

Mauro Rossi

Two platitudes about interpersonal comparisons

Working paper

Original citation:

Rossi, Mauro (2006) Two platitudes about interpersonal comparisons. CPNSS working paper, vol. 2, no. 2. The Centre for Philosophy of Natural and Social Science (CPNSS), London, UK.

This version available at: <http://eprints.lse.ac.uk/23898/>

Originally available from [Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science](#)

Available in LSE Research Online: February 2010

© 2006 The author

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

Two Platitudes About Interpersonal Comparisons

Mauro Rossi

LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

1. Introduction

Theoretical reflection about interpersonal comparisons is centred around three questions¹:

- (1) What is the object to be interpersonally compared?
- (2) Are interpersonal comparisons empirically meaningful?
- (3) Are interpersonal comparisons possible?

The first question is conceptually prior. It is necessary to select an object to compare, before asking whether its comparison across different individuals is empirically meaningful or possible. In this paper, I shall take utility as the object of interest. More precisely, I shall define utility as the value of a function that represents the intensity of individual preferences. When speaking about interpersonal comparisons of utility (IUCs), therefore, I shall consider interpersonal comparisons of preference strength.

The second question is concerned with the issue as to whether or not IUCs are empirically meaningful. I shall adopt Christian List's definition, according to which IUCs are empirically meaningful if and only if they are determined by the empirical evidence². On the other hand, IUC are empirically meaningless if and only if they are either (i) underdetermined by the empirical evidence; or (ii) indeterminate³.

The question whether IUCs are empirically meaningful is different from the question whether IUCs have a factual basis. The reason why is that even if the empirical evidence is not sufficient to determine IUCs, and, thus, leave them underdetermined, it may still leave no doubts about the existence of a fact of the matter concerning the interpersonal comparability of preference strengths.

¹ A similar list of questions is proposed by ELSTER, J. and J, ROEMER [1991], p. 1.

² LIST, C. [2003], p. 229.

³ A theory, or a statement, T is determined by a set of observation sentences F if F implies T. A theory, or a statement, T is underdetermined by a set of observation sentences F if T is consistent with, but not determined by, F. A theory, or a statement, T is indeterminate if it is underdetermined by F and there is no independent fact of the matter. See LIST, C. [2003], p. 232.

Instead, this is denied by the indeterminacy thesis. Thus, I shall say that IUCs have a factual basis only if they are not indeterminate.

The third question is often confused with the previous one, but is different. IUCs may not have a factual basis, but they may well have another, non-factual, basis. Therefore, it can still be possible to make them⁴. In general, IUCs are possible if and only if there exists *a* basis to make comparative judgments. On the other hand, IUCs are impossible if and only if there is neither a factual nor a non-factual basis to make them.

The focus of this paper is on both the second and the third questions. I shall proceed as follows. In the next section, I shall explain why IUCs appear to be empirically meaningless. The issue is pressing because people seem to make IUCs in everyday life, without any particular difficulty. In section 3, I shall present an example of one such situation and examine two platitudinous observations that the example suggests: the first concerns the role of IUCs in decision-making; the second concerns their role in explaining people's behaviour. One may try to prove that IUCs are empirically meaningful by showing that the assumption of comparable preference strengths helps explain both platitudes. In section 4, I shall claim that this strategy fails. Section 5 discusses an alternative solution, based on a reading of Davidson's position. In section 6, I shall propose some alternative ways of how it is still possible to account for the platitudes about IUCs. Finally, I summarize my results and conclude in section 7.

2. The 'standard picture'

Suppose we want to reach a judgment based on utility considerations, where utility is a representation of preference strength. The 'standard picture' follows three sequential steps⁵, dealing with:

- (1) the determination of individual preferences and their measurement through a utility function
- (2) the interpersonal comparison of utilities
- (3) the judgment

The first step is governed by two sets of axioms. Choice axioms fix the conditions for inferring the existence of preference relations from observed choice behaviour. Preference axioms provide additional constraints: typically, they require that each individual has complete and transitive

⁴ See LIST, C. [2003].

⁵ DAVIDSON, D. [2004] and FLEURBAEY, M. and J., HAMMOND [2004] offer similar, although not identical, reconstructions.

preferences. These are necessary and sufficient conditions for having a preference ordering. The corresponding set of axioms is richer when the choice is made under conditions of uncertainty. Together with the ordering axioms, it typically includes an Archimedean and an independence axiom.

If individual preferences satisfy those axioms, then they can be represented through a utility function. When choices are performed under conditions of certainty, preferences can be represented through an ordinal utility function, that is, a utility function unique up to a monotonic transformation. When choices are made under conditions of uncertainty, instead, preferences can be represented through a cardinal utility function, which is unique up to a positive linear transformation. Measurement leads to the formation of profiles of utility functions, that is, of n tuples of $\{U_i\}$, for any individual $i = 1, \dots, N$.

