

I

Preferences, Comparative Evaluations, and Reasons

I.I. WHAT ARE PREFERENCES?

When English-speakers talk about preferences, they do not need to use the words “prefer” or “preference.” The waiter or waitress may ask Jack, “Which would you like?” rather than “Which would you prefer?” We often ask about preferences by asking people which alternative they would choose, which they like better, or which they think would be better.

There appear to be four main concepts of preference:

1. *Enjoyment comparisons.* When I say that my dog Itzhak prefers ground beef to dry dog food, or that my mother prefers yellow roses to red ones, I am talking about what gives them more enjoyment. In this sense, an agent such as Jill prefers x to y if and only if Jill enjoys x more than she enjoys y . Let us call this an *enjoyment comparison*. It may compare either overall enjoyment or enjoyment in some particular regard.
2. *Comparative evaluations.* When a politician such as Margaret Thatcher expresses a preference for one policy over another, she is not reporting her enjoyment. She is talking about which policy she judges to be superior. In this sense, an agent such as Thatcher prefers x to y if and only if she judges x to be better than y in some regard. Let us call this a *comparative evaluation*. Comparative evaluations can be *partial* – implying a ranking with respect to some specific criterion – or *total* – implying a ranking with respect to every relevant consideration. From an ideological perspective, Thatcher might object to government provision of health care, while from a political perspective, she might favor bolstering the National Health Service. When, as in this case, an agent’s partial evaluations conflict, the agent must adjudicate among the different considerations if he or she is to construct a total ranking. Making a total comparative evaluation is more cognitively demanding than making an enjoyment comparison, and there is more room for mistakes.

A total comparative evaluation takes into account *every* consideration the agent judges to be relevant. Unless there is a blunder in execution, total rankings determine choices. I shall argue that preferences in economics are for the most part, and should be, total comparative evaluations.

3. *Favoring.* When people say that affirmative action calls for racial preferences, they mean that it favors racial minorities – that is, it gives them a better chance to be hired by firms or admitted by universities. There is no implication that those who implement these preferences like members of these minorities better or judge them to be superior. Favoring does not imply either an enjoyment comparison or a comparative evaluation.
4. *Choice ranking.* Finally, when a waiter asks Jill whether she would prefer the soup or the salad, he wants to know simply which Jill chooses. In this sense, Jill prefers x to y if and only if when she knows she is faced with a choice between x and y , she chooses x . Call such preferences *choice rankings*.

The four senses of “preference” often come apart: what counts as a preference given one definition does not count as a preference given another definition. For example, having recently seen the results of a cholesterol test, Jill might judge that, all things considered, the nonfat vanilla frozen yogurt is a better choice than the chocolate fudge ice cream she enjoys more. Because of laziness, habit, or mistake, Jack can choose one alternative while believing that another is better. Preference relations may be of very different kinds. Enjoyment comparisons and comparative evaluations are mental attitudes, whereas favoring and choosing are actions.

Translations of “preference” and “prefer” into other languages do not always have all of these meanings. For example, German-speaking friends tell me that the German equivalent is not used to express choices, while Chinese-speaking friends tell me that there is no good translation at all of “preference” into Chinese. These linguistic claims do not, of course, show that German-speakers are incapable of talking about what they choose. They just use different words from the ones they employ to make comparative evaluations. Nor are the Chinese incapable of comparing enjoyments, of making comparative evaluations, or of reporting what they favor or choose. But the fact that the counterparts to “prefer” and “preference” in other languages do not have all four of these meanings underlines the fact that there are four concepts here, not just one. I conjecture that English permits the use of the same word for all four concepts because they all involve favoring in some way, whether in feeling, judgment, or action.

Whether they use the word “preference” or some word in another language that roughly corresponds to it, people are constantly talking about their preferences, considering what their preferences ought to be, and inquiring about each other’s preferences. Sometimes preferences require little thought.

Enjoyment comparisons require sensitivity to how one is feeling, rather than judgment. They are more like introspective reports than judgments, which may sometimes require extensive deliberation. A serious legislator's preferences among alternative social policies may depend on months of study. Even my son's early preferences reflected deliberation on his part. It may be hard to know what one prefers.

I.2. OVERALL AND TOTAL COMPARATIVE EVALUATIONS

This book focuses on preferences as *comparative evaluations*, and it takes the alternatives that people compare to be actions, their consequences, and states of affairs. I take claims about preferences among objects or properties to be elliptical statements of preferences among states of affairs. Thus, for example, to prefer Gore to Bush is to make a comparative evaluation of the states of affairs in which Gore or Bush is elected.

