



DISCUSSION ARTICLE

Subjective total comparative evaluations

Daniel M. Hausman 

Center for Population-Level Bioethics, Rutgers University, Room 400, 112 Paterson St., New Brunswick, NJ 08901, USA

Email: dhausman@cplb.rutgers.edu

(Received 02 March 2022; revised 21 September 2022; accepted 04 October 2022; first published online 05 May 2023)

Abstract

In *Preference, Value, Choice, and Welfare*, I argued, among other things, that preferences in economics are and ought to be total subjective comparative evaluations, that the theory of rational choice is a reformulation of everyday folk-psychological explanations and predictions of behaviour, and that revealed preference theory is completely untenable. All three of these theses have been challenged in essays by Erik Angner (2018), Francesco Guala (2019) and Johanna Thoma (2021a, 2021b). This essay responds to these criticisms and defends these three theses.

Keywords: Preference; belief; choice; revealed preference theory; total subjective comparative evaluations

In *Preference, Values, Choice, and Welfare*, I argued that implicit in the formal conditions economists impose on preferences and in the use of preferences to explain and predict economic behaviour and outcomes is a concept of preferences as subjective total comparative evaluations. In this essay, I explain why this thesis concerning the concept of preferences is important and comment on apparently serious criticisms from Erik Angner (2018) and Johanna Thoma (2021a, 2021b).

In arguing that economists take preferences to be subjective total comparative evaluations, I am not criticizing economists. Although explaining and predicting choices in terms of preferences, beliefs and constraints has limitations, it is a coherent and sensible way to proceed that serves many of the purposes of economists. I am critical of what many economists say about preferences, which conflicts with their practice, even though, fortunately, it has little effect on how they do economics.

Section 1 lays out briefly the concept of preferences I attribute to economists, the evidence in support of this attribution, and my critique of competing accounts of the nature of preferences. Section 2 presents and responds to Erik Angner's recent criticisms of my claims concerning preferences. Section 3 responds to arguments

of Johanna Thoma defending revealed preference theory and denying that decision theory employs a mentalistic conception of preference. It also accepts an important criticism from Francesco Guala. Section 4 explains why it is important to clarify the concept of preferences rather than leaving preferences to be defined implicitly by formal conditions and the explanatory and predictive practices of mainstream economics (which is the only variety of economics discussed here).

1. Total Subjective Comparative Evaluations

The axioms of ordinal utility theory are the core of positive economic theory, and they also constitute a fragmentary theory of rationality. Economists sometimes place additional constraints on preferences, but the axioms of ordinal utility theory are central. Although there are differences in details, formulations of ordinal utility theory rely on the following five axioms concerning preferences. These axioms define a model of rational choice that serves, to some degree of approximation, as a model of actual choice. The axioms are more compactly stated employing the notion of ‘weak preference’ where P weakly prefers x to y if and only if P prefers x to y or P is indifferent.

1. (*Completeness*) The preferences of agents are complete: For any two alternatives x and y , an agent weakly prefers x to y or y to x .
2. (*Transitivity*) The preferences of agents are transitive: For any three alternatives x , y and z , if the agent weakly prefers x to y and y to z , then the agent weakly prefers x to z .
3. (*Context independence*) The preferences of agents do not change from context to context, and whether an agent prefers x to y does not depend on what other alternatives are available.
4. (*Stability*) The preferences of agents remain largely unchanged during the period concerning which economists attempt to explain or predict choices.
5. (*Choice determination*) The preferences of agents determine their choices: If an agent knows which alternatives are feasible – that is, which alternatives can be chosen – and an agent chooses an alternative x , then there is no feasible alternative that the agent prefers to x , and if the agent knows of some feasible alternative x that he or she prefers to all others, then the agent chooses x .

These axioms define a model of choice. They do not define preferences, which can, of course, fail to satisfy any of the axioms. Incomplete preferences are still preferences. The axioms do however constrain the concept of preferences: preferences must be the sort of thing of which these axioms could be true.