In the second step, the utilities of different individuals are compared. There are two main kinds of IUCs, namely, interpersonal comparisons of utility levels and interpersonal comparisons of utility differences. For any two individuals i and j , and for any four options x, y, w, z , interpersonal comparisons of utility levels are judgments of the form: $U_i(x) = U_j(y)$, while interpersonal comparisons of utility differences are judgments of the form: $U_i(x) - U_i(y) = U_j(w) - U_j(z)$. If the judgment concerns a decision involving two (or more) individuals, a decision rule typically establishes the relevant kind of comparison to be made.

The question whether interpersonal comparisons of utility are empirically meaningful arises at this stage. The problem is that choice behaviour is not sufficient to determine the interpersonal comparability of different people's preference strength. Consider the following case. Suppose there are two individuals, i and j , and four options x, y, w, z . Individual i ranks the options in the following way: $xRyRwRz$. Instead, individual j has the following ranking: $wRzRxRy$. Suppose we measure their preferences on a (interval) zero-one scale, such that we assign the value 1 to the most preferred option and the value 0 to the worst option. Then, we can assign a value that represents the intensity of their preferences for the other options, relative to the best and worst in each individual's ranking.

Suppose we get that $U_i(y) = U_j(x) = 0.5$. Can we conclude that individual i prefers option y with the same strength with which individual j prefers option x ? In other words, is the empirical evidence sufficient to determine the interpersonal comparisons of individuals i 's and j 's utility (levels)? The answer is negative. Even if the scale of measurement is identically normalized, the evidence is not enough to tell whether preferences that have the same utility values, but belong to different individuals, have the same intensity. The measurement is relative to the best and worst option in each individual's preference ranking. The empirical evidence is consistent with the case that i

prefers the most preferred option with intensity ten times greater than j . In other words, there are two incompatible theories which are consistent with the evidence: a theory T_1 , maintaining that utilities are interpersonally comparable and, thereby, concluding that i and j have the same utilities; and a theory T_2 , maintaining that utilities are not interpersonally comparable and, thereby, concluding that i and j have not the same utilities, despite them having the same numerical value. It follows that IUCs are, at best, underdetermined; at worst, indeterminate.

3. Two platitudes

Theoretical scepticism contrasts with the apparent ease with which people make IUCs in everyday life. For instance, Jeffrey describes the following situation:

“Shall we open the can of New England clam chowder or the can of tomato soup, for the children’s lunch? Adam prefers the chowder; his sister Eve prefers the other. Their preferences conflict. But it is acknowledged between them that Adam finds tomatoes really repulsive, and loves clams, whereas Eve can take clam chowder or leave it alone, but is moderately fond of tomato soup. They agree to have the chowder.”⁶

Here, we can see a decision problem, which involves considering facts about other people. The decision problem arises because people advance incompatible requests, as a consequence of holding conflicting preferences. The goal is the choice of one option over the other (e.g. chowder over tomato). The solution to the decision problem makes appeal to a normative principle, according to which the decision ought to be based on people’s preferences; in particular, on how the parties’ preference intensity compares. In the example, the choice is a consequence of the agreement between the parties involved, but it could well have been the consequence of the decision of a judge, having to adjudicate the parties’ conflicting claims.

The example can be given either a descriptive or a normative reading. The descriptive reading takes the example to show that people actually make IUCs. The normative reading takes the example to show that people should make IUCs to reach a decision. In this paper, I shall be concerned with the first interpretation only. The question to ask, then, is the following: why do people make IUCs?

In Jeffrey’s example, the solution to the decision problem is reached by comparing the intensity with which the parties prefer one option over the other. The agreement is possible only because “the children are convinced that Adam’s preference for clam over tomato exceeds Eve’s preference for

⁶ JEFFREY, R. [1974], reprinted in JEFFREY, R. [1992], p. 182.

tomato over clam”⁷. Thus, the example suggests a first observation: when a choice affects other people, decision-making requires making interpersonal comparisons (e.g. of preference strength, in this case). I shall call this, the platitude about ‘the normative role of interpersonal comparisons’.

Furthermore, in Jeffrey’s example, interpersonal comparisons of preference strength are derived from empirical evidence of various kinds, namely, choice behaviour and expressive reactions in similar choice situations. Choice behaviour and expressive reactions allegedly provide evidence for the ascription of comparable preference intensities. Thus, the example suggests a second observation: interpersonal comparisons help explain people’s behaviour. By ascribing comparable degrees of preference strength, we can make sense of why the parties show different behavioural responses. In turn, the explanatory role of the comparability assumption is what justifies taking the parties’ conduct to be evidence for the ascription⁸. I shall call this, the platitude about ‘the explanatory role of interpersonal comparisons’.

