – that is, with what he calls "the intentional stance", but with "the *personal* stance". He writes as follows:

> One adopts the intentional stance towards any system one assumes to be (roughly) rational, where the complexities of its operation preclude maintaining the design stance effectively. The second choice, to adopt a truly moral stance towards the system (thus viewing it as a person), might often turn out to be psychologically irresistible given the first choice [to adopt the intentional stance], but it is logically distinct... We might, then, distinguish a fourth stance, above the intentional stance, called the *personal stance*. The personal stance presupposes the intentional stance (note that the intentional stance presupposes *neither* lower stance) and seems, to cursory view at least, to be just an annexation of moral commitment to the intentional. [Dennett (1981a), p. 240]

Unfortunately, Dennett does not explain in detail how should we understand the concept of "a truly moral stance" or "the personal stance", although he provides with some conditions of personhood that any personal stance should meet [Dennett (1981b)]. We do not know if the truly moral stance is Aristotelian, Kantian, Hobbesian, Utilitarian or a mixture of some of these. He only says that if we choose to adopt the personal stance, this is a pragmatic choice, made in the light of our aims and desires, and that, consequently, we are probably neither right nor wrong if we fail to treat a system, even a person, morally [Dennett (1981a), p. 241]. If this is so, then the moral point of view is still threatened, since we can, at least in principle, choose to adopt a different stance instead – the design, the physical or the intentional stance – without hampering our explanations of behaviour except, perhaps, to the extent that we make them more cumbersome.

I think these considerations make it clear that cognitive science, as it is con-

ceived by several influential philosophers of mind, coheres with the fundamental intuition in the philosophy of science, according to which the propositions of natural science have no axiological content or consequences. If we choose to treat a system from an ethical or personal point of view, this cannot be a choice that was forced upon us by any considerations pertaining to the ontological commitments and presuppositions of cognitive science. It depends exclusively upon our personal interests. Cognitive science and ethics are two quite different domains, because the objects that are the subject matter of cognitive sciences are premoral items that, at most, are *associated* with the intentional stance, and, possibly on occasion, the personal stance; but it cannot be compelled by the conclusions of scientific psychology to adopt the personal stance towards ourselves or others. In the final analysis, it seems that in order to have a natural science of behaviour without abandoning mentalism, we need to concede that the application of expressions such as "morally right" and "morally wrong" to behaviour has no empirical justification.

I think we can reinforce this conclusion by briefly, and somewhat cursorily, considering two typical strategies that the philosophy of mind has used to characterize mental states. These strategies are distinct, but closely interrelated. Their common presupposition is that mental states, inasmuch as they are conceived as efficient causes of behaviour, should be conceived in a morally neutral way. If our behaviour is the function of our beliefs and desires, and if a scientific treatment of these inner states is possible, then beliefs and desires cannot have any moral properties.

The first strategy to strip mental states of any moral significance is to regard them as functional states of the brain. According to the computational version of functionalism, mental processes are like programs running in the hardware of a computer. But a consequence that many philosophers of mind have drawn from this thesis is that mental processes are real but, in principle, inaccessible to consciousness. And unconscious mental states that are in principle inaccessible to consciousness cannot have moral properties. How can it be right or wrong to be in such a state, if one cannot know whether one is in it or not? Such states can have functions, which are defined by their causal roles, but they cannot be good nor bad in any moral sense of these terms.

The second strategy consists in defining mental states in a morally neutral way. Searle, who has consistently criticized the idea that there are unconscious mental states that are in principle inaccessible to consciousness, has argued that all intentional mental states have a direction of fit [Searle (1992), pp. 151-173]. Beliefs have a mind/world direction of fit. If our beliefs are false, then we have to change them but do not need to change the world. Desires, on the contrary, have a world/mind direction of fit. If our desires are not satisfied by the way the world is, then we will tend to change the world in order to satisfy them [Searle (1983)]. But in doing so, we will not take into account whether these desires are morally legitimate. The con-

tent of your desires can be anything whatsoever. Your behaviour and your disposition to change the world will be explained by your desires in conjunction with your beliefs. In this way, desires, as they are described by Searle, are neutral from an ethical point of view. In fact, Searle, like many other philosophers of mind, endorses a Humean conception of desire. Desires are not expressions of a set of moral rules and values that the agent has internalized and organized in a given way [MacIntyre (1988), p. 76]. Desires are just nude facts that the world satisfies or fails to satisfy. If this is the notion of desire that cognitive sciences defend, then it is a premoral idea.