Typically, only the first two of these five claims about preferences are presented as axioms, but all these claims are implicit in the use of ordinal utility theory. These axioms are far from obvious truths. As economists readily concede, there are counterexamples galore: people cannot compare all the myriad alternatives, people’s rankings of large sets of alternatives are rarely perfectly transitive, context often influences rankings, in situations of flux preferences may change rapidly, and people may choose an alternative that is ‘good enough’ even if

better choices are available. Yet, apart from context independence, these axioms might appear to be reasonable exaggerations in many practical applications of economic theory. For example, individuals buying cars may have some familiarity with all the readily available alternatives and, among the short list they do not rule out (on the basis of price, qualms about reliability and so forth), the individual's preferences may be complete and transitive. There may be only one relevant context and hence no problem with context independence. During the period of deliberation, the relevant preference ranking may be stable and, subject to availability and cost, the choice is determined by the agent's preferences. Economics is an *inexact science*. Its fundamental generalizations are not universal truths, but should be understood as statements of tendencies or as in some sense approximations (Hausman 1992).

At first glance, it seems that context independence rules out such ordinary preferences as preferring bread to cereal for breakfast when butter is available, but cereal to bread when butter is not available. But, if one defines the objects of choice so that, for example, bread when butter is available is a different alternative than bread when butter is not available, context independence may also appear to be a reasonable exaggeration of the modest stability of preferences across contexts. Section 3 returns to difficulties concerning context independence.

The ordinal representation theorem proves that when people's preferences satisfy the first two axioms and some further technical conditions, then they can be represented by a continuous utility function that is unique up to a positive order-preserving transformation (Debreu 1959). The further technical axioms needed to prove this theorem are not trivial, but they add nothing to the interpretation of preference, and I shall not discuss them.

Axiom 5 is often stated as the claim that agents maximize utility. Stating it this way makes it sound as if utility is some ultimately desirable thing that agents seek – some special good whose pursuit motivates all action. But utility is not an object of desire. It is not something people seek or accumulate. All that axiom 5 says is that choices are determined by preferences, beliefs, and, implicitly, constraints – that agents always go as far up their preference ranking as they think they can. Since ordinal utility theory is typically applied to circumstances in which it is assumed that agents have perfect knowledge of the alternatives among which they choose, axiom 5 is often stated without any explicit reference to the agent's knowledge about the availability of alternatives. But it is obvious that agents may choose inferior alternatives when they are unaware that better ones are available.

Axiom 5 has two important implications concerning the interpretation of preferences. First, it implies that, given beliefs, preferences motivate action. If Jill knows that x is feasible, and she prefers x to all other feasible alternatives, then Jill does something: She chooses x . Preferences determine choices only within reality *as seen by the agent*. With different beliefs, the same preferences lead to different choices. Second, since preferences *determine* choices, they must take into account *everything* relevant to choice apart from beliefs and constraints. If, as in the everyday usage of the word, 'preference', as 'overall evaluations', Jill's preferences leave out something that influences her choices, such as moral commitments, then her preferences will not determine her choices. Axiom 5

rules out this possibility. Axiom 5 demands that preferences express total rather than partial evaluations and are intrinsically motivating.

I maintain that the first three axioms have three other important implications concerning preferences. First, they imply that preferences, unlike wants or wishes, are comparative. Unlike wants, preferences rank alternatives.

Second, completeness and transitivity, coupled with the implication of axiom 5, that preferences must rank alternatives in terms of everything the agent regards as relevant to choice, apparently place enormous cognitive demands on agents. Though it would be possible to possess a complete, transitive and context-independent ranking of all alternatives with respect to everything relevant to choice merely by virtue of having a remarkably finely tuned gut, the only plausible explanation of the possession of complete, transitive and context-independent preferences or some reasonable approximation thereto is that people have somehow or other (off-stage as it were) successfully adjudicated among myriad competing considerations. A complete and transitive preference ranking of the complex alternatives people face would be a remarkable intellectual achievement. This observation provides a strong reason to regard the axioms as idealizations and to recognize people's actual preference rankings will *not* satisfy the axioms, rather than to conclude that people have superhuman evaluative capacities.

