Toward a cognitive experimental economics

Marco Novarese
Centre for Cognitive Economics
Dipartimento di Scienze Giuridiche ed Economiche - Università del Piemonte Orientale
e-mail: marco@novarese.org

For the last few years, economic literature has shown a new stream of analysis, at the crossroads of the heterodox tradition and cognitive sciences. A new approach is born, called Cognitive Economics, since attention is focused on the study of individual and organizational learning, seen as a key factor in shaping social phenomena (Rizzello, 2000).

This new stream of analysis aims at refounding economic methodologies, starting from more realistic hypotheses than those at the basis of the mainstream. Therefore, empirical observation of real men and women's learning and decision making is necessary. Experimental economics, as the main way to gather data on these aspects, is essential for the development of such an approach. At the moment, though, there is still a lack of methodological discussion on how to perform experiments.

This chapter aims at analyzing the relation between experimentalism and cognitive approach in economics. As it will be shown, cognitivism implies and requires methodological and procedural peculiarities - as compared with the mainstream - in the realization of experiments and in the analysis of the relative data.

These differences depend, first of all, on the fact that, though it is deep-rooted in social sciences, this can be considered a new-born discipline with obvious different necessities from those of an older science. At the moment, in fact, empirical analysis should here be directed mainly towards the gathering of empirical evidence, instead of comparing alternative models to find the most fitted one. The importance of this descriptive interest is also related to the goal of understanding reality, and not just building models with predictive capacity.

One of the main questions suggested by the cognitive literature, a question which raises many theoretical and empirical problems, is the heterogeneity of human agents. Ceteris paribus hypotheses cannot be simply posed and accepted, but should be tested and verified. Differences among subjects should, in fact, be studied and understood. But this goal requires new instruments and new kinds of analysis.

For example, if individuals differ, a similar experimental treatment can bring to different behaviours. A macro analysis could then lead to conclude that a variable is not relevant, because positive and negative effects can compensate each other at a mean level. But the "mean agent" could make no sense.

Some new devices and tools are available to carry on the study. New kinds of experiments, psychological tests and interviews have been recently developed and tested. Also the interaction between experimentalism and computational economics and its simulations can be very useful.

Simulations with artificial agent is a growing and promising area of research; but the heterogeneity of human agents makes it difficult, at the moment, to perform such models, without looking at the real world. The risk would be, again, that of a science disconnected from reality.

The following discussion is necessarily preliminary. At the moment there are many questions opened and few definite answers, but some promising results give courage in the work started.

This chapter aims also at raising a necessary methodological discussion on the general experimental methods. Some of the problems which emerge while conducting experiments based on a cognitive approach can, in fact, be relevant in more general terms.

This work is organised in the following way. The first paragraph proposes some methodological and historical notes on experimental economics and singles out a specific new-born cognitive

stream of studies. Some peculiarities seem to characterize papers that can be related to this approach.
Then a series of methodological and procedural problems and results are discussed through the presentation of two experiments conducted at the Centre for Cognitive Economics, in Alessandria. The first one is designed to study individual learning. It helps us demonstrating that a detailed analysis of each player's individual data is necessary and useful. The second one is directed to investigate team learning and to show two possible tools, i.e. psychological tests and interviews.
Paragraph 3 proposes some considerations on the relation between simulations with artificial agents and experiments. Conclusions follow.

1. Experimental economics and Cognitive Economics, some historical and methodological notes

1.1. An Historical perspective

Usually three main areas of research are supposed to characterize the development of experimental economics (see, for example, Roth, 1993):
- experiments on markets,
- experiments on human reasoning and decision-making (mainly focused on von Neumann and Morgenstern's theory of expected utility),
- experiments on the theory of games.

Novarese and Rizzello (1999) individuate a fourth, very recent and - at the moment - less developed, stream: the one focused on individual and organisational learning and, more generally on cognitive economics. The main papers to recall now are Cohen and Backdayan (1991), Egidi and Narduzzo (1997), Novarese and Rizzello (2002).

This classification is obviously arbitrary, as it often happens in this case. Consider two examples, showing how a different classification could be used.
- The theory of games has been used to model market mechanisms or individual decision-making and learning, i.e. the issues that define the other areas.
- Cognitive Economics is strictly related to important experiments classified in the stream on decision making (see Rizzello, 2000), such as Allais (1953), or Kahnemann e Tversky (1979 and 1992).

Yet, the proposed classification is useful to evidence origins and specificity of the different disciplinary areas that gave birth and favoured the development of experimental economics, influencing its methodological evolution. In this sense, for example, the theory of games had a fundamental role in stimulating the realization of experiments and the debate on them, during the 50's and thanks to the Rand Association (see Innocenti, 1995).
Similarly, it seems then useful to separate the fourth stream from the others, because of its novelty, in terms of methodological and procedural aims and problems, and also because of the hypothesis under investigation.

1.2. A methodological perspective

Even if sometimes the mainstream has considered it as a minor discipline, experimental economics has often been, directly or indirectly, put at the service of the neoclassical approach and it has also affected its methodological development. Different examples could be proposed. Consider these ones, here. Other notes will be presented later.
John Hey (1991), one of the most important scholars in this field, states that experimentalism allows testing economic models under ceteris paribus hypotheses. Then Economics can test its models just as sciences such as Physics or Chemistry do.

Hey and Darnanoni (1988) show that, even if subjects use rules of thumbs to take decisions, in an experiment on consumer behaviour, they make the optimal choiche, "as if" they were able to maximize. The theoretical model under analysis seems then perfectly able to survive the test of the real world. Experiments are used to test this fact[1].

In introducing one of the first papers of the new approach, Egidi and Narduzzo (1997) underline a different point: experiments make it possible to gather empirical evidence (for example on learning processes) on aspects otherwise almost impossible to observe (in the field). One of the main difficulties is the tacitness and opacity of mental mechanisms for the individual himself. For this reason, there are problems in the use of standard psychological experimental tools (like protocol analysis, for example) and different tools are required.