The joint effect of these two strategies can be reinforced by another consideration. As we have seen, cognitive science conceives of mental states as being detached of moral properties. For cognitive science, mental states simply are efficient causes of behaviour. Only in this way can they be taken seriously as objects of investigation by natural science. However, it could be argued, as Aristotelian philosophers surely would do it, that desires will never have normative value if they are not essentially characterized by the object that is desired and by the desirability of this object within the life of a rational individual and of a certain community. This assertion amounts to the thesis that desires can only have normative value if they are explained in terms of final causes and of the desirability of these causes in each particular situation. That desires only have efficient powers means that they are essentially non-moral items that cause behaviour in a non-moral sense. This picture is coherent because cognitive sciences should have explanatory value not only for human beings but also for animals and, perhaps, for certain complicated machines as well. And the behaviour of these two latter kinds of objects does not have moral characteristics, at least, *prima facie*.

For these various reasons, the rehabilitation of mental states by cognitive science has not been accompanied by a rehabilitation of the idea that the explanation of behaviour is intrinsically moral or that it has moral implications. Just the opposite. Any moral property of a mental state will fall outside the purview of science, and therefore outside the scientific conception of the world. Hence, the items that appear in the old scheme "belief + desire *cause* behaviour" lack any moral significance.

## III. Gauthier's Contractualism and Psychological Theory

If this is so, how can we draw moral conclusions from these theoretical foundations? How can cognitive science make a contribution to moral inquiry? And if it cannot make any contribution at all, does it pose a threat to ethics? I shall answer these questions briefly, and in specific connection with Gauthier's liberal contractualism.

Gauthier defends a subjectivist and relativist conception of value. By "subjectivism", I mean the thesis that value is dependent on the agent's appetites or preferences; and by "relativism", in this context, I shall mean the thesis that value is relative to each individual. According to Gauthier, since it is always possible for two individuals to attach a different value to the same state of affairs, even though they are both fully informed and fully reflective and experienced, it follows that the deepest roots of the differences between evaluations do not lie in knowledge but in desire [Gauthier (1986), pp. 49-50]. And it also follows that neither one of two different evaluations is morally better than the other: on the contrary, desires and preferences are all, in principle, on an equal footing from a moral point of view. Gauthier completely subscribes Hume´s thesis according to which it is not more rational to desire the annihilation of the world than to have a scratch in one finger. Accordingly, desires are non-moral or pre-moral items that set the stage for different evaluations. In this respect, both liberal contractualism and the philosophy of mind share a very important thesis, namely, both of them defend a premoral conception of desire.

The protagonist of Gauthier's theory, "economic man", is at the outset a non-moral being who is equipped with desires and preferences characterized in this way. Apart from these desires and preferences, economic man also possesses a "natural endowment" that Gauthier defines as "what one brings to society, to market and cooperative interaction" [Gauthier (1986), p. 220]. This "natural endowment" should not be the result of coercion. Rational bargaining and cooperation are only possible if this requirement is satisfied. The task of moral theory would then be to explain how moral considerations would arise from a base-line – a state of nature, in Hobbes's sense – which is defined by the natural endowment of individuals, and their desires and preferences.

