



## Introduction

# Neuroeconomics: Present and future

*The minds of men are mirrors to one another.*

[David Hume, *A Treatise on Human Nature*, Section V, Part II]

### 1. What is neuroeconomics doing?

This issue of *Games and Economic Behavior* collects a set of papers that apply the concepts, methods, and technical tools of neuroscience to economic analysis. This is what has by now come to be called neuroeconomics (NE). If one wants to understand what NE is, then the most useful way is probably to look at what NE does in concrete research, so we invite the curious reader to choose one of the articles and begin to read. But if one is questioning the method or even the usefulness of this line of research, then an introduction may be the right place for a discussion. In particular this is true if one is trying to understand what this developing field of research is trying to accomplish in the future. The main content of this introduction will be an attempt to provide a possible answer to this question.

In a different paper (Glimcher and Rustichini, 2004) Paul Glimcher and I have tried to provide our view on what neuroeconomics is technically, what methods it uses, and how researchers in the area are in general planning to deal with the classical themes of economics, decision theory and game theory in the first place. A different view is presented in Camerer et al. (2005).

In summary, I think the following is the main point. At the very least, neuroeconomics provides new data in addition to those we have available from theoretical, empirical, and experimental research on human behavior. This is the set of psychophysiological data (for example, the galvanic skin response, which gives a rough measure of the visceral response to a stimulus, or the heart rate), and the imaging analysis of brain activity, measured in several different ways (MRI and PET). We think that neuroeconomics is much more than this, but this seems an indisputable fact. Now, a common criticism that is raised to neuroeconomics is the following. With this method we now know, for example, the specific regions in the brain that are active when some behavior is observed. This information may be very interesting for a neuroscientist, who aims at reconstructing the brain processes underlying behavior. But what does it add to the understanding of economic behavior? The task of an

economist is to establish useful predictions on which human behavior will follow given certain incentives, preferences, and feasibility constraints. This set of parameters, that is available to the economist analyzing the situation, defines the input, and behavior is the output. What, if anything, do we need to know of the intermediate process? The answer so far has been nothing. Let us revisit the reason for this answer.

Contemporary decision theory and game theory (which provide the foundation of economic theory) have been resting so far on the method of *as if* modeling of human behavior. The idea of the method is well known. A model of the process underlying human behavior may be useful to the economic theorist in formulating the predictions. An example of these models is provided by our representations of preferences. This model does not need to be, or even be related to, the actual process. In fact, for several reasons, such as simplicity, generality, and independence of individual variability, this model may be a condensed, reduced form of the real process. It may be, in other words, factually false. The real test of the quality of the model is the predictions it provides.

The theory of subjective expected utility (SEU) is an illuminating example of this method. The entire theory is a tautological equivalence between a complete description of the behavior of an individual, conveniently summarized in a finite list of axioms, and a representation of these preferences, the SEU representation. This representation can be taken to be (although, usually, it is not) the description of a process. The subject has to choose between two different acts. He or she summarizes the value of each of these two acts, according to the following steps. First, he or she forms a subjective belief over the states, then evaluates the value of the consequence of an act in each state, and finally, integrates the value as the SEU formula prescribes. These are the steps of an algorithm that a machine can implement, and the human brain is one such machine. From the point of view of the *as if* theory, this interpretation may be suggestive, but is completely irrelevant to the main task.

As we noted, the axiomatic system is a complete description of choices in a given environment. Keeping in mind this emphasis on completeness, how can this be reconciled with the task of predicting behavior? This seems to have been solved by assumption. How can a system to formulate predictions of behavior be useful to me, if I have to provide from the start the behavior in all possible choices, namely the answer to my very question? The answer must be that prediction has to be understood as prediction in new environments, similar to the ones for which the theory is built, but different enough to make prediction something more than the trivial task of checking an answer that we already have. It is prediction that we can call out of sample.

If one accepted the view that economics is only concerned with formulating predictions of human choice given preferences, incentives, and feasibility constraints, then the interpretation of this representation as an *as if* construction would be sufficient. The success of this method, however, was based on an assumption, it had to work. That is, the reconstruction of the actual decision process was not necessary if the single, unified, simple model was giving the correct predictions for a general set of conditions. SEU was such a model and its extension to a norm of behavior in games and competitive economies was its natural completion.

