Book Review Column

Does Theory of Rationality Need Folk Psychology?

Elias L. Khalil
Ohio State University, Mansfield

While the term “folk psychology” has derogatory origins, it is no longer used in that sense. It is usually defended by philosophers, like Jerry Fodor [1987], to highlight the importance of commonsense notions like belief, desire, meaningful goal, conscience, love, hope, and so on. Adherents of folk psychology usually oppose the modeling of the mind as a series of computational or neural networks. In contrast, eliminivist materialists, like Patricia Churchland [1986], charge that folk psychology, in the face of the rise of neuroscience, will eventually go the way of phlogiston, witches, and the geocentric theory. Marvin Minsky, another critic of folk psychology, puts it this way: “Though prescientific idea germs like ‘believe,’ ‘know,’ and ‘mean’ are used in daily life, they seem technically too coarse to support powerful theories; we need to supplant, rather than support and explicate them” [quoted in Flanagan, 1991, p.223].

The standard theory of rationality, without the help of neuroscience, has provided technically refined concepts like “preferences,” “constraint,” and “maximization” in order to supplant commonsense ideas like entrepreneurship, alertness, status seeking, worthy goals, pride, desire, and so on. Standard theory has not been eager to explain such folk terms. Instead, it has offered a set of axioms which could be logically consistent, situate choices in objective lexicon, and make deductive propositions amiable to mathematical manipulations.

I.

The book at hand sprang from a series of seminars organized around a graduate course on philosophical issues in economics at the University of Cambridge. The book has two well-represented, but unrelated themes: The minor one, relegated to the end of the book, is concerned with how agents make decisions under uncertainty and twisted rationality. For lack of space, I ignore the three essays which deal with that theme, viz., John Broome’s “Rationality and the Sure-Thing Principle,” Robin Matthews’ “Animal spirits,” and the editor’s “Keynes on the Rationality of Decision Procedures Under Uncertainty: The Investment Decision.”

The second theme of economics and morality, on which I focus, is highlighted by the editor [p.1] in the Introduction: “Rational economic man has sometimes been very narrowly conceived as a crudely calculating self-interested maximizer.” Do we need a radical upheaval in economics, along lines suggested by folk psychology, to accommodate observations that humans are neither exclusively self-interested nor crude maximizers?

Frank Hahn, in “Benevolence,” answers that we could still work within the standard theory. Hahn admits that benevolence is not disguised selfishness. But the variety of motives should not lead to the divided-self approach as supposedly advocated by Amartya Sen [1977]. Hahn [p.8] asks: “Why do you give the beggar ten pence and not ten pounds? Why do you visit your friend in hospital once a week and not daily?” He answers is that you must be comparing diverse tastes: “This essentially becomes a theory of the integrated personality: a man knows what he wants.” (Hahn then raises an unrelated subject; viz., it is better to channel benevolence activity through government and charity organizations, rather than make it part of private action.)

Amartya Sen, in “Beneconfusion,” focuses exclusively on Hahn’s theses. Sen denies that he advocates a divided-self approach in the 1977 article. Sen [p.13] clarifies that Hahn seems to have missed the major thesis, viz., “whether an integrated person’s preferences can be seen in terms of his or her self-interest.” That is, the 1977 article was not concerned with whether agents can consistently order their tastes between self-interest and benevolence — which they could. Rather, the ordering among diverse objective cannot be
simultaneously claimed by the standard general equilibrium theory to represent the person’s self-interest or individual well-being. As Sen puts it:

A person who chooses on the basis of a systematic preference ordering that takes note of self-interest as well as other objectives may well possess an integrated personality, but his or her behaviour would not typically coincide with one of self-interest maximization [p.14].

David Collard, in “Love is not Enough,” goes beyond his earlier book, *Altruism and Economy*. In that book he argued that altruism is so widespread that it warrants including the concern for others in the utility function. Here he skillfully develops the Kantian rule of ethics in relation to how to channel altruistic sympathies. He distinguishes between the categorical imperative, which includes goals like “apartheid is an evil system,” and the hypothetical imperative, which includes instruments like “one should boycott South African goods.” Collard argues that love is not enough especially with regard to the hypothetical imperative: The person must have at least some intuition (hopefully assisted by social theory and information) about the outcome, before he boycotts South African goods.

In his article, “Exploitation, Justice and the Rational Man,” Alan Ryan objects to the agenda of analytical Marxism of John Roemer and Jon Elster who try to save Marx on the basis of methodological individualism. For Ryan, rational theories have to appeal to a theory of justice to determine if a certain relation is exploitative. For Marx, there is simply no such an ideal realm of values; principles of justice are social products of the particular formation. Thus, to develop a theory of exploitation based on rational property rights, which is difficult to accomplish, necessitates the unfortunate abandonment of Marx.

