IT SIMPLY DOES NOT ADD UP: TROUBLE WITH OVERALL SIMILARITY*

Comparative overall similarity lies at the basis of a lot of recent metaphysics and epistemology. It is a poor foundation. Overall similarity is supposed to be an aggregate of similarities and differences in various respects. But there is no good way of combining them all.

I. SIMILARITIES AND DIFFERENCES

Similarity is relative: things are similar in one respect but different in another. And it is comparative: some things are more similar to each other, in a given respect, than are other things. This much is quite straightforward. The idea behind comparative overall similarity has been that some things might be more similar than other things—but simply so, not in any particular respect—somehow as a result of similarities and differences in *several* respects. This is not at all straightforward, because overall similarity is supposed to be some sort of aggregate. It is supposed to be the result of adding up similarities or weighing them against differences or combining them in another way.

Comparative overall similarity, I shall argue, does not meet the demands that philosophers make of it. At root, the trouble is that, in general, greater similarity in one respect will not make up for less similarity in another respect. For this reason, we will see, there can be no combining of the various similarities and differences of things into useful comparisons of overall similarity. Before going any further, though, let us stop and see what depends on this.

As a first example, take the question, "How could things have been different?" One theory has it that people and other ordinary things have counterparts in other possible worlds. That is how, for example, you could have grown up in a traveling circus: you could have done so because some counterpart of yours did in fact grow up in one. Now, counterparthood is a matter of overall similarity. Your counterpart is someone who resembles you more closely, overall, than do others in his possible world.[1]

* For discussion, I thank Johan van Benthem, Ruth Chang, Max Cresswell, Matti Eklund, Steve Kuhn, Barry Loewer, Ray Martin, Dan Moller, Georges Rey, Denis Robinson, Michael Slote, Mark Wunderlich, and everybody at Rosewind Farm.

[1] David Lewis, "Counterpart Theory and Quantified Modal Logic," this JOURNAL, LXV, 5 (March 1968): 113–26.

0022-362X/10/0709/469–490

Or, consider how things undergo change. This is tied up with overall similarity if, as many think, ordinary objects persist from one time to the next by having a series of temporal stages, and change by having stages that are dissimilar. It is thought that what conjoins the successive stages—that is, what makes them parts of one temporally extended thing—is not only their causal connectedness but also their overall similarity.[2] This explains why a pile of decayed planks, discarded in the process of preserving the ship of Theseus, is not itself the original ship, although the required causal connection is there. The pile's stages do not resemble earlier stages of the original ship as closely, overall, as do the stages of the preserved ship.[3]

There is yet more work cut out for overall similarity. Some say that a counterfactual conditional sentence is true if some possible world in which both the antecedent and consequent are true is more similar, overall, to the world of evaluation, than is any world in which the antecedent is true but the consequent is false.[4] Accounts of causation,[5] the direction of time,[6] knowledge,[7] and intentionality[8] in turn depend on counterfactuals. Verisimilitude, or comparative likeness of false theories to the truth, has been thought to be an aggregate of likenesses in respect of truth and of content.[9]

Nelson Goodman once complained of the many ways in which he thought similarities have failed philosophy. Similarity judgments, he observed, require not only the selection of relevant properties but also the weighting of their importance. Importance is a volatile matter,

[2] Lewis writes that temporal parts "are united as much by relations of causal dependence as by qualitative similarity" in *On the Plurality of Worlds* (New York: Blackwell, 1986), p. 218. According to Robert Nozick, temporal identity is a matter of "not merely the degree of causal connection, but also the qualitative connection of what is connected, as this is judged by some weighting of dimensions and features in a similarity metric." See his *Philosophical Explanations* (Cambridge: Harvard, 1981), p. 37.

[3] The ancient Athenians are said to have preserved the ship in which Theseus returned by taking away old planks as they decayed and replacing them with new ones. I assume that the pile of discarded planks is a continuer that, but for the presence of the preserved ship, might itself have been the original ship of Theseus.

[4] See Lewis, *Counterfactuals* (Cambridge: Harvard, 1973); and Robert Stalnaker, "A Theory of Conditionals," in Nicholas Rescher, ed., *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, 2 (Oxford: Blackwell, 1968), pp. 98–112.

[5] Lewis, "Causation," this JOURNAL, LXX, 17 (Oct. 11, 1973): 556–67.

[6] Lewis, "Counterfactual Dependence and Time's Arrow," *Noûs*, XIII, 4 (November 1979): 455–76.

[7] Fred Dretske, "Conclusive Reasons," *Australasian Journal of Philosophy*, XLIX, 1 (May 1971): 1–22; Nozick, *op. cit.*, p. 321.

[8] Jerry Fodor, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge: MIT, 1987).

[9] Risto Hilpinen, "Approximate Truth and Truthlikeness," in M. Przelecki, K. Szaniawski, and R. Wojcicki, eds., *Formal Methods in the Methodology of the Empirical Sciences* (Dordrecht: Reidel, 1976), pp. 19–42.

however, varying from one context to the next; so, he argued, it cannot support the distinctions that philosophers would rest on it.[10] Many specific difficulties have since arisen with philosophy built on comparative overall similarity, despite such misgivings.[11] But let us return to the neighborhood of the observation about weighting properties. There lurks real trouble.

The trouble comes to light when we ask just how to combine similarities and differences in various respects. In fact, no one has had any real idea! There are only metaphors, however promising these might seem. David Lewis draws a comforting analogy to vector addition, with talk of "resultant" similarities.[12] Robert Nozick conjures an image of the judicious balancing of similarities against differences when he speaks of the "weights" of relevant properties.[13] Everyone seems to picture a space, framed by the dimensions of comparison, in which similar things are close together and dissimilar things are far apart.

I shall argue that all these metaphors are false. We cannot add up similarities or weigh them against differences. Nor can we combine them in any other way. Goodman was right to be skeptical. No useful comparisons of overall similarity will result.

My first main point in support of this conclusion will be that there really does have to be a balance of similarities if there are to be useful overall similarities. Greater similarity in one respect will have to make up for less similarity in another respect. Section II asks how to combine similarities and answers in terms of supervenience. Then it considers several ways of combining similarities without weighing them and shows that each fails a reasonable requirement. One of these requirements is that there should not be a *dictator*, that is, a critical respect of similarity that excessively influences overall similarities.

