[Published in *The Philosophical Quarterly* 60(241):754–763, 2010, DOI: 10.1111/j.1467-9213.2009.648.x.
The definitive version is available at www.blackwell-synergy.com.]

Conflicting Reasons in the Small-Improvement Argument

Johan E. Gustafsson and Nicolas Espinoza

The small-improvement argument is usually considered the most powerful argument against comparability, viz the view that for any two alternatives an agent is rationally required either to prefer one of the alternatives to the other or to be indifferent between them. We argue that while there might be reasons to believe each of the premises in the small-improvement argument, there is a conflict between these reasons. As a result, the reasons do not provide support for believing the conjunction of the premises. Without support for the conjunction of the premises, the small-improvement argument for incomparability fails.

The small-improvement argument is an influential line of reasoning often employed in the contemporary debate on incomparable values.[1] Joseph Raz famously called it 'the mark of incommensurability', and more recently it figured as an integral part in Ruth Chang's attempt to establish a fourth value relation called 'parity'.[2] In this paper we argue that the small-improvement argument fails.[3]

Roughly, the small-improvement argument goes like this: suppose two alternatives *a* and *b* are such that neither is rationally preferred to the other. Ordinarily, you would then assume that the two alternatives are rationally equi-preferred. However, if one of the alternatives is improved, if ever so slightly, and it turns out that this slightly improved alternative *c* is still not rationally preferred to the non-improved alternative *b*, the original alternatives *a* and *b* cannot have been rationally equi-preferred to begin with. For if they had been rationally equi-preferred, then any small improvement would have tipped the scale in favour of the improved alternative.

1. See, e.g. R. Chang, 'Introduction', in R. Chang (ed.), *Incommensurability, Incomparability, and Practical Reason* (Harvard UP, 1997), pp. 1–34, and 'The Possibility of Parity', *Ethics*, 112 (2002), pp. 659–88; J. Gert, 'Value and Parity', *Ethics*, 114 (2004), pp. 492–510; R. Wasserman, 'Indeterminacy, Ignorance and the Possibility of Parity', *Philosophical Perspectives*, 18 (2004), pp. 391–403; N.-H. Hsieh, 'Equality, Clumpiness and Incomparability', *Utilitas*, 17 (2005), pp. 180–204; E. Carlson, 'Incomparability and Measurement of Value', in K. McDaniel et al. (eds), *The Good, the Right, Life and Death* (Aldershot: Ashgate, 2006), pp. 19–44; M. Peterson, 'Indeterminate Preferences', *Philosophical Studies*, 130 (2006), pp. 297–320; N. Espinoza, 'The Small Improvement Argument', *Synthese*, 165 (2008), pp. 127–39; W. Rabinowicz, 'Value Relations', *Theoria*, 74 (2008), pp. 18–49.

2. J. Raz, *The Morality of Freedom* (Clarendon Press, 1988); Chang, 'The Possibility of Parity'.

3. There are both preferential and axiological versions of the argument. The received view, however, seems to be that rationally required preferences and value judgements are closely related, and according to the popular fitting attitudes and buck-passing account of good, the one can even be analysed in terms of the other (see Rabinowicz, 'Value Relations'). In this paper we discuss the preferential version of the argument.

Since still neither alternative is preferred to the other and they are not equi-preferred, the conclusion is that they are incomparable.

The objection we raise against the small-improvement argument is that while there might be reasons to believe each of its premises, there is a conflict between these reasons. As a result the reasons do not provide a reason to believe the conjunction of the premises. Without support for the conjunction of its premises, the small-improvement argument for incomparability fails.

In § 1 we present the small-improvement argument and its premises in detail. In § 2, we put forward a plausible condition which reasons for the premises of an argument must satisfy in order to support the argument's conclusion. In § 3, we argue that the reasons offered in support of the premises in the small-improvement argument do not satisfy this condition.