But the deviations from the norm established by the SEU seem to be many and various, both in empirical and in experimental tests. The same happens in game theory and its

applications. For example, it is widely accepted that individuals behave in a competitive, price taking behavior when the number of individuals is larger than a given threshold (let us say, ten subjects), while their behavior is different when the number is smaller.

This fact has produced, in past years, a large class of models trying to provide a description of human behavior facing uncertainty. The vast majority of these theories come with a uniqueness result (up to trivial transformations), hence, each one claims to be a complete and exclusive description of human behavior. This proliferation has now produced a problem, even from the point of view of the strict *as if* interpretation of the theory. Economists have to produce useful prediction of human behavior. Now we have several models competing, each claiming to be the unique predictor. Which one do we choose?

A possible claim is that they are all, each within a specific class of environments, correct and individuals use the different models in different situations. For instance, in the example we were giving previously of the price taking vs the strategic behavior, the two models are both useful, and people switch from one to the other. This is a reasonable answer, but it then poses a completely new question, as challenging as the initial one, do we have a theory on which one will be used? Presumably, we are not allowed to choose among the models we have available by picking the one that predicts better the observed behavior. This is necessary to satisfy correct scientific standards, but even more because the prediction is really interesting only, as we have said earlier, as prediction out of a sample. It is well known that Friedman, in his fundamental methodological paper (Friedman, 1953) went as far as allowing the economic theorist to hold, for the same unit, two *contradictory* assumptions when analyzing two different issues. In Section V, he discusses the behavior of the cigarette industry with respect to two different policies: an increase in the federal cigarette tax and price controls. He states that the outcome of the policies can be described by a theory assuming perfect competition (in the case of the first policy) and oligopolistic structure (in the case of the second). Formulating a successful prediction seems too easy if the choice of the assumptions is allowed to depend on the prediction. I think many economists would consider this legitimate only because there is the presumption that a more general theory, including the two just mentioned as special cases is possible; and they would feel uneasy if such a more general explanation were found to be impossible. This theory would be necessary, for instance, to formulate predictions facing a completely new policy, with an outcome as yet unknown, this is, after all, the real purpose of economics.

Understanding the process is the way to recover the uniqueness of the model of human behavior that we do not seem to have. In other words, the question of whether the algorithm is a description of the process can no longer be postponed or ignored. Neuroeconomics can provide this unified model. Many of the papers in the special issue address, in one form or another, this question, and try to provide an answer.

We think, however, that neuroeconomics has a wider scope than a useful test of existing theories, or even new theories. To this more ambitious project we now turn.

## 2. What will neuroeconomics do?

The more ambitious aim of neuroeconomics is going to be the attempt to complete the research program that the early classics (in particular Hume and Smith) set out in the first

place: to provide a unified theory of human behavior. Arguing in support of this statement requires first a reconsideration of the work of the founding father of modern economics.<sup>1</sup>

### 2.1. Adam Smith and modern economics

A striking difference separates the two early classics of modern economics, Adam Smith's *Enquiry into the Nature and Causes of the Wealth of Nations* (*WoN*) (1904) first published in 1776, and Ricardo's *On the Principles of Political Economy and Taxation* (*PPET*) (1821) first published 41 years later in 1817.

The *Enquiry* was considered by Adam Smith, through his life, only as one of several components of a larger inquiry into human nature. In this, he was following the path set by Hume, who laid at the foundation of the four sciences of logic, morals, criticism, and politics an understanding of human nature.<sup>2</sup> The *Enquiry* was in fact the second in a body of work that included three parts. The first was *The Theory of Moral Sentiments* (*TMS*) (Smith, 1790), the second the *WoN*, and the third a treatise on the theory of jurisprudence that was never completed. This great plan was still in the mind of the elder Smith in the Advertisement to the sixth edition of the *TMS*, written in 1790, the year of his death.<sup>3</sup> It is clear from this ambitious master plan that his two great works, *TMS* and *WoN*, were both parts of a unified inquiry into the nature of human society.