The next main point is that there is no balance of similarities. Section III illustrates the idea of a balance of dimensions with a spatial

[10] Nelson Goodman, "Seven Strictures on Similarity," in Goodman, ed., *Problems and Projects* (Indianapolis: Bobbs-Merrill, 1972), pp. 437–46.

[11] For difficulties with Lewis's treatment of *de re* modality, see Fred Feldman, "Counterparts," this JOURNAL, LXVIII, 13 (July 1, 1971): 406–09; Allen Hazen, "Counterpart-Theoretic Semantics for Modal Logic," this JOURNAL, LXXVI, 6 (June 1979): 319–38; and Michael Fara and Timothy Williamson, "Counterparts and Actuality," *Mind*, CXIV, 453 (January 2005): 1–30. For difficulties with the interpretation of counterfactuals in relation to comparative overall similarities, see Jonathan Bennett, "Counterfactuals and Possible Worlds," *Canadian Journal of Philosophy*, IV, 2 (December 1974): 381–402; Kit Fine, "Critical Notice, *Counterfactuals*. By D. Lewis, Oxford: Blackwell, 1973," *Mind*, LXXXIV, 335 (July 1975): 451–58; and Paul Horwich, *Asymmetries in Time* (Cambridge: MIT, 1987).

[12] Lewis, "Counterpart Theory."

[13] Nozick, *op. cit.*, p. 33.

example. Then it argues that, in general, greater similarity in one respect will not make up for less similarity in another. Similarities are incommensurable when they are merely ordinal and we cannot meaningfully say how much more or less alike things are, but they also are incommensurable when they are cardinal, and we can.

Sections II and III avoid technicality in order to develop an intuitive sense of the trouble with overall similarity. As a result, the argument is less rigorous than you might wish. It considers only a few representative ways of combining similarities. You might wonder whether some other way is better. Also, because the discussion remains informal, there is room for unwanted assumptions to slip in unnoticed. You might wonder whether some such interloper is the real troublemaker.

Section IV gives skepticism about overall similarity a precise sense and a completely rigorous justification. After making matters from the previous sections technically explicit, a reinterpretation of Kenneth Arrow's theorem of social choice shows that a relation of comparative overall similarity must always have a dictator if it supervenes on similarities in several respects.[14]

If all this is so, why then has overall similarity seemed such a promising foundation for philosophy? Perhaps this is because we imagine that our everyday thinking depends on it. For instance, you might have thought that we implicitly compare overall similarities when sorting things into categories. Since we often agree among ourselves about what is what, it is perhaps only natural to suppose that, in categorization, we latch onto genuine overall similarities and differences of things. Certainly, it feels as if we are onto something real.

However, there is another explanation for our agreement. Presumably, there is an innate psychological basis for categorization that does not vary greatly across our species. We are bound to find ourselves agreeing quite a bit, given that we all categorize in much the same way.

Whether the psychological mechanisms of categorization reveal genuine overall similarities is another matter, though, and they need not do so at all. Overall similarities are involved in categorization according to one influential proposal.[15] But things are similar or different, in the relevant sense, only indirectly, through the mediation of mental representations that pick out some features as salient. Such

[14] The relevance of social-choice theory for this topic has gone largely unremarked, but Williamson touches on it in "First-Order Logics for Comparative Similarity," *Notre Dame Journal of Formal Logic*, XXIX, 4 (Fall 1988): 457–81.

[15] See Amos Tversky, "Features of Similarity," *Psychological Review*, LXXXIV, 4 (July 1977): 327–52.

mediated similarities might play a role in our everyday thinking even if things are not in themselves similar or different, independently of how we represent them to ourselves.

Moreover, categorization might not depend on any overall similarities, whether or not they are mediated by representations. When you judge someone to be drunk, for instance, because he has jumped fully clothed into a swimming pool, you need not have done so by establishing an overall likeness to other drunks. Instead, you might have *explained* what happened, drawing on a more-or-less implicit theory of human behavior and the effects on it of too many drinks.[16] However categorization feels from the inside, so to speak, it need not rely on relations of overall similarity among things. Perhaps it is our intuitive sense of similarity and difference that depends on our ability to categorize and not the other way around.

II. THERE MUST BE A BALANCE

I now shall argue that the metaphor of weighing similarities is to be taken quite seriously. Greater similarity in one respect will have to make up for less similarity in another, if there are to be useful overall similarities.

To get started, we will need some understanding of what it is to combine similarities. It should be compatible with adding them up and weighing them against differences as well as with other suggestive metaphors for what is involved: weaving similarities together, or what have you. Fortunately, we can make do with very little understanding. We will proceed with the idea that overall similarities supervene on particular similarities, comparisons of overall similarity being the same whenever all comparisons of particular similarity are the same. In section IV, we will have a completely precise formulation. Meanwhile, an example will illustrate.

Imagine looking over the preserved ship of Theseus on a fine day. In a corner somewhere, you notice the pile of decayed planks that have been removed over the years. You judge, perhaps, that the preserved ship resembles the original ship of Theseus more closely, overall, than the pile does. Soon afterward, back for another look, you find both ship and pile to be just as you left them. Neither has become in any way more like the original ship of Theseus, and neither has become less like it. What supervenience requires is that the overall

[16] Compare Gregory L. Murphy and Douglas L. Medin, "The Role of Theories in Conceptual Coherence," *Psychological Review*, XCII, 3 (July 1985): 289–316, see p. 295. For further discussion, see Ulrike Hahn and Michael Ramscar, eds., *Similarity and Categorization* (New York: Oxford, 2001).

comparison remains unchanged. On both occasions, your judgment ought to be the same.

Do not be misled about the supervenience of overall similarities by the slack between them and their expressions in thought and language. Lewis used the idea that counterfactual conditionals are contextual in order to explain how someone can meaningfully assert either of a contrary pair. Suppose that in an ordinary conversation someone claims:

> If Caesar had been in command [in the Korean war], he would have used the atom bomb.

