*CRÍTICA, Revista Hispanoamericana de Filosofía.* Vol. 46, No. 136 (abril 2014): 27–41

# TRANS-WORLD CAUSATION REVISITED

Axel Arturo Barceló Aspeitia
Instituto de Investigaciones Filosóficas
UNAM
abarcelo@filosoficas.unam.mx

SUMMARY: In a recent paper, García-Ramírez (2012) has argued that Lewis' counterfactual analysis of causation has the undesirable consequence of making trans-world causation possible. In this paper I argue, against García-Ramírez, that the possibility of trans-world causation cannot be derived from Lewis' account of in-world causation, since there is no way of extending Lewis' closeness relation among worlds into a similar closeness relation among pairs of worlds that is neither trivial nor ad-hoc.

KEY WORDS: resemblance, closeness, possible worlds, counterfactuals, Lewis

RESUMEN: En un artículo reciente, García-Ramírez (2012) ha argumentado que el análisis contrafáctico de la causalidad de Lewis tiene la indeseable consecuencia de hacer posible la causalidad transmundana. En este artículo argumento que, contrario a lo que García-Ramírez sostiene, la causalidad transmundana no se deriva de la teoría de Lewis de la causalidad intramundana, ya que no se puede extender la relación de cercanía entre mundos de Lewis a pares de mundo de una manera que no sea trivial o ad hoc.

PALABRAS CLAVE: similitud, cercanía, mundos posibles, contrafácticos, Lewis

## 1. *Introduction*

In a recent paper, Eduardo García-Ramírez (2012) has accused Lewis' (1973) counterfactual account of causation of entailing the possibility of trans-world causation. Given that one "may think that it is a metaphysical fact about causation that there cannot be trans-world causation, [this] may seem like a terrible consequence of counterfactual theory" (García-Ramírez 2012, p. 82). In finer detail, García-Ramírez's reductio argument against Lewis goes as follows:

1. The counterfactual account of causation: according to Lewis, an event $C$ in a world $W$ causes another event $E$ in the same world $W$ if and only if at the closest worlds to $W$ where $C$ does not occur, $E$ does not occur either.

2. If $C$ causes $E$ in world $W$ if and only if $E$ does not occur at all the closest worlds to $W$ where $C$ does not occur, then an event $C$ in a world $W_C$ would cause an event $E$ in a different world $W_E$ if and only if $E$ does not occur at the second world of all the closest world-pairs to $<W_C, W_E>$ such that $C$ does not occur at their first worlds.

3. There are events $E$ and $C$, and worlds $W_C$, $W_E$ such that $C$ occurs in $W_C$, $E$ occurs in $W_E$ and for all the closest world-pairs to $<W_C, W_E>$ such that $C$ does not occur at the first world of the pair, $E$ does not occur at the second world of the pair.

4. From the previous premises it follows by a couple of Modus Ponens that it is possible for an event in a world to cause an event in a different world, i.e. trans-world causation is possible.

5. Presumably, trans-world causation is impossible.

6. Therefore, we must abandon the counterfactual account of causation.

García-Ramírez presents his argument as a reductio against the conjunction of Lewis' counterfactual account of causation and his brand of modal realism. However it is clear that, as the reconstruction above shows, his argument does not require that possible worlds be concrete. Consequently, I will take his argument as a reductio of Lewis' theory of causation, period.

In (1986), Lewis himself had already considered a similar objection, but concluded that, even if we wanted to extend his analysis of counterfactuals to make sense of trans-world cases in something like the way suggested in premise 2 of García-Ramírez's argument, the most natural way of doing so would not yield any non-trivial trans-world counterfactuals, and instead would make all claims of trans-world causation false. In other words, Lewis had already anticipated García-Ramírez's objection, and replied by challenging premises 2 and 3. According to Lewis, the above argument is unsound, for it is based on an equivocation in the predicate "being closer to". For Lewis, the closeness relation relevant for in-world causation is a (three-arguments) relation among worlds, not pairs of worlds, and furthermore, there is no way of extending his closeness relation to develop an analogous relation for pairs of worlds in a way that makes both premises 2 and 3 true, i.e. there is no way of defining a closeness relation among pairs of worlds that is both the most natural extension of the closeness relation governing ordinary in-world causation (so that premise 2 is true) and does not make all relevant counterfactuals trivially false (so that premise 3 is true as well).