Individuals can evaluate alternatives in specific regards (i.e., a Bush presidency is better from the perspective of affluent taxpayers), or they can make *overall* or *total* evaluations. In everyday usage, preferences are typically "overall" comparative evaluations. In an overall evaluation, agents compare alternatives with respect to *most* of what matters to them rather than in some specific regard or with respect to everything that matters to them. In an overall comparative evaluation, in contrast to a total comparative evaluation, people regard some of the factors that affect their evaluation of alternatives as *competing with* preferences rather than as influencing preferences (Reynolds and Paris 1979, p. 356). For example, consider the following two statements concerning Jill:

1. Jill drank water rather than wine with dinner, despite preferring to drink wine, because she promised her husband she would stay sober.
2. Jill drank water with dinner because she preferred to do so. But for the promise she made her husband to stay sober, she would have preferred to drink wine rather than water with dinner.

The difference between these claims is illustrated in Figure 1.1. In Figure 1.1a, Jill's promise has no effect on her preference ranking. Instead, like a constraint, it rules out her preferred alternative and, jointly with her preferences, determines her action. In contrast, in Figure 1.1b, Jill's promise reverses her preferences between water and wine and affects her action exclusively via changing her preferences. The two sentences (1) and (2) might describe the same state of affairs, and everyday usage permits both of them, although (1) sounds to my ear somewhat more natural than (2). Standard usage favors taking feelings and tastes as influencing preferences, while regarding cognitively sophisticated evaluative factors, including especially moral considerations, as competing with preferences. But there are no hard and fast rules. One bad reason to favor formulation (1) is if one mistakenly thinks of moral

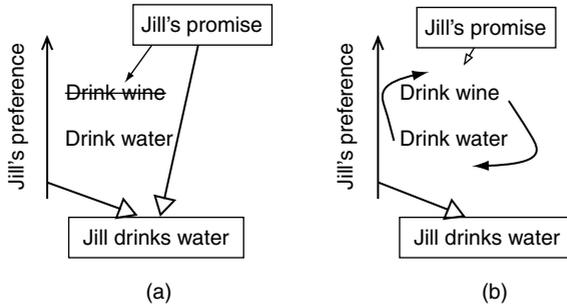


Figure 1.1. Promising and preferring.

obligation as something external to the self – like a ball and chain preventing people from doing what they truly want to do rather than as part of what determines who people are and what they value. A somewhat better reason to adopt (1) is that in their deliberations, people often treat moral considerations as constraints rather than as factors to be balanced against others. In any case, ordinary usage permits moral requirements both to compete with preference as in (1) and to influence preferences as in (2).

In everyday speech, it is natural to speak of preferences as subjective attitudes that determine how one ranks alternatives, rather than as identical to the ranking, but nothing important is at stake if I simplify and identify preferences with the rankings they determine.

To sum up: This book is concerned with preferences as comparative evaluations. Comparative evaluations can be partial, overall, or total. When people express their preferences, they are typically expressing an *overall* comparative evaluation – that is, a comparative evaluation in terms of almost all of the considerations that they find relevant. In contrast, preferences in economics are *total* comparative evaluations. As I have already noted, people sometimes use the terms “prefer” and “preference” in other ways.

1.3. PREFERENCES, REASONS, AND FOLK PSYCHOLOGY

People think about what to do and explain why others do things in terms of *reasons*. To explain why Jill went to the library, Jack might say, “Because she wanted to borrow a copy of *War and Peace*.” In deciding whether to go to the library, Jack might think, “Jill is there. It’s great to spend time with her.” Jack explains Jill’s trip to the library in terms of Jill’s desire to borrow a copy of *War and Peace* and, implicitly, her belief that the library has a copy. In considering what he has reason to do, Jack cites what he believes to be a fact and what he takes to be valuable. In explaining Jill’s actions, Jack employs “folk psychology,” the implicit theory that people use to explain, predict, and rationalize human actions in terms of beliefs and desires. The explanations and predictions that economists give are refinements of folk psychology. Most psychologists characterize the decision making of human beings from an early

age as conforming to folk psychology, and most philosophers also endorse the explanatory and predictive merits of belief-desire psychology. But there are controversies in each field. Some philosophers have gone so far as to question whether beliefs and desires exist.¹

