

## THEORY OF REASONING BY GOALS

\*

NICOLA D'ALFONSO

Independent Scholar, Italy  
nicola.dalfonso@hotmail.com

*The purpose of this paper is to explore the possibility of introducing a new theoretical model of deductive reasoning into the psychology of reasoning. This new theoretical model, which we will call here the theory of reasoning by goals, has as its main property to make every characteristic of deductive reasoning depend on the goal employed by an agent to be able to reason. The theoretical context within which this goal must be framed is therefore that of the agency, here understood in a generic way as the ability that each individual has, in his or her role as agent, to pursue his or her own ends through the achievement of specific goals.*

*Keywords: deductive reasoning; agency; goals; reasoning by goals.*

## **1. Introduction**

In the first part of this paper the theory of reasoning by goals will be presented. The initial paragraph of this first part will be dedicated to explaining how the goals of the agency should be understood, making it clear why the theory of reasoning by goals can be a theory of reasoning without having to be an agency theory. The next paragraph will be dedicated to show how the line of thought that led to the development of the theory of reasoning by goals is that according to which theoretical reasoning and practical reasoning are similar modes of reasoning. It being understood that in this field the role of practical reasoning will be limited to what has just been said, that is to say, to being that something whose characteristics have inspired the development of the theory of reasoning by goals. The aim of this paper is in fact to explore the possibility of introducing a new theory in the field of deductive reasoning and not to deal with practical reasoning. The remaining paragraphs will instead be dedicated to describing the main characteristics of the theory of reasoning by goals.

In the second part of this paper the theory of reasoning by goals is compared with the main theories of deductive reasoning within the psychology of reasoning. The initial paragraph of this second part will be dedicated to a brief presentation of these theories, indicating the limits that prevent each of them from being a unitary explanation of all the characteristics of deductive reasoning. The next paragraph will be dedicated to explaining the difference between these theories and the theory of reasoning by goals, indicating what allows the latter to be a unitary explanation of all the characteristics of deductive reasoning. The remaining paragraphs will instead be dedicated to show how the characteristics of deductive reasoning emerged from the main experiments in the literature are in line with the explanations that the theory of reasoning by goals can provide them.

## **2. Goals of the agency**

In this paper, the goals of the agency are defined as those mental tools that allow a human agent to identify the way in which he or she can pursue his or her ends. What is important to underline about the notion of goals of the agency just presented is its ability to apply to the main types of agencies that we can find in the literature. Let us see a short list of them here.

There is the agency that is based on the notion of intention. According to that agency, every agent has a proactive attitude, called intention, which enables him to perform actions, called intentional actions.

The intention can be understood (Dretske, 1988) as the combination of an agent's desire and the belief through which that agent establishes how to pursue that desire. An example is as follows:

*(Desire) I want to avoid getting wet.*

*(Belief) I will avoid getting wet by taking an umbrella.*

*(Action) I take the umbrella.*

In this theoretical context, the goals of the agency will be constituted by these beliefs since the role they play is precisely to indicate to the agent how to pursue his desire. With reference to the example just shown, the goal of the agency will be constituted by the belief that taking the umbrella is a way for the agent to avoid getting wet.

The intention can also be understood (Enç 2003) as an autonomous state of mind that cannot be traced back to desires and beliefs, whose role is to causally produce the actions that a given agent uses to pursue his or her own end. In this theoretical context the goals of the agency will have to be identified within the mental representations of the mental state implementing the agent's intention. And specifically in that part of mental representations that indicate to the agent how he can pursue his end.

There is the agency that is based on volition. According to this agency, an agent's ability to perform actions derives directly and irreducibly from his volition.

The volition can be understood (Nathan, 1992) as the practical decision of an agent who is responsible for determining what to do. This notion of volition intends to fill the typical gap of the decision-making process, explaining why the decision does not seem to be an result automatically achieved by an agent starting from reasons that may be conflicting with each other, but rather it is an act that requires the agent's introspective participation in identifying the reasons that must prevail.

The volition can also be understood (Searle, 2001) as the ability to move from decision to action. This notion of volition intends to fill the typical gap of action initiation, explaining why the action does not seem to be a result automatically achieved by an agent starting from the decision to do something, but rather it is an act that requires the agent's introspective participation in identifying when it is time to act.

The volition can also be understood (McCann, 1998, p.140) as the ability to bring an action to completion. This notion of volition intends to fill the typical gap in the implementation of the action, explaining why the action does not seem to be a process that an agent performs automatically, but rather is a fluid and continuous activity that requires the agent's introspective participation in keeping the execution aligned with the intention.

Finally, the volition can be understood (Zhu, 2004) as a unitary phenomenon that has as different moments the three notions of volition seen previously. What makes these three notions of volition a unitary phenomenon is the fact that they are made up of mental actions able to satisfy our intuitive idea of acts of will.

In all these cases we can attribute to volition the intrinsic ability to bring an agent to the realization of its actions, regardless of whether these actions are the consequence of the causal history of that agent (Lowe 2008). In this theoretical context the goals of the agency will be identified as part of the content of the volition. And specifically as that part of the content that indicates to the agent how he or she can pursue his or her end, be it the need to make a decision, to start an action or to bring an action to completion.

There is the agency that is based on the causal action of the agent. According to this agency (Clarke, 2003), an agent's ability to perform actions derives directly and irreducibly from the agent himself. In other words, the agent is recognised as having the intrinsic capacity to perform those actions which are the direct consequence of his causal history. In this theoretical context the goals of the agency will be constituted by the intrinsic ability of the agent to identify how it can pursue its own end.

We can therefore conclude that within the theory of reasoning by goals it is possible to refer to the goals of the agency in a completely generic way without having to engage in one particular notion of agency rather than another.

### **3. Development of the theory of reasoning by goals**

In order to understand the line of thought that has guided the development of the theory of reasoning by goals, it is necessary to give a brief definition of what practical reasoning and theoretical reasoning are.

Practical reasoning can be considered a reflection on actions that have not yet happened, which is done in order to decide which actions would be the best ones to take or should be taken, and consequently it is a reflection on issues of value and what would be desirable. This reasoning is performed by taking a personal point of view and therefore requires agents to reason first-hand about the practical situation in which they find themselves.

Theoretical reasoning can instead be considered a reflection on events that concern the world, which is done in order to explain the events that have already occurred and to predict the events that may happen, and consequently it is a reflection on matters of fact and their explanation. This reasoning is performed by taking an impersonal point of view, and therefore in principle can be performed by any human agent.