> We are committed to showing why an individual, reasoning from non-moral premises, would accept the constraints of morality in his choices. [Gauthier (1986), p. 5]

Gauthier's well-known answer is that moral considerations are rational constraints on the maximization-seeking behaviour of economic man in an environment in which the conditions of a perfect competitive market do not hold because there are free-riders and externalities, i.e., goods that an individual can enjoy without "paying" for them. Given that Gauthier identifies rationality with economic rationality, and this latter with the mathematical calculus that decision theory provides, it follows that rational moral principles stem from, and are justified by, a 'computation' that applies the tools of rational decision theory to a base-line that is prior to any moral assumptions.

This being so, we can see that cognitivism lies at the root of liberal contractu-

alism as a condition and justification of the human psychology that this theory endorses. It is clear that both cognitivism and liberal contractualism defend a pre-moral conception of human desire that has its origin in the Hobbesian and Humean tradition that both theories share. Cognitive science has done Gauthier a tremendous favour, because it has provided his moral theory with a scientific background. The main achievement of cognitive science, in this regard, has been to strip mental states entirely of moral properties. Furthermore, the computational and syntactic themes in cognitive science, support the notion, which Gauthier evidently shares, that the human mind – or, in Gauthier's case, the mind of "economic man" – is a purely formal mechanism, characterized by logical and mathematical operations. According to classical computational and syntactic theories within cognitive science, the human mind is a computational device, and mental processes consist in the manipulation of representations according to algorithms which are defined in purely syntactic terms. Behaviour is the causal outcome of these processes [Cfr. Stich (1983) and Fodor (1987)]. Analogously, Gauthier claims that "economic man" is "conceived as an independent centre of activity, endeavouring to direct his capacities and resources to the fulfilment of his interests. He considers what he can do, but initially draws no distinction between what he may and may not do" [Gauthier 1986, p. 9]. In order to fulfil his interests and satisfy his desires in an environment in which there are parasites and externalities, economic man will apply the formal calculus of rational decision theory to his interaction with other individuals like him. Economic man is therefore defined as a "mechanism" which applies a formal calculus, a computation, to his preferences and desires without any regard to their morality. This is the process of rational choice.

It could be argued that this partial match between certain strands in cognitive science and Gauthier's liberal contractualism does not indicate a defect in this moral theory. For any moral theory that claims our rational allegiance should be in general harmony with the well-established results of cognitive science and, in turn, with their philosophical basis. If not, it can, for this reason, be suspected of involving a fallacy. Goldman, for example, has argued in this fashion. In particular, he has held that scientific psychology "may impinge on prescriptive ethics, *viz*., by setting constraints of realism or feasibility" [Goldman (1993), p. 640]. He spells out this assertion in the following way.

> A moral code that is psychologically unrealizable by human beings, or just too demanding for people to satisfy, might be rejected on meta-ethical grounds. Not all moral theorists would accept this constraint ... Nevertheless, it is plausible to impose a constraint like Owen Flanagan's Principle of Minimal Psychological Realism. [Goldman 1993, pp. 640-641]

Flanagan's Principle says this: "make sure when constructing a moral theory or

projecting a moral ideal that the character, decision processing, and behaviour prescribed are possible... for creatures like us" [Flanagan (1991), p. 32]. Gauthier's theory probably passes Flanagan's test, whereas if Gauthier is right, Rawls's does not, because the condition of equality that the original position requires is practically unattainable for most people.