One may argue that the two platitudes implicitly support the thesis that IUCs are empirically meaningful. Let us take for granted that the assumption that preference strengths are interpersonally comparable is justified only if we have a reason to make this assumption. In our case, there seems to be a clear epistemic reason, namely, that the comparability assumption helps explain both platitudes about IUCs. However, the argument is mistaken. If people really make IUCs both for explanatory and normative purposes, the correct inference to draw is that they *believe* that preferences are interpersonally comparable in terms of strength. However, the fact that they have such beliefs does not imply that there is a fact of the matter about IUCs. IUCs may even be indeterminate. This position would configure an ‘error theory’ about IUCs. We need an independent reason to think that there is a fact of the matter grounding people’s judgments about how their preferences compare in terms of strength

4. Two strategies

Choice behavioural evidence is insufficient to determine IUCs. As seen above, IUCs are empirically meaningless if and only if they are either underdetermined by the empirical evidence or indeterminate. However, if choice behaviour is not the only empirical evidence available, the possibility remains open that IUCs are simply underdetermined and that, by considering further evidence, they can be shown to be empirically meaningful. The literature explores two strategies of this kind. They both advocate the use of other behavioural proxies (e.g. verbal and/or expressive

⁷ *Ibidem*.

⁸ See DAVIDSON, D. [1986], reprinted in DAVIDSON, D. [2004], p. 65.

reactions) in order to provide empirical foundations for IUCs. They differ only with respect to how the additional evidence is supposed to imply comparability.

The first strategy configures a first-person approach, that is, an approach that aims at reducing inter-personal comparisons to intra-personal comparisons of utility. The literature offers significant instances in the work of Harsanyi, Arrow and others⁹. The starting point is the collection of information about all the behavioural proxies that are thought to be relevant either as causes or expressions of people's preferences towards alternative options. Then, the evidence is interpreted in the light of an introspective exercise. Introspection takes the form of a thought experiment, which engages the interpreter in the construction of an extended preference ranking. Extended preferences range over pairs of options and agent-types and they can be represented through an extended utility functions. The idea is that, if every individual has the same extended preference ranking and certain assumptions connecting extended to simple utilities hold, it is possible to bring people's actual utilities on the same scale.

The second strategy configures a third-person approach. It does not make use of introspection, but tries to connect the measurement of utility to the measurement of more objective empirical proxies. More precisely, it aims at establishing a functional relation between choice behaviour and other proxies. The literature offers significant instances in the work of Waldner and, more recently, List¹⁰. The first step consists in selecting proxies, whose measures are interpersonally comparable. If there exists a function relating empirical proxies to utilities and if it is unique for all individuals, then it possible to take proxy measures as arguments and transform them into interpersonally comparable utility measures. The goal of the second step, then, is to find a functional relation, unique for all individuals.

Both strategies try to show that IUCs are indeed empirically meaningful by extending the set of available evidence. Nevertheless, their success is conditional upon the adoption of a non-empirical principle. The first strategy solves the problem of interpersonal comparisons by means of thought experiments. However, this solution works only if one accepts the following principle: if the individuals are alike in all relevant respects, then they have identical extended preference strengths and, thereby, co-scaled utilities. The second strategy tries to find a unique function relating objective proxies to subjective utilities. However, uniqueness is possible only if one accepts the following principle: if the individuals are alike in all observable respects, then they have an identical function, relating objective proxies and subjective mental states and, thereby, co-scaled utilities. Although different, I shall group these principles together and refer to them as 'Harsanyi's principle'.

⁹ HARSANYI, J. C. [1955] and [1977], ARROW, K. J. [1977].

¹⁰ WALDNER, I. [1972], LIST, C. [2003].

The fact that both strategies crucially depend on the adoption of a non-empirical principle suggests one important consideration. Even if we broaden behavioural evidence, IUCs remain fundamentally underdetermined by empirical observations. This may not be too worry, provided we can avoid indeterminacy. After all, scientific theories are often underdetermined by the available evidence, but this does not necessarily mean that they are indeterminate, at least if other considerations can be advanced to think otherwise. In our case, the adoption of ‘Harsanyi’s principle’, if justifiable, may give us at least one reason to think that, although IUCs are undetermined by the empirical evidence, there still is a fact of the matter about preference strength comparability. I shall take the justification to be the following: Harsanyi’s principle is justified only if it leads to a theory that explains more evidence or that explains the same evidence in a more parsimonious or simpler way than a competing theory without Harsanyi’s principle. Can the adoption of Harsanyi’s principle be vindicated along these lines? Most importantly, does this really offer a reason to think that IUCs are not indeterminate?

5. Against Harsanyi’s principle

The first objection that one can raise is the following. The justifying argument under consideration makes reference to some pragmatic advantages that a theory embracing Harsanyi’s principle has over a theory that does not. Parsimony and simplicity are certainly theoretical virtues. As such, they have a pragmatic value. For instance, they can play a relevant role in theory choice, in a case of empirical underdetermination. However, it is questionable whether they have epistemic value as well. That is, it is questionable whether they can be used to infer something about how the world is. Therefore, one may argue that, even if Harsanyi’s principle can be justified along the lines suggested above, it can be used to reject IUCs indeterminacy, that is, it Harsanyi’s principle does not give any reason to think that there is a fact of the matter about preference strength comparability. Although important, in what follows I shall leave this objection aside.

Suppose we want to explain individual behaviour. Let us consider a theory T_1 that assumes that utilities with the same value are interpersonally comparable and a theory T_2 that assumes that utilities with the same value are not interpersonally comparable. First, let us consider explanatory power. Does T_1 explain more aspects of individual behaviour than T_2 ? The answer is negative. The power of both theories is the same, despite the fact that they differ with respect to the assumptions made about the comparability of preference strengths. The fact that T_1 allows one to making IUCs does not add anything to the power of the explanation.