Deductive reasoning is part of theoretical reasoning. Specifically it can be considered that particular theoretical reasoning in which the agent is called to identify a proposition which is the correct conclusion that can be deduced from propositions describing possible states of the world.

Although theoretical and practical reasoning seem to be very different from each other, various proposals can be found in literature that highlight their great affinities. Let us see some of these contributions.

According to Moran (2001) theoretical reasoning can be considered a reflection on normative and not factual issues, which is performed in first person in order to establish what is best to believe about how the world is made. From this point of view theoretical reasoning and practical reasoning are equivalent and differ only because they deal with different norms: practical reasoning deals with norms relating to action, theoretical reasoning with norms relating to belief.

According to Bratman (1987) theoretical and practical reasoning do not differ in their consequences. It is not true that the former produces changes in one's mental states and the latter in one's body movements because, in reality, in both cases there is a change in attitudes. In the case of theoretical reasoning attitudes are related to beliefs, while in the case of practical reasoning attitudes are related to intentions..

According to Broome (2013) theoretical and practical reasoning can be understood as an inferential process that takes into input attitudes and produces in output the formation or modification of other attitudes. Also here in the case of theoretical reasoning attitudes are related to beliefs, while in the case of practical reasoning attitudes are related to intentions.

According to Korsgaard (1996) even practical reasoning like theoretical reasoning can fail. And if in the theoretical field failing means reaching incorrect conclusions from certain premises, in the practical field it means failing to perform the actions that have been recognized as the best (weakness of will).

This paper can be seen as a further contribution to support the affinities between practical and theoretical reasoning. In this case the affinity would be given by the fundamental role that can be attributed to the goals of the agency. The practical and theoretical reasoning would be in this perspective of the actions that agents perform through the identification of precise goals.

That this applies to practical reasoning is nothing new. It is what Walton, Reed and Macagno (2008), for example, argue in showing that goals are necessary to select the actions of practical reasoning. However, this paper goes one step further by showing how the role that goals play within practical reasoning is the same role that goals play within deductive reasoning.

However, since the aim of this paper is not to argue in favor of the affinities between theoretical and practical reasoning, such an affinity will not be argued here in an organic way but simply suggested. And to suggest it, it will be shown how all the characteristics of the theory of reasoning by goals can be directly taken

from those that can be found within a concrete case of practical reasoning, such as the one used by an agent when playing tennis.

#### 4. Necessity of the goals

When we have to pursue an end such as playing tennis, that end answers the question: "What end are we pursuing?". Because the end we are pursuing concerns what we are doing but not how, it is not the kind of information we use to coordinate our movements. What coordinates our movements when we play tennis are the goals that answer the question: "How can we achieve our end in the particular situation we are in at the moment?"

This is because the basic movements at our disposal (this applies to us human agents but would also apply to non-human agents) are much more elementary than those required to achieve our ends. In other words, playing tennis is too general a concept to recall specific movements such as quickly stretching the left leg, rotating the ankle to give a certain direction to our left foot, widening the right arm, twisting the back to the right and so on. The specificity of these basic movements can be better connected to a type of information that concerns more closely the motor aspects of our body. And this type of information is precisely that which is provided by our goals, precisely because they are concerned with showing us how we can achieve our end. So, for example, when we see the ball arrive in our court, what allows us to select and coordinate our basic movements will not be the too general information constituted by the end of "playing tennis", but by the particular goal of "moving to get closer to the ball so that we can hit it with the racket".

Within the theory of reasoning by goals, taking inspiration from what we have just seen, it is considered the basic operations available to the mind to reason deductively too elementary to connect to what the agent is doing. Rather, it will be the way the agent will have chosen to do what he is doing that will allow him to select such mental operations, this being a type of information closer to the functionality of our mind. For this reason the theory of reasoning by goals gives the agent both a end that indicates the action to be pursued and an goal that indicates the way to perform it.

More specifically in the case of an agent who has to execute a task that requires him to reason deductively, his end will be given by that task. But "executing that task" will be too general information for him to use it to select and coordinate the basic operations available to his mind to reason deductively. He will therefore need to use a goal that tells him how such an end can be pursued.

## **5. Acquisition of the goals**

When an agent finds himself hitting a tennis ball with his racket for the first time, only two things can happen. Either he believes that the task can be accomplished in any way that appears suitable for purpose, or he believes he must accomplish it in a way that meets additional requirements, perhaps those he can learn from a specific master.

In the first case, he will believe that the way to hit the ball will be given by that combination of movements that appeared to him suitable for the purpose, having allowed him to hit the tennis ball with the racket. And so he will ultimately take as his goal an improvised shot that involves a grip of the racket that is the one he will have used in that circumstance, and an impact between the ball and the strings of his racket that is the one he will have used in that circumstance.

In the second case he will believe that the way to hit the ball will be given by that combination of movements that not only allowed him to hit the tennis ball with the racket but was in accordance with the instructions given to him by his tennis master. And so he will ultimately take as his goal a well-proven shot that involves a very precise grip of the racket, and a very precise impact between the ball and the strings of his racket.

Within the theory of reasoning by goals, taking inspiration from what we have just seen, it is considered that only two things can happen when an agent finds himself for the first time using deductive reasoning to perform tasks. Either he believes that such a task can be accomplished by any conclusion suitable for the purpose and therefore able to be a possible solution of his reasoning, or he believes he must accomplish it through a conclusion that in addition to being a possible solution of his reasoning meets additional requirements, perhaps those that he can learn from a specific master.

Goals of the first type will be defined as naive goals, while goals of the second type will be defined as educated goals.

Having established this, it is clear that in a society such as ours, in which people learn to reason freely, that is, without learning any specific model of reasoning, we will have to expect that most of them will find themselves reasoning by resorting to naive goals, in other words, settling for the first conclusions will appear to them to be fit for purpose. And therefore conclusions with the sole characteristic of being possible solutions of their deductive reasoning, without having had to meet further requirements capable of ensuring their correctness.

## **6. Effectiveness of reasoning**

Our ability to play tennis depends on how we have learned to coordinate and follow in sequence the various movements we use to hit the ball. And since it is the goal pursued by our agency that determines what movements we are making moment

by moment, we will have no difficulty in agreeing that our effectiveness in playing tennis will depend precisely on the goal we choose to use from time to time during the game. By making it clear that for each game situation there will always be a better goal than all the others, and precisely the one able to make us realize the most appropriate movements to face it. This can be said in another way. That is to say that if we take into consideration different tennis players, those who are better at it will be those who, all other things being equal, can count on better goals. And since it is more likely that a goal learned by a master is better than one learned by improvisation, it is reasonable to expect that among the best tennis players there will be more of those with educated goals than those with naive goals.

