Rational Choice, Social Identity, and Beliefs about Oneself

Fernando Aguiar
_Institute for Advanced Social Studies–Spanish Council for Scientific Research, Cordova_

Andrés de Francisco
_Complutense University of Madrid, Spain_

Social identity poses one of the most important challenges to rational choice theory, but rational choice theorists do not hold a common position regarding identity. On one hand, externalist rational choice ignores the concept of identity or reduces it to revealed preferences. On the other hand, internalist rational choice considers identity as a key concept in explaining social action because it permits expressive motivations to be included in the models. However, internalist theorists tend to reduce identity to desire—the desire of a person to express his or her social being. From an internalist point of view, that is, from a viewpoint in which not only desires but also beliefs play a key role in social explanations as mental entities, this article rejects externalist reductionism and proposes a redefinition of social identity as a net of beliefs about oneself, beliefs that are indexical, robust, and socially shaped.

**Keywords:** beliefs; externalism; identity; internalism; rational choice theory; rationality

The theory of rational choice can be interpreted from two epistemological approaches: an internalist one and an externalist one. According to the first, the agent’s preferences and beliefs are mental states that constitute

_Received 9 October 2008_

_Authors’ Note:_ An earlier version of this article was presented to the International Sociological Association (ISA) Research Committee on Rational Choice at the First ISA Forum of Sociology, Barcelona, 5 to 9 September 2008. We are grateful to José Antonio Noguera and to the participants in the debate for their helpful comments. We are also grateful to two anonymous referees from _Philosophy of the Social Sciences_. They all forced us to introduce changes that improved the final outcome, but none of them are responsible for any remaining shortcomings.
reasons for action, while the second solely considers the preferences revealed in the action and certain behavioral hypotheses, neither of which entails delving into the minds of individuals. Although both approaches are neutral with respect to formal models of rational choice, they diverge when the paradigm is expanded to incorporate motives that are expressive of or derived from the identities of individuals. In this article, we argue that the internalist conception of rational choice is the only interpretation that can be extended to include identity-based preferences or desires. However, this proves to be insufficient since at the same time we hold that people’s social identity has less to do with preferences or desires than with their beliefs about themselves and the world. These beliefs can be distributed into practical syllogisms whose conclusions are reasons for action in the form of desires to act or obligations to act. In the last part of this article we explore the nature of identity beliefs, that is, their indexation, robustness, and emotional dimension, among other properties, with a view to demonstrating the difficulties involved in incorporating all of these complex dimensions in the paradigm of rational choice, even when the model is interpreted from an internalist approach. Furthermore, given the essentially linguistic and cultural nature of practical identity syllogisms, we underscore in the conclusions the importance of analyzing language to understand this aspect of human behavior that appears to respond to what individuals claim to be.

Two Approaches to Rational Choice: Internalist and Externalist

The so-called canonical paradigm of rational choice theory (Bell, Raiffa, and Tversky 1988, 18) explains all social action in terms of the decisions made by an individual or a group of individuals in a given context, the consequences of those decisions, and an assumption about the individuals’ reasons for having made the decision, namely that individuals will try to maximize their benefits (whatever that benefit may be). For a choice to be made, the agent must face a set of possible options (her feasible set). It is assumed that given her feasible set, the agent will choose the option that leads to—or she believes leads to—the best outcome. If she has full information about the outcome of her decisions, she will find herself in a situation of certainty and be able to maximize her utility. If the information is incomplete, however, she will only be able to maximize the expected utility in a context of risk and uncertainty. In this situation, she will attribute an objective or
subjective probability to the outcome of her action. This probability is the belief she has about the results of her action.

Formal decision theory holds that if a person chooses X instead of Y, she prefers X to Y, but it does not consider the nature of individual preferences, nor why people prefer some things to others. From a formal perspective, the only thing that matters is that these preferences satisfy certain basic criteria of logical consistency, particularly transitivity and completeness. If these requisites are violated, it will be impossible to know what the people prefer; their preferences cannot be ranked or ordered, and the decision theorist will assume that the people did not choose rationally.

This formal foundation of rational choice theory gives rise to two distinct interpretations regarding the behavior of individuals: one that is internal to the individual and another that is external. The internalist interpretation of rational choice holds that the theory describes “what is actually going on inside us when we reason” (Satz and Ferejohn 1994, 73). But this internalist approach also assumes that mental entities—the desires and beliefs on which individual preferences are based—are causally linked to the individual’s decision given that they provide reasons for action: beliefs and desires are mental states that motivate the individual and causally explain her action (Williams 1981; Boudon 2003; Landa 2004; Hedström 2005, 2006). If a person chooses X instead of Y, it is because she believes that X best satisfies her desires. These internal beliefs and desires, which are subjective, prompt her action.1

From the internalist perspective, rational choice theory is both explanatory and normative at the same time (Hausman 2001, 320; Landa 2004). By relying on individual beliefs, desires, and preferences and the role they play as reasons for action, the theory not only explains individual behavior—why individuals choose as they do—but also justifies these decisions from a normative standpoint insofar as they are the best ones that an individual can make to maximize her interest, provided, of course, that she has sufficient information and is not deceiving herself (Davidson 1980, 21; Elster 1989, 25). Nevertheless, this conception would prove to be inconsistent if it were not founded on an explicit psychological realism regarding the mental entities or states that constitute the basic pieces of any explanation.