Third, I argue that completeness, transitivity and context independence imply that the 'alternatives' among which people have preferences cannot be ordinary options such as eating one or two scoops of ice cream. Whether Jack prefers two scoops of ice cream to one depends on whether he's eaten a large dinner before, whether his stomach is upset, whether he is currently on a diet, whether the two scoops are likely to drip on his elegant suit and so forth. Thus, Jack cannot be said to have a preference for one scoop or two. To have preferences that satisfy the axioms, the alternatives over which preferences are defined must be complete states of the world such as, for example, two states of the world at a particular moment in time that differ only with respect to whether Jack eats one or two scoops of ice cream. Clearly, this is an idealization, not a surprising conclusion concerning how fine-grained the objects of preferences are.

To have complete and transitive preferences over such alternatives requires fantastically complete knowledge, and when choices involve uncertainty, as they almost always do, then specifying the alternatives among which people have preferences is still more complicated. I shall return to these issues when I discuss the role of beliefs in the formation of preferences and in the relationship between preferences and choices.

Despite the demands of completeness, transitivity, context independence, stability and choice determination, the axioms of ordinal utility theory are usually regarded as very weak. They say nothing substantive about what sorts of things people prefer, and they could as easily be satisfied by people who consistently seek pain and suffering for themselves as by those who consistently pursue their own interests. Positive economic theory supplements the axioms of ordinal utility theory with axioms concerning the content of preferences, such as the claim that people prefer more commodities to fewer.

The axioms governing preference imply that preferences are action-guiding rankings of complete states of the world in terms of everything relevant to choice, the possession of which may be extremely demanding cognitively. Although the axioms may appear to leave open several competing interpretations of preferences, Hausman argues that there is only one plausible interpretation, which economists favour, despite what some economists sometimes say.

Despite his philosophical sophistication, Amartya Sen mistakenly maintains that economists employ mainly two concepts of preferences. According to Sen, '[T]he normal use of the word permits the identification of preference with the concept of being better off, and at the same time it is not quite unnatural to define "preferred" as "chosen"' (Sen 1977: 329). But 'A prefers x to y ', neither entails nor is entailed by '[A believes that] x is better for A than y '. There is no contradiction in Abigail preferring that the Republicans regain a majority in the Senate, while being rightly convinced that who controls the Senate in 2023 will have no effect on her interests or that a democratic Senate would be better for her.

The other definition of preferences that Sen attributes to economists, revealed preference theory, links preferences tightly to choices – too tightly, I maintain. To maintain that agents prefer x to y if and only if agents never choose y from any set including x limits preferences to those alternatives among which individuals can choose, which makes it impossible to define games, which require assigning utilities to outcomes among which agents cannot choose. In a two-person game, suppose that I choose first and that my choice determines the outcomes among which you can choose. Which outcome obtains among those which my choice did not rule out is up to you. What is best for me depends on my preferences over the outcomes that you choose. Revealed preference theory implies that I have no preferences among those outcomes, because I have no choice to make among those outcomes. Moreover, my choice of strategy depends on what I think you would choose among the outcomes that my choice would make available to you, but you cannot reveal a preference among outcomes that my choice makes unavailable to you. Moreover, if preferences are limited to the immediate objects of choice, then all it takes to explain or predict a choice of x is the proposition that the agent prefers x to the other available alternatives. In that case, explanations and predictions are all trivial. The answer to all questions of the form, 'Why did David choose x ?' would be 'David preferred x to the available alternatives'.

In addition, revealed preference theory mistakenly ignores the possibility that an agent may choose y from a set of alternatives including x , *despite preferring x* , because the agent falsely believes that x is not available. Without information about beliefs there is no way to infer preferences from choices or choices from preferences. Preferences, like beliefs, are subjective states that causally determine choices only in conjunction with belief.

2. Responses to Erik Angner's Criticisms

In this section, I focus on a criticism Erik Angner makes in his essay, 'What Preferences Really Are'. He argues that the central thesis of *Preference, Value,*

Choice and Welfare, that preferences in economics are total subjective comparative evaluations, is mistaken. Angner maintains that preferences are not ‘subjective judgments to the effect that something is better than something else all things told’ (2018: 660), and that economists should not employ this notion of preferences, or indeed any other definition of preferences.

