

PHILOSOPHY OF ECONOMICS

A VALUE-LADEN SCIENCE?

PREFERENCES Rational choice theory is *rooted in folk psychology* (Reiss, 2013, 29), insofar as the preferences and expectations invoked to predict and explain choices, are kinds of desires and beliefs. Folk psychology explains actions, that is, behaviour caused by the agent (whether intentionally or unintentionally), in terms of reasons for action, that is, beliefs and desires. In particular, in philosophy of mind there is the widespread (Davidsonian) view that beliefs and desires explain the action because they cause it. Analogously, rational choice theory explains choices between possible courses of action in terms of preferences towards the outcomes of the actions and (if the outcomes are probabilistic, given the actions) expectations towards such outcomes.¹ Preferences and expectations are “rational” causes of the action, in the sense that if one wants to infer the reasons for behaviour from (observed) behaviour, and thus explain behaviour, one must make assumptions on the agent’s system of beliefs and desires, which amount to “formal” requirements of rationality, such that, the explanandum behaviour is not random (Reiss, 2013, 31). In particular, preference rankings are assumed to be complete, transitive, and continuous.

One may distinguish between preferences as *partial*, *total*, and *overall* evaluative rankings (cf. WEEK 3). The first are relative to one respect. The second are not so relative, and factor in all considerations that bear on choice. The third leave out considerations, such as norms and rules, which may or may not conflict with preferences in determining choices.² Economists are concerned with preferences in the second sense. However, the very possibility that preferences in the second

and third sense may not coincide makes an exclusive focus on preferences in the second sense (as, e.g., in revealed preference theory) debatable (see again WEEK 3). Certainly, economists must estimate preferences from data accessible to them, namely observed choices. This entails that they may end up using revealed preferences in explaining economic phenomena. At the same time, “true” preferences are neither reducible to choices nor useless in the explanation of behaviour. In the same way that beliefs and desires are essential in the explanation of behaviour (cf. objections to behaviourism in philosophy of mind), so are preferences and expectations essential in the explanation of choice.

For one thing, if one reduced preferences to choices, one could not make sense of preferences over states of affairs, which one will never be in the position to choose from (34). Moreover, if one maintained that economics can ignore non-revealed preferences, one would exclude *a priori* that economic phenomena exist (e.g., involuntary unemployment, asset bubbles) that require explaining in terms of “counter-preferential choices”, namely choices, which cannot be explained by revealed preferences and expectations, but are partly due to inattentiveness, weakness of will or false beliefs (35). And even if one denied the existence of counter-preferential choices, one would still need to allow a mismatch between preferences and choices for another reason. Economist must presuppose the *stability* of preferences in order to use choices to estimate preferences, and thus to provide a rational explanation; but many patterns of choice—at least *prima facie*—appear to violate the axioms of the the-

¹The key difference is that preferences, contrary to desires, necessarily involve *comparative* evaluations.

²Notice that, when Reiss writes that “People can also rank alternatives overall or “all things considered””, he is using the “or” as introducing an exclusive disjunction (between total and overall rankings) rather than an explication (of “overall” as “all things considered”). This is consistent with the distinction in (Hausman, 2012).

ory, most notably transitivity; therefore, Reiss argues, one must allow that agents make mistaken choices given their rational preferences (*ibid.*). Either way, revealed preferences aren't enough.

For another thing, in situations of uncertainty over the outcomes of choice, choice cannot be explained without reference to expectations; if one accepts unobservable mental entities such as expectations, there's no point—so Reiss argues—in denying the existence or the importance of non-revealed preferences, on the ground that they are unobservable mental entities (35-6).

PLAUSIBILITY OF THE AXIOMS The truth of rational choice theory hinges on the plausibility of its axioms, in particular transitivity and completeness. Reiss considers two possible defences: (1) arguing that the axioms are normatively accurate, such that preferences *ought to* satisfy them if they are to count as rational (whether or not they do in fact satisfy them); and (2) arguing that the axioms are descriptively accurate, and thus useful in predicting and explaining *actual* choices (38). He then discusses the two possibilities in turn, in relation to the validity of ordinal choice theory, which describes decision making under certainty.

Let us consider option (1) first. Transitivity is typically defended by invoking the “money-pump” argument. Intransitive preferences may be exploited, and thus it is not rational to entertain them. If one prefers x to y , y to z and z to x , one would—if the opportunity arises—pay to trade x for y , y for z , and z for x , incurring a loss. This argument, Reiss notes, is both too weak and too strong. On the one hand, one may have intransitive preferences and yet avoid being money-pumped. For instance, one may not act on the preferences, as it happens if one is not trading, or one may entertain transitive preferences only during trades, but revert to intransitive preferences otherwise. On the other hand, the argument is too strong. One may have different rankings at different times, and thus be dynamically inconsis-

tent and yet be rational at all times. In that case, one may incur money-pumping without violating the axioms.