As soon as the person has spoken, Lewis argues, we rush to help him to have spoken the truth. We evaluate his utterance by using a relation of overall similarity among possible worlds that attaches greater importance to similarity in respect of the knowledge of weapons common to commanders in Korea. With this accommodation, worlds in which Caesar has a modern knowledge of weapons are more similar to our actual world, overall, and the speaker's utterance is true. If, on the other hand, he says:

> If Caesar had been in command, he would have used catapults,

we instead attach greater importance to historical knowledge, and this becomes the true utterance. According to Lewis, we evaluate differently in the two cases because we evaluate with different relations of overall similarity.[17] If he is right about this, the truth of a counterfactual that reveals the overall similarity of worlds need not supervene on their particular similarities, even though, we should suppose, overall similarity itself does supervene.

Consider now some ways of combining similarities without weighing them against differences. Comparing overall similarities is completely straightforward in some rather special cases. We will judge that one person resembles you as closely as another does, overall, if he *dominates*, which is to say that he resembles you as closely in every respect. The overall comparison is easy because there is no need for weighing. It does not depend on how much similarity in the one respect goes for how much dissimilarity in the other—nor on whether there are any such rates of exchange at all.

In cases of dominance, comparative overall similarity is just as transparent and dependable as familiar mathematical notions of similarity, such as congruence among geometrical figures and isomorphism among structures. If overall comparisons can be made

[17] Lewis, *Counterfactuals*, p. 67.

only in such cases, however, then comparative overall similarity will not be of much use in philosophy.

For one thing, you will lack counterparts. Take your spitting image. He walks like you, and he talks like you. He resembles you as closely as can be, except for this: he grew up in a traveling circus. It is hard to imagine a more likely candidate, and yet he will not qualify as your counterpart—not if he lives in any normal possible world and similarity in respect of origins is relevant. Normally, there will be other candidates whose origins are more like yours, and your spitting image will not stand a chance against even the least likely of these. He fails to outdo them in overall likeness to you because, being in the one way less like you than they are, he does not dominate them. It is likely that none of them dominates *him*, either, but we cannot allow that to qualify him as your counterpart. That would make it too easy to qualify. You would wind up with too many counterparts.

In some special cases, then, all comparisons of particular similarity align. One candidate dominates the other and is more similar to you, overall. But it will not do for these to be the only cases in which overall comparisons are available. Something will have to close the comparability gaps.[18]

Your spitting image does not quite dominate the other candidates, but he is not far off. One might suppose that a candidate resembles you as closely, overall, if he *nearly* dominates, that is, if he resembles you as closely in nearly every respect.

This supposition will close comparability gaps because, as with full dominance, what counts is just the proportion of respects of greater similarity. No weighing is called for. Also, it will account for many intuitive similarity judgments. Still, it is not a suitable foundation for philosophy. One complication is the vagueness of "nearly every": just how close to complete agreement among the various dimensions must we come in order for composition to occur? The real problem, though, is that there are incoherent results.

Consider a case patterned on Condorcet's "paradox" of voting. Three candidates compete to be your counterpart. In one respect, Alfie resembles you more closely than Bozo does, and Bozo resembles you more closely than Coco does; in another respect, Bozo is most like you, followed by Coco and then Alfie; and, in some third respect, Coco is most like you, followed by Alfie and then Bozo. Let these dimensions be the only relevant ones:[19]

[18] The standard assumption is that comparative similarities are *connected*: one of any two things resembles you at least as closely as the other one does.

[19] Or, let the candidates resemble you equally in all other respects.

	Increasing Resemblance to You ←
First Respect:	Alfie, Bozo, Coco
Second Respect:	Bozo, Coco, Alfie
Third Respect:	Coco, Alfie, Bozo

Alfie resembles you more closely than Bozo does in every respect but one; so, no matter where the threshold for composition is set, short of full dominance, Alfie nearly dominates Bozo, and Alfie resembles you more closely than Bozo does, overall. Likewise, Bozo resembles you more closely than Coco does, overall. Coherence requires the comparisons to be transitive;[20] in particular, it requires that Alfie resembles you more closely than Coco does, overall; but, by this reckoning, he does not. On the contrary, since Coco resembles you more closely than Alfie does in the second and third respects, the comparison between them comes out the wrong way around. To the extent that our intuitive similarity judgments track near dominance, in a range of possible cases, they are very much the worse for it, because they are incoherent.[21]

For another try at combining similarities, suppose we somehow rank the respects of similarity in order of their importance. Then we can obtain overall comparisons by alphabetic composition, allowing each successive relation of comparative similarity in some respect to refine the result of putting together its predecessors, by breaking ties. This closes comparability gaps, and there is no weighing. For example, more similarity in a lesser respect will never make up for less similarity in a more important one.

You might wonder how to rank the respects of similarity. One idea has been that their relative importance is revealed by our counterfactual judgments. To see how, consider a well-known objection to the theory that the truth of a counterfactual depends on the truth of its consequent in all most-similar antecedent worlds.[22] It easily can be imagined that during the nuclear alert of 1973:

> If Nixon had pressed the button, there would have been a nuclear holocaust.

This is puzzling if it requires that some possible world in which he pressed the button and set off a holocaust is more like our actual world,

[20] The standard assumption is that comparative similarities are *weak orders*—that is, connected and transitive. See Lewis, *Counterfactuals*, p. 48.

[21] Someone might try to save the idea of comparative overall similarities by saying that sometimes composition does not occur and that this is such a case. And he might say that it also does not occur with the profiles in the proof of the theorem in section IV. But saying so would be unwise. There is nothing funny about this case or about the possibilities that those profiles represent. We may expect composition to occur here, if it ever does.

[22] See Fine, *op. cit.*, p. 452; and Bennett, *op. cit.*

overall, than is any world in which he pressed the button but no holocaust followed. You would have thought that, for any given world in which life as we know it was wiped from the face of the Earth, there always will be another, more like our actual world, in which Nixon pressed the button and nothing very much happened. Some miraculous little glitch saved the day. His moment of truth came and went, and he sat there trembling for a good long time. After he pulled himself back together, though, life went on pretty much as it actually did.