1. The philosophical debate on the internal nature of reasons for action has revolved around the position of Bernard Williams (1981). According to Williams, the only way to satisfy the conditions of truth of the proposition “X has a reason R to do A” is by appealing to what he calls the subjective motivational set of agent X. Although Williams’s theory clearly has Humean roots, he diverges from Hume as he does not consider that only desires have motivational force. A subject’s motivational set includes desires, emotions, personal loyalties, projects, commitments, and so on (Williams 1981, 105).
in terms of rational choice (Hausman 1995, 100). For the internalist, “beliefs,” “desires,” and “preferences” are not mere labels, nor simple theoretical constructions that aid us to understand human conduct, nor constructs that have no real correlation to what is going on in our minds, but are instead real events in the intentional processes of the human mind. As Boudon (2003, 3, emphasis added) formulates it, “Any action is caused by reasons in the mind of individuals (rationality).” Preferences, desires, and beliefs thus form the building blocks of folk psychology defended by internalism. Given that these preferences can be ranked and represented as utility functions, the internalist interpretation of rational choice mathematically transforms this folk psychology into a theory of utility maximization (Hausman 2000, 114).

The externalist, on the other hand, categorically rejects each of these psychological assumptions. The externalist interpretation of rational choice theory rests on an instrumental conception of the basic terms of the theory since it considers that the external, observable conduct of individuals is the only thing that can be taken into account. What goes on inside of each of our heads is of no interest to the social scientist; it is a black box that need not be opened to obtain genuine explanations (Börgers 1996). “Beliefs” and “preferences” (which are not desires) are mere labels that we use to describe individual action, but they are not real mental representations, or even if they are they add nothing to the explanation: “Reference to preferences as representations of something ‘inside the head’ . . . is not necessary” (Dowding 2007, 12; also see Satz and Ferejohn 1994, 76). To explain individual action, it is therefore unnecessary to account for the individual’s reasons for action. If a person chooses X instead of Y, we say that she “prefers” X to Y where that preference is revealed by the individual when deciding about a mere binary relation (Samuelson 1948; Sen 1971). Beliefs are interpreted as mere probabilities that a decision will lead to certain outcomes rather than others. Thus, given a person’s preferences and beliefs, it is assumed that she will act as if she is maximizing her interest (Friedman 1953; Hardin 1982). The externalist conception, then, denies that rational choice theory is a normative theory. According to the externalists, it is simply a descriptive theory that explains actions instrumentally but does not rely on mental representations or the internal reasons for action. Reasons are always external; the researcher attributes or assigns them to the individual, and they are determined by the underlying parameters of the social structure—the market, the family, the political party, social norms, the state, and so on—in which the individual interacts with others (Satz and Ferejohn 1994, 77).

While the formal component of rational choice theory is the same for both externalists and internalists, it is the interpretation of that component
that differs. For the internalists, beliefs, desires, and preferences are actual mental entities having motivational force, whereas for the externalists they are little more than instrumental concepts that describe external conduct, that is, the observable behavior influenced by contextual variables. However, these two divergent interpretations of rational choice theory do not always achieve different results: this divergence in the theory’s foundations need not influence the models or the outcome of research. But this is not always so. In contrast to what is often said in other spheres of sociological theory about the problem of social identity, rational choice theorists do not share the same view. Indeed, in this case a divergent interpretation of the theory’s foundations also leads to a different position with regard to the possibility of including identity in the models. As we see below, externalists and internalists approach the question of identity in different ways. While the former deny its utility, the latter attempt to include it in their models.

Rational Choice and Social Identity

Before turning to the question of how the internalist and the externalist interpretations of rational choice deal with social identity, let us take a moment to explore the very concept of identity, as it is sufficiently ambiguous to merit a definition.

An Initial Instrumental Approach to Social Identity

The literature on social identity is so vast that one might be tempted to believe that it is founded on a perfectly well-defined concept that is widely accepted among researchers. Yet this is not so, and in many cases social identity not only is vague, ambiguous, and enigmatic (Fearon 1999, 1) but also is made to do a great deal of work (Brubaker and Cooper 2000, 8; MacInnes 2004). The concept of identity has been employed to analyze noninstrumental, expressive modes of action that assumedly aid us in understanding participation in social movements (Melucci 1994). It has been used in turn to explain the fragmented nature of the contemporary self (Giddens 1991) and to reject that the contemporary self is fragmented (Rinken 2000). Identity has been interpreted as a fuzzy set of roles (Montgomery 2000), as a source of individual security and stability (Woodward 2003), and as a process of self-categorization that creates meaningful group boundaries that, depending on the context, may exacerbate the contrast with other groups through stereotyping and prejudice (Turner et al. 1987). And these are but a few of the examples chosen at random.
In spite of its enormous heterogeneity, social identity can be defined as “(a) a social category, defined by membership rules and allegedly characteristic attributes or expected behaviors or (b) a socially distinguishing feature that a person takes special pride in or views as unchangeable but socially consequential” or both things at the same time (Fearon 1999, 36). As a social category and a socially distinguishing feature, identity is, as well, a process of self-categorization and self-understanding (Turner et al. 1987), of identifying with others, and of commonality, connectedness, and groupness (Brubaker and Cooper 2000, 14-21). These are processes that permit the members of a group or collective to mutually recognize each other as such and define their social self (i.e., to respond to the questions “Who am I?” and “Who are we?”; Stryker 1980, 385; Callero 1985, 205; Gusfield, Johnston, and Laraña 1994; Calvert 2000). This process of self-categorization and mutual recognition, which is a social-cognitive process, shapes the perceptions and motivations of individuals and can result in increased intragroup cooperation and/or intergroup conflict. For example, pride at belonging to a club—or any other identity group—may prompt us to defend our club against other clubs or groups even when this involves a personal cost to ourselves (Orbell et al. 1988). Likewise, prejudice toward another group may lead us to confront it when it would be more rational and beneficial to cooperate. It is generally thought that identity is a motivational factor that explains actions focused inward to the group—usually cooperative action—and outward—which tends to be conflictive. Rational choice theory has no difficulties in explaining conflict—in terms of a conflict of interests—but given the theory’s fundamental assumption of self-interest, it encounters serious problems for explaining cooperation and, in general, any behavior that does not wed with the selfish dictates of private utility maximization. From the behavior of the median voter to the selfless and self-sacrificing revolutionary activist, from individual action undertaken in the name of social class or race to the action that puts group welfare before individual welfare, there is a whole rainbow of behaviors that standard rational choice theory

2. We lean on James Fearon (1999) because he proposes a useful synthetic definition after analyzing a representative set of identity definitions. (Fearon’s text is a draft. We thank the author for allowing us to cite it.)