Additionally, there may be good reasons for having intransitive preferences in certain contexts. One may be offered to choose between an orange and a small apple, and between a large apple and an orange, and prefer the former to the latter in both cases; however, if one were offered to choose between a small and a large apple, one may—rationally—choose the small apple, if one is guided by a norm of good social behaviour.³

Differently from transitivity, completeness is more difficult to test. It is hard to distinguish between a lack of preference, or incompleteness (no rank) and indifference (equal rank) (cf. WEEK 3). Yet, the distinction is important, for Reiss: indifference (contrary to incompleteness) is subject to the so-called small-improvement argument. If one is indifferent between two options (having to pay \$1m and letting someone die), one of which is morally debatable (e.g., letting someone die), a small incentive (having to pay slightly less than \$1m) could suffice to tip the balance in favour of the morally debatable choice. But intuitively, the morally debatable option should not be subject to this argument—matters of life and death are *not commensurable* with pecuniary matters. One may resist the conclusion of the argument by allowing for incompleteness.

Perhaps to the extent that preferences are used for explanations, this lack of justification does not matter too much. In a decision situation one is often forced to choose among alternatives, even in the absence of good reasons to go one way or the other. Perhaps economists are mainly interested in giving accounts for such situations. But, still, that is not a justification for the completeness axiom as an axiom of rationality. (40)

Turning to option (2), on the one hand “preference reversal” (which we'll discuss more at length in WEEK 6), as observed in psychological experiments, seems to constitute evidence of violations of transitivity. Economists seem to take

³Here, however, it seems legitimate to wonder whether this is evidence of intransitive preferences, rather than of a conflict between transitive preferences and rules/norms, which do not influence the preference ordering.

the experiments seriously enough as to modify their theory. (Regret theory was proposed as an alternative.) On the other hand, it is observed that sometimes agents choose between two alternatives probabilistically rather than deterministically. This does not conclusively establish that completeness is violated—observed choices are equally compatible with deterministic preferences switching back and forth (no violation) and with stochastic preferences (violation). Yet, the latter seems a more plausible explanation of what is going on than the former (42).

In sum, the axioms may not be accurate, whether in a normative or in a descriptive sense.

THE ECONOMIST’S DILEMMA Reiss argues that economists face a dilemma between regarding rationality as a formal constraint on choice, as defined by the axioms of the theory, and regarding it as a substantive constraint. He introduces the dilemma against the backdrop of alleged violations of the axiom of Strong Independence in the context of expected utility theory.

Expected utility theory describes decision-making under risk or uncertainty. It views the alternatives between which the agent chooses as prospects, or lotteries, whose expected utility is calculated by weighing the utility of the possible outcomes of the action (or states of the world) in each prospect by their probability of occurrence (given the choice of a particular prospect). In order for the preferences to be representable by an expected utility function, the theory imposes the axiom of Strong Independence:

Strong Independence. If $y = (x_i, x_j; p, 1 - p)$ and $x_i \sim y_i$, then $y \sim (y_i, x_j; p, 1 - p)$.

In words, any component of a prospect can be replaced by another prospect to which the agent is indifferent, and the agent will be indifferent between the original prospect the new prospect. Relative to an expected utility function, one may define different attitudes towards risk—neutrality, seeking and avoidance—depending on

whether playing the lottery is as desirable as, or more or less desirable than, receiving the expected value of the lottery for sure.

Expected utility theory, for Reiss, is *more explanatory* than ordinal choice theory. The latter explains the choice of a bundle of goods x in the presence of another bundle y simply by mentioning some basic and given preference for x over y . The former, by contrast, can offer a “more nuanced story”, if interpreted as saying that “it constructs preferences over prospects from preferences over outcomes, given a risk attitude” (45). Many economists would say that preferences over prospects, too, are given. For Reiss, by contrast, it is better to interpret the theory as taking preferences over outcomes as given and those over prospects as derived from them. This would do justice to the idea that “people will have more stable and basic preferences over things they ultimately care about” (46), and form preferences between prospects based on what they care about most, namely the outcomes. As a result, the theory could explain the choice of a prospect x in the presence of a prospect y in terms of more fundamental economic entities, namely the preference for the outcomes in x and y , weighed by their expectation and one’s attitude towards risk.

Now, the explanatory power of the theory depends on the validity of its axioms. Allegedly, experimental evidence would show that Strong Independence is sometimes violated. I describe below a counterexample in a situation of decision under risk, namely the Allais paradox (see table).⁴

Probl	Choice	Payoffs		
		S1 (1%)	S2(89%)	S3 (10%)
I	A1	\$1,000	\$1,000	\$1,000
	A2	\$0	\$1,000	\$5,000
II	A3	\$1,000	\$0	\$1,000
	A4	\$0	\$0	\$5,000

In the experiment, agents are asked whether they would choose A1 or A2 in Problem I, and A3 or A4 in Problem II. No assumption is made as

⁴A similar point could be made with reference to the Ellsberg paradox, which also involves uncertainty (48).

regards the desirability of the payoffs in either problem: payoffs in A1 (respectively, A3) may be more or less desirable than the payoffs in A2 (resp. A4). Whatever the preference towards the payoffs, the axiom of Strong Independence entails that equal outcomes, namely \$1000 with a 89% chance in Problem I and \$0 with an 89% chance in Problem II, should not matter to the choice in either problem (43). Once the “irrelevant” alternative is removed, the remaining prospects are pairwise identical: A1 is the same as A3, and A2 is the same as A4. Therefore, for choices to reflect preferences—as described by expected utility theory—agents should choose A1 (resp. A2) in Problem I if and only if they choose A3 (resp. A4) in Problem II.