Within the theory of reasoning by goals, taking inspiration from what we have just seen, it is recognized to the goals the ability to determine the effectiveness of deductive reasoning. Making it clear that for every deductive reasoning there will always be a better goal than all the others, and precisely the one able to make us perform the most appropriate mental operations to deal with it. This can be said in another way. That is to say, if we take into consideration different agents engaged in deductive reasoning, those who are better at it will be those who, all other things being equal, can count on a better goal. And since it is more likely that a conclusion that satisfies further requirements is better than one that is merely a possible solution to the reasoning faced, it is reasonable to expect that among the agents who are best at reasoning there will be more agents with educated goals than those with naive goals.

## **7. Goals as end of action**

When we set ourselves a goal to play tennis, such as, for example, "moving to get closer to the ball so that we can hit it with the racket", the movements it allows us to perform stop at the precise moment when we have satisfied it. In other words, the composite movement that allows us to get closer to the ball does not continue indefinitely, but stops as soon as we get as close to it as required by our goal. At that moment, our goal exhausts its function and then decays allowing us to move on to another goal that we will need to deal with the new game situation that has come to be created, such as "hitting the ball in a certain way with our racket".

Naturally, our action of playing tennis continues goal after goal until we reach a goal at which the game ends, as happens when the last point has been played. In such a case it will not only be the last goal that we will have achieved that will decay, but also our end, allowing us to move on to another end that will serve to satisfy the new need that we will see emerge at that time.

Within the theory of reasoning by goals, taking inspiration from what we have just seen, it is recognized to the goals the ability to decay automatically when they exhaust their function, that is, when the agent performs the mental operations necessary to satisfy them. This means that as soon as the agent identifies the conclusion of his deductive reasoning, his goal decays and with it also the end that that reasoning sought to pursue.

In order to understand the consequences of the characteristic just outlined, we must refer to the naive goal, and therefore to cases in which the agent is willing to recognize as a conclusion any possible solution to his reasoning. This is because such a conclusion, not being able to count on additional requirements that can ensure its correctness, will always have the possibility of proving wrong. It follows that when an agent faces deductive reasoning with the naive goal he would do better to verify that the conclusion reached is actually correct. The problem is that as soon as the agent identifies a conclusion that he considers a possible solution to his reasoning, the reasoning in question comes to a halt, and the agent, thinking he has found the very solution he was looking for, will no longer have any incentive to verify its correctness.

## **8. A posteriori identification of the goals**

The first times we find ourselves playing tennis, the analysis of how we can pursue this end and therefore ultimately the identification of which goals to adopt is an integral part of our reflection, and as a result it is something we are still aware of. This occurs both when we try to learn how to play tennis by improvising and when it is a master to show us which movements to make.

Later on, when we become familiar with the goals to be adopted in various game situations, we no longer need to actively reflect on them. What happens at that moment is that we lose awareness of those goals because once they come to connect to specific game situations, they acquire the ability to automatically activate themselves in correspondence of such situations. So when we see the ball coming into our court at a certain distance from us, we no longer need to ask ourselves what goal to adopt because it manifests itself automatically and we simply follow it.

Although the goals are not part of our thinking when we have been playing tennis for a long time, except in cases where we face new game situations, we are still able to trace them, at least in the immediacy of the movements we made, when our short-term memory still allows us to reconstruct the game situation in which we found ourselves and the movements we made.

The fact is that when our movements satisfy the goal by which they are guided, we feel that the situation we find ourselves in is exactly the one we wanted to be in, and

where we feel ready to deal with our next goal. And it is thanks to this awareness that in mentally reconstructing our action we are able to identify the moment in which our movements have stopped being what we were doing and have instead become what we were pushed towards, indicating what our goal was.

For example, all we have to do to recognise our goal of getting closer to the ball when we play tennis is to reconstruct our movements in search of that precise moment when we reached the very situation we wanted to be in before moving on to the next goal. And consequently identifying that there was a moment when we moved closer to the ball as we wanted to and it needed to be, in order to be able to deal with the next goal of hitting it.

Within the theory of reasoning by goals, taking inspiration from what we have just seen, it is recognized to the goal the ability to activate itself automatically, thus coming out from the reflection of the agent engaged in reasoning. In other words, it is argued that once an agent has become familiar with deductive reasoning in general or with a certain type of deductive reasoning, he will be able to achieve it without having to ask himself how. This goal will in fact succeed in activating itself automatically, determining the characteristics that a solution must have in order for the agent to consider it as the conclusion of his reasoning. If it is the naive goal to activate itself, the agent will be pushed towards a conclusion that will have the sole characteristic of being a possible solution to his reasoning. If, on the other hand, it is a specific educated goal to activate itself, the agent will be pushed towards a conclusion which, in addition to being a possible solution to his reasoning, will also prove to be able to satisfy further requirements.

Always taking inspiration from what we have seen as valid for the game of tennis, within the theory of reasoning by goals is recognized to the agent the ability to trace back to the goal he has set himself reasoning, at least in the immediacy of reasoning. It is a question of attributing to the agent the ability to reconstruct the mental operations he himself performed in the short term, so as to identify the moment in which an information has ceased to be only an information, but has turned into the conclusion of his reasoning. Establishing in relation to which operation this step has occurred is in fact all that is needed for the agent to determine whether at its conclusion it was required to be only a possible solution to his reasoning (naive goal), or whether it was required to satisfy any further requirements (educated goal).

## **9. New theoretical and experimental perspectives**

The theory of reasoning by goals provides an interpretation of deductive reasoning that makes it particularly suitable for experimental analysis. For the simple reason that such an interpretation not only recognizes to the experimenters multiple possi-

bilities of intervention on the subjects of the experiments, but it also clarifies which are the most neuralgic aspects of deductive reasoning to be checked. In order to better illustrate what these possibilities and aspects are, it has been decided to present below some possible types of experiments, which, however, are to be understood as generic and principled indications, rather than as actual ready-to-use experiments.

A first type of experiment may be aimed at identifying the goals that agents employ in performing their reasoning. It is a question of finding out who among them adopts the naive goal, and who on the other hand an educated goal that will have to be identified. In order to achieve this, it is not enough to assign the various agents a task that requires them to reason deductively, but a questionnaire must also be drawn up. In this questionnaire, agents may be asked to indicate whether they recognised as such the conclusion they reached when they realised that it was a possible solution to their reasoning (naive goal) or when it proved capable of satisfying particular requirements (educated goal), specifying in that case which requirements were involved. Or, more prudently, if we want to avoid that the subjective interpretation by the agent of what is a "particular requirement" could lead him to answer in an inaccurate way, we could ask him to simply indicate from what he understood that the conclusion he reached could be the right solution, and if he relied on some particular criteria to arrive at that conviction.