This conclusion is more evident if we consider one of the goals in the explanation of human behaviour, namely, the goal of accounting for individual behaviour in terms of its causes. In our case, this means explaining behaviour in terms of individual preferences and their (causal) property of strength. Typically, empirical evidence offers a ground for positing entities and properties that can causally explain that very same evidence. Although non-causal properties can have pragmatic relevance, the information coming from empirical evidence is typically used to refine the ascription of causally relevant properties. Moreover, the explanatory power of a theory is based on the extent to which the entities and properties postulated by the theory are able to causally account for the empirical evidence.

The property of being comparable in terms of strength is not a causal property of preferences¹¹. That is, comparability plays no causal role in accounting for individual behaviour. Thus, the collection of further empirical evidence is used in two ways. First, it helps improve the understanding of the content of an individual's preferences. Second, it helps refine the ascription of relative strength between preferences for different options. It is not used to ground interpersonal comparability. The fact that comparability does not play any causal role makes clear why the explanatory power of two theories differing only as far as the assumptions made about interpersonal comparability of preference strength are concerned is the same. The theories share the same causal properties. Therefore, they have the same explanatory power. Indeed, IUCs do not add anything to the explanatory power of a theory about individual behaviour.

Second, let us consider parsimony. I shall take parsimony to be defined with respect to the number and/or kinds of properties postulated. Thus, the most parsimonious theory is the one that explains the evidence with the least number and/or kinds of property assumptions. However, the assumption that utilities are interpersonally comparable does not lead to a more parsimonious theory. Consider again the case when $U_i(y) = U_j(x) = 0.5$. Before we saw that the empirical evidence is consistent with two incompatible theories, T_1 and T_2 . On the one hand, T_1 holds that utilities are interpersonally comparable and, thereby, that i and j have the same utilities. Call the attitude conveyed by T_1 one of 'optimism'. On the other hand, T_2 holds that utilities are not interpersonally comparable and, thereby, that i and j have not the same utilities. Call the attitude conveyed by T_2 one of 'scepticism'.

Suppose now that the empirical evidence is extended so to include not only choice behavioural observations, but all the relevant behavioural proxies. Suppose that the individuals show exactly the same behaviour and that, again, $U_i(y) = U_j(x) = 0.5$. If T_1 and T_2 are the only theories at issue, it is hard not to conclude that the first is more parsimonious than the second. It characterizes the

¹¹ It can still be explanatorily relevant if pragmatic considerations are accepted as having epistemic value. See below for discussion.

individuals' behaviour by assuming that their preferences share the same, comparable property of strength. In other words, it does not postulate any hidden difference in preference strengths when the empirical evidence is identical for both individuals. This is precisely what Harsanyi's principle recommends. If so, it looks like the adoption of Harsanyi's principle does indeed lead to a more parsimonious theory, and is, therefore, justified.

However, there is another theory, T_3 , that is compatible with the empirical evidence. T_3 registers the fact that the individuals' utilities are numerically identical, but does not make any assumptions as to whether and how they compare. In other words, T_3 does not take any position about IUCs. Call the attitude conveyed by T_3 one of 'neutrality'. Neutrality is a legitimate position, because the assumption of interpersonal comparability is simply not required for the explanation of individual behaviour. Since comparability plays no role, it is possible to remain agnostic as to whether people's preference strengths are comparable or not.

If what we care about is just parsimony, then the question to ask is the following: is it more parsimonious to have a theory that does not postulate any differences between individuals' utilities or to have a theory that does not postulate anything at all? Strictly speaking, the latter is more parsimonious than the former both with respect to the number and to the kind of properties postulated and, therefore, it should be favoured. I admit that we might have conflicting intuitions here. However, the fact that the issue cannot be easily solved is enough to reject parsimony as giving a conclusive reason to believe that IUCs have a factual basis.

Third, let us consider simplicity. It is generally difficult to define what simplicity amounts to. One account reduces simplicity to parsimony. In this case, the previous remarks apply. Alternatively, another account constructs simplicity as elegance, which, in turn, is defined as the ease with which a theory favours computation or decision-making. Is a theory (T_1) that assumes interpersonal comparability simpler than either a sceptical (T_2) or a neutral (T_3) theory? To answer, we need to bear in mind the purposes for which such a theory can be used, namely, the explanatory and normative purposes.

Suppose we use T_1 to explain individual behaviour. The assumption that utility is interpersonally comparable plays no role in accounting for individual behaviour. Thus, it does not make computation any easier. It follows that T_1 cannot be deemed simpler than T_2 or T_3 , at least as far as the criterion of simplicity is defined as the ease in the calculation.