By the time Ricardo's *Principles* was written, economics had become a separate discipline, with no need for a preliminary investigation into the fundamental characteristics of human nature. It is well known that in Ricardo's view political economy was focused on a very specific issue:

To determine the laws which regulate this distribution, is the principal problem in Political Economy. (*Preface, PPET*)

This is an issue that would not require or even find useful any investigation into human nature. It is important to note here that the Ricardo does not even mention, or discuss, the dramatic shift in the scope of the analysis that has occurred between his analysis of economics and the one of Adam Smith, although Smith is clearly, and explicitly, his reference point, as made clear in the Preface to the *Principles*.

<sup>1</sup> A similar analysis, an explicit discussion of the work of Smith on sympathy and the connection with modern neuroscience is given by Rizzolatti and Craighero (2005).

<sup>2</sup> "Here then is the only expedient, from which we can hope for success in our philosophical researches . . . to march directly to the center of these sciences, to human nature itself" [Preface to *A Treatise of Human Nature* (Hume, 1739)].

<sup>3</sup> "In the last paragraph of the first edition of the present work, I said, that I should in another discourse endeavour to give an account of the general principles of law and government, and of the different revolutions which they had undergone in the different ages and periods of society; not only in what concerns justice, but in what concerns police, revenue, and arms, and whatever else is the object of law. In the *Enquiry Concerning the Nature and Causes of the Wealth of Nations*, I have partly executed this promise; at least so far as concerns police, revenue, and arms. What remains, the theory of jurisprudence, which I have long projected, I have hitherto been hindered from executing" (*TMS*, sixth ed., Advertisement).

## 2.2. *The theory of moral sentiments*

The fundamental question that is addressed in the *TMS* can be formulated as “How can human societies, based on people who are fundamentally selfish, function properly?” The question was one of the most important in the tradition of the school of the Scottish philosophers. The starting observation of the school was that man is usually found not in isolation, but in society. Hence there must be something that ensures the stability of social arrangements. Since there was no illusion about the fundamentally selfish nature of man,<sup>4</sup> some corrective factor must intervene.

The method used by the school to determine this factor finds as fundamental an analysis of human nature, this in turn is based on the evidence provided by introspection. In the general chapter entitled “Characteristics of the School,” McCosh in his classical treatise states that the school “employs self-consciousness as the instrument of observation.”<sup>5</sup> The discussion that follows makes it clear, however, that this restriction, and the exclusion from the evidence of the observation of “brain and nerves” (which is the equivalent of modern-day neuroeconomics) was more a necessity imposed by the limitations of the state of psychology of the day, rather than a preferred choice.<sup>6</sup> One may wonder whether Adam Smith, were he working today, would not be a neuroeconomist.

On the basis of this method, in the *TMS*, Smith provides his answer to the fundamental question of what makes human society stable. The cornerstone around which the theory of moral sentiments is built is the concept of sympathy. The very opening words of the first chapter (*On Sympathy*) of the treatise make it clear:

How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. Of this kind is pity or compassion, the emotion which we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner. (*TMS*, I.I.1)

A few points have to be emphasized. First, the starting point of the entire discourse in the *TMS* is sympathy. Consider for example the theory of the *propriety* of the actions and affections of other people. Actions or affections are not appropriate if they satisfy a set of norms. Rather, they are if they correspond to the affections that the circumstance creates in the spectator.<sup>7</sup>

---

<sup>4</sup> At least there was no illusion of the sort in Adam Smith. The Scottish school has noticeable exceptions to this statement, first of all in the work of Francis Hutcheson. But the statement holds for the majority of the philosophers in the school.

<sup>5</sup> In McCosh (1875, p. 8).

<sup>6</sup> “There is nothing in the method, or in the spirit, or the cherished doctrines of the school tending to discountenance or disparage a painstaking experimental investigation of the parts of the bodily frame most intimately connected with mental action” (McCosh, 1875, p. 9).

<sup>7</sup> “To approve of the passions of another . . . as suitable to their objects, is the same thing as to observe that we entirely sympathize with them; and not to approve of them as such, is the same thing as to observe that we do not entirely sympathize with them” (*TMS*, II.I.20).

Second, it is clear that sympathy for Smith is a process based on simulation. Sympathy is not the output of a process of cognitive understanding of the situation of the others, but of a process that consists of reproducing what we would feel in the same situation, and recreating in one's mind the affection that we would feel.