1. The small-improvement argument

One of the core premises of the small-improvement argument was, as far as we know, first introduced by Leonard J. Savage:

> If the person really does regard **f** and **g** as equivalent, that is, if he is indifferent between them, then, if **f** or **g** were modified by attaching an arbitrarily small bonus to its consequences in every state, the person's decision would presumably be for whichever act was thus modified.[4]

Let 'P' denote the preference relation and 'I' the indifference relation (equi-preference). Savage's proposal can be seen as an instance of PI-transitivity:

> *PI-transitivity.* $\forall x \forall y \forall z((xPy \wedge yIz) \rightarrow xPz)$.

This is just to say that if y and z are equi-preferred, and x is preferred to y, then x is preferred to z. PI-transitivity may be viewed either as a normative requirement or as an empirical hypothesis. The latter is decidedly false, but the former is still up for debate.

The small-improvement argument is generated by combining PI-transitivity with the following kind of example, due to Ronald de Sousa:

> … the case of the fairly virtuous wife. I tempt her to come away with me and spend an adulterous weekend in Cayucos, California. Imagine for simplicity of argument that my charm leaves her cold. The only inducement that makes her hesitate is money. I offer $1,000 and she hesitates. Indeed she is so thoroughly hesitant that the classical decision theorist must conclude that she is *indifferent* between keeping her virtue for nothing and losing it in Cayucos for $1,000. […] The obvious thing for me to do now is to get her to the point of clear preference. That should be easy: everyone prefers $1,500 to $1,000, and since

4. L. J. Savage, *The Foundations of Statistics* (New York: Wiley, 1954), p. 17.

> she is indifferent between virtue and \$1,000, she must prefer \$1,500 to virtue by exactly the same margin as she prefers \$1,500 to \$1,000: or so the axioms of preference dictate. Yet she does not. As it turns out she is again 'indifferent' between the two alternatives.[5]

The conclusion de Sousa draws from this is that the alternatives a = 'lose virtue for \$1000' and b = 'keep virtue' are *incomparable*, that is, a is not preferred to b, and b is not preferred to a, and a and b are not equi-preferred. This implies incomparability, which is simply the negation of comparability, defined as

Comparability. $\forall x \forall y (xPy \vee yPx \vee xIy)$.

How does he reach this conclusion? Let c = 'lose virtue for \$1500'. Then according to the story the wife has the following preferences:

The virtuous-wife preferences. $\neg(aPb) \wedge \neg(bPa) \wedge cPa \wedge \neg(cPb)$.

She does not prefer a to b, she does not prefer b to a, but she does prefer c to a. By the assumption that she is indifferent between a and b it should then follow from PI-transitivity that she prefers c to b. However, she does not prefer c to b. By *modus tollens* it then follows that she is not indifferent between a and b. But if she is not indifferent, and she does not prefer either of a or b to the other, we must conclude that a and b are incomparable.

Chang ('The Possibility of Parity', p. 669) has an analogous axiological example:

> Suppose you must determine which of a cup of coffee and a cup of tea tastes better to you. The coffee has a full-bodied, sharp, pungent taste, and the tea has a warm, soothing, fragrant taste. It is surely possible that you rationally judge that the cup of Sumatra Gold tastes neither better nor worse than the cup of Pearl Jasmine and that although a slightly more fragrant cup of the Jasmine would taste better than the original, the more fragrant Jasmine would not taste better than the cup of coffee.

The structure of the small-improvement argument is illustrated more clearly in the light of the following trilemma:

The comparability trilemma. The following three statements cannot all be true (this can be proved trivially):

1. The virtuous-wife preferences are rational
2. PI-transitivity is rationally required
3. Comparability is rationally required.

5. R. de Sousa, 'The Good and the True', *Mind*, 83 (1974), pp. 534–51, at pp. 544–5.

Advocates of the small-improvement argument offer a reason to believe (1) with the virtuous-wife example, and there are reasons to believe (2) from money-pump arguments. Furthermore, it follows logically from the comparability trilemma that if the conjunction of (1) and (2) is true, (3) is false. Having a reason then to believe the conjunction (1) $\wedge$ (2) would imply having a reason to reject comparability.