In contrast, Tony Cramp, in “Pleasures, Prices and Principles,” argues for the abandonment of utilitarian rational theories from “explicit Christian allegiance.” Cramp situates the articulation of the goals of the agent within ethical norms of stewardship, which includes the adherence to principles, the discounting of profits arising from excessive exploitation of resources, the containment of destructive competition, sacrificial giving, and the resistance of temptations. Unlike Thomas Schelling, Cramp does not view the challenge of the weak will as the contest among equally different selves. Rather, following St. Paul, William Law, and Kant, the countering of greed expresses the will of reason or what Adam Smith [1976] called the impartial spectator (conscience). Thus, for Cramp [p.63], “rational action is something to be achieved by struggling against desires, rather than by yielding to them.”

II.

Do folk notions about temptations, principles, benevolent actions call for a radical reconstruction of the standard theory of rationality? I think, as Collard has shown, that neoclassical theory can handle benevolence well. I contend that critics who have focused on the concern for others have been barking up the wrong tree. Neoclassical theory could be challenged on more serious counts, of which I discuss two. Namely, one may more fruitfully raise, first, the existential deliberation of goals as advanced by folk psychology and, second, how pride/shame fit into the ordering of options.

In order to clarify the two challenges, let me state first how benevolence has been treated in two ways by neoclassical theory successfully, i.e., without trivializing it as sophisticated egoism or self-interest. First, Gary Becker’s [1976] Rotten Kid Theorem shows that altruism (of a parent) arises from including someone else’s welfare (the offspring) in the agent’s utility function — parallel to the “inclusive fitness” hypothesis of sociobiology à la Edward Wilson. That is, the tradeoff is between your consumption of a banana vs. the pleasure of seeing someone else consuming your banana. Of course, this gives rise to infinite regression if the beneficiary’s pleasure constitutes also the knowledge about the motive of your charitable act. Aside from this problem, which could be neutralized by restricting the magnitude of the interaction [Becker, 1974, pp.1080-1081n], Becker’s Rotten Kid Theorem raises the issue that the agent is still not acting out of principle or conscience — but rather out of egocentrism.

To be accurate, however, egocentrism should not be confused with disguised egoism because the altruist, in this story, truly enjoys the pleasure of others; i.e., he does not expect other pay offs in the future. But Becker’s altruist is egocentric: As Adam Smith [1980, p.317] pointed out in his critique of Hobbes’s egocentric theory of benevolence, the actor according to Hobbes judges the pleasure of the beneficiary from the actor’s ego-centered station, rather than from a third, impartial station.

The second neoclassical solution worth noting is Robert Frank’s [1988] work which seems to escape from the egocentric charge leveled at Becker’s Rotten Kid Theorem. Frank offers a theory which equips the actor with conscience. He posits that if there is a mutant for honesty in a population
which could be detected by signals like sweating, facial twitching, and grimacing — others would like to cooperate with the carrier of the mutant in order (if they are honest themselves) to ensure their rewards or (if they are dishonest) to exploit the carrier. Therefore, there is a tendency for the honest gene to replicate faster than the dishonest one. In fact, the honest gene could become 100% dominant if the cost of checking the truthfulness of the signal is zero. That is, the ratio of honest members in a population is a negative function of the cost of checking deception.

In this light, for Frank, there is no need to assert that people act according to their commitment because it makes them feel “good” in the same sense as the consumption of a banana in Becker’s approach. However, Frank’s non-egocentric solution engenders another problem. Namely, people are either born honest or dishonest. This invites biological reductionism from which Becker’s Theorem is free. In Frank’s moral theory, there is ironically little moral choice. A more related critique of Frank’s biological reductionism is offered by Jerry Evensky [1992, pp.28-29]. Frank’s approach, which poses a one-to-one relation between the honesty gene and actual action, ironically implies that there is no use for signals. Signals like blushing and sweating are only useful when an honest person is doing an unconscionable act because they indicate internal conflict. However, they cannot reveal the dishonest agents when acting unconscionable or the honest agents when acting conscionable. Thus, aside from the cost of checking signals, signals cannot be relied on to distinguish between the honest and dishonest.

But, it is possible to eschew such an internal problem in Frank’s biological explanation by assuming that agents develop a taste for conscience via socialization, or simply possess such a taste like the case with other tastes. Thus, Frank’s approach, like Becker’s, shows how neoclassical theory is capable of handling benevolence without reducing it to sophisticated egoism.