However, it is observed that agents choose A1 over A2 in Problem I and A4 over A3 in Problem II. The observation of these choices cannot be taken as evidence that the agents are maximizing utility—because there is *no utility function*, which is consistent with the preference of A1 over A2 *and* of A4 over A3.⁵ How should the result of the experiment be interpreted? One’s reaction will depend on whether one takes expected utility theory to be normative or descriptive as a theory of decision-making under risk.

If one takes the theory to be *normative*, one can maintain that the agents’ actual choices are irrational because they violate the axiom of independence. The payoffs of the irrelevant alternative, which are the only difference between the two problems, do make a difference—although they shouldn’t. So, no revision of the theory is called for. Hausman (1992), for instance, takes the normative force of expected utility theory (irrational choices are costly, so agents learn to “correct” deviations from the theory) as also having a bearing on the explanatory problem (there may be no better theory because the deviation from

the theory is unstable, given the agents’ tendency to correct themselves) (218-19). However, for Reiss, in the Allais paradox the choice may be rational also upon reflection, and thus stable:

If I were to end up in state S3 after choosing A2, I will regret my choice a great deal. I could have had a good amount of money for sure. I chose to gamble and lost. That was silly. In the choice between A3 and A4, the odds that I end up with nothing are overwhelming anyway. I’d consider myself lucky if I did win but not winning wasn’t silly. Quite the contrary, it would have been unreasonable to forfeit a good chance of a considerable higher gain for a minimally smaller chance of losing. I would not regret my choice. (47-8)

Alternatively, if one takes the theory to be *descriptive*, one must put the blame somewhere on the set of the premisses on which the derivation depends, which might lead to revising the theory.⁶ Here, however, the problem arises that the evidence underdetermines which of the premisses is to blame.

Any apparent violation of an axiom of the theory can always be interpreted as any of three things:

- the subjects’ preferences genuinely violate the axioms of the theory;
- the subjects’ preferences have changed during the course of the experiment;
- the experimenter has overlooked a relevant feature of the context that affects the subjects’ preferences. (49)

Reiss argues that only if, in addition to the axioms, further assumptions are made on the *stability* of the preferences during the time observations are collected, and on the *invariance* of the preferences to irrelevant changes in the decision-making context, may one infer that there is a violation in the axioms. From this observation, he infers that the economist faces a dilemma:

He can either stick with the “formal axioms” of completeness, transitivity, Strong Independence and so on and refuse to assume the principles of stability and invariance. But then rational-choice theory will be useless for all explanatory and predictive purposes because

⁵Notice, in this regard, that risk-aversion, being consistent with expected utility theory (47), does not qualify as a possible explanation of the choices in the Allais paradox.

⁶As the quote illustrates, regret considerations seem to influence the decision. Regret theory modifies expected utility theory by giving up independence. Another option is to give up completeness (as in Levi’s theory; cf. Hausman 1992, 219-21).

people could have fully rational preferences that constantly change or are immensely context-dependent. Alternatively he can assume stability and invariance but only at the expense of making rational-choice theory a substantive theory, a theory laden not just with values but with the economist's values. The economist then has to decide whether [...] framing a problem one way or another may reasonably affect someone's preferences; what relevant alternatives are; whether, to what extent and what social norms may matter; [...] and so on. (52-3)

The dilemma is between viewing rationality as (i) a normative and sufficient determinant of choice, as described by the axioms—which would be useless for prediction or explanation, given the possibilities of taste changes and context sensitivity; or (ii) as a descriptive constraint, which in conjunction with the assumptions of stability and invariance does predict and explain choices, but in a way that is dependent on the “values” of the scientists, namely on what they—rather than the

agents themselves—view as rational.

Max Weber, who believed in a clear distinction between facts and values, thought that social sciences can be objective even though, being sciences of human behaviour, they are concerned with values. The idea is that the agents' values may be taken as given, and their consequences may be studied by scientists in a descriptive way. Horn (ii) of Reiss' dilemma casts doubt on the realizability of Weber's ideal of objectivity and thus on the possibility of economics as a positive science.

References

- Hausman, D. (1992). *The Inexact and Separate Science of Economics*. Cambridge University Press.
- Hausman, D. M. (2012). *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.
- Reiss, J. (2013). *Philosophy of Economics. A Contemporary Introduction*. New York: Routledge.