

# Degrees of Preference and Degrees of Preference Satisfaction

MAURO ROSSI

*Université du Québec à Montréal*

The standard view holds that the degree to which an individual's preferences are satisfied is simply the degree to which the individual prefers the prospect that is realized to the other prospects in her preference domain. In this article, I reject the standard view by showing that it violates one fundamental intuition about degrees of preference satisfaction.

It is commonly admitted that well-being comes in degrees. If we take an individual's well-being to be constituted by the satisfaction of her (rational and well-informed) preferences, then the degree of that individual's well-being is constituted by the degree to which her preferences are satisfied. One question immediately arises: what are degrees of preference satisfaction? The standard view defines the degree to which an individual's preferences are satisfied as the degree, or intensity, to which the individual prefers the prospect that is realized to the other prospects.<sup>1</sup> According to this view, to judge that an individual's preferences are satisfied to degree  $n$  is thus equivalent to judge that the individual prefers the realized prospect to degree  $n$ .<sup>2</sup> I shall call this the 'intensity view' of degrees of preference satisfaction. In this article, I am going to show that the 'intensity view' should be rejected because it violates one fundamental intuition about degrees of preference satisfaction.

<sup>1</sup> See, for instance, John Harsanyi, *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations* (Cambridge, 1977); Richard Hare, *Moral Thinking: Its Levels, Method and Point* (Oxford, 1981); James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford, 1986).

<sup>2</sup> Throughout this article, I shall speak of preferences as roughly synonymous with desires. Importantly, I shall take the degree of preference for one prospect to be the degree to which the individual desires that prospect. For this understanding of degrees of preference, see also Philip Pettit, 'Preference, Deliberation and Satisfaction', *Royal Institute of Philosophy Supplement* 81 (2006), pp. 131–54; and Richard Bradley, 'Comparing Evaluations', *Proceedings of the Aristotelian Society* 108 (2007), pp. 85–100.

## I. BASIC NOTIONS

Let us clarify the meaning of the basic notions involved in the debate. Let us start with the notion of preference. Preferences are binary relations, i.e. relations between two items. The items that figure in a preference relation belong to the individual's preference domain. To put it simply, the items in the preference domain are the prospects that the individual faces. More formally, let  $i$  be an individual considering a finite set of prospects  $A = \{x, y, \dots, w, z\}$ . The individual  $i$ 's preference for  $x$  over  $y$  can be represented as  $x R_i y$ , for any  $x, y \in A$ . When is a preference satisfied? We can say that an individual's preference is satisfied if and only if the world is like the individual prefers it to be.<sup>3</sup> Equivalently, we can say that an individual's preference is satisfied if and only if the prospect that the individual prefers is realized. More formally, an individual  $i$ 's preference for  $x$  over  $y$  is satisfied if and only if  $x$  is the case, for any  $x, y \in A$ .

Before proceeding, it is worth emphasizing one important distinction. In ordinary language, we speak of preference satisfaction in two different ways. On the one hand, we say that a preference is satisfied when the preferred prospect is realized, in conformity with the definition presented above. This notion of preference satisfaction refers to a relation of correspondence between the object of the individual's preference and the world. On the other hand, we sometimes say that a preference is satisfied when the realization of the preferred prospect is accompanied by a feeling of satisfaction. In this sense, satisfaction is a pleasurable experience that occurs to the individual as a consequence of the realization of the preferred prospect.

If we want to remain within a preference satisfaction theory of well-being, however, the latter definition should be discarded. The reason is the following. If one conceives preference satisfaction as a pleasurable experience of satisfaction, then the resulting account of well-being is that well-being is constituted by pleasurable experiences of satisfaction. But this amounts to a hedonist theory of well-being. Now, the preference satisfaction theory is supposed to provide a distinct and, in its defenders' intentions, better account of well-being than hedonism.<sup>4</sup> If we want to preserve the distinction, we need to characterize preference satisfaction in accordance with the first definition presented above. Accordingly, the satisfaction of an individual's

<sup>3</sup> See also Daniel Hausman, 'The Impossibility of Interpersonal Utility Comparisons', *Mind* 104 (1995), pp. 473–90, at 473.

<sup>4</sup> One reason is that the preference satisfaction theory is immune from the 'experience machine' objection against hedonist theories of well-being. For this objection, see Robert Nozick, *Anarchy, State and Utopia* (New York, 1974).

preference depends on the realization of the preferred prospect, not on the feelings of satisfaction experienced by the preferring individual.