A second type of experiment may be aimed at finding out if it is possible to induce the agents to reason by adopting a precise educated goal decided by the experimenters. It is a question of whether and how the goal that one agent has acquired through the practice can be temporarily set aside in favour of another. In order to achieve this, we will have to limit ourselves to those agents for whom we have verified the use of the naive goal, and assign them a task that requires them to reason deductively. These agents will then have to be divided into two groups, so that each group can be given different indications about the task to be performed, which in its basic form will remain the same for both of them.

The first group of agents will have to be told that the task assigned to them is such that its solution must necessarily satisfy very specific requirements, which will be outlined to them. It is tested here if an agent who finds a possible solution to his reasoning will stop to reason as his naive goal requires, or instead he will be pushed to verify that the solution identified by him satisfies the additional requirements that have been outlined to him, making them act in accordance with the educated goal wanted by the experimenters.

The second group of agents will be asked to answer additional questions in writing. These questions will serve to oblige these agents to make those checks on their

conclusion that only the correct conclusions would be able to satisfy. It is tested here whether, by including as part of the task assigned to the agents the verification that their conclusion satisfies specific requirements, it is possible to oblige them to reach the correct conclusion, making them act in accordance with the educated goal wanted by the experimenters.

A third type of experiment may be aimed at finding out if it is possible to make the agents acquire the ability to reason by adopting an educated goal decided by the experimenters. It is a question of whether and how the goal that one agent has acquired through practice can be replaced by another. In order to achieve this, we will have to limit ourselves to those agents for whom we have verified the use of the naive goal, and assign them a task that requires them to reason deductively. These agents will then have to be divided into various groups, so that before performing this task each group can be subjected to a different teaching method through which to try to learn a given model of reasoning. It is thus tested whether, by training agents to consider as conclusions of deductive reasoning the only conclusions that satisfy the requirements decided by the experimenters, it is possible to make them acquire the corresponding educated goal, so that they can use it spontaneously when they will be called to perform the same task previously prepared for all. Naturally, this task will be followed by a questionnaire to enable the experimenters to determine which goal these agents have actually adopted.

The teaching methods assigned may differ from each other for various factors, such as the number of training sessions, the length of such sessions, and the type of training given. Different teaching methods may be assigned to groups composed of the same type of agents, but also identical teaching methods may be assigned to groups of agents that differ in the number of agents, age of agents, sex of agents, profession of agents. It is thus tested which factors influence the learning of precise teaching methods, as well as which teaching models are the best and which are the easiest to learn.

A fourth type of experiment may be aimed at discovering the universality and effectiveness of the educated goals. It is a question of identifying whether there are educated goals that can be applied to any kind of deductive reasoning, and whether there are some that can guarantee high performance. In order to achieve this, we will have to limit ourselves to those agents for whom we have verified the use of specific educated goals, and assign them a large number of tasks that require them to engage in various types of deductive reasoning. It is thus tested the universality and effectiveness of these educated goals.

Experimenters may also decide to verify the universality and effectiveness of educated goals prepared by them. In order to achieve this, they will have to limit

themselves to those agents for whom they have verified the use of the naive goal, and assign them a large number of tasks that force them to adopt these educated goals during their deductive reasoning.

## 10. Example of application

In order to better understand the theory of reasoning by goals outlined so far, it may be useful to apply it to a concrete example of reasoning.

Suppose an agent is asked to consider the following premise valid:

*If it rains, I'll take an umbrella..*

and to complete the following statement:

*If it does not rain, then ...*

In a case like this the end of the agent will be to pursue the task assigned to him. But it will not be this end that will determine the characteristics of his reasoning but rather the goal he will use to reason.

If the agent in question has learned to reason through the naive goal, he will simply identify the first conclusion he considers to be a possible solution to his reasoning. As he has no other indication on how to proceed, he will not make his conclusion dependent on the satisfaction of any other particular requirement.

In such a situation, our agent may consider as a possible solution that in the absence of rain it is not necessary to take an umbrella. Simply because it has always been his desire to shelter from the rain that made him take the umbrella. In other words, if there is no rain, he will also lack a valid reason to take the umbrella, and consequently he will choose to complete the previous statement in the following way:

*If it doesn't rain, then I won't take an umbrella*

The fact that in this particular example the naive goal led our agent to find a wrong solution does not mean that this should always happen. In other words, there is nothing to prevent the naive goal from leading the agent to a conclusion that, in addition to being a possible solution to his reasoning, it is also a correct solution. It may happen, for example, that an agent immediately realizes that there may be reasons to take the umbrella even in the absence of rain, perhaps just out of caution because of the cloudy sky. And so he comes to consider as a possible solution that in the absence of rain there may be both reasons to take and reasons not to take the umbrella. Consequently he will choose to complete the previous statement in the following way:

*If it doesn't rain, then nothing can be said about the umbrella.*

What must be understood in the case of a naive goal is that whatever conclusion an agent finds himself considering as a possible solution to his reasoning, it will be in his correspondence that he will stop reasoning, considering his end as pursued. And this means that he will not be pushed to perform any further verification, thereby incurring the possibility of error.

In the case where the task in question is proposed to an agent who has learnt to reason through a specific educated goal, he will not merely identify the first conclusion that he will consider to be a possible solution to his reasoning. But he will take as his conclusion only that conclusion which satisfies very specific requirements. Let us suppose that these very precise requirements are constituted by the decision of the agent to consider valid only what is within his imagination of the premises.

Since imagining the initial premise means imagining taking the umbrella while it is raining, there will be no part of that imagination to represent the absence of rain. Therefore, the only conclusion that such an agent will have at his disposal will be to complete the assigned statement in the following way:

*If it doesn't rain, then nothing can be said about the umbrella.*

The fact that in this specific example an educated goal has led our agent to find the correct solution does not mean that this should always happen. In other words, there is nothing to prevent an educated goal from leading the agent to the wrong solution. It may happen, for example, that an agent has an educated goal that limits his conclusions to only those that can be reflected in his life experience. And so if he has used the umbrella always and only in the presence of rain, he will find himself completing the previous statement in the following way:

*If it doesn't rain, then I won't take an umbrella*

making a mistake.