Suppose now we use T_1 to take a decision affecting the interests of two or more individuals. Undoubtedly, a theory that assumes that IUCs are empirically meaningful considerably simplifies decision-making. However, it seems to me that, in the case at stake, the justification of interpersonal comparability based on simplicity ceases to be purely pragmatic and becomes rather close to a

normative justification. We do not count as simpler a theory assuming interpersonal utility comparability merely because it leads to *a* decision; rather simplicity is valuable to the extent that it leads, or it favours reaching, decisions that are considered to be fair or even-handed. If this is true, however, it is fairness, or even-handedness, which provides the ultimate justification for the assumption that utilities are interpersonally comparable. In other words, assuming that people's utilities are interpersonally comparable is justified only insofar as that helps us to reach fair or even-handed results. Simplicity plays a merely instrumental role. Therefore, the basis for IUCs is neither pragmatic nor factual, but normative. The argument based on simplicity does not give reason to think that IUCs have a factual basis.

The conclusion is the following. The need for a non-empirical principle shows that both strategies fail to prove the empirical meaningfulness of IUCs. However, if one such principle, namely, Harsanyi's principle, is justifiable, we have at least one reason to think that IUCs are not indeterminate. We explored three possible justificatory arguments appealing to pragmatic considerations. None of them offers conclusive grounds to think that IUCs have indeed a factual basis.

6. Davidson's solution

The conclusion reached in the previous section shows that no amount of empirical evidence will make IUCs empirically meaningful. Behavioural evidence leaves IUCs, at best, underdetermined; at worst, indeterminate. The fact that no argument based on empirical evidence helps establish that IUCs have a factual basis does not entail yet that there can be no argument at all in support of this thesis. One may offer an *a priori* justification. For instance, it is possible to reformulate Davidson's position in such a way to offer an *a priori* reason to think that preference strengths are interpersonally comparable.

According to Davidson, in order to attribute mental states, the interpreter projects his own standards of rationality and values on the agent¹². By means of this projection, the interpreter establishes a comparison between his mental states and the agent's. Hence, the agent's preferences become - from the start - comparable, in certain relevant dimensions. More specifically, these dimensions include the role, the content and the structure of the agent's mental states.

First, the interpreter must assume that the agent's mental states play a role similar to the role that his mental states play in his mind. This means that the agent's preferences, desires and beliefs possess a dimension of strength. Second, *ceteris paribus*, the interpreter believes, values, desires *p*

¹² DAVIDSON, D. [2004], p. 67.

if and only if the agent believes, values, desires p . This means that similar causes generate the same beliefs, values, desires, etc. Third, the agent's mental states obey standards of rationality similar to the interpreter's ones. Minimally, this means that the agent's preferences are consistent in a specified way.

The projection establishes a comparison in the sense that it allows ascribing differences between the interpreter's and the agent's mental states across these dimensions. The interpreter can attribute to the agent irrational preferences, when they do not satisfy the standards of rationality recognized by the interpreter. Moreover, the interpreter can relate the agent's evaluative attitudes to objects different from those that provide the content of his evaluative attitudes. Clearly, differences can be tolerated up to a certain extent, that is, up to the point where the agent's behaviour remains intelligible. However, this is enough to establish interpersonal comparisons between the interpreter's and the agent's mental states.

The principle of charity offers the basis for such interpersonal comparisons. Charity is required in order to optimize agreement between the interpreter and the agent, so as to make understanding possible. It is an *a priori* principle, which is neither discovered, nor normatively chosen¹³. It follows that what justifies comparability of preferences in the relevant dimensions is neither an empirical, nor a methodological reason; but rather, an *a priori* reason.

One may read Davidson as offering an *a priori* argument for thinking that IUCs have a factual basis.¹⁴ The interpreter's attribution endows the agent's preferences with a dimension of strength. Since interpretation establishes a comparison, it also makes people's preferences interpersonally comparable in terms of strength from the start. Allegedly, then, we have an *a priori* reason to think that IUCs have a factual basis. The possibility of been interpreted is constitutive of one's having mental states. Since interpretation implies comparability, it follows that people's utilities, when suitably normalized, are already on the same scale.

Is an argument of this kind correct? In order to answer this question, let us start by considering how the principle of charity can be formulated. Since Davidson does not offer a clear definition, I shall follow rather closely Lepore and Ludwig's analysis¹⁵. According to them, the principle of charity can be defined in various ways. For the purpose of this paper, I shall focus on two of their formulations only.

Agreement: *Ceteris paribus*: the interpreter believes p iff the speaker believes p .

¹³ *Ibid.*, p. 73.

¹⁴ For instance, WEINTRAUB, R. [1998] reads Davidson in this way.

¹⁵ See LEPORE, E. and K. LUDWIG, [2005], pp. 189-190

Charity: For any speaker S , time t , belief b , *ceteris paribus*: b is a belief of S 's at t about and prompted by S 's environment iff b is true.

It is clear that Charity and Agreement differ considerably. For instance, Agreement does not refer to truth in any way, but it simply a condition about shared beliefs. Despite the differences, one may argue that at least one version of the principle of charity is sufficient to prove that IUCs have a factual basis, along the lines suggested by the previous reading of Davidson.

One strategy to assess this position consists in, first, assuming that the principle of charity is indeed sufficient for having interpersonally comparable utilities and, then, trying to see whether or not there are counterexamples to the conclusion following from that premise. One possible argument is the following.