## II. THE STANDARD VIEW

Typically, an individual has preferences over several prospects. What can we say about the satisfaction of her preferences when one of these prospects is realized, but not the one that she prefers the most? Let us consider an example. Let  $i$  be an individual considering a finite set of prospects  $A = \{x, y, \dots, w, z\}$ . If  $i$ 's preferences are complete and transitive, they form an ordering, or ranking, of prospects. Suppose that  $x R_i y R_i \dots w R_i z$ . If  $i$ 's preferences satisfy the other axioms of expected utility theory, they can be represented by an expected utility function  $u: A \rightarrow \mathfrak{R}$ , unique up to a positive affine transformation of the following sort:  $u(\cdot) = \alpha u(\cdot) + \beta$ , for some  $\alpha > 0$  and  $\beta \in \mathfrak{R}$ . Suppose now that  $y$  is the realized prospect. Are the individual's preferences satisfied?

It is quite intuitive to think that the individual's preferences are satisfied *to a certain degree*. That is, it is intuitive to evaluate the individual's situation in terms of degrees of preference satisfaction. According to a widespread view, the degree to which an individual's preferences are satisfied is simply the degree of importance of the realized prospect, relative to the other prospects in the preference domain. As the relative importance of each prospect is measured by the expected utility function  $u$ , the degree to which individual  $i$ 's preferences are satisfied is equal to the utility value of  $y$ , i.e. the realized prospect. To arrive at the 'intensity view' of degrees of preference satisfaction, one needs only to add the claim that expected utility represents the degree, or intensity, of an individual's preferences.<sup>5</sup> If this is the case, the relative importance of the various prospects within the individual's preference ranking is just the intensity with which the individual prefers each of these prospects to the others. It follows that the degree to which individual  $i$ 's preferences are satisfied is equal to her degree of preference for  $y$ , i.e. the realized prospect, which is numerically represented by the utility value of  $y$ .

## III. ABSOLUTE VERSUS INTERVAL SCALES OF MEASUREMENT

I want to argue that the 'intensity view' is not the correct account of degrees of preference satisfaction. My argument will be the following.

<sup>5</sup> See, for instance, Harsanyi, *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*; Pettit, 'Preference, Deliberation and Satisfaction'; and Bradley, 'Comparing Evaluations'.

We have independent intuitions about what the correct measurement of degrees of preference satisfaction should look like. Now, if degrees of preference satisfaction are to be equated to degrees of preference, then the measurement of degrees of preference has to match the intuitively correct measurement of degrees of preference satisfaction. However, there are reasons to think that the two do not match. If this is the case, it follows that degrees of preference satisfaction cannot be equated to degrees of preference.

Let us come back to our previous example. If the individual's preference domain  $A$  contains a finite number of prospects, her preference ranking has both an upper and a lower limit. In our example, the upper limit is given by  $x$ , i.e. the prospect that the individual prefers the most, whereas the lower limit is given by  $z$ , i.e. the prospect that the individual prefers the least. If preferences are *bounded* in this sense, then we can understand the idea of degrees of preference satisfaction in a very intuitive and natural way: the individual  $i$ 's preferences are *fully satisfied* if and only if the realized prospect is at the top of the individual's ranking; the individual  $i$ 's preferences are *fully frustrated* if and only if the realized prospect is at the bottom of the individual's ranking; the individual  $i$ 's preferences are *satisfied to an intermediate degree* if and only if the realized prospect is between the prospect at the top and the prospect at the bottom of the individual's ranking. Let us call this the 'basic intuition' about degrees of preference satisfaction.

The 'basic intuition' suggests that degrees of preference satisfaction should be measured on an *absolute* scale. An absolute scale is a cardinal scale of measurement. As a cardinal scale, it is determined by fixing two points, the zero and the unit. What makes it an absolute scale is that both the zero and the unit are 'absolute'. In simpler words, this means that both the zero and the unit have a natural, non-arbitrary, understanding. Degrees of preference satisfaction meet the conditions for being measured on an absolute scale of this sort. Indeed, if preferences are bounded, the degree of preference satisfaction ranges from 0 to 1. The value 0 corresponds to the situation when the realized prospect is the prospect at the bottom of the individual's ranking, in which case the individual's preferences are fully frustrated; the value 1 corresponds to the situation when the realized prospect is at the top of the individual's ranking, in which case the individual's preferences are fully satisfied.