What we must understand in the case of an educated goal is that whatever conclusion an agent finds himself considering as a solution to his reasoning, it will be a solution capable of satisfying very precise requirements. These will be the only guarantee of its correctness, since even in this case, once identified a conclusion, our agent will stop reasoning, considering his end as pursued. This means that he will not be pushed to perform any further verification.

In essence, the only way to be sure that an agent reaches the right conclusion of a deductive reasoning is for that agent to be guided by the correct educated goal.

In the absence of such a guarantee, whether or not an agent will come to the right conclusion will depend on many factors, such as the objective difficulty of finding the correct solution, the presence of additional beliefs or information that can make one solution more credible than another, and so on.

If we want to use this task to identify the goals that the various agents in a given group use to reason, we should prepare a simple questionnaire to be submitted to them after completing the usual statement. Within this questionnaire, agents may be asked to indicate from what they have understood that the answer they have given might be the right solution, and whether they have relied on any particular criteria to arrive at that belief.

The agent who completed the statement as follows:

*If it doesn't rain, then I won't take the umbrella*

simply because he only ever took the umbrella to get out of the house in the rain, he would presumably find himself giving an answer like this:

*I came to that conclusion from the fact that the umbrella is always taken in order to shelter from the rain, and I came to think this without having followed any particular criteria*

which highlights a naive goal because it appeals to nothing more than what is needed to find a possible solution to the reasoning, just as pointed out by the agent himself.

If the agent had completed the statement in the following way:

*If it doesn't rain, then nothing can be said about the umbrella*

because used to drawing his own conclusions by limiting himself to consider valid only what is within his imagination of the premises, he will presumably find himself giving an answer of the type:

*I came to that conclusion from the fact that nothing in the premise binds to any use of the umbrella when it is not raining, and I came to think this because I consider valid only what the premises allow me to imagine*

which highlights an educated goal because by invoking constraints it appeals to something more than what is needed to find a possible solution to the reasoning. The conclusion chosen therefore satisfies further requirements, just as pointed out by the agent himself.

If we want to use this task to find out whether it is possible to induce agents to reason by employing an educated goal decided by us, we must first select those

agents for whom we have verified the use of the naive goal, and then divide them into two groups.

In the first of these two groups we will give an indication that the correct ways to complete the following statement:

*If it doesn't rain, then...*

are those that describe something that can be found within the imagination of the following initial premise:

*If it rains, I'll get the umbrella.*

At the second of these groups we will first ask to make a drawing that represents the above-mentioned initial premise, and only then will they be allowed to complete the usual statement, and only with the description of something that can be found inside that drawing.

In this way we will be able to verify whether, in order to bring the agents to adopt the educated goal decided by us, it is sufficient to give them a verbal indication (first group) or whether the verifications that it involves should be made part of the task to be performed (second group).

Naturally, the educated goal in question, which is to limit the conclusions of the agents to what they are able to imagine of the initial premise, can be taught to the agents even before they face the task we are examining here. In this way we will be able to establish whether and what teaching has been effective, and what kind of teaching can be considered the best.

Or we can complement the task we are examining here with many other tasks that require the agents to reason deductively, and use them to verify the universality and effectiveness of various alternative educated goals. Not only the educated goal that we have just seen, but also that of limiting the conclusions of the agents to what they are able to find in their experiences, or to many other possible goals. The important thing is to be able to design these tasks in such a way that they will induce agents to use precisely the educated goals that we have chosen to test.

## **11. Analysis of the different theories of deductive reasoning**

Various theories have been proposed in the literature in order to be able to explain the deductive reasoning, or at least some particular aspect of it. However, none of these theories, as we will see shortly, has been able to provide a unitary explanation of all the characteristics of this type of reasoning.

Before proceeding with a concise review of these theories, one more aspect should be clarified. Since the theory of reasoning by goals is proposed within the psychology of reasoning, the theories presented here will be limited to this precise field. In other words, we have chosen not to make an overview of how deductive processes have been explained within the various disciplines that have dealt with human reasoning also to avoid an unnecessary dispersion with respect to the aims of this paper.

In the context of the theory of formal rules (Rips, 1983) it is believed that the human mind has its own rules of inference that follow those of formal logic. In other words, it is believed that the human mind possesses a so-called "natural logic" that each individual unconsciously applies during deductive reasoning to succeed in reaching the correct conclusion. Clearly, such a theory is not able to explain why there are logical rules that people have more difficulty in applying correctly than others, nor is it able to explain why the same logical rule can be applied with different degrees of effectiveness depending on the type of content involved. In the same way, such a theory cannot explain why the effectiveness in applying the logical rules correctly can change in the presence of premises not necessary for finding the correct conclusion.

In the context of the theory of mental logic (Braine & O'Brien, 1998) it is believed that the human mind has its own schemes of inference that have developed with evolution. In other words, it is believed that the human mind is in possession of its own "mental logic" which, having developed for evolutionary purposes, has ended up rewarding short, direct and immediate schemes of inferences that are somewhat distinct from those typical of formal logic. In a subsequent phase due to the mass diffusion of intellectual work, alongside these main inference schemes, support schemes of a more complex nature have developed.

Clearly, such a theory can be considered a more advanced version of the theory of formal rules, also because by attributing to people a natural logic different from the formal one it is certainly able to better explain their logical errors. However, just like the theory of formal rules, the theory of mental logic is not suitable to explain the influence of content type on people's ability to apply the same logical rules correctly.

In the context of the theory of mental models (Johnson-Laird & Byrne, 2002) it is believed that the human mind is not able to recognize in a natural way the syntactic structure at the basis of logical rules. In other words, it is believed that in order to arrive at the correct solution of a deductive reasoning, the human mind is forced to represent the meanings involved and to manage them through a precise descriptive analysis. Using this theory it becomes possible to explain why some logical rules are easier to apply correctly than others. Typically, the logical rules that will be easier to apply correctly will be those that require fewer representations and are easier to han-

dle at the descriptive level. Moreover, this theory also explains why the effectiveness in applying the same logical rule correctly can depend on the content dealt with. Typically, the easier it will be to represent and mentally describe the content of a given logical rule, the more effective it will be for people to apply it correctly. Where, on the other hand, such a theory is not useful is in explaining why the effectiveness of people in applying the logical rules correctly changes in the presence of premises that are not necessary for finding the correct solution.

In the context of the theory of pragmatic schemes (Cheng & Holyoak, 1985) it is believed that people, being in contact with certain contexts, end up developing reasoning schemes in the form of obligations and permissions. In other words, it is believed that in such contexts they learn what can and cannot be done. This theory can be used to explain why people can apply logical rules correctly when concrete and family contexts are involved. While it is not able to provide any explanation of people's ability to apply the logical rules correctly when abstract contexts are involved, with respect to which they have not been able to develop any previous reasoning scheme.