Lewis responded to this example by arguing that there are many relations of overall similarity, corresponding to different rankings of the various dimensions, and that the similarities implicit in our counterfactual judgments do not need to be the same ones that our explicit similarity judgments reveal.[23] He then proposed a ranking that he took to be correct insofar as it makes the right conditionals true. In the same vein, Nozick argued that we can discover the ordering of dimensions in our identity judgments.[24] They both had in mind what we might call *revealed* overall similarities, the weights or priorities of dimensions being implicit in which counterfactuals and identities we take to be true and which false.[25] This is how we might hope to rank the respects of similarity. If *priorities* are what we discover, not weights, then the revealed similarities might be alphabetic orders.

However, alphabetic orders are unsuitable no matter how the dimensions are ranked. This is because they are:

> *Dictatorial.* There is a critical respect of similarity such that whenever some things are more similar in this one respect than are some others, their overall similarity is at least as great.

[23] Lewis, "Counterfactual Dependence."

[24] Nozick, *Philosophical Explanations*, pp. 34–35.

[25] Given that the revealed similarities strikingly disagree with our ordinary sense of similarity and difference, you may wonder how we could ever have become attuned to them. As Horwich writes in *Asymmetries in Time* (p. 172):

> [T]hese criteria of similarity might well engender the right result in each case. However, it seems to me problematic that they have no pre-theoretic plausibility and are derived solely from the need to make certain conditionals come out true and others false. For it is quite mysterious why we should have evolved such a baroque notion of counterfactual dependence. Why did we not, for example, base our concept of counterfactual dependence on our ordinary notion of overall similarity?

A further problem with Lewis's ranking is that it in fact does not engender the right results. See Adam Elga, "Statistical Mechanics and the Asymmetry of Counterfactual Dependence," *Philosophy of Science*, LXVIII, 3, Supplement: Proceedings of the 2000 Biennial Meeting of the Philosophy of Science Association. Part I: Contributed Papers (September 2001): S313–24; and Barry Loewer, "Counterfactuals and the Second Law," in Huw Price and Richard Corry, eds., *Causation, Physics and the Constitution of Reality* (New York: Oxford, 2007), pp. 293–326.

The critical respect is the one with first priority.[26]

Dictatorship is pernicious. Lewis warned of its excesses: "respects of similarity and difference trade off. If we try too hard for exact similarity…in one respect, we will get excessive differences in some other respect."[27] The problem with dictatorship is precisely that it enforces trying too hard for exact similarity in the critical respect. In the metaphor of balancing, there is no judicious weighing of similarities against differences. Greater similarity in the critical respect simply locks up the balance, preventing it from tipping the other way no matter which differences pile up on the other side. In terms of a similarity space, as things become closer in the critical dimension, they can become only closer overall, no matter how distant they become in other dimensions. Overall similarities are the "resultant" of a multitude of particular similarities and differences only in a tortured sense of the word.

Dictatorship not only offends against the very idea of overall similarity. It also compromises philosophical theories that build on it. Take for instance Lewis's counterpart-theoretic account of *de re* modality. We should expect it not to be committed to the doctrine that things have some of their attributes essentially, independently of how they are specified. Lewis thought that it was.[28] But dictatorship imposes essentialism.

Under dictatorship, there is some critical respect of comparison that trumps the others. Consider any one of your candidate counterparts who is even slightly unlike you in this respect. He will not qualify as your counterpart—not if he shares his world with another candidate who is more like you in the critical respect; under dictatorship, this other candidate must resemble you at least as closely overall.[29] To qualify as your counterpart, a candidate must resemble you in the critical respect at least as closely as his competitors do. Your counterparts are bound to resemble you, in this respect, as closely as can be.

[26] Alphabetic composition actually produces a more severe dictatorship in which things more similar in the critical respect are not merely as similar overall but indeed are more so. I consider the milder sort here because this is the one that returns in section IV.

[27] Lewis, *Counterfactuals*, p. 9.

[28] Lewis writes that "a suitable context might deliver an antiessentialist counterpart relation—one on which anything is a counterpart of anything, and nothing has any essence worth mentioning." See "Postscripts to 'Counterpart Theory and Quantified Modal Logic'," in *Philosophical Papers*, vol. 1 (New York: Oxford, 1983), p. 43. Essentialism is more natural in other accounts of *de re* modality, such as Saul Kripke's in *Naming and Necessity* (Cambridge: Harvard, 1980).

[29] I assume that there is at most one of you in the world in question: you have at most a single counterpart there, who resembles you strictly more closely than all other candidates do.

There is related trouble with counterfactuals. Following Lewis, let us accommodate the claim that Caesar would have used the atomic bomb, by letting similarity in respect of modern knowledge have first importance. Under dictatorship, this respect of similarity is critical. Now consider a possible world in which Caesar's knowledge was completely modern. This world presumably is more similar to our actual world, in the critical respect, than is any world in which his knowledge was not completely modern; so, in the context created by the speaker's claim, it is as similar to our actual world, overall. That is, in this context, no world with an incompletely modernized Caesar is more similar to our actual world, overall, than is the world with the completely modernized Caesar.

Now there is a problem. Intuitively, you can agree with the speaker that Caesar would have used the bomb if he had been in command in Korea, while thinking to yourself that he also, as the need arose, would have used catapults, pila, and other kinds of weapons—even ones that have been forgotten over the millennia:

> If Caesar had been in command, he also would have used long-forgotten weapons.[30]

Under dictatorship, though, you would be mistaken. The truth of this sentence requires that some world in which Caesar was in command and used long-forgotten weapons is more similar to our actual world, overall, than is any world in which he was in command but did not use them. A Caesar who used long-forgotten weapons, though, is an incompletely modernized one. And, as we have seen, when the speaker's utterance is accommodated, no such world is more similar to our actual world, overall, than is the world with the completely modernized Caesar, who did not use long-forgotten weapons because he did not know the first thing about them.

This shows that, just as Lewis warned, trying too hard for exact similarity in one respect can only lead to excessive differences in other respects. Even if one dimension of similarity is most important, the other dimensions still should count as well. There has to be a balance.