By the imagination we place ourselves in his situation, we conceive ourselves enduring all the same torments, we enter as it were into his body, and become in some measure the same person with him, and thence form some idea of his sensations, and even feel something which, though weaker in degree, is not altogether unlike them. (*TMS*, I.I.2)

Third, sympathy is first and foremost in the direction of the misfortunes of others: “The word sympathy, in its most proper and primitive signification, denotes our fellow-feeling with the sufferings, not with the enjoyments, of others” (*TMS*, I.III.1). This different attitude of sympathy is something that Smith takes as a given, commonly observed fact. A subtle observer of human dispositions, he knows that sympathy can be produced by both grief and joy, but “there is however this difference between grief and joy, that we are generally most disposed to sympathize with small joys and great sorrows” (*TMS*, I.II.32). Several reasons may justify this asymmetry. One, for example, is that pain is a stronger emotion than pleasure, and is therefore more likely to induce an effect in others. In several steps of his analysis Smith touches upon the most likely explanation, envy, but he seems reluctant to adopt it. For example: “The obvious observation . . . is that our propensity to sympathize with sorrow must be very strong, and our inclination to sympathize with joy very weak. Notwithstanding this prejudice, I will venture to affirm that when there is no envy in the case” the opposite holds (*TMS*, I.III.5).

The last fundamental aspect of Smith's view is that sympathy arises from the observation of the *event* involving the other person, rather than from the observation of his emotional or affective display following the event. Sympathy proceeds from the observation of an actual fact, for example the pain inflicted on someone else, and recreates the internal state that presumably follows this fact. From this inference the observer can potentially infer useful information on the action that the observed person will choose. As we are going to see, this is an important difference from modern developments in the analysis of sympathy, where the starting observation is that of an act or of an affective reaction, and the inference we make is to the internal state that is likely to have preceded this external display.

Although “upon some occasion sympathy may seem to arise merely from the view of a certain emotion in another person . . . this does however not hold universally, or with regard to every passion” (*TMS*, I.I.7). Some emotions, he observes, do not excite any sort of sympathy, or may induce the opposite: this happens for instance when we observe an angry man without any knowledge of the reason for his anger. But also the sympathy for emotions, such as joy, or grief, that make identification easier is extremely imperfect until we know their cause: “Sympathy does not arise from the view of the passion, as from that of the situation which excites it” (*TMS*, I.I.10).<sup>8</sup> A further proof of the claim is the

<sup>8</sup> The same view is shared by Hume. Describing how our sympathy for a person of merit falling into misfortune develops, he observes that “we form a notion of his condition, and carrying our fancy from the cause to the external effects, first conceive a lively idea of his sorrows” (*THN*, Book II, II.vii).

observation that we may feel for another a passion that the other is unable to feel. The examples Smith proposes are the “blush for impudence and rudeness of another, though he himself appears to have no sense of the impropriety of his own behavior” (*TMS*, I.I.10) or the anguish we feel of the view of a poor wretch who lacks reason, and “laughs and sings . . . and is altogether insensible of his own misery” (*TMS*, I.I.12).

The body of the *TMS* is an elaboration of these main points. If the initial question was, “What corrects the selfishness of human nature to provide a stable social arrangement?” then the answer provided in the *TMS* is that the desire for approval of the fellow human beings restricts men to moral behavior. In turn, this desire for approval is grounded on the innate affect of sympathy.

### 2.3. From *The Theory of Moral Sentiments* to *The Wealth of Nations*

The general terms of the thesis presented in the *TMS* are known, I think, to most economists. Still, this part of the work of Adam Smith, his work as moral philosopher, is considered separately, and disconnected, from his main work as an economist, and is largely ignored by economists. How did this happen, if he thought that on the contrary the two components of his work were both essential steps in the understanding of the way human societies work?

The success of the explanation provided in the *WoN*, based on the idea of the Invisible Hand, is probably the very reason that the unified program set down by Hume and the Scottish philosophers and defended in the two main works of Adam Smith has been *de facto* abandoned, and economics has followed the more restricted path set by Ricardo. Society, this seems to be the implication of the Invisible Hand theorem, can function well even if there is no social inclination in human nature, at least as long as property rights are well defined and guaranteed. No additional support from special properties of human nature is needed.