2. Assumption of other conjuncts

However, to have a reason to believe (1) and a reason to believe (2) is not necessarily to have a reason to believe the conjunction (1) $\wedge$ (2). For example, you may have a reason to believe that Sally is at home because her car is in the driveway, and you may have a reason to believe that she is not at home because all the lights are out, but you would not therefore have a reason to believe that she is both at home and not at home at the same time. Of course, someone might object that this example is not analogous to the comparability trilemma, since Sally's being at home and not being at home are logically inconsistent, while (1) and (2) in the comparability trilemma are logically consistent. An example that does take this into account is the following, where (i), (ii) and (iii) are logically inconsistent, but (i) and (ii) are logically consistent:

(i) $u \rightarrow a$
(ii) $u \rightarrow \neg a$
(iii) u.

Let 'u' denote your favourite moral theory, and let 'a' denote that a certain act is permissible. Suppose you have a reason to believe (i) that the act is permissible according to u, and you also have a reason to believe (ii) that the act is not permissible according to u. Obviously it does not then follow that you have a reason to believe (i) $\wedge$ (ii), that the act is both permissible and not permissible according to u. If you do not have a reason to believe (i) $\wedge$ (ii), you do not have a reason to reject (iii).

This, of course, has a crucial bearing on the small-improvement argument, since a reason for (1) and a reason for (2) might not provide a reason for (1) $\wedge$ (2), which, as previously noted, is needed in order to establish incomparability. So while there certainly are arguments in which the reasons for the individual conjuncts provide a reason to believe the conjunction, the question is whether the reasons to believe (1) and the reasons to believe (2) do actually provide a reason to believe (1) $\wedge$ (2) in the small-improvement argument.

Although it is difficult to give a complete account of the circumstances under which a set of reasons for individual conjuncts combine into a reason to believe the conjunction as well, there is at least a plausible necessary condition, the assumption of other conjuncts:

AC. A collection of reasons to believe the individual conjuncts of a conjunction provides a reason to believe the conjunction only if they are reasons to believe each conjunct under the assumption that the other conjuncts are true.

The following cases may provide a feel for the intuition underlying this condition:

α It is 5:30
β. Your watch is broken.

The reason to believe α is that your watch says 5:30, and the reason to believe β is that the hands of your watch do not move. These reasons violate (AC), because the fact that your watch says 5:30 is not a reason to believe that it is 5:30 under the assumption that your watch is broken. Obviously, these reasons do not provide a reason to believe the conjunction $\alpha \wedge \beta$. This is a case which fails to satisfy (AC) because the evidential relevance of the reasons for one conjunct is rendered problematic by assuming the other conjunct. An example with a slightly different structure is when the assumption of other conjuncts implies that the reasons for a conjunct are false:

γ. Jones is good at mathematics
δ. Jones failed the maths exam.

The reason to believe γ is that Jones passed the maths exam, and the reason for δ is that Jones said he failed the maths exam. These reasons violate (AC) because, obviously, it is false that Jones passed the maths exam under the assumption that Jones failed the maths exam. That Jones passed the maths exam is therefore not a reason to believe that Jones is good at maths under the assumption that Jones failed the maths exam.

3. Reasons to believe (2) under the assumption that (1)

We now proceed to examine whether the reasons to believe (1) and the reasons to believe (2) support the conjunction (1) $\wedge$ (2). They do not support the conjunction if the reasons for either (1) or (2) fail to satisfy (AC). Although one might argue that the virtuous-wife preferences in de Sousa's example are less intuitively compelling when considered under the assumption that (2), we focus on reasons to believe (2) under the assumption that (1). We shall show that the most commonly offered reason for (2) does not support (2) under the assumption that (1).