We have considered several ways of combining similarities without weighing them against differences, and we have seen that none has a satisfactory result. Composition in the case of dominance is good as far as it goes but leaves too many comparability gaps. Composition in

[30] Be sure to keep this thought to yourself, or you will spoil my example! As soon as you speak up, your accommodating partners in conversation will see to it that you have spoken the truth, by evaluating your utterance in another context, using a different relation of overall similarity.

the case of near dominance fills some of the gaps but has incoherent results. Alphabetic composition imposes dictatorship. There are other ways, but, as we will see in section IV, they are no better. Suitable comparative overall similarities will result, if at all, on the balance of similarities.

III. BUT THERE IS NO BALANCE

I now shall argue that, in general, greater similarity in one respect will not make up for less similarity in another.

It is instructive to contrast similarities with spatial dimensions. Suppose that one person stands closer to you than someone else does. Let him take a single step to the north or south. How far to the east or west should he then move if his relative distance from you is to end up the same as it was to begin with? There is an obvious answer to this question. Any change will do that keeps him on the relevant indifference curve, which, in this case, has a particularly simple form: it is the circle around you that is defined by his starting position. Spatial dimensions are commensurable. In a range of cases, a change in one dimension will make up for a change in another.

We might conceive of corresponding spatial similarities. The overall spatial similarity of two locations, we might say, varies inversely with the great-circle distance between them. It is a function of their similarities in respect of longitude and latitude, which likewise depend on differences in these dimensions. Such spatial similarities are commensurable because the underlying spatial dimensions are commensurable. They inherit their indifference curves from them.

However, the spatial analogy is false. John Maynard Keynes took similarities as an example in making a related point about probabilities:

> [A] book bound in blue morocco is more like a book bound in red morocco than if it were bound in blue calf; and a book bound in red calf is more like the book in red morocco than if it were in blue calf. But there may be no comparison between the degree of similarity which exists between books bound in red morocco and blue morocco, and that which exists between books bound in red morocco and red calf.[31]

The point I take from this is that there is no trading of similarities in respect of the color and the kind of leather of a binding. A book bound in blue morocco bears some overall likeness to a book in red morocco. You can decrease this likeness by changing from morocco to calf, while keeping the color the same. But you cannot regain the original overall likeness to the book in red morocco by subsequently changing the color

[31] John Maynard Keynes, *A Treatise on Probability* (London: Macmillan, 1921), p. 36.

of the calf binding from blue to red. More similarity in respect of color will not make up for less similarity in respect of the kind of leather.

The example involves a dimension that is perhaps merely ordinal. In respect of the kind of leather, the book in blue morocco is more similar to the book in red morocco than is the book in blue calf, but there might be no saying how *much* more similar it is. Perhaps bindings of the same leather are similar in this respect, while those of another sort are different, and that is all there is to it. However, it seems that, in general, there also is no balancing of similarities measured on a cardinal scale.

Take, for instance, similarities in respect of weight and temperature. They might be cardinal, since the underlying quantities are. Let one person resemble you more closely, overall, than someone else does. And let him become a bit less like you in respect of his weight, by gaining a little. Now answer these questions: How much warmer or cooler should he become to restore the original overall comparison? How much more similar in respect of his height? What about his income or his wisdom or hairstyle? That there might be factual answers to these questions is hard to believe.

You might wonder whether the apparent incommensurability is really just a matter of vagueness. True, there is no saying exactly how much more similarity in one respect can be exchanged for less similarity in another, but there might be a rough rate of exchange even so. Indeed, that is just what we should expect, under the assumption that identities, *de re* modal claims, and counterfactuals reveal overall similarities. Normally, what we say is a bit vague. When we utter these sentences, there remain in play several relations of comparative overall similarity, each striking its own balance among similarities and differences. Each relation has a claim to be the right one for the interpretation of what we have said, but none has an exclusive claim. Taken separately, these relations might embody precise rates of exchange among similarities. Taken collectively, they embody none. At best, the context supports rough rates of exchange.

Lewis made a virtue of this vagueness: "[C]omparative similarity is not ill-understood. It is vague—very vague—in a well-understood way. Therefore it is just the sort of primitive that we must use to give a correct analysis of something that is itself undeniably vague."[32] It is the contextual resolution of vagueness that enables him to explain how the contrary counterfactuals about Caesar in Korea can both hold true, each in the context arising from its own utterance.

[32] Lewis, *Counterfactuals*, p. 91.

The problem is that, in general, there do not appear to be rough rates of exchange any more than there are precise ones. Were we aware of any, we should be able to say at least something about them. This we could do by using suitably vague language, as when, before the haggling begins, we can express our rough sense of what things are worth by saying that "only a little" of one will be needed in exchange for a given quantity of another or that, on the contrary, "quite a lot" or "a vast amount" will be needed, as the case may be. When someone has become less like you in respect of his weight, though, we cannot say that he will need to become "only a little" or "quite a lot" or "vastly" more similar in respect of his temperature in order to regain his earlier overall likeness to you. Nor does it seem to be ignorance about the case that keeps us from saying—what could we possibly be missing? As far as we can tell, there are no rates of exchange here.

David Wiggins wrote that values are incommensurable when "there is no general way in which [they] trade off."[33] Similarities seem to be more radically incommensurable than this. There is no general formula for the expression of rates of exchange, such as the circular indifference curves in the case of spatial dimensions. Similarities do not seem to trade off even in highly particular ways, with rates of exchange varying from case to case in complicated ways that defy general description.

Sometimes, perhaps, we should not expect to discover rates of exchange but may make them up to suit ourselves. Consider a speedster built with some salvaged parts. Could it really be "Little Bastard," the very car that James Dean wrecked? If not, how many more original parts would there have to be for the reconstruction to be authentic? How much more causal continuity with the original car would there have to be? Perhaps these are questions for car buffs and their lawyers to settle to their own satisfaction. If this means weighting or prioritizing dimensions and stipulating thresholds for authenticity, perhaps the weights, priorities, and thresholds are theirs to attach and stipulate as they see fit. It is up to them to make up the fact of whether this is "Little Bastard."

But identities, *de re* modal possibilities, and counterfactuals cannot in general depend on made-up similarities.[34] I take it that you are

[33] David Wiggins, "Incommensurability: Four Proposals," in Ruth Chang, ed., *Incommensurability, Incomparability, and Practical Reason* (Cambridge: Harvard, 1997), p. 59.