Angner summarizes his criticism as follows:

The “only plausible way” for agents to come up with total comparative evaluations satisfying such axioms, in Hausman’s view, is for the agent “to adjudicate between competing considerations” – including partial comparative evaluations (2012, 17–18). A preference ordering satisfying the axioms therefore “must be the outcome of an un-modeled process of exhaustive comparative evaluation” (18). In this way economists are – and also should be, if they want preferences to be transitive and complete – committed to a particular theory of preference formation and change, in which preferences result from a series of cognitive operations on, among other things, the relevant partial evaluations. (Angner 2018: 117)

The quotation begins by stating my view that if individuals were able to have complete and transitive preferences (which they are not), they would need to carry out complex cognitive evaluations of the merits and drawbacks of alternatives. How else could individuals possibly generate a transitive ranking of the immense number of alternatives among which they can choose? But this is not a story that economists need to tell. They do not model processes that would lead people to possess complete and transitive preferences, and I do not maintain that they should.

If the task is to explain or predict choices among alternatives, then it is often convenient to use a stripped-down model whereby choice depends on constraints, beliefs and preferences among the available alternatives and their consequences. Whether such a model is useful is an empirical question, whose answer depends on the success of the predictions and explanations economists offer. There is obviously a great deal that can be said about the intricate processes whereby emotions, drives, antipathies, habits, fears and so forth enable some sort of ranking of the desirability of alternatives that can be depicted (with the help of considerable idealization) as a preference ordering. But this psychological inquiry is often irrelevant to the questions that economists are concerned with, and economists employ preferences among the immediate objects of choice and their consequences in their explanations and predictions without any further concern with human psychology. This abstraction from psychological detail is not always successful, and behavioural economists have had more to say about the determinants of preferences.

Notice that the simple story whereby choice is determined by constraints, beliefs, and preferences depends on preferences among the consequences of choices and beliefs about their probabilities, not just on preferences among the immediate objects of choice. Otherwise, there would be no interesting structure to the theory of choice. All there would be to say in explaining or predicting choice is that individuals choose whatever they most prefer from among those alternatives

they believe to be available. To avoid this trivialization, economists endorse (often implicitly) accounts of how preferences among alternative actions depend on preferences over their possible consequences and beliefs concerning how probable those consequences are.¹ This additional structure within the theory of choice expands the range of items over which individuals have preferences, but the factors determining choices remain preferences, beliefs, and constraints. Beyond linking preferences among actions to beliefs and preferences concerning their consequences, the typical mainstream economist responds to questions about the formation of preferences by in effect saying, 'That's not my department'.

Angner does not interpret *Preference, Value, Choice, and Welfare* this way. He ends the quotation above by attributing to me the view that 'economists are – and also should be, if they want preferences to be transitive and complete – committed to a particular theory of preference formation and change, in which preferences result from a series of cognitive operations on, among other things, the relevant partial evaluations'. In order to possess a complete and transitive preference ordering, individuals would have to carry out an unbelievably complex weighing and combining of the various specific considerations bearing on the values of alternatives. Angner concludes, mistakenly, although not unreasonably, given some of what I say at the beginning of Chapter 10, that I believe that economists are and ought to be committed to a theory of how people are able carry out this extremely complicated adjudication. But my view, on the contrary, is that there is no such theory, since people are *not* able to form such preferences. Attributing complete and transitive preferences to individuals is an idealization that is a reasonable approximation in circumstances where the alternatives and context are limited. Economists have no general theory of how individuals accomplish the impossible task of generating preference rankings among all possible alternatives that are genuinely complete and transitive. Apart from connecting preferences among alternative actions to preferences among their consequences and beliefs about the probabilities of those consequences, economists take preferences as fixed starting points, whose origins and modification are not to be investigated further.