3.1 Forcing money-pumps

The most common, and possibly strongest, reason in the literature to believe (2) is that if it did not hold we could be rationally exploited in a so-called money-pump.[6] The structure of the argument is very simple. Suppose Alice prefers x to y, and is indifferent between y and z, but does not prefer x to z. This constitutes a violation of PI-transitivity. Under the assumption of comparability it follows that if Alice does not prefer x to z she must either prefer z to x or be indifferent between them. In both cases it is easy to show how she may become a money-pump.

In the case where she prefers z to x (xPy, yIz and zPx), she will presumably be willing to pay a small sum of money m_1 to exchange y for x, and a small sum m_2 to exchange x for z. Furthermore, we can assume that she will be willing to exchange z for y if given a small amount of money m_3 such that $m_3 < m_1 + m_2$. Then if she starts out with x she will pay m_2 to get z, and once she has z she will want to switch to y if given m_3. In possession of y she will now pay m_1 to get x, finishing where she started but with less money. In the case where she is indifferent between z and x (xPy, yIz and zIx) she will, as before, be willing to pay a small sum of money m_1 to exchange y for x. But in this case she will need to be given money in two steps instead of one in order to be persuaded to switch where she is indifferent. She will have to receive a small amount of money m_4 to switch x for z and a small amount of money m_5 to switch z for y, where $m_4 + m_5 < m_1$. As before, she will then finish where she started but with less money.

Since there are no further cases under comparability that violate PI-transitivity, and the agent is exploited in both of them, this seems to be a strong argument for (2). But it is to be noted that the money-pump argument assumes comparability. We have pointed out that in order to satisfy (AC) and provide support for the conjunction (1) $\wedge$ (2), the money-pump argument must be a reason for (2) under the assumption of (1). The problem is that according to the comparability trilemma, (2) under the assumption of (1) implies that comparability is false. This means that comparability can no longer be assumed in the money-pump argument if it is to provide support for (1) $\wedge$ (2). But, as we shall show, the money-pump argument does not work unless comparability is assumed.

Under incomparability, if Alice prefers x to y, is indifferent between y and z, but does not prefer x to z, she can in addition to the cases above also violate PI-transitivity if she finds x and z incomparable. But the pump strategy used in the cases of preference and indifference does not work in this case. Let '#' denote the incomparability relation, defined as $x\#y$ iff $\neg(xPy) \wedge \neg(yPx) \wedge \neg(xIy)$. If Alice finds z and x incomparable (xPy, yIz and $z\#x$), she will

6. See, e.g. D. Davidson, J. McKinsey and P. Suppes, 'Outlines of a Formal Theory of Value, I', *Philosophy of Science*, 22 (1955), pp. 140–60, at p. 146; H. Raiffa, *Decision Analysis: Introductory Lectures on Choices Under Uncertainty* (Reading: Addison-Wesley, 1968), p. 78; S.O. Hansson, 'Money-Pumps, Self-Torturers and the Demons of Real Life', *Australasian Journal of Philosophy*, 71 (1993), pp. 476–85; W. Rabinowicz, 'Money Pump with Foresight', in M. Almeida (ed.), *Imperceptible Harms and Benefits* (Dordrecht: Kluwer, 2000), pp. 123–54.

as before presumably pay an amount of money to switch from y to x. She can presumably be rationally persuaded to switch z for y given a small sum of money. But a small sum of money might not persuade her to switch x for z if she finds these alternatives incomparable. Unlike the case where she was indifferent between x and z, where a small improvement of z necessarily made the improved z-alternative preferable to x and thus persuaded her to switch, an improved version of z might not be rationally preferred to x in the case where she finds x and z incomparable. In other words, when the alternatives are incomparable, there is no guarantee that the money needed to persuade Alice to choose one alternative rather than the other will be sufficiently small. If a sufficiently small amount of money does not make her prefer z to x, she is not rationally required to switch. Therefore the money-pump is blocked in this case.