3. In social psychology experiments to study if group identity explains cooperative behavior, it has been shown that conversations among individuals reinforce group identity—the sense of belonging to the group, of being “one of us” and not “one of them”—as well as promoting cooperation. Furthermore, even in experiments in which a group of strangers is assigned an arbitrary identity trait (e.g., those who have a red symbol versus those who have a yellow one), this trait generates group identity (Orbell, van de Kragt, and Dawes 1988).
would claim to be incomprehensible because of their irrationality and hence would exclude them from the explanatory sphere.

In contrast, identity theories assume that individual or collective social action is governed by “particularistic self-understandings [of a noninstrumental nature] rather than by putatively universal self-interest” (Brubaker and Cooper 2000, 6; Cohen 1985) and can therefore move social agents to perform actions that transcend their own interests and that are inexplicable from the standpoint of orthodox rational choice. Thus, it seems that identity provides an interesting heuristic tool that can serve to bridge certain explanatory gaps in the rational choice paradigm, which is otherwise very powerful. For this reason, in what follows we explore the question of whether and how it is possible to unify the two approaches to social action or, to put it another way, if it is possible, and how, for rational choice theory to incorporate (some of) the assumptions of identity theories. However, as we stated above, this will depend on whether rational choice theory is interpreted from an internalist or an externalist approach.

External Reasons, Rational Choice, and Identity

The externalist conception of rational choice has sound arguments for rejecting both the criticisms of identity theorists as well as the use of the concept of identity. Yet as we see in the following section, their arguments are not conclusive. The rejection of identity theories stems from the fact that the causal relationship between identity and action is an obscure one, and on many occasions identity theorists accept the theory as fact rather than providing evidence for it, attributing an explanatory power to identity that it lacks unless it is made explicit (MacInnes 2004, 533; de Francisco and Aguiar 2005, 13). For this reason, externalists often view explanations for social action in terms of supposed gender, ethnic, or class identities as being little more than just another way of saying that individuals share interests and preferences with the members of a group (Hardin 1995, 10; Laitin 1998). From the viewpoint of external reasons, it is held that the theories of social identity needlessly multiple the reasons for action. The externalist would say that identity must rely on the individual’s internal processes, on why the

4. According to Hardin, individuals identify those who share their interests, but identifica-

tion is not the same as identity: “If we did not have identifications, that is, commitments, it

would not matter so much that we have the quasi-objective identities we have—I as an Anglo-

Saxon-Celtic-Huguenot-Hillbilly-Texas-American, you as Tutsi, Serb, or whatever” (Hardin

1995, 7). Identity would be none other than preferences and norms or rules of cooperation. In

this same sense, and also from an externalist perspective, see Bicchieri (2002).
social agent states that she is together with others, on her way of understanding herself, on the image that she has of herself, that is, on reasons for action that are accessible only to the social agent herself. The social scientist—the externalist would add—lacks this privileged access and can rely only on the observable behavior of individuals embedded in the core of social structures. But unlike preferences, identity is not revealed in the action; it has to be overtly expressed. For the externalist, however, resorting to verbal affirmations about what individuals say they are involves making post hoc justifications about the social action, justifications that do not explain it. That the social scientist attributes an identity to an individual is not justified either if we accept that the only thing we have to rely on is the action of individuals or their aggregate effects.5

If the identity were revealed in the action, it could be reduced to mere preferences (revealed) about possible options: to join or not to join a cause, to go or not to go on strike, to demonstrate or not to demonstrate. As an essentialist and internalist concept of identity, it adds nothing to the explanation of these social events; it is enough to analyze the revealed preferences. Those who claim, for example, that a female worker takes part in a strike because of her class identity do not explain anything as they are resorting to an obscure internal entity of the individual—the identity or the conscience—without an apparent causal force on the action. If a person or a group of people declare that they take part in a collective action to promote equality, the externalist theorist of rational choice would say that the agents are attempting to satisfy their preferences (their taste for fairness) but that the action does not involve any type of internal process or definition of what one is—“I am egalitarian”—together with others. The action can be explained in the instrumental manner the externalist wishes (Friedman 1953; Coleman 1990).

According to this methodological approach, sociological theories of identity explain little with many assumptions: identity neither explains social action nor is a clear concept. At the same time, it is held that the internalist theory of rational choice adds needless assumptions: the externalist believes that the psychological realism of the internalist is unnecessary and that the formal theory of instrumental rationality is sufficient to explain all types of social phenomena. The external reasons—costs and benefits—are enough to construct hypotheses; it is not necessary to delve into the minds of individuals, their internal reasons for action, or their identity.