In the context of evolutionary theory (Cosmides, 1989) it is believed that the mind, and consequently also reasoning, are the product of natural selection and adaptation to the environment. In other words, it is believed that reasoning has no logical nature but has evolved through those strategies that have allowed people to solve the problems posed to them by the natural and social environment. This theory can be used to explain why people can apply logical rules correctly when contexts related to their social and natural environment are involved. While it is not able to provide any explanation of the ability to correctly apply logical rules when contexts that have not been selected to resolve are involved.

In the context of the theory of heuristics and bias (Pollard, 1982), it is believed that information that is particularly evident, significant and easy to obtain ends up taking on excessive importance compared to all the others, thus altering people's ability to apply the logical rules correctly. In other words, this theory explains the errors of reasoning when the conclusions reached by people are determined by information that they have easily obtained or to which they have recognized particular evidence or significance. While it does not provide any explanation for the errors of reasoning that have occurred in all other circumstances.

In the context of dualistic theory (Evans, 2003) it is believed that there are two systems that govern reasoning. The first system governs intuitive reasoning related to beliefs, and it comes into play when the answer people seek depends on something they have already learned and they can retrieve from memory. The second system governs abstract reasoning and hypothetical thinking, and it comes into play when

the answer people seek requires reflection and the construction of new models. Naturally these two systems can compete with each other and when this happens their conflict can result in the so-called bias of belief that occurs when some belief acquired in the past prevails over abstract reasoning. In other words, this theory explains the errors of reasoning when the conclusions reached by people are determined by beliefs that contrast with the logical nature of deductive reasoning. While it does not provide any explanation for the errors of reasoning that occur in all other circumstances.

In the context of the Bayesian paradigm (Oaksford & Chater, 2007) it is believed that reasoning has a probabilistic and not logical nature. In other words, it is believed that the conclusions reached by people through reasoning are those they consider most likely, regardless of their logical validity. This characteristic makes the Bayesian paradigm particularly suitable for explaining all those apparently absurd behaviours in which logic and probability conflict. As it happens, for example, when the conclusion reached by people in a given reasoning is modified by additional premises that are completely influential on the logical level, but that evidently contribute to make one solution appear more likely than the other. The fact that probabilistic considerations can be invoked in a very large number of contexts allows the Bayesian paradigm to explain the characteristics of deductive reasoning in a large number of different circumstances. That said, not even the Bayesian paradigm can be considered a unitary explanation of all deductive reasoning because, at least in principle, there can be reasonings for which no consideration of a probabilistic nature is admissible. Also because in the structure of deductive reasoning there is nothing that makes it necessary to use considerations of a probabilistic nature.

## **12. Peculiarities of the theory of reasoning by goals**

Let us now try to analyse how the theory of reasoning by goals is situated with respect to the theories described above.

Since the theory of reasoning by goals makes the effectiveness of deductive reasoning depend on the goal the agent pursues by reasoning, it goes beyond the naive idea that the mind possesses its own formal logic as argued by the theory of formal rules, or its own mental logic as argued by the theory of mental logic. Instead, it argues, in line with the other theories seen here, that the information that an agent puts as a conclusion of his deductive reasoning are those having very precise characteristics, or to which it is possible to attribute very precise characteristics. But for the other theories seen here such information is that which the agent is able to represent mentally, as argued by the theory of mental models, or that which he is able

to access more easily, as argued by the theory of heuristics and bias, or that which concerns obligations and permissions familiar to him, as argued by the theory of pragmatic schemes, or that which concerns the natural and social contexts in which human beings have evolved, as argued by the evolutionary theory, or that related to his beliefs and hypothetical thinking, as argued by the dualistic theory, or that whose correctness he considers more likely, as argued by the Bayesian paradigm theory, for the theory of reasoning by goals the information that an agent puts as conclusion of his deductive reasoning is the information that satisfies his goal. And therefore that information which appears to him to be capable of constituting a possible solution for his reasoning, and which possibly also meets additional requirements.

Another difference between the theory of reasoning by goals and other alternative theories is its ability to provide a unitary explanation for all deductive reasoning. For the simple reason that when an agent has to make deductive reasoning he never reaches an immediate and automatic conclusion, but needs to process the information in his possession, and to continue to do so until he finds a conclusion to which he recognizes everything he needs to be the solution he is looking for. This can be said in another way. That is to say that whatever reasoning an agent has to perform, it will always and in any case be subject to the explanation that gives of it the theory of reasoning by goals.

Once we have clarified how the theory of reasoning by goals sits in relation to the other theories seen here, there is a further question to be addressed. It is a question of seeing what explanation it is able to give for the fact that agents facing deductive reasoning find themselves, in most cases, to reach conclusions in line with the rules of logic and with what they consider most likely. The explanation for the theory of reasoning by goals in this regard is that if the conclusion of a deductive reasoning is determined by the fact that it is considered adequate for the purpose, it is inevitable that it has precisely the above characteristics. This is because the judgement that an agent develops on a given conclusion necessarily reflects his conception of the world, and since the conception of the world that we human beings have is a reasonable and rational conception, we will hardly be pushed to consider as a possible solution of a deductive reasoning a conclusion out of all logic and probability.

### **13. Dependence of reasoning on the type of task**

One characteristic of the deductive reasoning that can be explained by the theory of reasoning by goals is its dependence on the type of task. Where dependence on the type of task here means that there are types of tasks that people can perform correctly more frequently than other types.

This characteristic of deductive reasoning has been proven by a great many experiments. In particular, we can mention those experiments that use conditional syllogisms to measure people's ability to recognize as correct the use of the logical rules of Modus Ponens and Modus Tollens and as incorrect the use of the logical rules of denying the antecedent and affirming the consequent.