So under incomparability, there are violations of PI-transitivity where the agent cannot be pumped for money. Thus the money-pump argument is not a reason to believe (2) under incomparability. Since (2) under the assumption of (1) implies incomparability, the money-pump argument is not a reason for (2) given the assumption of (1). Therefore (AC) implies that the money-pump argument does not provide a reason to believe (1) $\wedge$ (2).

Before closing this section we shall discuss two objections. First, you might object that the preferences xPy, yIz and $x\#z$ together make the agent vulnerable to something very similar to a money-pump.[7] Suppose Alice starts out with z. Since yIz, she should rationally be willing to trade z for y plus a small amount of money m_1. Further, since xPy, there is a sum $m_2 > m_1$ such that she should be willing to trade y and m_2 for x. Thus she ends up with x but with $m_2 - m_1$ less money. But since she does not prefer x to z, there seems to be something wrong with a set of preferences that rationally obliges her to pay in order to trade, in two steps, z for x. It might seem, then, that the incomparabilist could defend (2) by a version of the money-pump argument. Admittedly, such preferences would be a sign of irrationality if Alice also held that $zP(x + m_1 - m_2)$. She would then have swapped, in two steps, to an alternative she rates worse than the alternative she started with, which arguably is irrational. But it does not seem plausible that an agent who holds xPy, yIz and $z\#x$ is rationally committed to hold that $zP(x+m_1-m_2)$. Rather, if comparability is rejected, then it seems plausible that an agent who holds xPy, yIz and $z\#x$ is rationally permitted to hold that $z\#(x + m_1 - m_2)$. But it does not seem to be a sign of irrationality to swap z for $x + m_1 - m_2$ if one holds that $z\#(x + m_1 - m_2)$. Thus this version of the money-pump argument fails to show that PI-transitivity is rationally required.

Another possible objection is that one may replace (2) with the principle

2*. PI-transitivity* is rationally required

7. This objection was raised by an anonymous referee.

where PI-transitivity* is defined as

$$PI\text{-transitivity}^{\star}.\ \forall x \forall y \forall z((xPy \wedge yIz) \rightarrow (xPz \vee x\#z)).$$

Since (1), (2*) and (3) are also inconsistent, the small-improvement argument remains logically valid when premise (2) is replaced by (2*).[8] But replacing (2) by (2*) would render our argument above doubtful. This is because it hinges on the fact that the money-pump argument cannot support (2) under the assumption of (1), since an agent with violating preferences of the type xPy, yIz and $z\#x$ cannot be exploited in a money-pump. This, however, is irrelevant if one employs (2*) instead of (2), since preferences of this type do not violate (2*).

In response to this objection we argue that it is also problematic to support (2*) by a money-pump argument. For example, suppose an agent S violates PI-transitivity* with the preferences aIb, bIc and cPa. In order to exploit these violating preferences, one needs at some point to make S rationally required to swap b for a. As previously explained, the standard strategy is then to offer S a slightly improved version $a + m$, that is, a with a small sum of money m. But it does not follow that S is rationally required to prefer $a + m$ to b unless it is excluded that S is rationally permitted to hold that, for instance, $b\#(a + m)$. One could remedy this problem by assuming (2), but then nothing would have been gained by replacing (2) with (2*) in the small-improvement argument.

3.2 *Non-forcing money-pumps*

To the argument in § 3.1 one might object that there are two kinds of money-pumps, viz forcing pumps in which the agent is rationally required to switch in every step of the pump, and non-forcing pumps in which the agent is not forced but is merely rationally permitted to switch in each step of the pump. While we have shown that the forcing pump does not provide a reason to believe (2) under the assumption of (1), one could still argue that the non-forcing pump could provide such a reason.[9] In the case where Alice finds x and z incomparable (xPy, yIz and $z\#x$), we argued that a small amount of money would not necessarily persuade her to switch between x and z. But even though agents are not rationally required to switch between two incomparable alternatives it may well be the case that they are rationally permitted to do so. Thus under incomparability agents are susceptible to the non-forcing money-pump, and this would provide a reason to believe (2).