5. In other words, the externalist does not accept the utility of the concept of identity as either a category of practice (the individuals themselves refer to their identity) or a category of analysis (the social researcher assigns identities that may or may not coincide with those given by the social agents). For more on this distinction, see Brubaker and Cooper (2000, 4).
Internal Reasons, Rational Choice, and Identity

Yet by excluding internal reasons, a wide range of social events is left unexplained, and the externalist must deal with this limitation. The most well-known and widely studied case is that of the voter’s paradox. Why do people turn out to vote if the individual benefit to be gained from voting is negligible compared to the costs? As the political scientist Morris Fiorina (1976, 410) contends, “Expressive factors probably dominate instrumental factors as an explanation of turnout.” If we are to confront the challenge posed by identity theories without leaving the sphere of rational choice, we must consider how identity—the social category to which one belongs and/or the socially distinguishing feature in which a person takes special pride—is expressed as a further argument of a voter’s utility function. This function would be composed of instrumental elements (voting for the party that best represents my interests) and of noninstrumental elements that are expressive and serve purposes of identification (voting because it is a civic duty, voting to express what one is) (Calvert 2002, 593). For expressive voters, “voting is a means to express political beliefs and preferences and, in doing so, to establish or reaffirm their own political identity” (Schuessler 2000, 88). From the standpoint of internalist rational choice, the explanation for social phenomena proves to be very weak (albeit parsimonious) if it is restricted to assuming that people act as if their reasons were only instrumental and selfish. Mixed models à la Margolis (1982), those that include both instrumental and expressive motivations, have undoubtedly done much to advance the development of rational choice theory by permitting the incorporation of such issues as the agent’s social identity, issues that were excluded from the realm of conventional Beckerian or Olsonian models of rational choice. However, it would be wrong to assume that these modifications never affect the theory. As we

6. For an in-depth review of how rational choice theory has dealt with the problem of social identity in political science, see Calvert (2000).

7. It is questionable whether expressive motivations are actually excluded from the classic Olsonian models given that such motivations represent a sort of internal selective incentive. This is a solution that Olson himself considered and rejected. The person who gives money to a charitable organization states that she does so not because she believes that the money will have a large influence but because she obtains “an individual, noncollective satisfaction in the form of a feeling of personal moral worth, or because of a desire for respectability or praise” (Olson 1965, 160, n. 91). However, by considering any decision based on selective incentives of any kind as being rational, the theory becomes vacant of content because “this theory (or any theory) becomes correct simply by virtue of its logical consistency, and is no longer capable of empirical refutation” (Olson 1965, 160, n. 91). If an individual always maximizes her internal or external utility no matter what she does, the theory is worthless. This is a risk that internalist models run if they reduce all motives for action to a mere internal utility.
have already said, in the case of identity—which only the internalists seem to be interested in incorporating into rational choice theory—there has been a notable turnaround regarding some of the classic topics of rational choice:

If . . . voting for a candidate is expressively motivated—which is to say, the returns to the voter are existential rather than instrumental—then there is no collective action problem. While there is a public dimension to the instrumental efficacy of the vote (in that voters and non-voters alike partake in its electoral consequences), the existential consequence of voting is wholly internal. It is only the voter herself that can point to her electoral deed and legitimately draw her political identification from it: to be Democrat, she has to vote Democrat. It is not that the collective action problem has in any way been solved. It is simply that under expressive motivation, free-riding is no longer an appropriate analytical prism with which to view voting, as you cannot free-ride on expressive identification. (Schuessler 2000, 91)

Rational action based on identity (being a Democrat) can be understood only by relying on the voter’s internal reasons, which express what she is when she votes. We could say that Schuessler’s voter creates a practical syllogism “in her head” whose minor premise is existential: her own identity. The syllogism would therefore be as follows: Democrats vote for the Democratic Party (major premise), I am a Democrat (minor identitarian premise), so I vote for the Democratic Party (conclusion). Thus, the person who defines herself as a Democrat has an internal reason to vote for the Democrats, namely, her own practical coherence. By expressing her identity through voting, she does honor not only to what she is but also to her own rationality, given the meanings she assigns to certain events in the world—in this case, the fact of being a Democrat.

In contrast, to explain voting for externalist reasons, instrumental motives are sufficient: voters vote for the Democrats because it benefits them. Although this is important, it is insufficient to explain the voter’s paradox. People need expressive—internal—motivations to manifest what they are. But when one expresses what one is, one cannot be a free rider of oneself. This is Schuessler’s argument, an argument that cannot be wholly understood if we do not accept that his conception of rational choice is based on the existence of internal reasons (although the author does not refer to them as such).

8. For a similar argument on identity—understood as a set of roles—and on action, see James D. Montgomery (2000).
In their proposal to construct a utility function that permits economics to account for social identity, George Akerlof and Rachel Kranton do not refer to “internal reasons” either. But their analysis is an interesting essay that transports economic theory to the internalist terrain as the authors propose “a utility function that incorporates identity as a motivation for behavior” (Akerlof and Kranton 2000, 718, emphasis added). Departing from the empirical evidence on the weight of social identity in economic and non-economic conduct, Akerlof and Kranton were the first to introduce the concept of identity in an economic analysis. In Akerlof and Kranton’s proposal, the utility of individual $j$ is a function of $j$’s actions, of the actions of others, and of identity $I$ or $j$’s self-image. On one hand, identity is based on the social category that others attribute to a given person, while on the other, it is based on the “degree to which the traits $\varepsilon_j$ of $j$ coincide with the ideal social category that others attribute to him, which is indicated by prescription $P$” (Akerlof and Kranton 2000, 719). Social identity therefore involves accommodating one’s own image and thus one’s own actions to that ideal. Insofar as individuals internalize the code of conduct linked to a prescribed behavior, feelings of anxiety and cognitive dissonance may be evoked when violating the prescription. Likewise, the person may have a higher sense of self when accommodating the behavior prescribed by the reference group. Nonetheless, as Akerlof and Kranton point out, general agreement does not exist regarding social categories and prescriptions, thus providing the individual the option to choose her identity.