To better understand this type of experiment we take the following premise:

*If it rains, I'll take an umbrella..*

Recognizing as correct the use of the rule of Modus Ponens means that given the following statement:

*If it rains, then ...*

people are able to complete it in the following way:

*If it rains, then I'll take an umbrella.*

Recognizing as correct the use of the rule of the Modus Tollens means that given the following statement:

*If I don't take an umbrella, then...*

people are able to complete it in the following way:

*If I don't take an umbrella, then it won't rain*

Recognising as incorrect the use of the rule of affirming the consequent means that given the following statement:

*If I take an umbrella, then...*

people are able to complete it in the following way:

*If I take an umbrella, then nothing can be said about the rain*

refraining from completing it in the following way:

*If I take an umbrella, then it's raining.*

Recognizing as incorrect the use of the rule of denying the antecedent means that given the following statement:

*If it does not rain, then ...*

people are able to complete it in the following way:

*If it doesn't rain, then nothing can be said about the umbrella.*

refraining from completing it in the following way:

*If it doesn't rain, then I won't take an umbrella.*

In these experiments (Evans, 1977) it was found that people are more likely to recognize as correct the rule of Modus Ponens than to recognize as correct the rule of Modus Tollens or as incorrect the rule of affirming the consequent and denying the antecedent. A situation of this kind is explained, for example, in the theory of mental models, arguing that the ability to recognize a logical rule as correct or incorrect must be based on mental representations. In particular, according to Johnson-Laird and Byrne (2002) when the mental representations of an individual allow him to draw a conclusion, he will draw that conclusion without going any further. For this reason it will be easier for him to recognize as correct the rule of the Modus Ponens rather than that of the Modus Tollens, requiring the rule of the Modus Tollens one more representation than the rule of the Modus Ponens. While it will be more difficult to recognize as incorrect the rule of affirming the consequent and denying the antecedent since in these two cases the representations to be used will be even more numerous.

Now that we have that established, let us see why the theory of reasoning by goals can explain without difficulty the dependence of reasoning on the type of task. What we need to do is to agree that if agents who perform deductive reasoning do not have an educated goal that makes their conclusion dependent on additional requirements that guarantee its correctness, they will be subject to errors. In general, in the absence of proper instruction on how to reason, we will have to expect that most agents will be in precisely that situation, and therefore, using the naive goal, they will be destined to be content with the first conclusion that will appear to be a possible solution to their reasoning. In general, in the absence of adequate education on how to reason, we will have to expect most agents to be in the above situation. And therefore that, using the naive goal, they end up settling for the first conclusion that they will consider to be a possible solution to their reasoning. All this means not only that such agents will be subject to errors, but that these errors will occur more frequently as the complexity of the reasoning they are facing increases. The sense to give to this statement is that the more difficult the conclusion of a deductive reasoning is to obtain, the more difficult it will be for it to coincide with the first conclusion that agents will identify as a possible solution to their reasoning.

In this perspective the result of the experiments mentioned above would mean that recognizing as correct the rule of Modus Ponens is an easier task than the one

related to recognizing as correct the rule of Modus Tollens, and also the one related to recognize as incorrect the rule of affirming the consequent or the rule of denying the antecedent.

We can say in other words, referring to the conditional syllogisms mentioned above, that given the premise:

*If it rains, I'll take an umbrella*

the number of agents who recognize the following conclusion as correct:

*If it rains, then I'll take an umbrella*

is greater than the number of agents who recognize the following as correct:

*If I don't take an umbrella, then it won't rain*

or the number of agents who recognize the following as incorrect:

*If I take an umbrella, then it's raining*

or the number of agents who recognize the following as incorrect:

*If it doesn't rain, then I won't take an umbrella*

because in the absence of a model of reasoning that allows them to recognize as correct the rules of Modus Ponens and Modus Tollens and as incorrect the rules of affirming the consequent and denying the antecedent, the task easier to perform will be the first. And therefore it will be precisely this task that the greatest number of agents will be able to face even without an adequate goal to perform it properly.

#### **14. Dependence of reasoning on the content of a task**

One characteristic of the deductive reasoning that can be explained by the theory of reasoning by goals is its dependence on the content of a task. Where dependence on the content of a task here means that given the same type of task there are contents that allow a larger (or smaller) number of people to perform them correctly.

This characteristic of deductive reasoning has been proven by a great many experiments. In particular, we can mention those experiments that use classical syllogisms to measure people's ability to recognize them as correct in relation to the credibility of their respective conclusions.

To better understand this type of experiments we take the following syllogisms.

The first syllogism has a valid and credible conclusion, and therefore there is no conflict between logic and beliefs:

*Premise 1: No human being is immortal.*

*Premise 2: Some living beings are immortal .*

*Conclusion: Some living beings are not human beings.*

The second syllogism has an invalid and non-credible conclusion, and therefore there is no conflict between logic and beliefs:

*Premise 1: No human being is immortal.*

*Premise 2: Some living beings are immortal.*

*Conclusion: Some human beings are not living beings.*

The third syllogism has a valid but non-credible conclusion, and therefore there is a conflict between logic and beliefs:

*Premise 1: No ugly woman is desired by men.*

*Premise 2: Some disfigured women are desired by men.*

*Conclusion: Some disfigured women are not ugly women.*

The fourth syllogism has an invalid but credible conclusion, and therefore there is a conflict between logic and beliefs:

*Premise 1: No ugly woman is desired by men.*

*Premise 2: Some disfigured women are desired by men.*

*Conclusion: Some ugly women are not disfigured women.*

In these experiments (Evans, 1983) it was found that when the conclusions of syllogisms are credible most people recognize them as correct even if they are not really correct from the point of view of logical rules. While when the conclusions of the syllogisms are non-credible most people recognize as correct or incorrect the syllogisms that are really such by virtue of the logical rules involved.

A situation of this kind is explained, for example, in the dualistic theory, arguing that although people try to reason logically, thus employing the system of abstract reasoning, they are influenced by their beliefs because the system connected to such beliefs is able to conflict with the other and prevail. In particular, according to Evans (2002), the probability of resolving this conflict in favour of logic increase as people's cognitive abilities increase and decline as their age increases.

Now that we have that established, let us see why the theory of reasoning by goals can explain without difficulty the dependence of reasoning on the content of a task.. What we need to do is to agree, as we have done before, that in the absence of adequate education on how to reason, most agents end up settling for the first conclusion they will consider to be a possible solution to their reasoning. All this means not

only that such agents will be subject to errors, but that any belief that can make an incorrect solution a credible solution will actually lead them to make mistakes. This means that according to the theory of reasoning by goals, in correspondence with a reasoning whose content lends itself to misleading beliefs about what its solution may be, there will be an increase in the number of agents subject to errors, at least among those who do not have a correct model of reasoning.

In this perspective, the result of the experiments mentioned above would confirm that in the presence of misleading beliefs the possibility of reaching a wrong conclusion increases.

We can say in other words, referring to the syllogisms mentioned above, that if the following two statements:

*some living beings are not human beings*  
*some ugly women are not disfigured women*

are considered by most agents as correct solutions of their respective syllogisms, regardless of whether they are really correct according to the logical rules involved, it is precisely because they are credible as solutions.