The analysis of Smith (and Hume before him) was prophetic, and it will appear to be so even more after we compare it with what we know now. It leaves us, however, with the following fundamental unanswered questions. Is sympathy an innate, automatic feature of human nature? Or is it something that we learn? Does it focus on reconstructing the internal state of the other starting from the external events that are affecting him, or from the observation of his affective reactions? Is it really naturally biased in the direction of sympathy for the other’s misfortunes, and if so why? And why do we need sympathy? As we have seen, the early masters relied on introspection as evidence to detect facts and as a test to refute hypotheses. We can now look at their same questions with the new tools that neuroscience is providing to us.

## 3. Mirror neurons and mirror systems

The entire conceptual issue of sympathy has been recently (in the past 10 years or so) put into a new perspective by some development in neuroscience, most notably by the development of the idea of mirror neurons.

### 3.1. *Mirror neurons*

A sector of the central premotor cortex of the monkey controls hand and mouth movements. The sector is clearly identifiable by specific features of the cells composing it, and has been called F5.

An important functional property of this area is that most of the neurons are active (in the form of a discharge) in association with actions, such as grasping, holding, tearing, and manipulating objects. The activity of these neurons is not associated with any single movement constituting the action: for example, it is not associated with the initial movement of the action of grasping, but with the action in its entirety.

The remarkable discovery (Rizzolatti et al., 1988) is that some of the neurons in the F5 area discharge both when the subject (a monkey) performs the action and when the same subject observes another monkey performing that action. In view of this dual property of being active both when the action is performed and also when it is observed, these neurons have been called mirror neurons (MN). For a recent survey on past research on mirror neurons, see Rizzolatti and Craighero (2004). An interesting application of the theory is to the nature and development of language (beginning with Rizzolatti and Arbib, 1998). An overall view of the concept of mirror system is in Gallese et al. (2004).

In the years following the original discovery of mirror neurons, several experimental facts have been established, which shed additional light on the meaning of the MNs.

First, mirror neurons generalize: for example, they discharge in response to the observation of a specific action irrespective of whether the subject performing it is close or far. Second, MNs respond to the action (for example, grasping) not to the object being grasped. The presentation of an object in isolation does not induce any activation, even if it does when being grasped. This fact does not change if the object is useful (for example, if it is a peanut for a monkey). Similarly, the MNs do not respond to the observation of a hand mimicking the movement of the object.

MNs usually have a strong congruence between the observed and executed actions that activates them. If a mirror neuron is activated by the observation of precision gripping then it is also activated by the execution of precision gripping and not, say, tearing.

MNs respond to an internally generated representation of the action, not to the physical reality of the action or the movement. This is illustrated by the experiment in Umiltá et al. (2001). In the experiment the action studied is gripping of an object. This includes the entire movement from a rest position to the object, together with the final act of gripping. There are two parameters that are being manipulated: hidden vs “full vision” and object versus nonobject. In the hidden condition, the final phase of gripping was hidden behind a screen; in the open condition, the entire action was in full view. In the object condition there was an object to grip, while in the no-object condition there was no object.<sup>9</sup>

The result of the experiment was that the MNs associated with the gripping action that the subject observed discharged both in the open and in the hidden condition, but only in the object condition. So the full action, including the grasping of the object, was needed to

---

<sup>9</sup> The subject (again a monkey) could differentiate between the object and no-object condition even in the hidden condition because the position where the object should have been was shown before the screen was lowered.



activate the neurons, independent of whether the action was partially hidden and hence not available to the visual system.

There is a wide agreement that MNs in the area F5 facilitate action understanding, through the simulation of the action in the premotor system of the observer. The fundamental element here is that understanding is produced by simulation: that is, by the activation in the observer of the same brain region that produces the action.

All the evidence considered so far proves the mirroring of a *motor* activity, as opposed to more general activities, and it is also confined to experiments with monkeys. The extension to human subjects has been difficult, since single-neuron recording techniques are not available for human subjects. Similar results have been confirmed, however, with different techniques. PET and MRI scanning are the principal ones: another clever technique is TMS, see Fadiga et al. (1995). This is not the place to enter into these technical details, but it is important to note that the fundamental result has been confirmed: there is an overlap of the brain structures that are devoted to the observation and to the execution of actions.