What emerges from Akerlof and Kranton’s proposal is that if the internalist interpretation of identity is excluded from utility functions, economics in particular and rational choice theory in general will be unable to explain certain social phenomena. As an example, both authors point out that the feminist movement’s struggle for equal rights has influenced the job market since, by transforming the social notions of femininity and masculinity, women have come to view working outside the home not only as an economic gain but also especially as a “gain in identity” (Akerlof and Kranton 2000, 719).9

9. In the first article in which they really apply their model empirically (Akerlof and Kranton 2002), the authors demonstrate that students’ performance depends on their identity, that is, the social category that is attributed to them (“crowd leaders,” “nerds,” or “burnouts”) and how they accommodate their self-image to this category. To maximize their utility, students must not only maximize their effort but also fit into a category. In a second study, they show how workers’ efficiency improves when they assume the values and principles of their company (Akerlof and Kranton 2005).
Needless to say, this is far removed from the methodological assumptions of Friedman and the externalist conception of identity and rational choice. *Identity-based preferences* provide reasons for action that can be understood only in terms of internal reasons and, therefore, by assuming a certain psychological realism. Staying at home involves not only external costs to women—to not earn a salary, to engage in certain tasks for reasons of gender, and so on—but also internal costs related to what they *are*, the way they conceive of themselves, their image understood as self-understanding. Women who are employed obtain not only external benefits—independence, income, relationships, and so on—but also internal benefits related to what it *means* to be a woman in their actual context. Working outside the home has external compensations that could lead to a rational maximization of utilities; but it also resolves a practical syllogism whose major premise would be “women today work outside the home” and whose minor premise would once again be existential and serve purposes of identification, “I am a woman.” Here, the rationality of action is measured in terms of correspondence with, or accommodation to, (socially) typified forms of behavior recognized by the agent. Akerlof and Kranton (2000, 719) spoke of coinciding “with the ideal social category that others attribute to him.” In short, we can say that identity forms part of the agent’s internal motives or reasons for action. It would seem then that the internalist approach to rational choice theory explains a larger number of social phenomena than does the externalist interpretation. However, as we see below, the internalist interpretation has its own limitations.

**Social Identity: Beliefs about Oneself and Rational Choice**

According to the internalist theory of rational choice, an action is rational if it is subject to a threefold process of optimization. First, the action must be the best means to fulfilling the agent’s desires given her beliefs about the means–ends relationship. Second, beliefs must be optimal given the information available to the agent. Finally, agents will attempt to optimize their information bearing in mind their previous beliefs about the cost of obtaining that information. If this optimization process fails in any way, the relationship between beliefs and desires will be blocked and the action will have no foundation in rationality (Elster 1989; Hedström 2005, 2006). Beliefs and desires are, as we have said, the building blocks of the explanatory arguments of action based on individual rationality.

The internalist, then, has two options for incorporating social identity into rational choice theory: by conceiving of identity either as a set of
Whenever rational choice has attempted to include identity, it has taken the first road. In this sense, it is understood that individuals not only want to maximize their benefits but also, as Schuessler argues, desire to express their selves, that is, to act in accordance with their social identity. Rational choice theory can therefore include the desire to express identity as yet another argument of the utility function. Voters not only would want the party that benefits them most (e.g., by cutting taxes) to win but also would desire to express their identity through the action of voting. The formal analysis must reflect both types of desires.

While this solution proves to be plausible from a formal viewpoint and constitutes a step forward for rational choice theory, from a conceptual perspective loose ends remain to be tied since the desire to express identity is not the same as identity, which is what is expressed (just as the desire to drink water when one is thirsty is not the same as thirst). What then is identity? As Akerlof and Kranton point out in a somewhat conceptually denser way, identity can be the result of social categorization as well as the image we have of ourselves. But what does the visual metaphor of “self-image” mean in this case? The image or idea that one has of oneself can be more complex than the mere desire to express it, which is what identity comes down to in some models. However, when making their concept operational, Akerlof and Kranton tend to reduce identity to identity-based preferences that are interiorized by the social agents, so that, once again, the very concept of identity is overshadowed by others. But if desires provide reasons for action based on beliefs in the internalist scheme of rational choice, is identity—regardless of how it is expressed—a type of desire?

Given the nature of desires, it proves impossible to define a person’s social identity as that person’s set of desires. Although desires can be involved in shaping one’s social identity (wishing to be something can be a strong reason for being it or attempting to be it), they cannot define it. If desire is considered a reason for action (x desires z and, given her beliefs, does action d to satisfy z), identity cannot be understood as a desire but rather as a source of desires. Desire in itself is not identity; it does not

10. Why not both, that is, a set of desires and beliefs? As we are going to see, we reject defining social identity as a set of desires.
11. If someone desires to be a pianist, at first glance we might think that this desire defines the person, defines (part of) her identity. However, the desire is not understandable without a set of beliefs that backs it and truly defines the person—her belief about her capacity to be a pianist, her belief about her capacity for working hard, and so on. Very much the same could be said of preferences. The fact that someone strictly prefers X to Y may be an expression of her identity but would not be identity itself.
define it insofar as desires have to do with how we want the world to be (Searle 1983, 54; Searle 2001), while social identity has to do with how the world is, that part of the world that is us.

While desires refer to how we want the world to be, beliefs have to do with how we believe the world is (whether or not we are right). We believe that the earth is round, we do not desire it to be round.\textsuperscript{12} We desire an orange juice, and we believe that what is placed before us is a real orange and not a wax or wooden one. We want a pay raise, and we believe that going on strike is the best way to achieve it. The difference between the mental state we call “belief” and the mental state we call “desire” lies in the \textit{direction of fit} with respect to the world (Anscombe 1957; Humberstone 1992): beliefs are directed at the world; they are representations about the world, and they tell us how it is (whether it is true or not). Desires, on the other hand, indicate how the world should be for it to satisfy us. Desires can be rational or irrational, but it makes no sense to say that they are true or false. Beliefs can be rational or irrational but also true or false. Moreover, the concepts of belief and truth are closely related in the case of both implicit and explicit beliefs (Falk 2004; Schwitzgebel 2006).\textsuperscript{13}

If it is not possible to consider social identity a set of desires, it fits into the internalist scheme of rational choice only by considering it a set of beliefs. But what type of beliefs are they? How are they shaped? What relationship do they have with truth? And what relationship do they have with social action? If beliefs are mental representations about how the world is, identity must be considered the persons’ set of beliefs about themselves shaped by beliefs about the concrete portion of the world to which they belong. Social identity—race, gender, class, and so on—will be a person’s set of beliefs about herself as, for example, a black working-class woman. In what follows, we explore some of the implications of this eminently \textit{epistemic} conception of social identity. These implications are summed up in the following four theses.