García-Ramírez takes Lewis' claims as setting up a challenge: to extend his counterfactual account of intra-world causation to cover trans-world causation in a way that is non-trivial and naturally follows from Lewis' original account. Most of the work on García-Ramírez (2012) is devoted to offering such an account. In this text, I contend that what he offers there is, at best, an incomplete sketch of how such an account could be given and, furthermore, that there is no way of completing it to satisfy the constraints of Lewis' challenge.

## 2. *Constraints on a Natural Counterfactual Theory of Trans-World Causation*

Before getting on to García-Ramírez's proposal, it is very important to be clear on what exactly is involved in Lewis' challenge. The first condition for developing an appropriate counterfactual theory of trans-world causation requires embracing what García-Ramírez calls the fundamental claim of a possible world account of counterfactuals: "Counterfactuals are true if it takes less of a departure from actuality to make the consequent true along with the antecedent than it does to make the antecedent true without the consequent" (Lewis 1973). Also, any account of trans-world causation that aims to be a natural extension of Lewis' intra-world account must also embrace the following two theses:

> Thesis 1: One pair of worlds $<X, Y>$ is closer to another pair $<A, B>$ than another pair $<W, Z>$, if $<X, Y>$ resembles $<A, B>$ more than $<W, Z>$ does.
>
> Thesis 2: There is causal dependence between an event $C$ in a world $W_C$ and another event $E$ in a different world $W_E$ iff if $C$ were not to occur in $W_C$, $E$ would not occur in $W_E$. (Lewis 1986)

Combining these two theses with the fundamental claim, we get the following counterfactual account of trans-world causation:

> An event $C$ in a world $W_C$ causes event $E$ in world $W_E$ if and only if, at the closest world-pairs to the pair $<W_C, W_E>$ where $C$ does not occur at the first world of the pair, $E$ does not occur at the second.

Consequently, Lewis' challenge is to define a relation of closeness between pairs of worlds that is both non-trivial (it must not make all trans-world causal counterfactuals vacuously false nor vacuously true) and the most natural extension of the closeness relation governing intra-world causation. In order to satisfy this second condition, the relevant closeness relation must also satisfy certain constraints, both in its form and its content. From a formal perspective, the closeness relation ought to be a three-place relation among pairs of worlds that is both weak and non-finite;[1] regarding its content, Lewis demands that it "be governed by the same sort of closeness that governs ordinary causal counterfactuals" (Lewis 1986, pp. 79–80). Presumably, this requires satisfying at least the following constraints:

> Conservativeness: It must not postulate more external relations among worlds or pairs of worlds that those already given in Lewis' theory.
>
> Homogeneity: Pair $<X, X>$ is closer to $<A, A>$ than $<Y, Y>$ if and only if world $X$ is closer to world $A$ than $Y$.

However, in order to meet Lewis' challenge, it is not enough to define a non-trivial closeness relation among pairs of worlds that is conservative and homogenous, or even less to show that developing such a relation is possible (even if showing so would already be a big blow to Lewis' theory, since he also makes the strong claim that "under a counterfactual analysis of causation, the causal isolation of worlds follows automatically" 1986, p. 78); it is also necessary that the closeness relation defined on pairs of worlds be the most natural extension possible of Lewis' original proposal. Otherwise, one could not justifiably claim that Lewis' theory has the possibility of trans-world causation as an undesirable consequence. In other words, in order to meet Lewis' challenge, one must warrant that the conditional in premise 2 is true, i.e. that one's closeness relation for pairs of worlds follows naturally from Lewis' own closeness relation for single worlds. There must not be other, trivial and more natural (or, at least as natural) ways of extending Lewis' relation. If there were many, equally natural ways of extending Lewis' closeness relation, some trivial and others not, the conditional in 2 would not be justified, and the Modus Tollens in 6 would not be valid (Lee Bowie 1979). Thus,

[1] Weakness: there can be ties, but any two pairs of worlds must be comparable with respect to a third, i.e. for any three pairs of worlds $<A, B>$, $<X, Y>$ and $<W, Z>$, either $<X, Y>$ is closer to $<A, B>$ than $<W, Z>$, $<W, Z>$ is closer to $<A, B>$ than $<X, Y>$, or $<X, Y>$ is as close to $<A, B>$ as $<W, Z>$.