Since the cognitive approach aims to be close to reality, experimental contests are here usually more complex. As Hey (1991) suggests, in fact, usually experimental economics deals with simple situation, because that is the requirement of the theories.

Take for example the studies of Allais (1953), and Kahnemann e Tversky (1979 and 1992). They are interested in criticising the theory of expected utility. This model represents, obviously, the main point of reference in their experimental setting. But it also becomes the cornerstone on which the performance of their alternative model (the prospect theory) is built and compared. Then individuals are called to face different kinds of (implicit or explicit) lotteries and of similar situations, in which they are generally perfectly informed on all relevant probabilities and numerical data necessary to choose. In the real world, these situations (that according to Knight, 1921, are characterized by risk, as opposed to uncertainty) are yet very uncommon.

Novarese and Rizzello or Egidi and Cohen's experiments are very different, as individuals build their knowledge in a context characterized by uncertainty and, even if well defined hypotheses are tested, there is no model according to which the laboratory framework is designed.

Other differences inevitably follow. In what I call "traditional experiments", attention is usually focused on the possible differences among treatment groups. In the other stream, even if there are treatments, there is also a strong descriptive stress. As subjects face complex tasks (and not just simple hypothetical market or game sets), and as experimenter aims at building models of the real learning processes and decision making (and not just to compare different theories), players' behaviour is seen as interesting in itself and analyzed also beyond treatment groups.

These practices are coherent with the recalled idea of experimental economics as a mean to gather data on aspects otherwise difficult to study. The researcher is interested in understanding why subjects behave in the way they do. So, attention is also focused on individual data and not only on macro phenomena.

Probably, these differences also reflect the different stages in the development of the two approaches and the resulting differences in their needs. On the one hand there is the mainstream, with long years of development, on the other hand there is a new-born discipline.

Following sections show some examples of analyses, results and problems related to the new "cognitive experimental economics".

It is useful to anticipate the main methodological problems faced and their major effects. These problems are suggested by both the empirical practice and the cognitive science itself. While they are probably more important in relation to the new approach, they have also a more general applicability.

Laboratory, even with all usual cautions, can hardly be a neutral place, in which individuals forget their usual lives (as sometimes, some experimentalists seem to hope).Their lives are not those of

---

[1] As in laboratory it is possible to satisfy ceteris paribus conditions, experimental economics could also be in contrast with econometrics, the discipline born to test economic models using data gathered from nature's, uncontrolled, experiments. Therefore experimental and field data are sometimes seen as contrasting.

neoclassical agents. On the contrary are lives full of ideas, beliefs, knowledge, and information. All of these elements differ in different players, as they are linked to and dependent on their personal attitude, experiences, studies, jobs and so on. And all of these differences enter the laboratory and guide players' behaviour[2].
If players were the agents of neoclassical economics, furthermore, experiments would be useless, as results would be, obviously, identical to those reached in the models!

As we will see, many different factors influence players' behaviour. All these factors cause loss of control and increase variability among players.
It should be underlined that people can react in different ways to the same treatments because of problems related to social experimentation, and also because of their personal differences, which experiments (luckily) cannot cancel.
As we will see in the following sections, aggregated data can sometimes be inevitably misleading, as they picture a mean player that is different from each real player. To understand how people behave, it becomes then necessary to study each individual case.
The following Herbert Simon's (1992, p. 20) methodological remarks on the analysis of the firm seem then useful also in relation to experiments:

> If you are trying to understand what firms are and how they operate, you will learn a lot from this kind of very detailed study of the processes of decision … Of course, we should not stop with five firms. Biologists have described millions of species of plants and animals in the world, and they think they've hardly started the job. Now, I'm not suggesting that we should go out and describe decision making in a million firm; but we might at least get on with the task and see if we can describe the first thousand. That doesn't immediately solve the aggregation problem, but surely, and in spite of the question of sampling, it is better to form an aggregate from detailed empirical knowledge of a thousand firms, or five, than from direct knowledge of none. But the latter is what we have been doing in economics for too many years

2. Learning, individual behaviour and methodological problems in experiments

2.1. An experiment on individual learning

The first example proposed is an experiment on individual learning. Additional notes and results on this experiment and on its theoretical framework are presented in Novarese and Rizzello (2002), on which this section is partly based.
Here the attention is focused mainly on its methodological aspects. In particular, two issues are dealt with:
- significant differences among individuals can arise even within the same treatment groups;
- an individual-based descriptive analysis can be very useful to understand macro results.

[2] A simple example is the well-known tendency of economics students to free-riding. They bring the mentality learned in their books into the experiments, behaving much more egoistically than other people (see, for example, Franck et al, 1993).
Consider also the following Andreoni (1995)'s statements, very representative of the mainstream experimental approach under many respects:
"laboratory experiments are designed to be neutral and to minimize social effects like kindness. Hence, regular public-goods experiments may already be eliminating a large amount of subjects' natural tendency to be cooperative …" (p.900) but "social and cultural propensities for kindness and generosity must clearly be very strong, and that such motives cannot easily be removed from experimenters simply by providing neutral environments and pledges of anonymity" (p. 892)

2.1.2. The experiment

Participants were given the following instructions:

A subject has to take a series of exams. Each of these exams should be evaluated as: very good, good, middle, bad, very bad. If the subject scores very good or good, he'll pass the exam. If he scores bad or very bad he will not pass the exam. If he scores middle he will have to repeat the exam.

Each exam is evaluated with a new system of score; it receives:
- a colour (black, blue, white or yellow)
- a shape (heart, circle, square or rectangle)
- a dimension (big or small)
Then, for example, an exam could be evaluated as blue - heart - small
- The connection between shapes-colours-dimensions and the final result of the exam is not known

Then the game worked in this way (through a specific software):
- the subject was presented the first combination of shape-colour-dimension,
- he had to choose one of the possible scores,
- he was informed whether his answer is right or not and which is the right solution;
- he was presented another combination;
- and so on for 231 rounds.

Between the combinations of information and the right answer there is a logical relation, stable for all rounds, described in Table 1[3].