I may venture and add that even Sen’s critique of general equilibrium theory, viz., that the ordering of diverse motives in the utility function undermines the maximization condition, is not necessarily valid. It is true that the maximization can no longer be carried out insofar as it expresses the self-interest of the agent. However, the maximization can be carried out insofar as it represents the interest of, to use Adam Smith’s [1976] folk psychology terminology, the impartial spectator. That is, Sen’s critique is valid only if it maintains that the diversified objectives (from self-interest to benevolence) of the utility function cannot represents the maximization of self-interest. But, the critique is invalid if it holds that the diversified-objective function represents the maximization of another person, the imagined impartial spectator who resides in the third station. In fact, the Adam Smith’s notion of a third station may show the road for a non-egocentric, anti-Becker conception of benevolence [see Khalil, 1990].

Thus, I may state categorically, especially in light of Smith’s impartial spectator, that the phenomena of benevolence is no longer a valid banner to wage against the standard theory of rationality. That does not mean there are no serious problems with the standard theory. As alluded above, there are two major problems, viz., first, the folk psychology challenge and, second, the pride/shame phenomena.

Concerning the first problem, Hahn argues: This [the comparability of tastes thesis] can be achieved in thinking of the individual not as in search of happiness, or respectability or physical well being but as having preferences over a suitable domain of alternatives [p.10].

This is not a semantics question, but rather an existentialist one. Should we view humans à la neoclassical theory as standing outside their tastes or goals and calibrating which option to choose in light of costs? Or should we conceive humans à la folk psychology as inseparable from their goals as much as they are inalienable, at least without great struggle, from their personality and neurotic traits. The question dividing the two views is not about instrumental rationality, but rather about end rationality.

To illustrate, Shaun Hargreaves Heap [1989] calls the concern about the end itself “expressive” rationality. For him, expressive rationality, which is similar to Cramp’s concern with Christian stewardship, is different from what he calls the “instrumental” and “procedural” kinds of rationality. In fact, Hargreaves Heap’s instrumental and procedural kinds are alternative approaches to the question of means: His instrumental kind is expressed in the standard neoclassical view of the agent as maximizer. In contrast, the procedural kind is represented, inter alia, in Herbert Simon’s view of the agent as satisficer or follower of rules of thumb. In contrast, Hargreaves Heap’s expressive rationality is about a different problem, viz., goals like self-respect and entrepreneurship:

Expressive rationality reflects our concern with making sense of the world... It is not
enough for a person to be guided by preferences; preferences must reflect an autonomous self... Autonomy and self-respect come through taking one’s preferences to be worthy of one’s self and questions of worth are notoriously difficult to answer. Consequently, we have a picture of the individual who is groping for what is worthwhile: she/he experiments with and discusses this or that [Hargreaves Heap, 1989, pp.4-5].

Concerning the second problem facing the standard theory of rationality is the failure of neoclassical theory, as well as its critics reviewed here, to differentiate between ordinary tastes and commitment. Ordinary tastes, what I call “constitutives,” are like the preferences for income and leisure, while commitment, what I dub “typological,” gives rise to the by-product feeling of pride or honor for selecting a wanted constitutive option. A typological taste includes the feeling of shame for, as Cramp has argued, succumbing to last minute temptations orrationally unwanted desires. That is, there is a difference between the utility derived from eating a healthy diet (longer life) and the pride derived from maintaining such an unappetizing course. In fact, the constitutive/typological distinction cuts across the difference between self-interest and benevolence. One derives utility, e.g., from quitting smoking or watching a significant other quitting smoking. Such a constitutive benefit seems to be different from the self-respect stemming from enduring the withdrawal period or the pride arising for helping the significant other through the ordeal.

If the proposed constitutive/typological distinction is valid, it may explain why self-respecting scientists would feel guilty (aside from the risk of punishment and the magnitude of benefit) for tampering with data in order to enhance the prospects of tenure and promotion. The constitutive/typological differentiation may also explain how an action might be considered shameful without resorting to ahistorical or natural law reasoning of supposed innate moral principles. For example, there might be nothing inherently wrong with the selling of sexual services or renting of uterus (like in modern surrogate motherhood). It becomes shameful if the agent judges (of course, under the influence of sociocultural tastes) that his or her body is part of the identity and, hence, is not for sale, but succumbs to the “seduction” of pecuniary or other benefits.

The question of Hargreaves Heap’s expressive rationality and the issue of pride/shame pose, I think, more formidable challenges to the neoclassical orthodoxy than the discovery that humans act benevolently in certain cases. The deliberation of what seems worthwhile objectives and the feelings of pride and shame, as much as the Chicago economists would like to liberate us from such relics, cry for explanations. With respect to pride, it certainly undermines the idea that the typological product lies along smooth indifference curves along with the taste which occasions it. However, the challenges of expressive rationality and the feeling of pride should not, as the advocates of folk psychology maintain, be the basis for throwing the standard theory of rationality to the tigers. In order to retain the rigor noted by Marvin Minsky, I think it is recommended to start from the standard approach and show how it could be refined in order to account for the commonsense notions of expressive rationality and pride. By all means, however, this is a long and challenging task.

References