- (a) The principle of charity is necessary for the interpretation of individual behaviour;
- (b) The principle of charity requires assuming comparable preference strength;
- (c) Therefore, the interpretation of individual behaviour requires assuming comparable preference strength.

Here, the conclusion is false. The interpreter's projection is certainly a necessary condition for the explanation of individual behaviour to be possible. It implies comparability in some relevant respects only, i.e. those respects that are necessary for the explanation of individual behaviour to be possible. The ascription of a causal property of strength to preferences is necessary, but their comparability in this respect is not. Therefore, the attribution of mental states does not ground IUCs.

One can offer an alternative argument based on the principle of charity, in order to justify the assumption of interpersonally comparable preference strength.

- (a) The principle of charity is necessary for reaching intersubjective agreement;
- (b) The principle of charity requires assuming comparable preference strength;
- (c) Therefore, intersubjective agreement requires assuming comparable preference strength.

Once again, the conclusion is false. Intersubjective agreement merely requires that the parties share the belief that preference strengths are interpersonally comparable. It does not require that there is a fact of the matter about their comparability. Therefore, the attribution of mental states does not ground IUCs.

There is another possible objection, which attacks the principle of charity more directly. As seen above, the principle of charity can be defined in two different ways. However, we can immediately see that one of the versions, i.e. Agreement, is not sufficient to ground interpersonal comparability. Agreement merely implies that people have shared beliefs, but remains silent as to whether those beliefs are true or not. In our case, it does not settle the issue of whether IUCs have a factual basis or not. Thus, we should adopt Charity as our preferred definition. Charity has a clear advantage over Agreement: it grounds intersubjective agreement on the truth of people's beliefs. Thus, it seems to better serve the goal of showing that IUCs have a factual basis.

However, Charity is not enough to conclude that IUCs are not indeterminate. Charity claims that, *ceteris paribus*, the interpreter and the speaker share true beliefs. Nonetheless, this means that *most* of the beliefs they share are true. The *ceteris paribus* clause serves to accommodate not only cases in which either the interpreter or the speaker do not appropriately respond to the environmental conditions and form erroneous beliefs; but also the case in which, although responding in the same way, they form beliefs that are false under theoretical scrutiny. Charity merely claims that this will not happen in most cases. However, it cannot justify the truth of specific beliefs or kinds of beliefs, at least in the absence of other independent considerations. Yet, this is precisely what is lacking in the case of IUCs. The conclusion is that the argument based on the principle of charity fails to show that IUCs have a factual basis.

7. Interpersonal comparisons in explanation and decision-making

The two platitudes about interpersonal comparisons show that people make IUCs both for explanatory and normative purposes. The goal of the previous sections was to see whether or not their comparative judgments are determined by the empirical evidence or, at least, whether or not they are not indeterminate. We examined two kinds of arguments. The first kind considers further empirical evidence; the second kind is based on *a priori* reasoning. We saw that IUCs remain underdetermined by the empirical evidence. Nevertheless, IUCs could avoid indeterminacy if either the adoption of Harsanyi's principle could be justified or the principle of charity could work in the intended way. However, neither condition is satisfied. If – and only if - we think that these are the only arguments available, then, the upshot is that IUCs are indeed indeterminate.

An alternative conclusion shares with the previous one the idea that IUCs are radically underdetermined by the empirical evidence, but diverges from it insofar as it remains neutral about indeterminacy. The reason for neutrality is not so much that there might be other arguments to show

that IUCs have a factual basis; rather, it is that other conditions might be necessary for indeterminacy that have not been discussed in this paper¹⁶.

The issue remains as to how we can explain people's beliefs about IUCs, if we do not have reasons to think that IUCs have a factual basis. Consider Jeffrey's example again. Let us assume that the parties truly believe to be making empirically meaningful comparative judgments. If we think that IUCs are indeterminate, we should conclude that people's beliefs are mistaken and are likely to result from the projection of other beliefs. The task is to defend an error theory of IUCs. If we remain neutral about indeterminacy, instead, we should suspend our judgments about the correctness of people's beliefs. The task is to explain why they have the beliefs they have when there is no reason to assume the existence that IUCs are factually grounded. In both cases, we need explain how IUCs are possible despite being empirically meaningless. As seen above, possibility requires that there be a basis for people's comparative judgments. It follows that our goal is to explain what basis is active in everyday life IUCs.

Let us consider the platitude about 'the normative role of interpersonal comparisons'. If IUCs are not empirically meaningful and are made for normative purposes, it is likely that their basis will come from normative considerations. I shall present three accounts. The first account holds that the parties formulate their judgments by making IUCs with a normative basis. A variety of normative conditions can be imposed to uniformly rescale people's utilities. Moreover, a variety of justifications can be offered. In general, however, the ultimate reason for adopting a particular solution is that it leads to a fair, or even-handed, outcome. People use factual information about individual preference strengths together with the assumption that preferences strengths are interpersonally comparable in order to reach a decision. If the decision is taken to be fair or even-handed, they have a (normative) justification for maintaining the comparability assumption. Such a basis guarantees the possibility of IUCs, despite their empirical meaninglessness¹⁷. Genuine agreement between the parties is ultimately the result of a convergence about the reasons for selecting a particular normative basis.