Table 1. Logical relation between combinations and results

Score	colour	shape	Dimension
very good	bright (white and yellow)	not angular shapes (circle and heart)	Big
good	bright (white and yellow)	not angular shapes (circle and heart)	Small
middle	dark (black and blue)	not angular shapes (circle and heart)	small or big
	bright (white and yellow)	angular shapes (square and rectangle)	
Bad	dark (black and blue)	angular shapes (square and rectangle)	Small
very bad	dark (black and blue)	angular shapes (square and rectangle)	Big

At the beginning of the session, players do not know the relation between information and results. So they have to find it.

As the game is based on a relatively small number of sequences of information (there are 4*4*2=32 different combinations of shapes-colours-dimensions, and the same combination appears many times during the game), subjects might try to remember them. In this case, the results of the experiment would be scarcely interesting, more artificial and conditioned by the fact of being in a laboratory. Therefore subjects were not explicitly told that there were fixed and repeated sequences (while they were informed of the existence of a logical relation between information and results). Remembering 32 sequences is, besides, not easy.

[3]. The categories of color and shape are not explicitly used in the presentation of the game to the players.

This experiment presents a few similarities with those belonging to the so-called "diagnostic task" series. Kelley and Friedman (1999), for example, asked the participants in their experiment to forecast the price of a good. The price was determined according to a linear stochastic process with two independent variables. Players know the model, but ignore its parameters. In each of the 480 rounds of the experiment, they have to forecast the price, after seeing the changing values of the independent variables. After each round they get information on the real value of the price.
Kelley and Friedman's experiment is built to test the least square models of learning (see, for example, Marcet and Sargent, 1989), where economics agents are expected to be able to estimate the parameters of a known models. According to their goal, then, Kelley and Friedman test the accuracy of their players' implicit estimates.
In Novarese and Rizzello (2002) there is a quite different idea of learning, seen - as in the psychological literature - as the human capacity to modify behaviour in a more or less permanent way, whenever new experience is acquired.
Then, as a first point, experimental subjects do not know the model used in the experiment (as in the real world usually happens), but they have to create their own knowledge. Moreover, subjects are not expected to understand the real relation in table 1, as, in our view, there is learning even if they develop a wrong system of rules. Even a wrong model of the world, in fact, allows them to face a the given environment.
Data are then analyzed in a different manner. There is no model to test and there is no analysis of the convergence on true parameters. The aim is mainly that of collecting empirical evidence and looking for eventual regularities among subjects, at the most guided by some theoretical hypothesis.

2.1.3. Results

The experiments were realised at the Centre for Cognitive Economics in Alessandria in October 2000. Sixty-four subjects (students in Law, with no Economics background) participated[4].

Three main aspects can be detected among subjects:
1) the tendency to confirm wrong answers to given combinations of information;
2) the tendency to change right answers, even after many right choices;
3) the tendency to develop (right and also wrong) rules.

These features characterize all subjects, but their specific content differs. In other words, not all subjects confirm the same wrong choice or change the same right one.

Table 2. An example of confirmation of wrong answers and change of right choice

round	colour	shape	dimension	correct result	subject's result
9	yellow	rectangle	Big	Middle	middle
16	**yellow**	**square**	**small**	**Middle**	**good**
33	**yellow**	**square**	**small**	**Middle**	**good**
43	yellow	rectangle	Big	middle	very good
50	**yellow**	**square**	**small**	**middle**	**good**
60	yellow	rectangle	Big	middle	very good
67	**yellow**	**square**	**small**	**middle**	**good**

[4] Similar results were obtained in a pilot experiment realized in March 2000 in Alessandria with a smaller group of subjects remunerated with money. The same experiment, but with a different frame, was replicated in July 2002. All main findings have been confirmed and new evidence has emerged

Table 2 helps to understand the first two points.
As said, a single combination of shape-colour-dimension appears many times during the game (never in two rounds running).
By taking into account all the answers a subject gives to a particular combination, we can count the number of wrong answers and the number of times wrong choices are confirmed.
For example, the sequence yellow-square-small (bold type in table 2) appears four times in the part of game showed. The participant in the experiment exemplified here confirms the same wrong answer after the three errors in the table. Then, we can say that he confirms 3 errors after 3 wrong choices (100% of confirmation). It's possible to calculate a similar rate for all the combinations faced by a player and to determine a mean individual value.
Table 3 shows the distribution of this rate among players[5] and the relevance of the phenomenon under exam.

Table 3: Distribution of the mean ratio of wrong answers confirmed by players

0%-25%	13
25%-50%	38
50%-75%	13
75%-100%	0
	64 players

Table 2 shows also the sequence yellow-rectangle-big. At its first appearance, player 1 gives a right answer (round 9) but in the following one he makes a wrong choice. Then we can say that (at round 43) he does not confirm a right solution.
Many other players show a similar behaviour in many cases and for different combinations. There are players who give a right answer to the same sequence for four or five (or even more) rounds and then change it, till the end of the game.
Table 4 shows, for the whole of players, the number of right answers "forgotten" in at least two following appearances. For example there are thirteen cases in which some players give the right answer to a given sequence in four subsequent rounds but later, in the following two (or more) appearances of the same information, they make a wrong choice.

Table 4: number of right answers "forgotten" in at least two following appearances related to the number of right answers in following appearances

1	332
2	85
3	27
4	13
5	7
6	5
7	0
8	2

It is useful, now, to concentrate on the last part of the game, and see if players develop regularities of behaviour and if they understand the rationale of the game.
As a first step, it is necessary to give a definition of rule (specific for this experiment). Take into account the third part of the game (the last 77 rounds), where many sequences appear three or four times.
For each sequence we can count how many times a subject gives the same answer:
- "rule 75" means that the subject gives the same answer 75 percent of the times;

[5] As shown in Novarese and Rizzello (2002), these results can hardly be explained with subjects' bad memory.

- "rule 100" means that the subject gives the same answer 100 percent of the times.
The idea is that if a player gives always or almost always the same answer when faced with the same sequence, he has probably developed a kind of routine[6].