In my view, this is too narrow. While quite sensibly avoiding providing any general theory of preference formation and modification, economists could say more about preference formation than that preferences among actions depend on preferences among their consequences and beliefs about those consequences and about their probabilities. I argued that

In some cases, what is of interest is the complicated interplay between desires, intention, self-command, and so forth, and [economic modeling of] the determination of actions by the beliefs and preferences of agents is of little

¹As a referee pointed out, standard consumer choice theory offers no account of the dependence of preferences among commodity bundles on preferences among the consequences of those bundles. Yet it escapes trivialization. It does so by explaining and predicting choices by prices and incomes as well as preferences. The immediate objects of choice are purchases: exchanges of money for commodities, not simply commodities. Preferences among purchases depend among other things on preferences on their consequences for the ability to make other purchases.

interest. In such cases, the standard model of choice is uninformative. In other cases, what is of interest will be the relations between preferences and beliefs and the actions that result from them, although something will still need to be said about what determines final preferences [among actions]. (Hausman 2021: 73)

Angner regards the thesis that preferences are total subjective comparative evaluations as entailing that economists model individuals as engaged in an immensely cognitively demanding process of combining and weighing various partial evaluations to generate a total evaluation. Although some of my wording lends support to this interpretation, consider specifically what it means to say that preferences are total subjective comparative evaluations. To say that they are evaluations is to say only that they are rankings that influence choices. It says nothing about the source, stability, context dependence, complexity or rationality of these rankings. To say that these rankings are subjective is to say that they rank alternatives *as they are individuated and understood by individuals* and hence that they are sensitive to beliefs about the properties, probabilities and consequences of alternatives. To say that they are comparative merely highlights the understanding of evaluations as rankings. Finally, to maintain that preferences are total evaluations is to maintain that all factors that influence choices other than beliefs and constraints do so via influencing preferences. That means that once economists know what the constraints are and what agents believe and prefer, they know what individuals will do (although in the case of indifference, there may be more than one possibility). This implies that a total evaluation depends on ‘all the considerations that [agents] find relevant’ (4).

Total rankings may be cognitively demanding, and they leave hostages to rational criticism. However, note that I deny that total rankings need to be well-informed or carefully considered:

A total ranking contrasts with a partial ranking, not with an ill-considered ranking. Examples of partial rankings would be Jill’s rankings of alternatives in terms of her own enjoyment or the expected benefits they will provide to her uncle. These rankings will not coincide with her total ranking unless all that mattered to Jill were her own enjoyment or the expected benefit to her uncle. Her total subjective ranking might be carefully thought out, or parts might rest on mere whim; and a ranking might collapse upon more careful reflection. Jill might be ignorant of important considerations or careless in canvassing relevant factors. Her total comparative evaluation need not be stable or rationally defensible, although it may be. (Hausman 2021: 34–35)

It seems to me that Angner has no quarrel with the claims that preferences are subjective, that they are evaluations, that they are comparative and that they are total. Conversely, I have no quarrel with Angner’s view that economists do not attempt to model preferences as determined by a complex adjudication of all the factors that are relevant to how individuals rank alternatives.

Angner makes five specific charges against my claims:

- (1) The structure of standard economic theory gives us little reason to think preferences are total comparative evaluations or that they are constructed from possibly myriad partial comparative evaluations.
- (2) The economists responsible for the postwar neoclassical synthesis treated preferences as primitives, explicitly and adamantly rejecting every effort to explain their origin and relationship to subjective phenomena.
- (3) Actual economic practice reveals little evidence of preferences being total comparative evaluations or being formed in a process of exhaustive comparative evaluation.
- (4) The notion that preferences are cognitively demanding total comparative evaluations has been fairly conclusively refuted by cognitive science buttressed by evolutionary considerations.
- (5) The models of preference formation and change that are favoured by practicing economists do not involve cognitive operations on a plethora of partial evaluations.

With the exception of the second, these criticisms all rely on Angner's reading of my book as defending an exaggeratedly cognitive view of preferences, and they require no further discussion here. With respect to the second criticism, what Angner claims is true of important post-war economists who were particularly concerned to minimize the role of subjective considerations in economics. However, I do not accept the authority of economists such as Paul Samuelson (1947), and would argue that they were mistaken to elide the differences between preferences and choices and to ignore the fact that their practice requires treating preferences as subjective.