Among those philosophers who grant that beliefs and desires help explain actions, most would argue that taking actions to be merely the causal consequence of beliefs, desires, and constraints falsifies the character of rational deliberation. From the first-person perspective of someone thinking about what to do, *reasons* are paramount. The question agents ask themselves is not “Given my beliefs and desires, what do I predict that I will do?” but “What *should* I do” or “What do I have most *reason* to do?” These may sound like moral questions, but let us set moral considerations aside. Questions about what one should do remain. To decide what to do, I try to be guided by what the facts are and by what is valuable, even though I know that my beliefs may be faulty and my desires may be nutty. That I happen to believe *P* and desire *X* is not decisive, because I can step back and question my beliefs and my desires. My awareness of a desire to do *X* does not automatically incline me to do *X* intentionally unless I can see some *reason* to do *X* (Quinn 1995; Scanlon 1998; Schapiro 2009). If I see no reason to do *X*, I will try to suppress my desire to do it. A fervent desire to eat mouse droppings sends people to a therapist rather than to their mouse-infested basement cupboards. Regardless of how value is to be understood, deliberation forces deliberators to regard values, like facts, as realities with which actions must come to terms, not as psychological crotchets.² From Jack’s perspective, the reason to go to the library is the *fact* that Jill is there and the *value* to him of being with her, not his psychological state. Although agents can step back and ask the explanatory question “Now, why in the world did I do that?” the fundamental question from a first-person perspective is the justificatory question “What do I have most reason to do?”

Social scientists are not usually asking the first-person question. They look at behavior from a third-person perspective, where explanatory and predictive questions are paramount. But if reasons govern deliberation, third-person accounts cannot ignore reasons. Moreover, there must be a close connection between the reasons that justify actions and the beliefs and desires in terms of which others predict and explain them. The need for a connection does not imply that beliefs and desires *be* reasons. Beliefs and desires can explain why facts and values constitute reasons for action without the beliefs and desires themselves being part of the reason (Schroeder 2007, chapter 2).

¹ “[O]ur commonsense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience” Churchland (1981), p. 67. See also Stich (1983).

² As discussed in Chapter 6, the agent may also have relatively settled plans or intentions that may guide his or her deliberation. See especially Bratman (1987, 1999, 2007a).

Beliefs are linked to reasons, because beliefs purport to provide agents with facts, and facts can be reasons. The connection between desire and value is less clear. It seems that there are almost always reasons (albeit not necessarily good ones) to act on even malicious and foolish desires, such as desires to pull a fire alarm or to have “one more for the road.” (But it is hard to see what reason there could be to eat mouse droppings.) Why should it be that there is typically reason to do what one desires to do? One plausible answer is that desire presents its object as in some regard valuable or “to-be-done” (Stampe 1987; Scanlon 1998; Schapiro 2009). On this view, desire motivates action and is *prima facie* evidence that there is reason to seek its object.

On this view of desire, there is an additional important distinction between desires and preferences. Desires provide *prima facie* evidence that there is reason to seek their objects. (Satisfying a whim may be reason enough if doing so is not otherwise ill-advised.) When those reasons appear not to exist, individuals take steps to purge themselves of their distorted desires. Preferences are not like this. As total or overall evaluations, they are already informed by reflection on what there is reason to do. They are the outcome of a comparative assessment, rather than inputs into deliberation like desires.

From a first-person perspective, facts and values rather than beliefs and desires count as reasons. If Jill is not at the library, then Jack has no reason to go there. If he believes she is there, he thinks he has a reason, but he is mistaken. From a third-person perspective, in contrast, the “facts” (as seen by the agent) are simply the agent’s beliefs and the “values” (as seen by the agent) are simply the agent’s desires. So from a third-person perspective, the agent’s reasons are his or her beliefs and desires. In addition, questions about the causes of behavior are, from this perspective, at least as important as questions about what justifies behavior. So the social scientist is interested in beliefs and desires (or beliefs and preferences) as both causes of behavior and reasons for behavior.

Figure 1.2 illustrates the contrasts between first-person and third-person accounts of Jack’s presence at the library. The left-hand side represents the situation as Jack sees it when he is deciding to go to the library. Facts and values are reasons for action. They serve as premises in deliberation. The right-hand side represents the situation as a third party sees it. Beliefs and preferences are both reasons for action and causes of actions (as well as causes of other beliefs and preferences). As a commentary on social science (and especially economics), this book will mainly adopt a third-person perspective, and, when speaking of reasons, I shall usually take them to consist of preferences and beliefs.

Reasons for actions do not always explain actions. Jill’s being at the library and the attractions of being with Jill are reasons for Jack to go to the library, whether or not they actually influence his behavior. It might be that what explains his going to the library is instead that Jack’s furnace is malfunctioning, and the library is nice and warm. When Jill’s presence does not influence

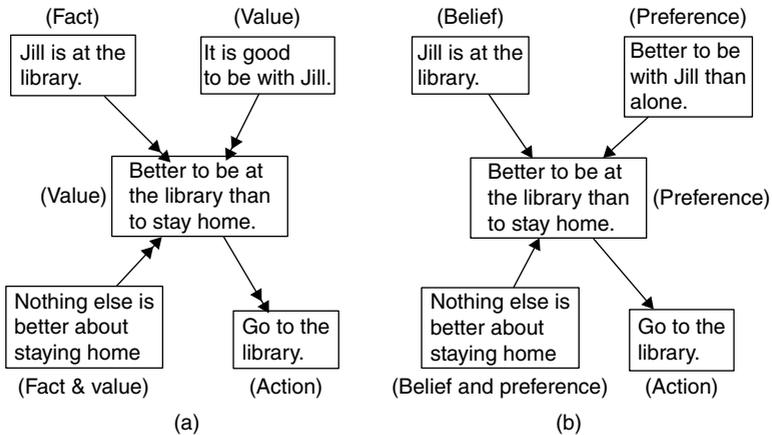


Figure 1.2. Reasons and causes.