There is wide heterogeneity among players, but most of them develop very routinized behaviour. At a mean level, 70 percent of the possible "rule 100" are developed. Some players develop all possible rules.
Does this mean that most of the subjects have understood the game? No, because players tend also to develop wrong rules (15 or 16 percent of the total), as shown in table 5.

Table 5 Percentage of rules (total and wrong) developed by experimental players, mean values*

	rule 75	Rule 100
total percentage of rules developed	75%	70%
Percentage of wrong rules developed	16%	15%

* Values are computed only for the sequences that appear at least four times during the whole game

It is interesting to see that the mean time taken to answer to "rule 100" (wrong or right) is lower than the overall mean time. The mean time for a reply to a rule is, in fact, about six seconds, while the overall mean time is about ten seconds. A shorter reaction time is one of the features that denote a routine (Cohen et al, 1996).

2.1.4. Understanding the empirical findings

Is it possible to understand why players behave in this, apparently strange, way? Why do they seem to act so irrationally? Are the three tendencies related?
To try to give an answer to such questions, it is necessary to look at each player. A generalized analysis is, in fact, incomplete, as the development of each experiment is very different from the others. An individual analysis, instead, in many cases seems to allow a full comprehension of the cognitive processes undertaken by the player.

Take into account the player in table 2. As seen, he confirms several times the wrong answer “good” to the sequence yellow-square-small. He does not confirm the right answer to the sequence yellow-rectangle-big. Table 6 helps us to understand the link between these facts. It shows the answer given by this individual to some selected sequences in the first 77 rounds (period I) and in the last 77 ones (period III).
For example, in the first period, the subject answers in all cases (100 percent) "very good" to the sequence yellow-circle-big. This sequence appears three times in the period (as we can see in the last column, where freq is equal to three). The coloured cell indicates the right answer.
The answer "good" (confirmed even if wrong) to the sequence yellow-square-small is coherent with the routines developed (from the very beginning) for the sequences yellow-circle (or heart)-small. On the contrary, the right answer (not confirmed) "middle" for the sequence yellow-rectangle-big is not coherent with the system of rules emerging in the last part of the game, when "square" and "rectangle" are compared to "circle" and "heart". The system of rules developed (also the part which is not shown here) is then wrong but has an evident and intelligible internal coherence.
Table 6 shows also answers to the sequence where the colour black appears. There is an evident analogy with sequences characterized by yellow. The sequences with blue (not shown) are similar to the black ones and the white ones are analogous to those with yellow[7].

[6] As the game is very long and repetitive, it seems reasonable to think that a subject who has developed a rule can make a mistake and give, in some cases, a different answer. Moreover, the rule could be in evolution and not perfectly defined at the beginning of what we define third period. That is the reason why we also use also a definition like "rule 75".
[7] It is evident, now, that the subject under exam is reasoning and not simply memorizing.

Table 6. A more detailed analysis of the behaviour of the player analyzed in table 2.

period	Colour	Shape	dim.	Very bad	Bad	Middle	Good	Very Good	freq
I	yellow	circle	Big					100%	3
			Small				100%		5
		heart	Big					100%	2
			Small				80%	20%	5
		square	Big						0
			Small				100%		4
		rectangle	Big			50%		50%	4
			Small						0
	black	circle	Big						0
			Small						0
		heart	Big						0
			Small						0
		square	Big	50%	50%				4
			Small	67%	33%				3
		rectangle	Big	75%		25%			4
			Small		40%	20%	40%		5
III	yellow	circle	Big					100%	3
			Small				100%		4
		heart	Big					100%	1
			Small				100%		3
		square	Big					100%	1
			Small				100%		4
		rectangle	Big					100%	3
			Small				100%		1
	black	circle	Big	100%					1
			Small		100%				2
		heart	Big	100%					2
			Small		100%				1
		square	Big	100%					3
			Small		75%	25%			4
		rectangle	Big	67%	33%				3
			Small		100%				4

This is just an example, but a similar system of rules emerges in many cases, and it is probably related to the order of appearance of the information during the experiment. In fact, at the beginning of the game there are many cases of bright-not angular shapes and of dark-angular shapes, i.e. the ones that receive a higher number of right answers. In this perspective, then, this result is coherent with the hypothesis of path-dependence in individual learning, stated at a theoretical level by Rizzello (2000) and by Egidi and Narduzzo (1996) in experiments on team organization[8].

Apart from those who understand the right logical system, there are players who develop a set of rules that is partially different from the one in table 6 (in other words, they build a different knowledge of this experimental environment), though they start from the same configuration of the game and on the basis of similar cognitive devices.
Consider in table 7 the final part of the game of another player.

[8] As some results could be context-specific, it is very important to find coherent results in different experiments.

Table 7. The system of rules developed by a player different from those in table 6*

period	Colour	Shape	dim.	Very bad	Bad	Middle	Good	Very Good	freq
III	yellow	circle	Big					**100%**	3
			Small				**100%**		4
		heart	Big					**100%**	1
			Small				**100%**		3
		square	Big				**100%**		1
			Small			**100%**			4
		rectangle	Big			33%	***67%***		3
			Small			**100%**			1
	black	circle	Big			**100%**			1
			Small		**100%**				2
		heart	Big		***50%***	50%			2
			Small		**100%**				1
		square	Big		**67%**	33%			3
			Small	50%	***25%***	25%			4
		rectangle	Big		**67%**	33%			3
			Small		**100%**				4

*The bold character shows the last answer given to each sequence

In table 7 there is more variance in the answers, but there are also some clear differences from table 6. Black (and blue, not shown in table) is generally associated with the answer bad (the last choices shown in the table in bold character enforce the idea that this is the path of convergence), while the sequences with a yellow are almost always understood.

Both systems of rules in table 6 and 7 are not optimal, as in some cases lead to wrong choices. Both of them have, yet, an important feature, as players' memory is bounded, like their mental energy: they are easy to remind and apply.