The second account holds that the parties' judgments stem from interpersonal comparisons, but not interpersonal comparisons of preference strengths. Although they believe themselves to be comparing preferences in terms of their strength, they are comparing preferences in terms of other considerations, e.g. their normative significance. In turn, normative significance expresses the value that the parties place on the satisfaction of one preference over another. If the satisfaction of one party's preferences is held to be more valuable than the satisfaction of the other party's preferences, then the decision is still based on a comparative judgment, whose object, however, is not preference

¹⁶ For a discussion of other conditions for indeterminacy see LEPORE, E. and K. LUDWIG, [2005], pp. 223-225.

¹⁷ Essentially, this is the position defended by Jeffrey in JEFFREY, R. [1971] and [1974].

strength. Agreement is the result of a convergence about the normative significance of the parties' preferences or, ultimately, about interpersonal comparisons of different objects.

It is easier to understand this account if one adopts a value-based theory of preferences. Accordingly, preferences are based on reasons, which are constituted by valuable aspects of the world. If the parties think that the reasons that individual i has for having certain preferences are better than the reasons that individual j has for having alternative preferences, then, if their judgments are based on individual preferences, they can reach agreement. One may assume that better reasons determine stronger reasons and that stronger reasons determine stronger individual preferences. Most importantly, one may *also* assume that, if reasons are interpersonally comparable in terms of goodness, so are preferences in terms of strength. In this case, agreement based on interpersonal comparisons of reasons implies agreement based on preference strength. It seems to me that people's reasoning often proceeds in this way. But the assumption about preference strength comparability needs a justification. As we have seen, the justification cannot be based on empirical observations. If there is one, it will be normative. More precisely, it will be one based on the normative significance of individual preferences.

Finally, the third account holds that the parties' judgments are grounded in no interpersonal comparisons at all. Although the parties believe they are making IUCs, their judgments depend on other features of their reasoning, which do not involve interpersonal comparisons of any kind. For instance, consider a decision rule such as the minimax relative concession rule. The minimax principle does not require any interpersonal comparisons, either of utility or of other objects. If the parties conform to this decision rule, then, the assumption that utilities are interpersonally comparable is entirely redundant. This means that agreement can be explained without appealing to the interpersonal comparability of preference strength, although the parties may well believe that agreement is the results of IUCs. Were they to abandon the assumption that utilities are interpersonally comparable, the relevant features of their reasoning would be such to lead them to the same conclusions. Once again, the basis of people's IUCs is normative: it lies on the fairness, or even-handedness, of the outcome, which, unlike the first account, is determined by features of the reasoning that have nothing to do with interpersonal comparisons.

Let us consider now the platitude about 'the explanatory role of interpersonal comparisons'. I shall present three accounts. The first account treats people's beliefs about IUCs made for explanatory purposes as different strategies in a game. Suppose there are two individuals, i and j . Suppose the utility values representing preference strengths are the same. They have three strategies, which consists in holding either an optimistic, or a sceptical or a neutral attitude towards the interpersonal comparability of their utilities. One may hold that communication puts a higher

prize on reaching agreement. In turn, agreement can be achieved if both parties adopt the same optimistic strategy towards IUCs, when they face the same body of evidence. Agreement is the only reason they have to assume comparability. Clearly, however, the mere fact that the parties can reach agreement by making a certain assumption does not, by itself, give them reason to think that the assumption is true. If the parties believe that this is the case, they merely have a believed reason to hold that IUCs have a factual basis. Indeed, their basis is simply conventional. IUCs are possible because certain assumptions are conventionally made that plausibly secure intersubjective agreement.

The second account reconsiders one of the arguments discussed above, namely, the argument appealing to pragmatic reasons. We saw that pragmatic considerations do not offer sufficient ground to think that IUCs have a factual basis. However, reference to parsimony or simplicity may well play a role in explaining how IUCs are possible. If, when facing the same body of evidence, the parties share the assumption that utilities are interpersonally comparable on the grounds that this is the most parsimonious or the simplest hypothesis to adopt, then they make IUCs on the basis of pragmatic considerations. However, the mere fact that the parties take pragmatic considerations to be a reason to hold a certain assumption does not, by itself, imply that those considerations actually offer a reason in favour of that assumption. If the parties believe so when the comparability assumption is at stake, they merely have a believed reason to hold that IUCs have a factual basis. The basis to make IUCs for explanatory purposes is only pragmatic¹⁸.

The third account reconsiders the argument based on the principle of charity. When making IUCs in order to explain people's behaviour, the parties take certain facts as a reason to think that preference strengths are interpersonally comparable. The idea is that the same environmental conditions prompt different individuals to form preferences in the same way and with the same strength. As seen above, this idea is the same conveyed by the first version of the principle of charity, namely, Agreement. If we add the condition that people's shared beliefs are mostly true, we get the second version, namely, Charity.