Nevertheless, it is one thing to say that our moral theories had better not contradict our psychological theories or their philosophical foundations, and quite another to accept these theories without trying to change or improve them as we move from the science of behaviour to ethics. Aristotelianism, for example, is incompatible with certain fundamental philosophical doctrines presupposed by cognitive science, as it stands. And if we accept these doctrines, we are therefore bound to distort the truths about ethics that Aristotle understood. For example, Aristotelian ethics includes the doctrine that desires are already expressions of our character, and of the norms of the cultures and groups to which we belong. Hence, they cannot provide "data that are neutral between rival moral claims" as liberal contractualism would pretend [MacIntyre (1988), p. 76]. If this is true, desires are not premoral items. On the contrary, they are expressions of our moral habits. Aristotelians are not bound to reject the general scheme 'belief + desire *cause* behaviour'. But they are bound to insist that the items which appear in this formal scheme have moral properties, and that ethics cannot consider beliefs and desires, as perhaps cognitive science can, in abstraction from these moral properties, since they are directly involved and fully implicated in moral theory. The psychological foundations of Gauthier's moral contractualism, by contrast, is thoroughly compatible with the philosophical foundations of cognitive science, just as it stands. Gauthier does not need to modify or enrich the conception of mental states advanced by cognitive science in the least. He can adopt this conception just as it is, and combine it with other specific theoretical elements – the concept of economic man, bargaining, etc. – in order to elaborate his moral theory. Thus, liberal contractualism and cognitive science form a unitary explanatory framework of human behaviour in a twofold dimension: the psychological and the ethical. For this kind of framework, the moral stance is simply the intentional stance, when we interpret the presupposition of rationality that defines the latter in an instrumental and economic sense.

## References

DENNETT, D.C. (1981a), "Mechanism and Responsibility", in *Brainstorms. Philosophical Essays on Mind and Psychology*, Cambridge, Mass., The MIT Press, pp. 233-255.

DENNETT, D.C. (1981b), "Conditions of Personhood", in *Brainstorms.*

*Philosophical Essays on Mind and Psychology*, Cambridge, Mass., The MIT Press, pp. 267-285

FLANAGAN, O. (1991), *Varieties of moral personality: Ethics and psychological realism*, Cambridge, Mass., Harvard University Press.

FODOR, J.A. (1987), *Psychosemantics. The Problem of Meaning in the Philosophy of Mind*, Cambridge, Mass., The MIT Press.

GARFIELD, J. L. (1988), *Belief in Psychology. A Study in the Ontology of Mind*, Cambridge, Mass., The MIT Press.

GAUTHIER, David. (1986), *Morals by Agreement*. Oxford: Oxford University Press.

GOLDMAN, A.I. (1993), "Ethics and Cognitive Science", in Goldman, A.I. (ed.), *Readings in Philosophy and Cognitive Science*, Cambridge, Mass., The MIT Press, pp. 623-646.

GONZÁLEZ-CASTÁN, O.L. (1998), "Can We Say When the Compensation of Social Differences Has Been Just?", *The Journal of Value Inquiry* 7, pp. 1-15.

GONZÁLEZ-CASTÁN, O.L. (1997), "El contractualismo liberal de Gauthier como falsa ideología", *Doxa* 20, pp. 457-476.

GONZÁLEZ-CASTÁN, O.L. (1996), "Protocolos para una rehabilitación no-naturalista de la psicología", *Diálogo Filosófico* 34, pp. 43-60.

GONZÁLEZ-CASTÁN, O.L. (1992), "Intencionalidad sin conciencia: Brentano, Searle y las Ciencias Cognitivas", *Revista de Filosofía*, Vol. V (3ª Época), pp. 99-118.

MACINTYRE, A. (1988), *Whose Justice? Which Rationality?*, Notre Dame, University of Notre Dame Press.

POPPER, K.R. (1982), *Objective knowledge. An Evolutionary Approach*, Oxford, Oxford University Press.

SEARLE, J.R. (1992), *The Rediscovery of the Mind*, Cambridge, Mass., The MIT Press.

SEARLE, J.R. (1983), *Intentionality. An Essay in the Philosophy of Mind*, Cambridge, Cambridge University Press.

STICH, S. (1983), *From Folk Psychology to Cognitive Science. The Case Against Belief*, Cambridge, Mass., The MIT Press.

WILLIAMS, B. (1985), *Ethics and the Limits of Philosophy*. Cambridge, Mass., Harvard University Press.

Óscar L. González-Castán
Dpto. Filosofía I (Metafísica y Teoría del Conocimiento)
Facultad de Filosofía
Universidad Complutense de Madrid
oscar.gonzalez@filos.ucm.es