\textsuperscript{12} In reality one can desire anything: that the world be flat or that the sun be blue. But here we want to focus on desires based on well-informed beliefs and not on pure imagination or fantasy (although these often play a key social role). See Boudon (1994) on the social importance of false beliefs.

\textsuperscript{13} Beliefs are implicit when we do not mentally represent the contents of the proposition that expresses the belief (e.g., we believe that the sun will rise tomorrow, but we do not think about it). When the opposite occurs, the beliefs are explicit.
1. The Beliefs That Shape Social Identity Are Explicit, Indexical, and Shared

If someone says, for example, “I am Christian,” she is expressing her religious identity through language. But that expression, like any other expression of identity, encompasses a complex stockpile of beliefs about the world. No one can declare herself to be Christian without holding a set of beliefs about what it means to be Christian. Thus, in general, a Christian person will believe in the existence of God, in the resurrection of souls, in universal sins; she will believe that it is her duty to go to mass on Sundays or be charitable to the needy and have compassion for those who suffer. She will believe this and many more things that constitute what it means for her to be Christian: “Christians—she will say—believe those things.” Identity, however, cannot be expressed solely through this set of beliefs. Indeed, another person may claim to share the same beliefs about Christians and not be Christian. Thus, she must have a further belief about herself, a clearly indexical, self-referential belief—*I* am X—that is linked to a set of explicit beliefs about what it means to be X. This indexical element of self-referential belief and the explicitness of the beliefs about the meaning of identity are essential if the individual is to construct practical syllogisms through which identity becomes a motivational factor and beliefs establish a causal or intentional link with action. “Christians do x, y, z,” “I am Christian,” ergo, “I do x, y, z.” If beliefs were merely implicit, or there were no first-person pronoun to assign them to, it would be difficult for identity (now solely a diffuse and unattached identity) to motivate action (Chai 2001, 83; Falk 2004, 238; Jervis 2006, 642). It is important to highlight the normative or prescriptive component of the practical identity syllogism: “I do x, y, z” is in reality “I *should* do x, y, z.” It is important to emphasize this normative component because practical syllogisms are often not respected and the individual does not do what they prescribe.¹⁴

On the other hand, the set of beliefs that make sense of a particular social self—Christian, to use the same example—can only be a *shared* set of beliefs. If every Christian who declared herself to be a Christian believed different things about what it meant to be Christian, it would be impossible to speak of a religious identity. In contrast, sharing meanings is the basis of intragroup identification and contributes to reinforcing the beliefs that the members of the identity group have about the world and about themselves.

---

¹⁴ One of the most widely studied phenomena in philosophy regarding the failure of the practical syllogism—from Aristotle to Davidson—is the *akrasia* or the weakness of will.
The indexation of beliefs can thus move from the first-person singular to the first-person plural, that is, from the I to the we. When this is the case, identity will be expressed in the form of collective action: “We, the Christians, do [should do] x, y, z.”

2. The Beliefs Involved in Social Identity Are Robust Subjective Beliefs

One does not go to bed at night a Christian and wake up in the morning an atheist to once again become a Christian over afternoon tea. Social identities tend to be stable and are not usually subject to unexpected or sudden changes. This is because they are based on robust beliefs:

The belief is robust because it is, by the subject’s lights, evidentially extremely well supported, so well supported that it stands near the core of her web of belief. Such contrary evidence as comes in would be seen as requiring adjustments not with the belief itself, but elsewhere in the subject’s web of belief. At the other end of the spectrum, however, a subject’s belief might be robust because she is completely irrationally disposed to cling to her belief. (Sayre-McCord and Smith 2003, 4)

Robust beliefs are stable even when the individual receives new information that allows her to reflect on them (when beliefs change easily under the impact of incoming information and reflection, they are fragile). At one extreme, there are beliefs—such as the belief in the existence of God—that transcend the available evidence and are maximally robust. In general, we can say that the more robust an individual’s beliefs, the more stable her social identity. Now, as the above text suggests, a robust belief is the product of all the evidence that is available to a person according to her own viewpoint. Social identity, therefore, is not always constructed through a rational process of accumulating evidence, that is, one in which the information that could falsify our belief is maximized. On occasion, even the most countervailing of facts will not cause people to renounce their social identity, to renounce what they believe to be. The subjective point of view is crucial to the formation of identity beliefs.

15. Sayre-McCord and Smith’s text is a draft. We thank the authors for allowing us to cite it.
16. For this reason identity conflicts—in the case of immigrants, for example—produce obvious cases of anomie in the sense of Durkheim. Changing one’s country is more or less easy; changing one’s identity is not: the web of beliefs about oneself and about one’s world that we call “identity” is too closely woven to substitute it rapidly for another. In the case of fragile beliefs, which produce fragile identities, it is a different matter altogether.
This has two far-reaching consequences. The first is that social identity is not always based on rational beliefs. The second is that the relationship between truth and belief, in regard to social identity, becomes very complex given the subjectivity of the beliefs involved. On one hand, subjects believe that their beliefs are true: no people can define themselves socially through a set of beliefs about themselves and the world—beliefs shared with others—and immediately afterward add that these beliefs are false. Like all genuine beliefs, identity beliefs are also influenced by the parallel *metabelief* in their truth—"I believe that my beliefs are true." But on the other hand, these are robust beliefs that easily stand up to the evidence or for which the degree of evidence needed to justify them is merely subjective.