Let us contrast the measurement of degrees of preference satisfaction with the measurement of degrees of preference. As we have seen above, if preferences satisfy the expected utility axioms, they can be represented by a utility function  $u$ , unique up to a positive affine transformation. This means that preferences can be represented on

an interval scale of measurement. An interval scale is a cardinal scale of measurement, whose specificity is that both the zero and the unit are fixed arbitrarily. Indeed, saying that the representation of preferences is unique up to a positive affine transformation is equivalent to saying that such a representation is invariant with respect to a linear shift of the zero and the unit. If this is the case, degrees of preference do not range absolutely between 0 and 1, like degrees of preference satisfaction. The prospect at the bottom of the individual's preference ranking can take values different from 0 (i.e. either positive or negative, depending on the transformation). Likewise, the prospect at the top of the individual's preference ranking can take values different from 1 (i.e. once again, either positive or negative, depending on the transformation). This suggests the possibility of a mismatch between the measurement of degrees of preference satisfaction and the measurement of degrees of preference.

#### IV. SAVING THE 'INTENSITY VIEW'?

Saying that a mismatch is possible is not equivalent to saying that a mismatch is inevitable. Indeed, we might simply be facing a problem of rescaling. One may reason in the following way. It is true that some admissible transformations may lead to a representation of degrees of preference that does not match the representation of degrees of preference satisfaction conforming to the 'basic intuition'. What we have to do in these cases is simply to ban those transformations. In other words, we have to recognize that only one of the admissible transformations of preference intensity is correct for the purpose of representing degrees of preference satisfaction. Such transformation is the 'zero-one' rule.<sup>6</sup> The 'zero-one' rule rescales degrees of preference by assigning the value 0 to the prospect at the bottom of the individual's preference ranking, the value 1 to the prospect at the top of her ranking, and values ranging from 0 to 1 to all the intermediate prospects. If we apply the 'zero-one' rule, the measurement of degrees of preference satisfaction in terms of preference intensity seems to conform to the 'basic intuition'.<sup>7</sup> One may wonder whether this is an ad hoc manoeuvre. Isn't it just arbitrary to single out one transformation amongst the otherwise admissible ones? One may reply that there is no harming arbitrariness here. If preferences are represented through an interval scale of measurement, the zero and the unit are fixed arbitrarily to

<sup>6</sup> See John Isbell, 'Absolute Games', *Contributions to the Theory of Games*, ed. A. W. Tucker and R. D. Luce (Princeton, 1959), pp. 357–96.

<sup>7</sup> Hausman makes a similar proposal, even though his goal is not to defend the 'intensity view' (Hausman, 'The Impossibility of Interpersonal Utility Comparisons', p. 480).

begin with. Nothing prevents us from setting them where it is more convenient and convenience greatly depends on the purpose that the representation of preference intensity serves – in this case, the correct representation of degrees of preference satisfaction.

This reply is acceptable only if the unit and zero of preference intensity are effectively arbitrary and can thus be assigned to prospects in a way that respects the ‘basic intuition’ about degrees of preference satisfaction. In what follows, however, I will argue, first, that, contrary to what most authors in the field believe, degrees of preference admit of a natural zero. This means that it is possible to represent preferences not only on an interval scale, but also on a *ratio* scale, i.e. a cardinal scale of measurement that has a natural zero combined with an arbitrary unit. Second, I will show that the natural zero of preference intensity does not coincide with the natural zero of preference satisfaction. If this is the case, an important result follows: there are no admissible transformations that allow us to rescale degrees of preference in such a way as to respect the ‘basic intuition’ about degrees of preference satisfaction. The conclusion is that the ‘intensity view’ should be rejected because it violates the ‘basic intuition’.

## V. THE ARGUMENT FROM THE ETHICALLY NEUTRAL PROSPECT

In order to represent preferences on a ratio scale, we need to fix the point that corresponds to the idea of a natural zero. Is there such a point? Most people are sceptical. Yet, there is a reason to think that such a sceptical attitude is unjustified. If we adopt a functionalist understanding of preferences, preferences can be characterized as mental states that lead an individual to choose one prospect over the others, in the presence of certain beliefs.<sup>8</sup> In this sense, intensity is simply a property of the individual’s preferences that contributes to shaping their causal role with respect to different prospects. But then the idea of zero intensity has a natural understanding: the degree of preference for a prospect is zero if and only if the desire for that prospect plays no causal role whatsoever in the determination of the agent’s behaviour.