[34] Even vehicle identities do not, at least for legal purposes. Instead, they depend on the possession of data plates. Basing them on stipulated similarities would not be better. "Little Bastard" then might turn out to be a car, having a temporal boundary coinciding with that of the speedster, but it also might not. Depending on what car buffs and their

unwilling to think that whether something is *you* can be a matter of more or less arbitrary decision or stipulation.[35] It is no easier to accept that some ruling about dimensions, weights, and priorities determines, for example, whether you could have been taller than you are or whether, had it been scratched, the match would have lit. These are not matters that we may settle to suit ourselves.

Admittedly, whether there are rates at which any given dimensions trade off will not be decided in the way that I have approached the matter here, by reflection on how we think and speak. It is an empirical question that, despite all I have said, remains open in the overwhelming majority of cases. In light of this, one might like to think of commensurability as a regulative ideal that guides us toward such rates of exchange as there are to be discovered. Sometimes, indeed, there are surprises: just over a century ago, few could have imagined that spatial and temporal dimensions might be commensurable, but now we know that the temporal order of events is relative to inertial frames and that, as well as spatial indifference curves, there are spatiotemporal ones.[36] Encouraged, one might hold out hope that similarities in different respects will turn out to be commensurable after all.

Time will tell. Meanwhile, the burden will remain on us to discover rates of exchange among similarities for each case separately. For my part, I do not expect that there is much progress to be made in this direction. That is just a hunch, but it also is anybody's hunch that a lot of them await discovery—so many that the idea of a balance of similarities will turn out to be realistic after all. Over the years, a great deal has been built on the notion of comparative overall similarity. The result is an impressive edifice covering large parts of metaphysics and epistemology. Its foundation is about as good as this second hunch.

IV. SIMILARITIES REALLY DO NOT ADD UP

I have argued that there is no good way of combining similarities and differences into useful comparisons of overall similarity. The discussion

---

lawyers decide, and which parts mechanics swap out, "Little Bastard" might come to another sudden end—not with a bang this time but almost imperceptibly when, with the removal of one too many of the original parts, the reconstructed speedster slips below the stipulated threshold for authenticity. In this case, "Little Bastard" will turn out not to have been a car at all but merely an initial temporal part of one, a funny sort of thing like one of Eli Hirsch's "incars" and "outcars" (see "Physical Identity," *The Philosophical Review*, LXXXV, 3 (July 1976): 357–89). This is puzzling: "Little Bastard" was a car if anything ever was.

[35] As Nozick points out in *Philosophical Explanations*, p. 34.

[36] I am told that, in relativity theory, points in space-time at a fixed interval from any given point describe a hyperbola.

was informal, though, and its conclusion remained less than fully secure. I promised a rigorous argument. The first part will make matters from the previous sections technically explicit. Then, a reinterpretation and slight generalization of Arrow's theorem of social choice will show that some respect of similarity always must be a dictator, if comparative overall similarity supervenes on similarities in several respects.

IV.1. Similarities. Comparative similarity is fundamentally a matter of two pairs of things: b resembles b^* as closely as a resembles a^*.[37] The trouble with overall similarity manifests itself in the binary relations that result when b^* and a^* are the same thing—you, for example. For simplicity's sake, we will continue with such relations and with examples having to do with counterparts: $a\mathrm{S}b$ will mean that b resembles you as closely as a does.

We will assume that these relations are *weak orders*:[38]

Connected. For every a and b, either $a\mathrm{S}b$ or $b\mathrm{S}a$;

Transitive. For every a, b, and c, if $a\mathrm{S}b$ and $b\mathrm{S}c$, then $a\mathrm{S}c$.

Connectedness makes the notion of a maximal overall similarity useful (compare the discussion of your spitting image in section II). Given connectedness, coherence requires transitivity as well.

IV.2. Similarity Profiles. These are representations of the similarities and dissimilarities from which comparative overall similarities have been thought to result. One profile concerns your candidate counterparts in one possible world; another concerns the candidates in another world. The domains of profiles may overlap, but they need not do so.[39]

Profiles represent both ordinal and cardinal similarities. Similarities are ordinal when one candidate is more similar to you than another but there is no saying how much more similar. Similarities are cardinal when we can assign proportions to differences—for example, when Alfie and Bozo differ in their resemblance to you twice as much as Coco and Dodo do. There is no need to sort out which similarities are ordinal and which are cardinal, provided that we can accommodate both kinds. No hiding any of the facts from which comparative overall similarity might be thought to result!

Measurement theory has resources for a uniform representation of ordinal and cardinal similarities.[40] Let a *similarity function* be a function

[37] For further discussion of logical aspects, see Williamson, "First-Order Logics."

[38] This is a common assumption. See for example Lewis, *Counterfactuals*, p. 48.

[39] This accommodates the idea that ordinary objects are confined to their own possible worlds.

[40] See for instance Patrick Suppes, "Theory of Measurement," in Edward Craig, ed., *Routledge Encyclopedia of Philosophy* (New York: Routledge, 1998), pp. 243–49.

from some things into real numbers; intuitively, it is a representation of the degree to which these things resemble you, either in some particular respect or overall, as the case may be. Similarity functions are equivalent if they represent the same facts, but what this means depends on whether the facts in question are ordinal or cardinal. One similarity function s is *ordinally equivalent* to another, t, if s is an order-preserving transformation of t, and s is *cardinally equivalent* to t if s is a positive affine transformation of t—that is, there are real numbers $\alpha > 0$ and β such that, for every object o in the domain, $s(o) = \alpha t(o) + \beta$. Here, α allows equivalent functions to use different units, while β makes the origin arbitrary. I assume, then, that any cardinal similarities are to be measured on an interval scale, not on a ratio scale with a fixed origin ($\beta = 0$). This seems right if, unlike mass or heat or other quantities measured on a ratio scale, similarity can neither accumulate nor be entirely absent. This assumption is important, though, and the measurement of similarities will be a good place to start any further investigation into the possibility of aggregating them.

A *similarity measure* is a maximal class of equivalent similarity functions with the same domain. It is ordinal or cardinal, according to the sort of equivalence. Any similarity measure S induces a relation of comparative similarity: $a\check{S}b$ means that, for some (equivalently, all) s within S, $s(a) \leq s(b)$. These induced relations are weak orders. If the domain is assumed to be finite, we may identify ordinal measures with the orders that they induce, but different cardinal measures can induce the same orders. Having distinguished induced orders from similarity measures, I will use 'S' for either and sometimes for both within the same sentence.