2.2. An experiment on team learning

This section describes the results of an experiment on team learning (see Novarese 2003 for more details and results). The experiment under exam is called "sum 10", and aims at analyzing if and how teams reach coordination in a situation of partially opposite incentives. "Sum 10" presents elements from games on both organizational learning and cooperation (trying to go beyond the simple traditional prisoner's dilemma).

Two aspects emerge:
- a strong effect of individual psychological traits (as measured by personality tests) on the behaviour in the experiment;
- the usefulness of the interview realized at the end of the experiment to understand players' behaviour, motivations, aims …

Both this points question, again, the ceteris paribus assumption.

2.2.1 The experiment

Team of three players are, anonymously and at random, built among participants. The game has 36 rounds. Teams are stable throughout game.
Each of the players has a set of numbers. This set remains unchanged in every round and is composed of the values: 0, 1, 3, 4, 10.
In every round each player has to declare one of the number in his set.

The numbers of the three people playing together are then summed.

According to the sum, each player receives a payoff, following this rule:

* if $S(i) = 10$, $I(i)=40 - D(i)$
* if $S(i) > 10$, $I(i)=30 - D(i)$
* if $S(i) < 10$, $I(i)=0 - D(i)$

where
$S(i)$= sum of the team I, of which player i is a member
$I(i)$= player i: individual payoff
$D(i)$: number declared by player i

Then subjects have to cooperate to reach a sum equal or higher than ten, by using the lowest possible number.
In treatment group one, players know the total number of rounds; while in treatment two they do not receive such information.

The players' behaviour can be influenced by many, interrelated elements. These considerations have general appliances to almost all experiments, but for some of them (as this one) they are probably more important.
- The first and obvious element is their actual comprehension of the experiment. Even if the game is not difficult, some players show very strange behaviours and sometimes, the interview realized after the game show their wrong comprehension or interpretation of the game (apparently this problem did not emerge in the example shown in the previous paragraph).
- Subjects can participate in the experiment with many and different motivations. The introduction of a remuneration (as prescribed in the most important handbooks, like Hey, 1991) is supposed to stimulate at least a reasonable degree of care, since subjects should have the same motivations which are presumed in the theory to test.
Sometimes this reward can be unnecessary. In fact, important results (for example Allais, 1956) have been reached in experiments with no reward. Sometimes students (the main experimental subjects) attend experiments with motivations different from monetary reward, for example their interest in the subject or just their curiosity. Besides, during the game they can be involved in the accomplishment of that peculiar task, independently of the reward. But these facts, inevitably, lead to a loss of control, at least in relation to the standard view of the laboratory. Besides, experiments are usually similar to games, and people like games. It is not impossible that they react and behave differently from experimenters' and theories' hopes (making more points as possible and getting the highest reward). On the contrary they can play just for fun or even to experiment the behaviour of their partners.
- The attitude toward other players, that can be seen in terms of dichotomies: altruism vs. egoism, free riding vs. giving, rationality vs. emotionality and desire of being reciprocated[9].
The attitude toward other members of the team (and, as a consequence, altruism or egoism) can depend on each subject's motivation in participating in the experiment. A player interested in "making money" will probably behave differently from one who is enjoying the game.
- Subjects should be able to coordinate their choices by communicating their own eventual strategies and/or understanding those of the others. Using the language of game theory, we could say that the experiment has many possible equilibria (like a game or a super-game). Subjects should coordinate to reach one of these.

[9] There is an extensive theoretical and experimental literature on this subject. It is not possible to analyze it here. For a review see Novarese (2002a and 2002b).

All such (different) behaviours and motivations emerge from the recalled interview conducted before the end of the experiment. Usually experimental papers seem to neglect this problem, though it has important effects on the behaviour during the game. We will analyse this issue in detail in the next paragraphs.

This experiment was realized in Alessandria, at the Centre for Cognitive Economics, in May 2000; 36 subjects (students in law) participated. The software used was developed by Swiee (Boero, 2001)

2.2.3. Psychological traits and behaviour in the experiment

The effects of personality traits on individual economic behaviour have been studied, for example, in relation with egoist/altruist behaviour by Boone et al (1999). They propose a survey and new empirical evidence leading to the conclusion that the personality of the players matters. In this and other papers, players' personality is analysed with appropriate psychological tests and related to the behaviour in the experiment (usually the Prisoner's dilemma or similar games). In many cases a relation between attitude to cooperation and psychological traits emerges.

A similar analysis is proposed here, connecting the behaviour in the game "sum 10" with individual extroversion, measured with Eysenck and Wilson' test (1975). This aspect of human personality, as recalled by Boone et al (1999), could influence behaviour in experiments on cooperation.

Box 1 describes Eysenck and Wilson's test. Box 2 presents the empirical analysis.
The most important result is that introverted individuals get a higher mean score in the game. The variable "extroversion/introversion" accounts for about 25 percent of the variance in the score among individuals. Then this individual psychological trait has an important role in determining individual (but also organizational) performance in this experiment.

Box 1

Eysenck and Wilson's test measures extroversion and introversion along seven dimensions, defined by the following dichotomies:
- dynamism/passivity
- sociability/unsociability
- boldness/caution
- impulsiveness/self-control
- expressiveness/inhibition
- practicalness/ reflectiveness
- irresponsibility/responsibility

For each of this dichotomy, the test, based on 210 questions, proposes a score ranging from zero to thirty. As specified by the authors, the individual score should be compared with those of a reference population. The absolute level of the variables has, in fact, for itself, low significance. According to the position in each of the seven distributions, it is possible to state whether an individual is extroverted or introverted.
Introversion is related to the predominance of high values of the first term of each of the dichotomies.