The fact that identity beliefs are in turn robust and subjectively true has often led collective movements that are firmly linked to questions of identity (especially ethnic or religious identity) to possess a component of irrationality and to continue to pursue their objectives in spite of repeated failures. The case of the so-called Basque liberation movement and the terrorist group ETA that is tied to the movement is revealing in this sense (Casquete 2003). That the movement has survived as long as it has can be explained only by what some Basques believe it means to be Basque, a belief that is as robust as it is subjectively true. Social identities are based on socially constructed, shared, and reinforced beliefs about the meaning of a certain social self, beliefs that do not have to be true, although it is essential that the social agents believe that they are. In fact, in many cases identity beliefs cannot be rigorously true or false as there is no empirical basis for their essentialist affirmations: there is no empirical fact with regard to “being Basque” that supports the universal affirmation “we Basques are like this or like that.” Yet those who hold that belief will hold it as being true. If the major premise of a practical syllogism is based on falsifiable beliefs, it can be rebutted rationally by providing new information. The subjective robustness of identity beliefs, however, means that the practical syllogism that leads to action cannot be refuted rationally from outside and is rational from inside, internally. Given the major premise, the agent is not irrational. However, this same major premise is based on nonfalsifiable beliefs—its semantics is a social construct—and therefore cannot be refuted by the same epistemic rationality. Surely this is one of the reasons why identity beliefs are not transformed by means of Bayesian learning, unlike the beliefs of rational choice models. The Bayesian models of learning (Breen 1999) assume that people act according to their prior beliefs and that they modify these beliefs to make room for others (posterior beliefs) in light of the outcome of their actions. In these models it is assumed that individuals assign a positive value to the objective truth of
their beliefs and the efficient accumulation of evidence, two assumptions that are often alien to identity beliefs.

3. Social Identity Embeds a Clearly Emotional Component

The above discussion does not exhaust the complexity of identity beliefs. Indeed, another dimension must be added, namely, that identity beliefs are “emotion laden.” Basically, what this means is that beliefs are accompanied by feelings. In her brilliant studies on classic ethics and Aristotle in particular, Martha Nussbaum (1986, 1994) emphasized that there are certain things—goods—that cannot be valued properly if our feelings about them are not adequate, that is, if our beliefs about such goods do not provide a given emotional charge. One cannot adequately conceive of the (relational) good of friendship or love if one does not believe that such a friend or lover is irreplaceable and does not fear losing her or him. Belief in the irreplaceability of the person who is the object of our erotic desire is essential if we are to properly value our love for her or for him, to understand love itself. If someone loves without this belief, Nussbaum would say, she does not know what it means to love. But since that belief embeds the fear of loss, which is passion, the conclusion is that we cannot understand love if we do not feel that fear.

This line of reasoning, which is currently supported by empirical research in cognitive science (Damasio 1994; Bechara and Damasio 2005), can be extended to say that identity beliefs are also accompanied by feelings and emotions, although not in a contingent or casual way but in a necessary and constitutive manner.17 Thus, we cannot perceive of ourselves as being Christians or Basques or Muslims if we do not hold certain feelings of pride or happiness or of mistrust and eventual hatred toward those who are unlike us (Smith 1999). On the other hand, as Nussbaum also points out in her interpretation and defense of stoic cosmopolitanism, there are emotional benefits in the fact of belonging: loyalties and identity ties (the more local the stronger the ties) provide the individual with an enormous amount of psychological security, they make a cozy nest of group certainties, they

---

17. It should come as no surprise that the externalist conception of rational choice rejects the explanatory value of emotions: “Explanations of ethnic conflict often invoke emotions. Unfortunately, explaining ethnic conflict as emotional may not be explaining it at all or may be explaining aspects of it given that it happens. The part we most need to explain is why the behavior happens, why such behavior is ethnically oriented” (Hardin 1995, 56). But to explain why the behavior happens it is often necessary to consider the emotional aspect, although this aspect does not explain everything. Remember that one component of Fearon’s definition is the feeling of pride that identity causes.
comfort the individual with direct recognition (Pizzorno 1986; Nussbaum 1997; Hogg and Mullin 1999). Thus, when we solve the practical syllogism of identity and act in a certain way, we not only gain the benefits of internal coherence but also laterally reinforce the emotional benefits of belonging. It goes without saying that although all of this seems to wrap human behavior in shadows of irrationality, it continues to be comprehensible.

Rational choice theory has overlooked emotions, in spite of the role that they play in providing pleasure, satisfaction, and utility (Elster 1996, 1386). It should come as no surprise that little effort has been dedicated to the study of the ties between identity and emotion. In the formal models of rational choice, beliefs and desires are taken as given, but emotions are not even mentioned; they simply do not exist. The fact that emotions have traditionally been considered irrational—and in this sense rational choice has been faithful to its own Humean tradition—has led to their exclusion from the realm of rational choice theory. Whatever the case may be, it seems clear that emotions are \textit{constitutive} of identity beliefs. Thus, if the internalist version of rational choice theory wants to include the former in its models, it must also include the latter.

4. Identity Beliefs Form Complex but Open Structures

Imagine that someone says “I am X” and “being X means that I smoke a cigar at 5:00 everyday.” If being X only meant that, my identity would be a simple one as it would motivate only one action. Yet social identities are not so simple. They are not so simple because they are based on complex belief structures. As we saw above, being Christian means very different but interconnected things. Indeed, empirical research may be able to even demonstrate the connection between identity belief structures, for example, that Christians tend to vote for the Right in elections. This could be because of the fact that Christianity tends to indirectly give rise to conservative beliefs about social reality. To put it another way, identity belief structures can have their own epistemic power and lead to other beliefs by establishing connections between an individual’s different identities. Obviously, this potential coherence can be thwarted by the coexistence of conflicting structures.