Non-Finiteness: the ordering need not be finite. For some pairs of worlds, there may be no single closest pair of worlds, i.e. there may exist a pair of worlds $<A, B>$ such that there is no pair of worlds $<X, Y>$ closer to $<A, B>$than any other pair of worlds $<W, Z>$ (different from $<A, B>$ and $<X, Y>$).

if 3 is false for at least one of the most natural ways of extending Lewis' account, then the possibility of trans-world causation cannot be said to be a consequence of Lewis' theory, making the resulting reductio argument unsound.

The existence of relatively natural extensions of comparative similarity of worlds for pairs of worlds that would allow for trans-world causation should not be more surprising or problematic than finding choices of similarity weighting among worlds that could result in there being as many backward causal relations as forward ones. Only if such metaphysical absurdities were entailed by Lewis' theory, could we speak of an actual reductio. What García-Ramírez needs to show, then, is not only that it is possible to extend Lewis's account so as to get cross-world causation, but that Lewis's account by itself already entails the possibility of trans-world causation. For this, he needs to show, not only that there is a non-trivial closeness relation among pairs of worlds that is both conservative and homogenous, but also that there are no other (at least as) natural ways of extending Lewis' closeness relation that are also conservative and homogenous, but not trivial. Unfortunately, García-Ramírez fails to show this stronger claim and thus fails to show that Lewis' account of causation actually entails the possibility of trans-world causation.

## 3. *García-Ramírez's Sketch of a Solution*

Eduardo García-Ramírez's proposal is based on a simple but commonly overlooked fact about how we ordinarily compare pairs of things, i.e. that when comparing pairs of things we have to consider, at least, two different kinds of respects of similarity:

> Member comparison: we must compare individual members against individual members.
>
> And
>
> Relational comparison: we must compare the pair against the pair, i.e., we need to take the relations that hold between the members of the first ordered pair and compare them against the relations that hold between the members of the second ordered pair.[2]

[2] García-Ramírez uses the more neutral names comparison 1 and comparison 2; since I will be referring to these all through the article, I have preferred to use more descriptive names.

This gives us two kinds of respects of similarity for comparing pairs of worlds. In the first respect, a pair of worlds $<X, Y>$ is more similar to pair $<A, B>$ than pair $<W, Z>$, if $X$ resembles $A$ more than $W$ does and $Y$ resembles $A$ more than $W$ does. From a relational perspective, in contrast, a pair of worlds $<X, Y>$ is more similar to pair $<A, B>$ than pair $<W, Z>$, if the relations that hold between $X$ and $Y$ are more similar to those that hold between $A$ and $B$ than those that hold between $W$ and $Z$. According to García-Ramírez, combining both sorts of comparison allows us to construct a closeness relation that is both non-trivial and the most natural. From this combined perspective, an ordered pair of worlds would be closer to a given pair of worlds "the more each individual member of the pair resembles its corresponding member and the more the relations that hold between the members of the resembling pair resemble those that hold between the members of the original pair" (García-Ramírez 2012, p. 77).

Of course, in order to compare pairs of worlds relationally, there must be inter-world relations upon which to base the comparison. Yet, given the constraint of conservativeness that must govern our closeness relation, we cannot appeal to any trans-world relation that is not already there in Lewis' original theory. Fortunately, there are enough relations in Lewis' theory to ground relational comparisons among pairs of worlds: relations of comparative similarity among possible worlds or RCS for short. García-Ramírez writes:

> Remember, counterfactual theory presupposes that possible worlds are comparable and, hence, that there are RCS among worlds. If an ordered pair is to be closer to a second one than a third one is, apart from individual resemblances across individual members, the RCS that hold between the members of the first one must also resemble the RCS that hold between members of the second one. (García-Ramírez 2012, p. 78)

Thus, if there is a privileged set of RCS with respect to which a world $X$ resembles $A$ more than $W$ does, $Y$ resembles $B$ more than $Z$ does, and the RCS between $X$ and $Y$ resemble those between $A$ and $B$ more than the RCS between $W$ and $Z$ do, then $<X, Y>$ will be closer to $<A, B>$ than $<W, Z>$. According to García-Ramírez's proposal, the closeness relation thus defined satisfies all the constraints in Lewis' challenge and therefore allows us to ground a counterfactual analysis of trans-world causation.

The problem with the proposal is that it falls short of delivering what it promises. In his (2012) article, García-Ramírez has not actually given us a closeness relation among pairs of worlds. At most, he has given us a substantial pointer about how such a relation could be constructed, but substantial work is still missing. As I will show in the remaining of the article, this a serious problem for García-Ramírez, for when we consider natural ways of developing his sketch, we always ends up with a proposal which is either trivial or not natural enough to justify the claim that Lewis' counterfactual theory entails the possibility of trans-world causation.