Box 2

Subjects were first assigned to one of the side of the distribution (above or below the appropriate median) for each of the seven dichotomies in box 1. Then subjects with values above the median in the distribution of at least four of the seven variables, were classified as "extroverted".
This procedure has obviously its limits (for example, does not account for possible differences in the variables that define extroversion) but it is coherent with the ideas of Eysenck and Wilson. Besides it allows us to build two groups composed of individuals with significantly different psychological traits.
The group of extroverted individuals was composed of twelve subjects (six from each of the treatment).
Different analyses are possible on the two groups. Here attention is focused on differences in the score (i.e. the main measure of performance) realized in the experiment.
Table 8a. and 8b. propose the results of the analysis of the variance aimed at comparing mean scores among extroverted and introverted subjects, controlling for the treatment group.
The following equation was then estimated:
score(i)= intercept+ a * introversion + b * treatment (I)

where:
- extroverted = 1 if individual is classified as "extroverted"; 0 elsewhere
- knowing_rounds = 1 if treatment group is 1 (subject knows the number of rounds); 0 elsewhere

Table 8a. Analysis of the variance for the estimation of equation (I)

	df	*sum of squares*	*mean square*	*F value*	*P value*
Model	2	434118.8	217059.4	4.792218	0.01
Error	29	1313530	45294.14		
Total	31	1747649			

Table 8b. Parameters estimation for equation I

	Coefficient	*Standard error*	*Stat t*	*P value*
Intercept	1079.76	60.66	17.79	0.0001
Introversion	-29.13	77.71	3.07	0.0045
knowing_rounds	-29.125	75.24	-0.39	0.7015
R squared=0.25				
standard error=212.8				

The group composed of extrovert individuals has a mean score significantly lower than the introverted one. On the contrary, there are no significant differences among treatments related to the information on the number of rounds.

At a first glance, this result might be surprising. Extroversion, in facts, is supposed to be associable to a tendency to cooperate. In this experiment, extroverted get worse results. It is yet necessary to observe that "sum 10" is also a game of coordination. Some attitudes of extroverted individuals (such as activity, impulsivity, irresponsibility …) can hinder the reaching of this goal. These individuals, in fact, can try to impose their own focal points and strategies, or react in a "stranger" way to others' choices. Introverted individuals are, on the contrary, more adaptive and more prone to understand the behaviour of other players (in certain respects, it is comparable to a rational

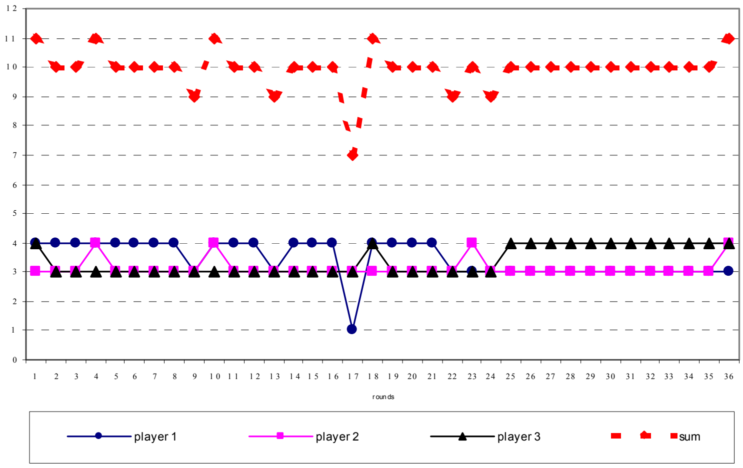
behaviour). Besides, as in this game there is no direct communication, extroversion might result less useful.
In other kinds of organizational tasks, personal psychological traits can have different effects on the performance.

2.2.4. Interviews on the behaviour in the game

After the end of the game "sum 10", players were required to describe, in a written interview, their strategies, their perception of the other members of their group (partners or rivals?) and of their behaviour.
While similar interviews are probably quite common among experimentalists, there are no standards in their realization and presentation. Normally, then, the results are not proposed at all in the papers[10].
Why? The reasons are probably related to the recalled applications and aims of experimental economics in the traditional approach. If scientists are interested in testing theories and their validity, moving from the paper to reality, interviews become useless. Their interest, in fact, is not focused on understanding players' behaviour, but on testing a theory or hypothesis in a more realistic environment.
The problem is that differences among players (even if not under analysis) can make the same test of theories more difficult, introducing variance and "errors of measure".
Consider the following examples.

Graph 1 shows one of the team playing the game, reporting individual choices in every round and the relative sum.
The team under exam gets a sum equal to ten only in a few rounds of the game. In other rounds they go above or below this number. In the interview, yet, all players express their satisfaction, in quite similar words: "we were able to reach ten in some of the rounds, without using costly strategies (i.e. strategies requiring to declare ten)". Then, these players' expectation were not those of reaching always the maximum goal. This is an important information that can be useful to explain their performance in relation to other groups, which were able to reach ten for almost all the game. These other players, in fact, show generally very different goals, declaring sometimes regret for the few rounds in which they have not been able to get the best result[11].

Graph. 1 The game played by one of the team in the experiment "sum 10"



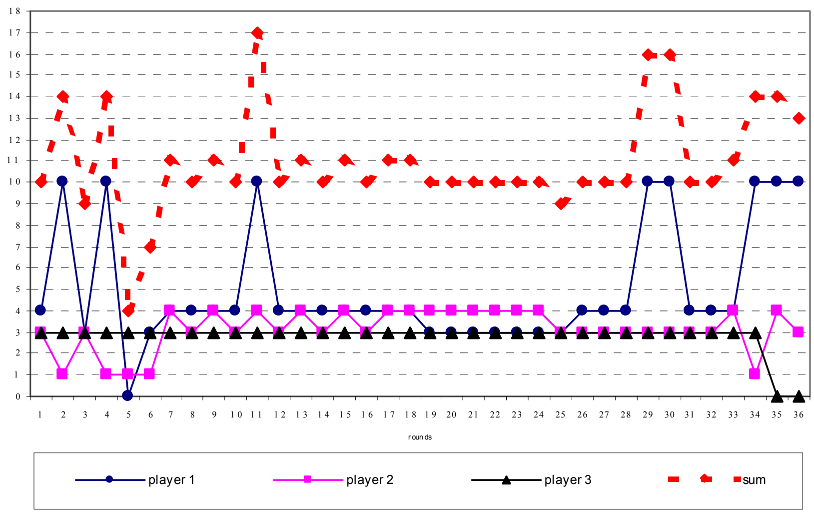
[10] Andreoni, 1995 (p. 898) talks about this kind of interview, but in a methodological paper.
[11] In the interview, player one in graph 1 explained also the reason for the choice at round 17. After having played three or four, he decided "to save some points". This apparent strange choice, thus, is not a signal sent to other members of the team, or a way to communicate new strategies, and it is not an error.