While if the following two statements:

*some disfigured women are not ugly women*  
*some human beings are not living beings*

tend to be considered the first as a correct solution and the second as an incorrect solution, in line with the rules of logic, it is precisely because not being credible as solutions will force us to take into account other factors of judgment. Among the factors of judgement, logical correctness should also be included.

## 15. Dependence of reasoning on the number of premises

One characteristic of deductive reasoning that can be explained by the theory of reasoning by goals is its dependence on the number of premises. Where dependence on the number of premises here means that the presence of extra premises is able to change people's ability to recognize a given logical conclusion as correct. While this should not happen because logic is monotonic, in the sense that adding extra premises to a valid argument should never affect the argument's validity.

This characteristic of deductive reasoning has been proven by a great many experiments. In particular, we can mention those experiments that use classical syllogisms to measure people's ability to recognize them as correct in relation to the number of premises involved.

To better understand this type of experiments we take the following syllogisms.

The first syllogism has a valid conclusion that starts from a minimum, but sufficient number of premises:

*Premise 1: If it rains, I'll take an umbrella.*

*Premise 2: It is raining.*

*Conclusion: I'll take an umbrella.*

The second syllogism is the same as the previous one, with the addition of two extra premises:

*Premise 1: If it rains, I'll take an umbrella.*

*Premise 2: It is raining.*

*Premise 3: The umbrella is broken*

*Premise 4: The rain is about to stop*

*Conclusion: I'll take an umbrella.*

In these experiments (Byrne, 1989) it was found that the adding of extra premises to valid syllogisms was able to significantly reduce the number of people who recognized their conclusion as correct.

A situation of this kind is explained, for example, in the Bayesian paradigm (Oaksford & Chater 2009), arguing that in everyday reasoning, non-monotonicity is the norm. Almost any conclusion in other words can be overturned if additional information is acquired and if this additional information makes a different conclusion more likely.

Now that we have that established, let us see why the theory of reasoning by goals can explain without difficulty the dependence of reasoning on the number of premises. What we need to do is to agree with what we have already seen in the previous paragraph. Namely that in correspondence with a reasoning whose content lends itself to misleading beliefs about what its solution may be, will increase the number of agents subject to errors, at least among those who do not have a proper model of reasoning. In this perspective, the result of the above experiments would confirm that the extra premises are actually able to make any conclusion credible or not credible as solution.

We can say in other words that if the following extra premises of the second syllogism:

*Premise 3: The umbrella is broken*

*Premise 4: The rain is about to stop*

reduce the number of agents who come to the following conclusion:

*Conclusion: I'll take an umbrella.*

is precisely because they make it less credible as solution.

## **16. Conclusion**

The possibility explored in this paper of using the theory of reasoning by goals to explain deductive reasoning can be considered promising. In favour of this theory there is both the fact that it provides predictions consistent with the main experiments performed in the psychology of reasoning, and that it can be considered a unitary explanation of all the characteristics of deductive reasoning.

Further investigation and development of this theory, especially from an experimental point of view, therefore seems desirable, if not necessary to expand research on deductive reasoning to important topics such as the identification of the best way to reason and the best way to teach reasoning.

## References

- Braine, M., O'Brien, D. P. (1998). *Mental logic*. Psychology Press.
- Bratman, M., (1987), *Intention, Plans, and Practical Reason*, Cambridge, Mass.: Harvard University Press.
- Broome, J., (2013), *Rationality through Reasoning*, Oxford: Wiley-Blackwell.
- Byrne R. M. J. (1989), *Suppressing Valid Inferences with Conditionals*, in "Cognition", 31, pp. 61-83
- Cheng P. W., Holyoak K. J. (1985), *Pragmatic Reasoning Schemas*, in "Cognitive Psychology", 17, pp. 391-416
- Clarke, R., (2003), *Libertarian Accounts of Free Will*, Oxford: Oxford University Press.
- Cosmides L. (1989), *The Logic of Social Exchange: Has Natural Selection Shaped how Humans Reason? Studies with the Wason Selection Task*, in "Cognition", 31, pp. 187-276
- Dretske, F., (1988), *Explaining Behavior: Reasons in a World of Causes*, Cambridge, MA: MIT Press.
- Enç, B., (2003), *How We Act: Causes, Reasons, and Intentions*, Oxford: Oxford University Press.
- Evans J. ST. B. T. (1977), *Linguistic Factors in Reasoning*, in "Quarterly Journal of Experimental Psychology", 29, pp. 297-306
- Evans J. ST. B. T. (1983), *On the Conflict between Logic and Belief in Syllogistic Reasoning*, in "Memory and Cognition", 11, pp. 295-306
- Evans J. ST. B. T. (2002), *Logic and Human Reasoning: An Assessment of the Deduction Paradigm*, in "Psychological Bulletin", 128, pp.978-96
- Evans J. ST. B. T. (2003), *In Two Minds: Dual-process Account of Reasoning*, in "Trends in Cognitive Science", 7(10), pp. 454-9
- Johnson-Laird P. N., Byrne R.M. (2002), *Conditionals: A Theory of Meaning, Pragmatics, and Interference*, in "Psychological Review", 109, pp.646-78
- Korsgaard, C., (1996), "Skepticism about Practical Reason", in her *Creating the Kingdom of Ends*, Cambridge: Cambridge University Press.

- Lowe, E.J., (2008), *Personal Agency: The Metaphysics of Mind and Action*, Oxford: Oxford University Press.
- McCann, H.J. (1998). *The works of agency*. Ithaca, NY: Cornell University Press.
- Moran, R., (2001), *Authority and Estrangement*, Princeton: Princeton University Press.
- Nathan, N.M.L. (1992). *Will and world: A study in metaphysics*. Oxford: Clarendon.
- Oaksford, M., Chater, N. (2007) *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.
- Oaksford, M., N. Chater (2009). *Precis of Bayesian Rationality: The Probabilistic Approach to Human Reasoning*, in *Behavioral and Brain Sciences* 32: pp. 69–120.
- Pollard P. (1982), *Human Reasoning: Some Possible Effects of Availability*, in “Cognition”, 12, pp. 65-96.
- Rips L.J. (1983) *Cognitive Process in Propositional Reasoning*, in “Psychological Review”, 90, pp.38-71
- Searle, J. (2001). *Rationality in action*. Cambridge, MA: MIT Press.
- Walton, D., Reed, C., Macagno, F. (2008). *Argumentation Schemes*. New York: Cambridge University Press.
- Zhu, J. (2004). *Understanding volition*. In “Philosophical Psychology”,17(2), pp. 247-273.