## 4. *Completing García-Ramírez's Sketch*

According to García-Ramírez's proposal, a pair of worlds is closer to another if it is more similar to it according to both member comparison (how each world in each pair resembles its corresponding member) and RCS-relational comparison (how the RCS that hold between each pair's members resemble the RCS that hold between the other pair's members). This gives us one way in which a pair of worlds can be closer to another than a third, but does not actually tells us how to decide, in general, whether a pair of worlds is closer to another than a third one or not. Notice that García-Ramírez presents his proposal in terms of a conditional of the form "$<X, Y>$ is closer to $<A, B>$ than $<W, Z>$, if ... ", not a biconditional. This gives us sufficient conditions for closeness, but no necessary ones. This makes García-Ramírez's proposal incomplete. It tells us what happens when a pair of worlds is more similar to another according to both kinds of comparisons, but it does not tell us what happens when they are not.

One might feel tempted to reply that I am being quite unfair to García-Ramírez, and that the missing conditional, even if not explicit, is obviously implied. This is just what I thought myself when I first read García-Ramírez's article. However, the problem is that if that were his actual proposal, it would make all trans-world counterfactuals trivially false, precisely for the reasons Lewis had already foreseen in his (1986), when he wrote:

> This makes sense, but not I think in a way that could make it true. For I suppose that the closeness of one world-pair to another consists of the closeness of the first worlds of the pairs together with the closeness of the second worlds of the pairs. We have to depart from $W_C$ for the first world of a closest pair, since we have to get rid of $C$. But we are not likewise forced to depart from $W_E$ for the second world of a closest pair, and what is so close to a world as that world itself? So the second

> world of any closest pair will just be $W_E$, at which $E$ does occur, so [the relevant counterfactual] is false. (Lewis 1986, p. 79)

Of course, García-Ramírez's proposal is not one where closeness of one world-pair to another consists of the closeness of the first worlds of the pairs together with the closeness of the second worlds of the pairs (that would be considering only member comparison, and thus ignoring the core of García-Ramírez's proposal). However, if we understood his proposal as containing bi-conditionals where he explicitly uses conditionals, Lewis' argument would still apply to it. Substituting the word "consists" with the word "requires" in the quote above would give us a more general argument that would trivialize any proposal, like García-Ramírez's, that takes member similarity as a necessary condition for the closeness of the respective pairs of worlds, as follows:

1. Assume, towards a contradiction, that trans-world causation is possible, i.e., that there are at least two different events $C$ and $E$ and two possible worlds $W_C$ and $W_E$ such that, if $C$ had not occurred in $W_C$, $E$ would not have occurred in $W_E$ either

2. From 1, it follows that at all the closest world-pairs to the pair $<W_C, W_E>$ such that $C$ does not occur at the first world of the pair, $E$ does not occur at the second world of the pair., i.e., for every world pair $<W_1, W_2>$ such that $C$ does not occur at $W_1$ and $E$ occurs at $W_2$, there is at least one other world pair $<W_3, W_4>$ closer to $<W_C, W_E>$ than $<W_1, W_2>$ such that $C$ does not occur at $W_3$ and $E$ does not occur at $W_4$.

3. Let $W_1$ be a world such that $C$ does not occur at $W_1$.

4. From 2 and 3, $<W_1, W_E>$ is a pair of worlds such that $C$ does not occur at $W_1$ and $E$ occurs at $W_E$.

5. From Modus Ponens of 2 and 4, there is at least one other world pair $<W_3, W_4>$ closer to $<W_C, W_E>$ than $<W_1, W_E>$ such that $C$ does not occur at $W_3$ and $E$ does not occur at $W_4$.

6. From García-Ramírez's definition of the closer-to relation, if $<W_3, W_4>$ is closer to $<W_C, W_E>$ than $<W_1, W_E>$, then $W_3$ resembles $W_C$ more than $W_1$ does, $W_4$ resembles $W_E$more than $W_E$ does, and the RCS between $W_3$ and $W_4$ resemble those between $W_C$ and $W_E$ more than the relation between $W_1$ and $W_E$ do.

7. In particular, from 5 and 6, $W_4$ resembles $W_E$ more than $W_E$ does.
8. Also, since $E$ occurs at $W_E$ and $E$ does not occur at $W_4$, $W_4$ and $W_E$ are different worlds.
9. But it is impossible for any world (different than $W_E$) to resemble $W_E$ more than $W_E$ does. QED.