A second extension of the research has tested whether the same principle applies to activities that are not just motor. Is there a mirroring of different activities, or even internal states, such as emotions? An experiment, see Wickers et al. (2003), shows that this is the case for a simple emotion, namely physical disgust. In the experiment, subjects were in the condition either of inhaling odorants (which could in turn be disgusting, pleasurable, or neutral) or of observing someone inhaling the same substances by watching their facial expressions in a short movie clip. A brain imaging analysis of the activation in the two cases isolated for each case a set of brain regions. The interesting finding is that the intersection of the two sets was nonempty, and consisted in large part of the anterior insula. This is the region in the brain that is usually associated with feelings of disgust (physical or social). The conclusion is that “as observing hand action activates the observer’s motor representation of that action, observing an emotion activates the neuronal representation of that emotion” (Wickers et al., 2003).

Similar experiments have been conducted for pain, see for example, Singer et al. (2004). In both cases, the results provide striking support for the main hypothesis, that there is a substantial overlap between the areas that are activated when we experience an emotion and when we observe someone experiencing that same emotion.

The conclusion of this excursus in recent neuroscience research is easily summarized. Smith (and Hume before him) had identified sympathy as a pervasive feature of human nature by the power of introspection. Evidence of this sort can easily be dismissed as weak, or biased. In addition, it says very little on the nature of sympathy. But there is now ample evidence that there is a deep reason for Smith’s intuition.

#### **4. Sympathy, introspection, and games**

It is now time to reconsider Smith’s (as well as Hume’s) view of sympathy in light of these recent discoveries.

First of all, the limit to the research that the Scottish school had is now being overcome: neuroscience is now adding to introspection the evidence provided by the analysis of brain activations. A fundamental result is already clear: sympathy has a basis in the way the

brain works, and is based on the sharing of networks (or even neurons) that do something and that observe something. For example, the mirror neuron is active when an action is executed and also when the same action is observed.

Second, Smith's theory is clearly a theory of sympathy with the emotions felt by the observed person. In his view, sympathy is first of all the ability to "place ourselves by the imagination in [someone else's] situation." In this view, reading someone's mind precedes the simulation of his emotions. This seems to be a related but distinct way of conceiving sympathy from the one here.

Third, the neural structures that produce sympathy (such as mirror neurons, or more generally "mirror systems") have quite likely a functional role. This is probably (as has already been suggested systematically in this literature) action understanding, at least when these structures are motor structures. More generally (as in the case of the insula activation in disgust, or insula and anterior cingulate in sympathy of pain) they probably produce understanding of the internal state of the other.

The very term "action understanding" may have at least two different meanings. One meaning is simply this: I am observing some movements of another person, and I want to reconstruct them into a coherent whole, which is called action, to "make sense of them." Using standard statistical theory of information processing, I can formulate this situation by taking the movement as an observed signal, and the action as the unknown parameter that I want to discover. A different interpretation is that action is really a sequence of different components, and I am interested in predicting the future ones on the basis of the ones I am currently observing. In both interpretations, a sympathetic system is useful to the observer. The second interpretation, however, makes it easier to understand the more general role of sympathy as simulation or inference over internal states of the others: in these cases, the functional role is ultimately prediction of future actions.

#### 4.1. An informational theory of sympathy

These facts point to sympathy as a component in a rich system of information processing. The thesis we suggest is the following. Sympathy is the process by which a subject who is observing a second person can internally reproduce the mental process of the observed person. It is an important element of mind reading because it provides information to the mind reader, see on this point Gallese and Goldman (1998). This reproduction is possible because the neural structure of the subject who is observing and of the observed are similar. The purpose is that of extracting information from the observed subject.

This general definition includes both types of sympathy: the "Smithian sympathy" that proceeds from the observation of the event and simulates the internal state of the observed individual, and the "mirror sympathy" that proceeds from the observation of the actions or displays of affections of the observed person. In both cases the intent is to acquire information on the internal state of the observed person.

Sympathy, however, is *not* the only element in the process of understanding of the motivations, intentions, and future actions of the others. It provides information, and hence it is useful in understanding the environment. But as with any other piece of information, it is even more useful if it is processed, and combined with a prior assessment of the same envi-

ronment. In other words, sympathy may be an affective state, but it is always sophisticated and not naive.