\footnote{In which it is inferred that not all beliefs about oneself constitute an identity belief. “I am dehydrated” is a belief about myself that is not an identity belief. Social identity can be conceived of as a complex system of beliefs about oneself and about the world, but, obviously, not all beliefs about oneself have to do with identity.}

\footnote{Social identity involves complex sets of beliefs, but given that a person may have a diversity of social identities, these sets can be contradictory, thus creating tension and uncertainty in the individual (see Aguiar and de Francisco 2002).}
However, given that belief structures are sets of meanings that form a shared semantics, they are subject to revision and reinterpretation. As we said above, there is no closed and definitive agreement about those semantics. Consequently, this permits the individual to negotiate internally to resolve her own conflict of interests or the conflict that arises as a result of incompatible prescriptions (Parfit 1984; Vescio et al. 1999). The agent may want to fit into the ideal social categories that constitute her identity, but she may lower her semantics—her demands—to satisfy other desires or needs; including that of belonging to—and being accepted by—other reference groups. She may also lower the emotional level of her identity beliefs with a view to maintaining other maximizing strategies, or she may reinterpret her identity beliefs to make them compatible with other beliefs. For example, voting for the left may involve the explicit reconsideration of what one understands by being Christian. Identity beliefs therefore form complex but open structures.

Conclusion

In this article, we have rejected externalist approaches to identity and examined internalists’ attempts to incorporate identity into rational choice models. In principle, this should pose no problems given that accommodating identity to socially typified behavior can have a symbolic utility (it reduces or eliminates the cost of deviation and also provides emotional benefits) and thus can be incorporated into an individual’s utility function—as the preference of a person for expressing her social being. We have also seen how this process is possible only from an internalist interpretation of rational choice theory as identity involves unavoidable mental processes and internal reasons that cannot be reduced to preferences revealed in choices.

Yet it is not such a simple task to include social identity in a rational choice model. It is not easy because, as we have argued, behavior motivated by identity ties and the behavior of rational choice models respond to two different rationales. In the case of the former, there is a logical internal coherence based on the accommodation (correspondence or coincidence) of behavior to the expectations prescribed by the ideal social category. In the latter case, it follows a logic of utility maximization given the beliefs and desires of the agent. The problem is that the two lines of logic not only are different but also do not have to be cumulative or compatible. Often, resolving the practical identity syllogism involves renouncing private utility maximizations. On the other hand, the two rationales may
become increasingly incompatible depending on the emotions involved in social identity, that is, in beliefs about oneself socially shaped. Emotions cannot be taken as given as they are constitutive of identity beliefs, and depending on their intensity they can block the maximizing motivations in the strict sense of the term. In social phenomena such as ethnic violence, religious fundamentalism, or nationalism, in which identity beliefs and intense emotions (fear, hatred, revenge) are completely intertwined, neither externalist nor internalist rational choice theories are well armed to explain social action because it is almost impossible to suppose that people are maximizing something. Moreover, we have seen how in less intense cases identity belief structures could be open to revision and reinterpretation, thus permitting individuals “to negotiate” both the meanings and the emotional component of their identity beliefs to resolve their internal conflicts.

Would it be possible, then, to reflect these complexities in a broader model of rational choice? Would this not involve the dangerous multiplication of assumptions and lead to a tremendous loss of methodological parsimony? While these questions serve to express our reservations, there is something that is unquestionable, namely, that our analysis of identity in terms of beliefs strongly points to the need for a linguistic analysis to gain a full understanding of human behavior (Asher 1986). Beliefs are mental events, and practical syllogisms are (or refer to) mental processes. However, “our sentences provide the only measure of the mental” (Davidson 2001, 77). When we speak of social identities, we refer to beliefs about the social world and about ourselves insofar as we are members of that world. We are speaking, then, of meaning, of shared meaning, of semantics that are socially constructed through language. In this case, what individuals say about themselves appears to be a central factor that must be analyzed if we are to understand their behavior. Identity beliefs are internal to the group, to the members of the group who claim to have them. The question about how can we broaden rational choice models turns then into the question about how can we broaden their theoretical foundations to include identity

---

20. This could translate into degrees of identification with the norms that are established in accordance with the roles with which the individual identifies (see Montgomery 2000, 264-66). Obviously, if the degree of identification is low, there is more room for alternative maximizing motivations.
21. Formally, it is always possible to argue that people are maximizing something but paying the high price of proposing a tautological proposition.
22. Aristotle himself viewed the practical syllogism as a logical ex post interpretation of a psychological or mental process (see the classic article by F. C. S. Schiller 1917), although the process is assumed to be internal.
beliefs. A first step would be to understand better the nature of beliefs about oneself and their indexical, social, emotional, and linguistic character, even if we have to go far beyond rational choice theory.

References


Philosophy of the Social Sciences


Fernando Aguiar (faguiar@iesa.csic.es) is a tenured scientist at the Institute for Advanced Social Studies–Spanish Council for Scientific Research (IESA-CSIC). His research interests cover rational choice theory, social identity, experimental philosophy, and republicanism. He has published on these subjects in European Journal of Sociology, Judgment and Decision
Andrés de Francisco (adefrancisco@cps.ucm.es) is a tenured professor at the Faculty of Political Science and Sociology in Complutense University of Madrid. He received his doctor in philosophy in 1989, and he currently works in the fields of methodology of the social sciences, social theory, political theory, and republicanism. He published “A Republican Interpretation of the Late Rawls” in Journal of Political Philosophy (2006), and his latest book is Ciudadanía y Democracia: Un Enfoque Republicano (Citizenship and Democracy: a Republican Approach; 2007).