3. Johanna Thoma's Critique of Mentalistic Theories of Preferences

Johanna Thoma argues that decision theory is not a regimentation of folk psychology, because decision theory does not and cannot employ what Hausman calls a subjective notion of preferences and Thoma calls a mentalist construal of preferences. Moreover, she maintains that, once it sheds its behaviourist raiment, revealed preference theory not only escapes the criticisms that I and others have made of it, but reveals itself as scientifically superior to views that regard choice as determined by belief and mentalist preference.

The reason why Thoma believes that decision theory is not an implementation of folk psychology is that the objects of preference and desire in folk psychology are coarsely individuated, unlike the objects of preference in decision theory. People prefer coffee to tea, hot mushroom pizza to cold sausage pizza, and so forth. But preferences among coarsely individuated alternatives are not context independent. I may prefer coffee to tea in the morning, but tea to coffee in the evening, except if I need to stay awake until late. Perhaps there is no context in which I would prefer cold sausage pizza to hot mushroom pizza, but the point is clear enough. Context-independent preferences must be individuated extremely

finely. Rather than coffee or tea, the alternatives must, as it were, build in all the features of context that affect my preferences between a cup of coffee now or a cup of tea. The objects of context-independent preferences will be the complete state of the world at time t with coffee versus the complete state of the world at time t with tea.

The problem is, Thoma argues, that we don't have mentalist preferences among such finely individuated alternatives. Since preferences in decision theory are context independent, their objects must be finely individuated and hence different from the objects of mentalist preferences. So preferences in decision theory are not mentalist, and decision theory, unlike folk psychology, does not explain choices by beliefs and mentalist preferences.

Thoma has more to say about the advantages of identifying preferences in decision theory with actual or possible choices, but I have said enough about her argument against interpreting decision theory as a formalization of folk psychology to permit a rebuttal. Thoma is correct that the objects of preferences and desires in folk psychology are coarse-grained and hence not context independent. Treating preferences as if they were context independent is thus a huge idealization that easily leads to false conclusions, but which is harmless if the preferences under consideration are not sensitive to any differences in the contexts that are relevant to a particular choice. In Thoma's example, she drinks coffee one morning, because she prefers its taste to tea, but drinks tea the next morning, because 'because she wanted to keep her nerves down for her important meeting, and believed the tea would keep her less nervous than the coffee' (Thoma 2021b: 914). If one wants to explain or predict her behaviour, one needs to refine the description of the alternatives to distinguish drinking coffee or tea when there is no issue of nerves from drinking coffee or tea when staying calm is important. The refined alternatives will still be coarse-grained and preferences among them are unlikely to be context independent, but their sensitivity to other aspects of context may not be germane to a particular application. There is no reason why one cannot then make use of decision theory, recognizing that the treatment of preferences as context independent is an idealization.

Moreover, the objects of behavioural preferences, that is, people's actual or hypothetical choices, are no more fine-grained than are the objects of mentalist preferences. The only choices Johanna makes are drinking coffee on one day and drinking tea on the next. Even if decision theorists are able to observe her beverage choices on many days, decision theorists are likely to be surprised when, for example, Johanna drinks tea on a stressless day rather than coffee because she has a sore throat.

Context-independence is an idealization. It is not true of mentalist preferences, and it is not true of behavioural preferences. It is an axiom of decision theory that enables inferences of choices across contexts from beliefs and preferences. The fact that mentalist preferences rank alternatives that are coarsely individuated is no argument for a behavioural interpretation of preferences, which are no more fine-grained.

Thoma also argues that preferences should be understood as either actual or hypothetical choices and that a version of revealed preference theory escapes my

criticism.² To defend revealed preference theory, she must respond to the difficulty that individuals may choose y , despite preferring x to y , if they falsely believe that x is unavailable. Francesco Guala provides a simple example: Tony and Vincent go to Pizza Bella Napoli. Vincent goes there because he prefers it to Pizza Vesuvio. Tony goes there despite preferring Pizza Vesuvio, because he falsely believes that Pizza Vesuvio is closed (Guala 2019: 385–386). One can infer choice from preference or preference from choice only if one knows what the individual believes (and that one can infer belief from choice or choice from belief only knowing what the individual prefers). So preference appears to be a subjective state that combines with belief to cause choice.