This argument shows that taking member comparison as a necessary condition for closeness requires from trans-world causation something impossible, i.e., that a world $X$ different from $Y$ may resemble $Y$ more than $Y$ itself. This means that the simplest way of completing García-Ramírez's sketch fails at providing us with a natural and non-trivial extension of Lewis' relation of closeness. Is there another, better way of doing it, one that gives us a most natural and non-trivial way of extending Lewis account into trans-world causation? In the following section, I will argue that no, there is not.

## 5. *How to Turn Two Similarity Relations into a Single Closeness Relation?*

As mentioned above, García-Ramírez tells us what happens when the members of a pair of worlds and the RCS relations between them are more similar to those of another. However, it is silent about what happens otherwise. In general, let $X$ and $Y$ and $Z$ be three pairs of worlds such that:

- The members of $X$ are more similar to those of $Y$ than to those of $Z$, but
- The RCS relations between the members of $Z$ are more similar to those that hold between the members of $Y$ than to those that hold between the members of $X$.

In order to develop a closeness relation amongst pairs of worlds that satisfies Lewis' challenge, one must give a principled answer to the question:

$$\text{Is } X \text{ closer to } Y \text{ than } Z \text{ or not?}$$

To know, the available options are the following:

(a) If the members of $X$ are more similar to those of $Y$ than to those of $Z$, then $X$ is closer to $Y$ than $Z$.

(b) If the RCS relations between the members of $X$ are more similar to those exhibited by the members of $Y$ than those exhibited by the members of $Z$, then $X$ is closer to $Y$ than $Z$.

(c) If $X$ is more similar to $Y$ than $Z$ according to the first comparison and $Z$ is more similar to $Y$ than $X$ according to the second, then neither is closer to $Y$ than the other.[3]

(d) If $X$ is more similar to $Y$ than $Z$ according to the first comparison and $Z$ is more similar to $Y$ than $X$ according to the second, then either can be closer to $Y$ than the other, depending on the case.

However, as I will argue now, all the available answers are either non-motivated and/or fail to meet the conditions of Lewis' challenge.

First of all, as we have seen, Lewis' argument in (1983) shows the triviality of any proposal that makes member similarity a necessary condition for closeness among pairs of worlds. Options (a) and (c), therefore, can be easily discarded as trivial. Option (b) fares even worse, since it entails that no event $C$ in a world $W$ can cause another event $E$, in the same world $W$, unless $E$ metaphysically entails $C$, making most intra-world counterfactuals false, as can be shown by the following argument:

1. Assume, towards a contradiction, that $C$ and $E$ are events such that $C$ causes $E$ in a world $W$, but $E$ does not metaphysically entail $C$.

2. Since $E$ does not metaphysically entail $C$, there is a world $X$ such that $C$ does not occur in $X$, but $E$ does.

3. Since, under any RCS, no world is more similar to itself than itself, the RCS that hold between any one world and itself cannot be different from the RCS that hold between any other world and itself.

4. From 3, the RCS-relational similarity between the members of $<W, W>$ is identical to the RCS-relational similarity between the members of $<X, X>$.

5. From 4, there is no pair of worlds RCS-relationally more similar to $<W, W>$ than $<X, X>$.

[3] Notice also that this is the way similarity is commonly formalized. For example, see Konikowska 1997.

6. From 5 and b, there is no pair of worlds closer to <$W$, $W$> than <$X$, $X$>.

7. From 6 and the constraint of Homogeneity, there is no world closer to $W$ than $X$.

8. From 7 and 2, it is false than in all the closest worlds to $W$ where $C$ does not occur, $E$ does not occur either.

9. From 8 and Lewis' fundamental claim, $C$ in $W$ does not cause $E$ in $W$. This directly contradicts assumption 1.

It seems therefore, that García-Ramírez must adopt perspective (d) where "we need to balance the different respects of similarity and weigh them with respect to their relevance for each case" (García-Ramírez 2012, p. 72). Instead of deciding which order of similarity is dominant, we can combine them both into a new closeness relation that takes both sorts of similarities into consideration. By making neither member nor relational similarity a sufficient condition for closeness, this fourth alternative seems to avoid the problems of the previously considered ones.