Assume there is a (perhaps contextual) finite collection of respects of comparison: 1, ..., n. A *similarity profile* $\vec{S}$ is a list ($\vec{S}_1$, ..., $\vec{S}_n$) of similarity measures, all on the same domain. Each $\vec{S}_i$ is a measure of the similarity to you, with respect to i, of each thing in the domain. The measures of a profile are ordinal or cardinal, according to their respects.

IV.3. Weights and Balance. It is commonly supposed that sometimes one dimension of similarity carries greater weight than another and that it then is possible to combine them. If Alfie resembles you more closely in some respect that carries greater weight, for instance, this supposedly can make up for his resembling you less closely than Bozo does in another respect that carries less weight. Then, on balance, Alfie is more similar to you overall. Whatever it means for dimensions to have weights, presumably things are more favorable for aggregation when they have them. Presumably, comparisons of overall similarity are possible then, if they ever are.

I have argued that often there are no rates of exchange among similarities. Sometimes perhaps there are some, but they are more or less indeterminate. We might suppose that, in general, there are *many* admissible outcomes of aggregation, corresponding to different ways of hypothetically weighting dimensions: the less determinacy there is, the more weightings agree with it, and the more admissible outcomes there are. But reducing indeterminacy to multiplicity in this way does not seem to bring us closer to an understanding of how similarities might add up. I shall now argue that they do not add up even in the most favorable case, in which everything possible has been done to weight them, so that for every profile $\vec{S}$ there is presumably a unique resultant measure S of comparative overall similarity.

IV.4. Supervenience. I distinguish between two notions. With ordinal supervenience, which of two candidates is more like you, overall, only depends on their comparative similarities in particular respects. With cardinal supervenience, distinctions that are invisible in these ordinal facts may count as well. Ordinal supervenience appears to be the stronger notion, because cardinal facts entail ordinal facts but not the other way around. We will formulate the ordinal assumption and obtain our result. Then, we will see that it still follows when cardinal supervenience is assumed instead.

IV.4.a. To begin, we must capture the idea that, with regard to comparative similarities to you, some candidate counterparts in one possible world are just like some other candidates in another world, in every respect. Let R be a similarity measure; let Δ be some things within the domain of R; and let f be a one-one mapping from Δ into the domain of another similarity measure S. R $\approx_{f,\Delta}$ S means that, for each a and b in Δ, $a\mathrm{R}b$ if and only if $f(a)\mathrm{S}f(b)$. For similarity profiles $\vec{R}$ and $\vec{S}$, $\vec{R} \approx_{f,\Delta} \vec{S}$ means that, for each i, $\vec{R}_i \approx_{f,\Delta} \vec{S}_i$. We assume:

> *Ordinal Supervenience.* For all profiles $\vec{R}$ and $\vec{S}$, for all pairs Δ of things from the domain of $\vec{R}$, and for all one-one mappings f from Δ into the domain of $\vec{S}$, if $\vec{R} \approx_{f,\Delta} \vec{S}$, then R $\approx_{f,\Delta}$ S.

This indicates that which of two candidates is more like you, overall, entirely depends on the ordinal facts of which is more and which is less like you in the relevant respects. Notice two things. First, only the candidates' *similarities* matter: like candidates shall be treated alike, no matter who they are. Second, only *their* similarities matter: they shall be treated alike no matter who else is in the running.[41] There is a further assumption:

[41] Ordinal supervenience is the analogue of Arrow's notion of "Independence of Irrelevant Alternatives," with a slight generalization that allows profiles to have different domains. Arrow named his notion for this second aspect.

Dominance. For every profile $\vec{S}$ and for every a and b in its domain, if $a\vec{S}_i b$ for all i, then aSb.

And here is a definition:

Dictatorship. Among the respects 1, …, n, there is a critical respect d such that, for every profile $\vec{S}$, if $\alpha\vec{S}_d\gamma$ but not $\gamma\vec{S}_d\alpha$, then $\alpha S\gamma$.

The critical d dictates overall similarities in the sense that, whenever some candidate γ is strictly more similar to you than is another candidate α, in respect of d, γ is at least as similar to you as is α, overall. Now, we have the following:

Theorem. If the similarity profiles and corresponding measures of overall similarity satisfy ordinal supervenience and dominance, then we have a dictatorship.

Proof. See the Technical Annex.

IV.4.b. Allowing overall comparative similarities to depend on cardinal similarities in various respects might be thought to be a way out of trouble, but it is not, if any cardinal similarities are measured on an interval scale.[42]

We will need a notion of cardinal supervenience. Let R be a measure of similarity, let Δ be some things within its domain, and let f be a one-one mapping from Δ into the domain of a measure S. $R \equiv_{f,\Delta} S$ means:

For each $r \in R$, there is some $s \in S$ such that $r|\Delta = s \circ f|\Delta$, and

For each $s \in S$, there is some $r \in R$ such that $s|f(\Delta) = r \circ f^{-1}|f(\Delta)$.

That is, up to the identification of candidates by f, the similarity functions of R, restricted to Δ, are the same as those of S; $\equiv_{f,\Delta}$ generalizes to profiles in the obvious way. Now, instead of ordinal supervenience, we assume:

Cardinal Supervenience. For all profiles $\vec{R}$ and $\vec{S}$ and for all suitable pairs Δ and mappings f, if $\vec{R} \equiv_{f,\Delta} \vec{S}$, then $R \approx_{f,\Delta} S$.

That substitution of cardinal supervenience for the apparently stronger ordinal supervenience is not a way to avoid dictatorship is the point of the following:

Consequence. If the similarity profiles and corresponding measures of overall similarity satisfy cardinal supervenience and dominance, then we have a dictatorship.

[42] Paul Samuelson conjectured that the introduction of cardinal preferences was not a way around Arrow's impossibility theorem. This was verified by, among others, Ehud Kalai and David Schmeidler in "Aggregation Procedure for Cardinal Preferences: A Formulation and Proof of Samuelson's Impossibility Conjecture," *Econometrica*, XLV, 6 (September 1977): 1431–38.