In response, Thoma argues, that this sensitivity to belief can be captured in the individuation of the objects of preference rather than in the agent's preferences. Thus, for example, the agent who chooses y from the set $\{x, y, z\}$ while preferring x to y (believing that x is unavailable) should be modelled as choosing from the set $\{y, z\}$, the set of alternatives the agents believes to be available, not as choosing from the actual set of available alternatives $\{x, y, z\}$. This proposal has the awkward consequence that one can no longer describe Tony and Vincent as making the same choice. Vincent chose 'eating at Pizza Bella Napoli when he could have eaten at Pizza Vesuvio' while Tony chose 'eating at Pizza Bella Napoli when he could not eat at Pizza Vesuvio'. The objects of their choices are no longer observable from their behaviour.

Insisting on this role for belief in determining the objects of preferences obviously introduces a mentalist element into choice theory and would not satisfy a behaviourist. But allowing this role for belief in the individuation of alternatives is, Thoma argues, consistent with a behavioural construal of preferences as actual or hypothetical choices among alternatives described in ways that are consistent with the relevant beliefs of agents and that encompass all of the aspects of alternatives that are relevant to the agent's choice. However, in order to determine what people believe, revealed preference theorists will be relying implicitly on mentalist preferences.³ Thoma can respond that economists do not need an account of belief attribution and hence need not worry about the fact that belief attribution relies on mentalist preference.

In addition to responding in this way to one of the main criticisms of revealed preference theory, Thoma argues in its defence that identifying preferences with choices enables economists to 'black-box' the factors that are responsible for preferences and that it streamlines the inference concerning future choices from observations of past choices. I deny the second of these reasons. Without knowing how preferences among alternative actions depend upon preferences among consequences of those actions *that are not objects of actual or hypothetical choices*, inferences concerning future choices will be stymied. For example, what explains and predicts how individuals vote are their preferences

²She explicitly leaves open whether to identify preferences with actual or hypothetical choices in her (2021a: footnote 2, p. 164).

³There is no way to draw any conclusions from people's actions concerning their beliefs, without premises concerning their preferences or desires. Even to draw the conclusion that Amy believes that P from Amy's saying, 'I believe that P ', I need to assume that Amy aims to tell the truth.

for who wins the election, which is not a matter of individual choice. Past voting among candidates P and Q is relevant to explaining and predicting current voting among candidates R and S only to the extent that past voting serves as evidence of preferences among the policies and personalities of the current candidates. Predictions and explanations of choices do not and should not rest exclusively on observations concerning past choices.

With respect to black-boxing or abstracting from the psychological complexities that influence choices, relying on preferences as total evaluations, which incorporate all the influences on choices apart from beliefs and constraints that the agent finds relevant, already abstracts from the multitude of psychological factors governing choices. However, a mentalist view of preferences does limit the attribution of preferences to those with the requisite mental capacities, and Francisco Guala has argued compellingly (2019), that preferences, as belief-dependent dispositions, may be attributed to non-human agents such as committees, other animals, or robots. When so attributed, preferences will have a different causal basis than is the case with people. So a mentalist view of preferences is only a view of human preferences. It is, in this way, less general than a view that identifies preferences with choices.

Granting Thoma all that can be said on behalf of this greater generality, her account is still in trouble. Here is her account of preferences:

when an economist ascribes a preference for an option *a* over an option *b* to an agent (be it a real or a model agent), this is meant to capture nothing more than that the agent does or would choose *a* over *b* from some specified choice set – for instance, that she does or would choose *a* over *b* when only those two options are available, or that she does or would not choose *b* in any choice situation where both are available. (Thoma 2021a: 164)

If preference is linked to actual choice (which seems not to be Thoma's view), then if the agent has never faced a choice between *a* and *b*, the first account denies that the agent prefers *a* to *b*, while the second account maintains that the agent prefers *a* to *b*. Both of these implications are unacceptable. From the fact that I have never made a choice between bran flakes or toasted whole wheat bread with orange marmalade, one cannot conclude anything about my preferences among these alternatives.