We saw before that Charity is not sufficient to justify the truth of specific beliefs or kinds of beliefs. Nevertheless, the parties may believe so. After all, as Davidson suggests, the principle of charity is neither chosen nor discovered, but is active as a pre-condition of interpretation. This means that the parties will proceed by assuming that they shared beliefs are true, even though specific belief-tokens or even belief-types might not be. If this is the case, they take the principle of charity as justifying the assumption that utilities are interpersonally comparable. According to our previous discussion, however, the principle of charity offers no actual reason for this conclusion. It

¹⁸ LIST, C. [2003] holds a similar position.

follows that the parties merely have a believed reason to hold that IUCs have a factual basis. However, the principle of charity explains how it is possible to make IUCs despite their empirical meaningfulness. Indeed, the principle of charity can be seen as the basis to make IUCs for explanatory purposes.

8. Conclusion

One problem concerning IUCs is whether they are empirically meaningful or not. The first goal of this paper has been to illustrate why the problem is so theoretically intractable. Choice behavioural evidence is not sufficient to determine IUCs. Therefore, they may be either underdetermined or indeterminate.

The debate is often introduced by means of examples of everyday IUCs. They can be used to derive two platitudinous observations about IUCs. The first is that, when a choice affects other people, decision-making requires making interpersonal comparisons; the second is that we need interpersonal comparisons to explain people's behaviour.

One way to explain why people make IUCs is by assuming that there is a fact of the matter about comparable preference strength. One strategy consists in collecting further evidence in addition to choice behavioural observations and showing that IUCs are empirically meaningful with respect to this broader evidence. However, this strategy relies on the adoption of a non-empirical principle, which I called Harsanyi's principle. I argued that IUCs remain underdetermined by the empirical evidence and that the arguments used to justify the adoption of Harsanyi's principle are unsuccessful.

Another strategy consists in defending the thesis that IUCs have a factual basis by means of an *a priori* argument. According to Davidson, for instance, the interpreter projects his own standards of rationality and values on the agent. By so doing, he establishes a comparison between his mental states and the agent's, in some relevant respects. The principle of charity offers the basis for such interpersonal comparisons. Nonetheless, I argued that Davidson's account cannot be used to justify the thesis that IUCs have a factual basis.

The upshot is that, if we think that there is no other contrary argument available, IUCs are indeterminate. Alternatively, one can adopt a weaker position, that maintains that IUCs are radically underdetermined by the empirical evidence, but remains neutral as to whether or not they are indeterminate.

The problem is how to explain everyday IUCs. I tried to show how an error theory of IUCs can be formulated that defend the possibility of IUCs without relying on their empirical meaningfulness.

I offered three accounts of the normative role of IUCs and three accounts of their explanatory role. If my analysis is correct, we can hope to explain everyday IUCs even if we do not assume that they have a factual basis.

References

- ARROW, K. J. "Extended Sympathy and the Possibility of Social Choice", *The American Economic Review*, 67 (1977), pp. 219-225.
- BROOME, J. *Weighing Goods: Equality, Uncertainty and Time*, Oxford: Blackwell, 1991.
- DAVIDSON, D. "Judging interpersonal interests", in ELSTER, J. and A., HYLLAND (eds.), *Foundations of social choice theory*, Cambridge: Cambridge University Press, 1986, pp. 195-211.
- "Interpersonal Comparisons of Values", in DAVIDSON, D. *Problems of Rationality*, Oxford: Oxford University Press, 2004, pp. 59-74.
- ELSTER, J. and J. E., ROEMER, "Introduction", in ELSTER, J. and J. E., ROEMER (eds.), *Interpersonal Comparisons of Well-Being*, Cambridge: Cambridge University Press, 1991, pp. 1-16.
- FLEURBAEY, M. and J., HAMMOND "Interpersonally Comparable Utility" in BARBERÀ, S. HAMMOND, J. AND C., SEIDL (eds.), *Handbook of Utility Theory*, Vol. II, Kluwer Academic, 2004.
- JEFFREY, R. "On Interpersonal Utility Theory", *The Journal of Philosophy*, 68 (1971), pp. 647-656.
- "Remarks on interpersonal utility theory", in STENLUND, S. (ed.) *Logical Theory and Semantical Analysis*, Dordrecht: D. Reidel, 1974, pp. 35-44.
- Probability and the art of judgment*, Cambridge: Cambridge University Press, 1992.
- HARSANYI, J. C. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility", *The Journal of Political Economy*, 63 (1955), pp. 309-321.
- Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*, Cambridge: Cambridge University Press, 1977.
- LEPORE, E. and K. LUDWIG, *Donald Davidson. Meaning, Truth, Language, and Reality*, Oxford: Clarendon Press, 2005.
- LIST, C. "Are Interpersonal Comparisons of Utility Indeterminate?", *Erkenntnis*, 58 (2003), pp. 229-260.
- WALDNER, I. "The Empirical Meaningfulness of Interpersonal Utility Comparisons", *The Journal of Philosophy*, 4 (1972), pp. 87-103.
- WEINTRAUB, R. "Do Utility Comparisons Pose a Problem?", *Philosophical Studies*, 92 (1998), pp. 307-319.