Jack's action, it does not explain it. Donald Davidson (1963) has argued compellingly that the difference between reasons that explain actions and reasons that do not explain actions is causal: Reasons that explain actions are also causes of the actions they explain. Reasons that do not explain actions may still serve to *justify* or *rationalize* actions, to show what makes actions rational. Jack's belief that Jill was at the library gave Jack an apparent reason to go to the library, whether or not it was responsible for his going.

Given that preferences (or the facts and values that ground them) serve as reasons for actions, theories of choice should show how they do so. *Theories of choice that purport to identify reasons for choice must be theories of rational choice.* As empirical theories, they must also describe choice behavior and its causes. Although the normative adequacy of a theory of action is a separate question from its empirical adequacy, normative considerations influence empirical theorizing.

1.4. MISCONCEPTIONS CONCERNING PREFERENCES

This book argues that preferences in economics are rankings that express *total* subjective comparative evaluations and that economists are right to employ this concept of preference. Modeling behavior in terms of preferences and beliefs is an attractive path for the other social sciences as well. My quarrels are less with how economists employ the concept of preferences than with how economists have described what they are doing. This is not to say that the practice of economists has been perfect. On the contrary, mistakes in characterizing economic practice have sometimes bled into the practice itself, and I shall accordingly criticize some of what economists do and suggest ways to improve both positive and normative economics.

I shall in particular be concerned with five problematic claims concerning preferences:

1. *Arbitrariness*. Preferences are matters of taste, concerning which rational criticism or discussion is impossible.
2. *Self-interest*. People's preference rankings coincide with and are determined by their rankings of alternatives in terms of their expected self-interested benefit.
3. *Revealed preference*. Preferences in economics can be defined in terms of choices, as in revealed-preference theory.
4. *Division of theoretical labor*. Economists have nothing to say about how preferences are formed and modified and need not concern themselves with these matters.
5. *Welfare as preference satisfaction*. Welfare is the satisfaction of preferences. What makes x better than y for some agent A is A 's preferring x to y .

These theses exaggerate sensible approximations or modeling strategies and distort important and defensible insights. Some, such as (2) and (3), are formulated explicitly and vigorously defended by some economists. Many economists regard claims such as (1), (4), and (5) as obvious.³ None of the five claims is defensible. At best, some (such as (2) and (4)) are reasonable approximations in some specific contexts. In many instances, correcting these misconceptions will merely help economists understand better the sensible things they are already doing. In other cases, correcting these misconceptions should improve economics.

1.5. CONCLUSIONS

Preferences lie at the core of mainstream economics because of their connection to welfare and to choice. On the assumption that people are rational, self-interested, and knowledgeable, their preferences will both explain their

³ But notice that (1) and (2) are inconsistent. What is best for me is subject to rational discussion. Later I shall discuss explicit formulations of (2), (3), and (5). As an example of (1) (the arbitrariness view), Mas-Colell et al. (1995) write in their authoritative graduate text, "The first [approach] treats the decision maker's tastes, as summarized in her *preference relation*, as the primitive characteristic of the individual" (p. 5, italics in the original). Lionel Robbins expresses the division-of-labor thesis (4): "Why the human animal attaches particular values in this sense to particular things, is a question which we do not discuss. That is quite properly a question for psychologists or perhaps even physiologists" (Robbins 1984, p. 86). More recently, Becker and Stigler phrase the division-of-labor view rather contemptuously: "On the traditional view, an explanation of economic phenomena that reaches a difference in tastes between people or times is the terminus of the argument: the problem is abandoned at this point to whoever studies and explains tastes (psychologists? anthropologists? phrenologists? sociobiologists?)" (Becker and Stigler 1977, p. 76).

choices and reflect what will benefit them. Understanding preferences better helps one understand the achievements and the challenges of economics.

A better understanding of preferences should be of use to other social scientists, too, many of whom face a choice between modeling human action as governed by preferences and beliefs or as driven by diverse motivating factors. This book should also clarify some of the relations between abstract work in action theory and empirical theorizing about human action. This inquiry concerning preferences aims to illuminate one path through the tangled terrain of human action, emotion, desire, belief, and well-being.