It follows that the understanding of sympathy requires a complete theory of information processing, where the informational input provided by the simulation of the internal processes of others is used as such, as another piece of information.

#### 4.2. Introspection in games

As an example, consider a player who is deciding what to do at a node in an extensive form game. His decision is obviously influenced by the expectation of what the other players will do at subsequent nodes. How can he form such expectations? One way is to formulate in his mind some hypothesis about the possible distribution of strategies in the population; or even more indirectly, about the distribution of parameters (such as preferences, or beliefs) that are relevant to shaping these decisions.

Sympathy suggests a different way: the player may introspectively consider what he would do (or, more indirectly, what he would think) at those nodes, and take that as useful information on what the others are likely to do or think. This is an application of “Smithian sympathy”; the player is considering the effect of an external event (finding oneself at a node) on another person, and is trying to infer the internal state as useful predictor of the action of the other player in that situation.

The hypothesis we have presented suggests that the true thought process of the player is going to be a combination of two elements. It will be influenced by the introspection, because sympathy is permanently active. But it will also be balanced by the assessment that the player gives of the general population, discounting the information that introspection gives because the same player may know that he is not representative of the general population: as we have said before, sympathy is not naive. These are the interesting themes for future research.

## References

- Camerer, C., Loewenstein, G., Prelec, D., 2005. Neuroeconomics: How neuroscience can inform economics. *J. Econ. Lit.* 43 (1), 9–64.
- Fadiga, L., Ferrari, P.F., Pavesi, G., Rizzolatti, G., 1995. Motor facilitation during action observation: A magnetic stimulation study. *J. Neurophysiol.* 73, 2608–2611.
- Friedman, M., 1953. The Methodology of Positive Economics. In: *Essays in Positive Economics*. Univ. of Chicago Press, Chicago, pp. 3–43.
- Gallese, V., Keysers, C., Rizzolatti, G., 2004. A unifying view of the basis of social cognition. *Trends in Cognitive Sci.* 8, 396–403.
- Gallese, V., Goldman, A., 1998. Mirror neurons and the simulation theory of mind reading. *Trends in Cognitive Neurosci.* 12, 493–501.
- Glimcher, P., Rustichini, A., 2004. The consilience of brain and decision. *Science* 306, 447–452.
- Hume, D., 1739. *A Treatise on Human Nature*. Edited by Selby-Bigge, L.A. Oxford Univ. Press.
- McCosh, J., 1875. *The Scottish Philosophy*. MacMillan, London.
- Ricardo, D., 1821. *On the Principles of Political Economy and Taxation*, third ed.
- Rizzolatti, G., Arbib, M., 1998. Language within our grasp. *Trends in Neurosci.* 21, 188–194.
- Rizzolatti, C., Craighero, L., 2004. The mirror–neuron system. *Ann. Rev. Neurosci.* 27, 169–192.
- Rizzolatti, C., Craighero, L., 2005. Mirror neuron: A neurological approach to sympathy. Manuscript.

- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., Matelli, M., 1988. Functional organization of inferior area 6 in the macaque monkey. *Exper. Brain Res.* 71, 491–507.
- Singer, T., Seymour, B., O’Doherty, J., Kaube, H., Dolan, R., Frith, C., 2004. Empathy for pain involves the affective but not the sensory components of pain. *Science* 303, 1157–1162.
- Smith, A., 1790. *The Theory of Moral Sentiments*, sixth ed.
- Smith, A., 1904. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edited by Cannan, E., fifth ed. Methuen, London (first published 1776).
- Umiltá, M., Kohler, E., Keysers, C., Gallese, V., Fogassi, L., Fadiga, L., Rizzolatti, G., 2001. I know what you are doing: A neurophysiological study. *Neuron* 32, 92–101.
- Wickers, B., Keysers, C., Plailly, J., Royet, J., Gallese, V., Rizzolatti, G., 2003. Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron* 40, 655–664.

Aldo Rustichini  
*Department of Economics, University of Minnesota,  
1035 Heller Hall, 271 19th Avenue South,  
Minneapolis, MN 55455, USA  
E-mail address: arust@econ.umn.edu.*