There is an easy way to dismiss this objection. Under PI-transitivity, the virtuous-wife preferences were $a\#b \wedge cPa \wedge \neg(cPb)$. According to these, she would have to have one of the

8. We would like to thank David Alm for suggesting this weakened version, although he did not mention it in this context.

9. See, e.g. J. Broome, *Ethics Out of Economics* (Cambridge UP, 1999), p. 156; M. Peterson, 'Parity, Clumpiness and Rational Choice', *Utilitas*, 19 (2007), pp. 505–13.

following preferences:

(I) $a\#b \wedge b\#c \wedge cPa$
(II) $a\#b \wedge bIc \wedge cPa$
(III) $a\#b \wedge bPc \wedge cPa$.

If the virtuous wife is rationally permitted to switch both between the alternatives about which she feels indifferent and the ones she finds incomparable, and she is willing to pay a small amount to switch to an alternative she prefers, then she is susceptible to the non-forcing money-pump in each of the cases (I)–(III). Thus when the reason to accept (2) is based on a non-forcing money-pump, this same reason implies that the virtuous-wife preferences under PI-transitivity are irrational. This, of course, contradicts the assumption of (1). A non-forcing pump thus cannot be a valid reason to believe (2) under the assumption of (1). Therefore (AC) implies that the non-forcing pump does not provide a reason to believe (1) $\wedge$ (2).

3.3 *Self-evidence*

Another reason to believe (2) could be that one finds it self-evident. To hold that (2) is self-evident is, however, problematic in face of strong counter-arguments, like the argument from unnoticeable differences. Suppose Alice is indifferent between c_0, a cup of coffee with no sugar, and c_1, a cup of coffee with one lump of sugar. Furthermore, suppose she prefers c_2, a third cup with two lumps of sugar, to c_0, but is indifferent between c_1 and c_2. These preferences violate PI-transitivity, but it is hardly self-evident that Alice is irrational in this case. For she may be indifferent between c_0 and c_1 because she cannot taste any difference between coffee with no sugar and coffee with merely one lump of sugar. Similarly, she may be indifferent between c_1 and c_2 because she cannot taste the difference between coffee with one lump of sugar and coffee with two lumps. She might, however, be able to taste the difference between coffee with two lumps and coffee with no sugar at all, and therefore prefer c_2 to c_0. This does not seem to be a self-evident case of irrationality.[10]

4. Conclusion

The small-improvement argument is considered the most powerful argument for incomparability (see, e.g. Chang, 'Introduction', p. 23). Its soundness is usually taken for granted; seldom

10. Similar arguments against transitive preferences based on unnoticeable differences have been put forward many times in the literature. See, e.g. R. D. Luce, 'Semiorders and a Theory of Utility Discrimination', *Econometrica*, 24 (1956), pp. 178–91; W. Quinn, 'The Puzzle of the Self-Torturer', *Philosophical Studies*, 59 (1990), pp. 79–90; Hansson, 'Money-Pumps, Self-Torturers and the Demons of Real Life'; A. Norcross, 'Comparing Harms: Headaches and Human Lives', *Philosophy and Public Affairs*, 26 (1997), pp. 135–67.

has it been subjected to critical scrutiny. In this paper we have argued that the argument suffers from a critical flaw. We have argued that it fails to establish incomparability, since there is a conflict between the reasons in support of the premises. We have shown that because of the conflict, the reasons in support of the individual premises do not provide a reason to believe the conclusion of the argument. This does not entail that comparability is rationally required, only that the small-improvement argument fails to establish incomparability.[11]

11. The authors (who are equal contributors) would like to thank David Alm, John Cantwell, Sven Ove Hansson, Martin Peterson, Wlodek Rabinowicz, and an anonymous referee for valuable comments on earlier versions.