This follows directly from the Theorem and from the fact that, perhaps surprisingly, cardinal supervenience and ordinal supervenience are equivalent. This is because, for all *pairs* Δ (although not in general):

$\vec{R} \equiv_{f,\Delta} \vec{S}$ if and only if $\vec{R} \approx_{f,\Delta} \vec{S}$.

The interesting part is "if." The basic idea of the demonstration is that any two points fall on a straight line and that any two straight lines with the same slope (both up or both down) are positive affine transformations of one another. This means that cardinal similarities, when restricted to pairs, might as well be ordinal similarities. Notice that this is where the assumption comes in that cardinal similarities are measured on interval scales.

MICHAEL MORREAU

University of Maryland at College Park and University of Oslo

TECHNICAL ANNEX: PROOF OF THE THEOREM[43]

An element μ is a *minimum* of relation R if, for each element *a* of the domain, μR*a*. Letting R* be the strict relation corresponding to R (*x*R**y* if *x*R*y* but not *y*R*x*), μ is a *strict minimum* of R if for each *a*, μR**a*. (There are analogous notions of *maxima*.) Take some finite set *A* with at least three elements and set aside one of them, *b*. Choose a series of *strict* profiles (all induced relations are strict) on *A* as follows: $\vec{Q}_0 = (\vec{Q}_{0,1}, \ldots, \vec{Q}_{0,n})$ is any strict profile such that, for each *i*, *b* is a minimum of the relation induced by $\vec{Q}_{0,i}$—that is, for every $a \in A$, $b\vec{Q}_{0,i}a$ but not $a\vec{Q}_{0,i}b$. Choose the next profile, $\vec{Q}_1$, such that its induced relations are just like those of $\vec{Q}_0$, except for $\vec{Q}_{1,1}$: *b* is a strict maximum of this relation.[44] Continuing in this way, we arrive finally at $\vec{Q}_n$; *b* is a strict maximum of each $\vec{Q}_{n,i}$.

Fact I. Let $\vec{Q}$ be any of the above profiles. Either *b* is a minimum of the resultant similarity measure Q, or *b* is a maximum of Q.

[43] This is a slight generalization of one of John Geanakoplos's "Three Brief Proofs of Arrow's Impossibility Theorem," *Economic Theory*, XXVI, 1 (July 2005): 211–15. The only real changes allow profiles to have different domains and to include cardinal as well as ordinal measures. The treatment of cardinal measures comes from Kalai and Schmeidler, *op. cit.*

[44] This construction and another, later one are objectionable, on an intended interpretation. In connection with counterpart theory, profiles represent possible worlds. On pain of begging the question against Lewis's views, we cannot find a series of worlds in which the very same things, represented by the elements of *A*, are organized differently, since worlds supposedly do not overlap. We can overcome this objection by introducing profile-world isomorphisms. The proof is more easily understood without the added clutter, though; once understood, it is clear enough what is needed.

Demonstration. For contradiction, suppose that b is neither a minimum of Q nor a maximum. Since Q is connected, there are a and c in A, such that cQ^*bQ^*a. Choose another similarity profile, $\vec{P}$, on A such that, for each i, c is ranked strictly above a, while the rankings of each relative to b are the same as in $\vec{Q}$ (simply switch the positions of a and c as needed). Let f be identity. Clearly,

$\vec{P} \approx_{f,\{a,b\}} \vec{Q}$, and

$\vec{P} \approx_{f,\{b,c\}} \vec{Q}$.

Therefore, by ordinal supervenience (where P is the resultant of $\vec{P}$), cP^*bP^*a; so, by transitivity of P, cP^*a. On the other hand, we have chosen $\vec{P}$ so that $a\vec{P}_i c$, for every i; by dominance, this delivers aPc. This is a contradiction.□

Fact II. There is some critical respect of comparison d such that b is a minimum of Q_{d-1}, and b is a maximum of Q_d.

Demonstration. By dominance, b is a minimum of Q_0 and a maximum of Q_n. Suppose there is some $d > 0$ such that b is a minimum of Q_{d-1} but not a minimum of Q_d. By Fact I, b is a maximum of Q_d. Otherwise, let $d=n$.□

We now will see that the critical respect of Fact II is a dictator: for any profile $\vec{S}$ and for any objects α and γ in its domain,

if $\alpha\vec{S}_d{}^*\gamma$, then $\alpha S\gamma$.

To this end, suppose $\alpha\vec{S}_d{}^*\gamma$. Let B be the domain of $\vec{S}$, and choose anything β that is not in B. Choose another profile $\vec{R}$ whose domain is $B \cup \{\beta\}$, such that the induced relations $\vec{R}_i$ satisfy, for all $\delta, \varepsilon \in B$:

For every i, $\delta\vec{R}_i\varepsilon$ if and only if $\delta\vec{S}_i\varepsilon$;

for every $i < d$, $\delta\vec{R}_i{}^*\beta$;

$\alpha\vec{R}_d{}^*\beta$ and $\beta\vec{R}_d{}^*\gamma$; and

for every $i > d$, $\beta\vec{R}_i{}^*\delta$.

$\vec{R}$ is just like $\vec{S}$, except that β ranks strictly above everything else in the orders before the critical dth order, between α and γ in this order, and below everything else in the remaining orders. Let f be identity, and note first that:

(1) $\vec{R} \approx_{f,\{\alpha,\gamma\}} \vec{S}$.

Let a be any element of A other than b, and let g be a mapping such that $g(a) = \alpha$ and $g(b) = \beta$; note also that:

(2) $\vec{Q}_d \approx_{g,\{a,b\}} \vec{R}$.

Let c be any element of A other than b, and let h be a mapping such that $h(b) = \beta$ and $h(c) = \gamma$; note also that:

(3) $\vec{Q}_{d-1} \approx_{h,\{b,c\}} \vec{R}$.

By Fact II, $aQ_d b$ and $bQ_{d-1}c$. By (2) and (3) above and by ordinal supervenience, $\alpha R \beta$ and $\beta R \gamma$; so, by transitivity, $\alpha R \gamma$. Finally, by (1) and ordinal supervenience, $\alpha S \gamma$. This completes the proof.