We can specify this idea further. Nozick suggests that a desire for a prospect plays no causal role whatsoever when the individual neither prefers nor dis-prefers a prospect, or, equivalently, when the individual is indifferent between an alternative and its negation. Formally, for any prospect  $p \in A$ ,  $p$  has zero utility if and only if  $p I_i \neg p$ , where  $I$  is

<sup>8</sup> See Pettit, ‘Preference, Deliberation and Satisfaction’, pp. 134–8.

the indifference relation.<sup>9</sup> Bradley appeals to the role of the ethically neutral prospect in a similar way.<sup>10</sup> The ethically neutral prospect is the prospect whose realization is a matter of indifference to the individual, no matter what else is the case.<sup>11</sup> More formally, for  $p$  to be an ethically neutral prospect, it is not enough that the individual be indifferent between  $p$  and its opposite  $\neg p$ . Instead,  $p$  is ethically neutral if and only if the agent is indifferent between the prospect ( $p \& q$ ) and the prospect ( $\neg p \& q$ ), for any  $p, q \in A$ .<sup>12</sup> If  $p$  is ethically neutral in this sense, then it is the case that  $u_i(p \& q) = u_i(\neg p \& q)$ , for any  $q \in A$ . This equivalence suggests that  $p$  does not make any difference to the individual's degrees of preference, or, which is the same, that it does not play any causal role in the determination of her behaviour. But then it is all the more natural to think that the ethically neutral prospect picks out the point where the individual's preferences have zero intensity, i.e. the point such that  $u_i(p) = 0$ . Thus, degrees of preference admit of a natural zero, which is fixed in correspondence with the ethically neutral prospect. As such, preferences can be represented on a ratio scale.

To disprove the 'intensity view' of degrees of preference satisfaction, we simply need to show that the natural zero of preference intensity does not coincide with the natural zero of preference satisfaction. Let us consider a finite set of prospects  $A$ , including the ethically neutral prospect  $p$ , such that:  $A = \{x, y, p, \dots, w, z\}$ . If  $i$ 's preferences are complete and transitive, they form an ordering, or ranking, of prospects. Suppose that  $x R_i y R_i p \dots w P_i z$ , where  $P$  is a relation of strict preference. The 'basic intuition' sets the natural zero of preference satisfaction in correspondence with the least favoured prospect, i.e.  $z$ . In other words, according to the 'basic intuition', the degree of satisfaction of  $i$ 's preferences is 0 in correspondence with  $z$ . By contrast, the natural zero of preference intensity is fixed in correspondence with the ethically neutral prospect. Now, given that  $u_i(p) = 0$  and that  $p P_i z$ , it follows that  $u_i(z) < 0$ . That is, the degree of preference for the prospect  $z$  is negative. Therefore, the natural zero of preference satisfaction does not coincide with the natural zero of preference intensity.<sup>13</sup>

<sup>9</sup> See Robert Nozick, 'Interpersonal Utility Theory', *Social Choice and Welfare* 2 (1985), pp. 161–79, esp. pp. 167–8.

<sup>10</sup> See Bradley, 'Comparing Evaluations', pp. 95–6.

<sup>11</sup> See Frank Ramsey, 'Truth and Probability' (1926), reprinted in *Philosophical Papers*, ed. D. H. Mellor (Cambridge, 1990), pp. 52–109.

<sup>12</sup> To clarify the notation: the prospect ( $p \& q$ ) is the prospect that both  $p$  and  $q$  are the case, while the prospect  $\neg p$  is the prospect that  $p$  is not the case. See also Bradley, 'Comparing Evaluations', p. 91.

<sup>13</sup> More strongly, preference intensity and preference satisfaction will not coincide *unless* every prospect is ranked the same. This is the only way for the ethically neutral prospect to occupy the bottom position in the individual's preference ranking and, thereby, for the natural zero of preference intensity to coincide with the natural zero

To make the implications more vivid, suppose that  $z$  is the realized prospect. According to the 'basic intuition', the individual's preferences are completely frustrated, since  $z$  is at the bottom of the individual's preference ranking. The degree of satisfaction of the individual's preferences is thus equal to 0. The 'intensity view' suggests instead that the degree of satisfaction of the individual's preferences is  $< 0$ . By so doing, it violates the 'basic intuition' and should therefore be rejected.<sup>14</sup>

rossi.mauro@uqam.ca

of preference satisfaction. Indeed, if there is at least one prospect  $x$  that is strictly preferred to the ethically neutral prospect  $p$ , then  $x P p$  entails that  $p P \neg x$ , for  $x, p \in A$ . The ethically neutral prospect  $p$  is not at the bottom of the individual's ranking and, therefore, preference satisfaction and preference intensity do not coincide. Thanks to Richard Bradley for drawing my attention to this point.

<sup>14</sup> I would like to thank Richard Bradley, Armin Schulz, Alain Voizard and two anonymous referees for their comments on previous drafts of the article.