Suppose, instead, that one identifies preference with hypothetical choice. Hypothetical choice is not choice. It is a prediction about what would be chosen. To make a prediction that I would choose the toast and marmalade rather than the bran flakes, the revealed preference theorist must attribute to me now a conditional disposition to choose the toast, if I were to believe that I faced a choice among them. Although dispositions need not be mentalist, a disposition to choose, conditional on a hypothetical belief, is a mentalist preference.

4. Why Does All This Matter?

In *Preference, Value, Choice, and Welfare*, I do not take the claim that preferences are total rather than partial comparative evaluations to be an economic thesis

subject to empirical testing or refutation. The claim instead expresses a modelling choice by economists, which is, I maintain, a sensible choice for certain purposes, including many of the explanatory and predictive purposes of economists, but a very poor choice if one wants to understand the processes of deliberation. Taking preferences to be total rather than ‘overall’ evaluations (which compete with other choice-determining considerations) is at odds with how most non-economists think about the factors that determine choices. However, provided that economists make clear how they are revising the pre-theoretic notion of preferences, they are free to reconceive of preferences in any way they wish – provided, of course, that their stated reconception of what constitutes a preference ranking actually matches their practice.

Reconceptualizations are not beyond criticism. Some, such as revealed preference theory, are inconsistent with the practices of economists and would make economics trivial or even impossible. Others, such as Sen’s view that there are many conceptions of preferences, fail to regiment economic practice and undermine the value of the simplified structure of the economic theory of choice. As I argue in *Preference, Value, Choice, and Welfare*, there are interesting things to say about the determinants of preferences without collapsing the distinction between economics and psychology. As useful as it is to simplify the relationship between preference and choice, there is no reason why economists must also simplify further what is already a bare-bones account of what determines and changes preferences. I do not think that there will ever be a canonical account of preference formation, but there are many specific factors to examine, such as linking preferences to their properties in multi-attribute utility functions.

The claim that preferences are total subjective comparative evaluations is not a straitjacket that I would place on economic practice. It is instead a way for economists to understand better what they are already doing and, in particular, to recognize that to escape trivialization, they must say something about how preferences among alternative actions – that is, what I call ‘final preferences’ – are formed and modified. Linking final preferences to preferences among the consequences of actions and beliefs about those consequences is a beginning, but more can be said. Theories of preference construction attempt to say more, but they also destabilize the explanatory structure of the standard theory of choice that economists accept. A great deal remains to be done.

References

- Angner E. 2018. What preferences really are. *Philosophy of Science* 85, 660–681.
- Debreu G. 1959. *Theory of Value*. New York: John Wiley & Sons.
- Guala F. 2019. Preferences: neither behavioural nor mental. *Economics and Philosophy* 35, 383–401.
- Hausman D. 1992. *The Inexact and Separate Science of Economics*. New York, NY: Cambridge University Press.
- Hausman D. 2021. *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.
- Samuelson P. 1947. *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.

Sen A. 1977. Rational fools: A critique of the behavioural foundations of economics theory. *Philosophy & Public Policy* 6, 317–344.

Thoma J. 2021a. In defense of revealed preference theory. *Economics and Philosophy* 37, 163–187.

Thoma J. 2021b. Folk psychology and the interpretation of decision theory. *Ergo* 7, 904–936.

Daniel M. Hausman is a Research Professor at the Center for Population-Level Bioethics at Rutgers University and the Herbert A. Simon Emeritus Professor of Philosophy at the University of Wisconsin-Madison. One of the founding editors of this Journal, Professor Hausman has worked on metaphysical, epistemological, methodological and ethical issues at the boundaries of economics and philosophy. Forthcoming books are *How Health Care Can Be Cost-Effective and Fair* and a second edition of *The Inexact and Separate Science of Economics*. In 2009, Professor Hausman was elected to the American Academy of Arts and Sciences. URL: <https://hausman.philosophy.wisc.edu/>.