Many problems and risks are related to the use of this kind of interviews. In relation to the example: subjects might have changed their level of expectation during the game, and in the interview they might have expressed only their final opinions (it would be necessary[12] to test satisfaction also before and during the game). Besides, some players' statements result completely useless. In other cases it is not easy to link interview and real behaviour (because people did not express their ideas clearly). Then, there could be a possible distortion in using only the part of the interview that seem useful.
But it seems undeniable that sometimes these statements give useful information and allow a better comprehension of players' behaviour in the game; therefore it would be absurd to ignore them.

See another example in graph 2. The main obstacle to the reaching of a stable equilibrium, for this team, seems the behaviour of player 1. He changes his strategies many times and break the equilibria. The interview gives interesting insights. His strange behaviour originates from the willingness to play and reach the goal in different ways and not just in one. He is not so interested in getting top scores, but he likes playing, and he tries to avoid repetitiveness.
In other cases, not shown here, the reasons at the basis of the difficulties in getting a stable equilibrium rely on the willingness of some players to reciprocate others. These players cannot accept that one of the other members of the team is forced to play a more costly strategy.

These examples show the variety of motivations and the usefulness of interviews (as, like psychological tests, they can help to better understand players' behaviour and also to create appropriate "control variables") and so the necessity to develop scientific criteria and appropriate methodologies to perform them (on this, see also Simon, 2000).

Graph. 2 The game played by another team in the experiment "sum 10"



3. Experiments and simulations

While there are relatively few papers focused on "cognitive experimental economics", in the last years there has been a growing diffusion of papers on agent-based simulations (Gilbert and Terna, 2000).
Sometimes the term experiment is used also for simulations with artificial agents, as they shows many features usually related to scientific empirical analysis. In fact, these models allow observing emergent unpredictable phenomena, starting from a population of evolving and learning agents.

[12] It would also be possible, but it would increase the management problems of the experiment.

Once the modeller has fixed initial conditions and built the environment, he has just to observe without intervention, as in a "culture-dish laboratory experiment" (Tesfatsion, 2002, p. 2).
The interrelation between experimental economics and agent based computational economics has increased because of the frequent use of common sets and games (the classical example are the tests of Axelrod, 1984, on prisoner's dilemma, widely carried out both with human and artificial agents).

As simulations are today easier to realize than experiments with human agents, and as they also show some advantages (for example there is no risk that players might misunderstand the rules of the game or have unpredictable motivations), it is not difficult to understand why they are so diffused. And the following question seems reasonable: why not using simulations instead of experiments?
The answer is easy: simulations are not the same as experiments, and artificial agents are not human agents.
Tesfatsion (2002) analyzes some of the possible risks and problems related to simulations. Two of them should be remembered here. The results obtained in simulations can be strongly affected by the representation of learning processes and by the parameters-values setting used. There are inevitably many risks. For example, as "many of these learning representations were originally developed with global optimality objectives in mind" (Testfation, 2002, p. 3), results could be the same as those of neoclassical models. In fact, if agents (real or artificial) were similar to the traditional homo oeconomicus, experiments would be useless, as the same results of the theory would constantly emerge.

Therefore, experiments and simulations are not alternative ways of gathering data. On the contrary, they can be usefully integrated. Tesfatsion (2002) suggests that experiments can help fixing parameters and decision models to be used in simulations.
As we have seen in the previous chapter, experiments always show unexpected results related to participants' behaviour. Some of the related heterogeneity can be considered as a noise and it could be useful to eliminate it. In other cases heterogeneity depends on individuals' peculiarity, and it is an unavoidable part of the world.
Only by means of a really empirically-founded research program these features can be recognised. If simulations want to start from reality, and not from a hypothetical world (even if it is different from the neoclassical one), the cooperation with experimental economics is necessary. Moreover, the analytical power of simulations can be fully expressed if it is used to account for the variety and complexity of the real world (as shown by experiments), which is impossible to manage through usual models.

On the other hand, simulations can also help to understand and read experimental data, comparing hypothetical agents and real ones (Tesfatsion, 2002).
But there is also another way of integrating experimental economics and simulations, exemplified by an application of "sum 10" performed at the Turin University, in collaboration with the Centre for Cognitive Economics and described in Novarese (2003). In this case artificial and human agents play together. Artificial agents are used to create specific situations and training environments for humans. For example, in the application recalled, a group of subjects played the first part of the game with two artificial egoist agents. Another group interacted with artificial altruist agents. In this way, it is possible to evaluate possible differences in the performance, by carrying out later a game in which all players interact together.

4. Conclusion

Cognitive economics imposes and suggests a series of methodological and instrumental novelties in the realization of experiments and in the analysis of data.

Some of these novelties interest mainly the works related to this approach; others have a more general appliance.

One of the main aspects that emerged here, in different ways, is a definite heterogeneity among individuals. This heterogeneity depends partly on the fact that experiments are artificial situations carried out in a laboratory and subjects react to this environment in different ways (in term of motivations and aims, for example).
Other differences are, instead, a reflex of the nature of human beings. Laboratory is not a neutral place. Subjects bring with them their personal and idiosyncratic knowledge, ideas, experience … The laboratory environment cannot eliminate these differences.
All these factors cause a loss of control, as they question the validity of the ceteris paribus condition. Subjects are, in fact, different, even when acting in identical experimental conditions.
If we are interested in testing whether a theory resists when we pass from the books to the real world, we can probably neglect these problems; but we have to remember them if we want to explain why people behave in a particular way.

Complex environments, instead of simple games, are also necessary in order to advance in the understanding of human behaviour.
Inevitably these experiments are sometimes less controlled and also more difficult to manage. Data gathered in this way are necessarily more difficult to read. Yet, such experiments give rise to new opportunities, like that of discovering unplanned events (if researchers are opened to look at them). As well known, also penicillin has been discovered thanks to an accident!
Often the apparent "noises", which emerge in an unexpected way, are very interesting. The satisfacing behaviour manifested by some experimental players could create problems, but it shows, again, the validity of Simon's model. Apparently strange choices make it very difficult to understand and read data, but are sometimes the result of innovative tendencies induced by boredom of repetitive behaviour (see Witt, 1993 for a theoretical analysis of this mechanism).

We have seen that an individual analysis of data, especially if guided by psychological tests and post-experiment interviews, can help understanding these phenomena. But these new methodologies and procedures need to be studied, fully understood and developed. More analyses and specific tests, will probably be necessary to make these novelties accepted by the scientific community, but their indications are already so strong that they cannot be neglected.

References

Allais, M. (1953) 'Le comportment de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école americaine', *Econometrica*, 21: 503-546.

Andreoni, J. (1995), 'Cooperation in Public-Goods Experiments: Kindness or Confusion? ', *The American Economic Review*, Volume 85, Issue 4: 891-904.

Axelrod, R. (1984), *The Evolution of Cooperation*, New York: Basic Books.

Boero, R. (2001), 'SWIEE a Swarm Web Interface for Experimental Economics', in F. Luna and A. Perrone (Eds.) *Agent Based Methods in Economic and Finance: Simulations in Swarm*, Kluver Academic Publishers.

Boone, C., De Brabander, B. and van Witteloostuijn, A. (1999), 'The impact of Personality on behavior in five Prisoner's Dilemma games', *Journal of Economic Psychology*, 20: 343-377.

Cohen, M.D and Backdayan, P. (1991), *Organizational Routines are Stored as Procedural Memory: Evidence from a Laboratory Study*, University of Michigan, mimeo

Cohen, M.D., Burkhart, , Dosi, G., Egidi, M., Marengo, L., Warglien, M., Winter, S. (1996), 'Routines and Other Recurring Action Patterns of Organizations: Contemporary Research Issues', *Industrial and Corporate Change*, 5 (3): 653-698.

Egidi, M., Narduzzo, A., (1997), ‘The emergence of path-dependent behaviours in cooperative contexts’, *International Journal of Industrial Organization*, 5: 677-709.
Eysenck, H.Y., and Wilson, G (1975), *Know your own Personality*, London: Temple Smith
Franck, R.H., Gilovich, T. Regan, D.T. (1993), ‘Does Studing Economics Inhibit Cooperation?’, *Journal of Economic Perspectives*, 7: 159-171.
Gilbert, N. and Terna, P. (2000), ‘How to build and use Agent-Based Models in Social Sciences’, *Mind & Socierty*, Issue 1, Vol. 1: 57-72.
Hey, J.D. (1991), *Experiments in Economics*, Oxford: Blackwell
Hey, J.D. and Dardanoni, V. (1988), ‘A Large Scale Experimental Investigation into Optimal Consumption under Uncertainty’, in *Economics Journal*, 98: 105-116.
Innocenti, A. (1995), *Le origini della Experimental Economics (1948-1959): una valutazione critica sull'evoluzione delle metodologie sperimentali*, working papers, Università di Siena.
Kahneman, D. and Tversky, A. (1992), ‘Advances in Prospect Theory: Cumulative Representation of Uncertainty’, *Journal of Risk and Uncertainty*. 5: 297-323.
----(1979), ‘Prospect Theory: An Analysis of Decision under Risk’, Econometrica, 47: 263-91
Kelly, H. and Friedman, D. (1999), *Learning to forecast Price*, University of California Santa Cruz, Economics Department working papers.
---- (1992), ‘Advances in Prospect Theory: Cumulative Representation of Uncertainty’, *Journal of Risk and Uncertainty*, 5: 297-323.
Knight, F.H. (1921), *Risk, Uncertainty and Profit*, London: The London School of Economics and Political Science
Marcet, A. and Sargent, T (1989), ‘Convergence of least squares learning mechanisms in self referential linear stochastic models’, *Journal of Economic Theory*, 48: 337-68.
Novarese, M. (2003), *Some experiments on team learning*, working papers, Dipartimento di Scienze Giuridiche ed Economiche, forthcoming.
---- (2002a), Altruismo ed egoismo in economia e nelle altre scienze sociali, Dipartimento di Scienze Giuridiche ed Economiche.
---- (2002b), Lealtà, apprendimendo ed innovazione nella teoria eterodossa dell'impresa, Dipartimento di Scienze Giuridiche ed Economiche.
---- and Rizzello, S. (2002), ‘A cognitive approach to individual learning: some experimental results’, *Revue d'Economie Politique*, Forthcoming
---- (1999), ‘Origin and Recent Development of Experimental Economics’, Storia del Pensiero Economico, n. 37: 201-234.
Rizzello, S. (1999), The Economics of the Mind, Edward Elgar, Aldershot.
Roth, A.E. (1993), ‘On the Early History of Experimental Economics’, in *Journal of the History of Economic Thought*, 15: 184-209.
Simon, H.A. (2000), ‘Bounded Rationality in Social Science: Today and Tomorrow’, *Mind & Society*, Issue 1, Vol 1: 25-40
---- (1992), Colloquium with H. A. Simon, in M. Egidi and M. Marris R. (Eds.) *Economics, Bounded Rationality and the Cognitive Revolution*. Aldershot, U.K.: Elgar.
Tesfatsion, L. (2002), Agent-Based Computational Economics, ISU Economics Working Paper No. 1, Department of Economics, Iowa State University

Acknowledgments
The experiments discussed in this paper were realized with the contribution of many people. Special thanks to Riccardo Boero, Salvatore Rizzello, Pietro Terna, and Cesare Tibaldeschi.