However, as promising as this option sounds, it is also far from satisfactory. Mostly, because the RCS relations that govern intra-world causation according to Lewis lack the kind of structure that could help us compose them into a single well-behaved weak order between pairs of worlds (Morreau 2010). Notice that in order to get a single closeness relation out of both member and relational similarity, it is necessary to compare how much one pair is more similar than another according to each respect of similarity, and then, assigning different weights to each respect of similarity, determine which one is the most similar pair overall. To do this, it would not be enough to know if one world is more similar to another (regarding a third), we would also need to know how much more similar it is. In other words, it would be necessary to introduce something like degrees of similarities among worlds. However, the closeness relation at play in Lewis' theory is little more than a weak order between worlds. Consequently, it can only tell us if one pair is closer to another than a third, but not how much closer it is.

Let me illustrate this point with a more familiar example. Suppose you are given a list of students in your class ordered by height. The list tells you who is taller than whom, whether one student is as tall as another, and nothing more. Notice that from the list alone, it is impossible to tell how much the first student on the list is taller than the last one. It is also impossible to tell whether or not the difference between the first student and the second is larger than the difference between the second and the third. Suppose now you wanted to create a new list ordering pairs of students regarding height; even though you know who is taller than who, this information is of little help to determine whether one pair of students is taller than another. Certainly, if you knew the exact height of each student, you could use this information to order the pairs of students. For example, you could just add the heights of each student so that one pair of students is taller than another if the combined height of the first pairs is higher than the other. Without knowing how tall each student is however, no similar strategy is available. Now, something similar happens when we talk of similarity relations between worlds. Lewis' closeness relation between worlds is just like an ordered list of worlds: it allows us to tell if one world is more similar to another (with respect to a third one), but not how much more similar it is. If one cannot tell how much more similar one world is to another than a third one, much the less can we add, weigh or compare such similarity with a different one.

The problem is not that it is impossible or even specially challenging to introduce something like degrees of similarity among worlds into Lewis' theory and then assign different weights to member and relational degrees of similarity. On the contrary, the real problem is that it is too easy. There are almost no constraints on how degrees may be introduced and how different respects of comparison may be weighted, consequently there are many ways in which we can do so, none of them more faithful to Lewis' theory than the others. The reason why this is a problem for García-Ramírez is that only some of them allow for trans-world causation, while others do not (Morreau 2010). Therefore, there is no principled reason to introduce degrees and weights in a way that allows for trans-world causation, instead of introducing them in a way that excludes trans-world causation.

Remember that, in order for a relevant trans-world counterfactual to be true, it must be possible for a pair of worlds $X$ to be closer to a given pair $Y$ than another $Z$, even if the members of $Z$ are closer to those of $Y$ than those of $X$. Suppose now that we have successfully introduced some degrees of similarity into our model so that the members of some pair $A$ are $n$ degrees closer to those of $B$ than those of $C$, but the members of $C$ are $m$ degrees closer to how similar the members of $B$ are than how similar the members of $A$ are. How do we combine these two opposing quantities $n$ and $m$ in order to get an overall distance among $A$, $B$ and $C$? There are infinite ways of doing so, assigning different relative weights to $n$ and $m$ and then combining them into a single quantity: some of them will give more relative weight to $n$ than to $m$, making all relevant trans-world counterfactuals false, and others will assign $m$ at least equal relative weight than $n$, making some relevant trans-world counterfactuals true. We can decide that member comparison is twice as important as relational comparison, so that a pair of worlds $X$ might be closer to another $Y$ than a third $Z$, even if the similarity between the members of this third pair $Z$ is itself more similar to that of $Y$ than that of $X$, as long as the members of $X$ are at least half as similar to those of $Y$ than those of $Z$. Or we might decide that, in order to trump relational comparison, member comparison must exceed it by at least some given number of degrees, or that it must be at least three times as much as it, or whatever. The possibilities are truly infinite. Whether or not trans-world causation is possible would depend on our choice. But Lewis' theory gives us no principled way of choosing, so any option we decide on would be equally ad-hoc.

In other words, even if it were possible to define some notion of distance between worlds consistent with Lewis' original model and then use it to construct a similarity relation that makes not all relevant counterfactuals false, doing so would still not help further the case against Lewis. Mostly, because this would require adding structure —degrees and weights— that is not there in Lewis' original proposal; and the more is added to his original proposal, the less the resulting closeness relation can be said to be a most natural extension of his. Remember that for the desired reductio to work, whatever is necessary for Lewis' theory of intra-world causation must be sufficient for trans-world causation; and while Lewis' theory does not require anything stronger than a weak order relation of closeness among worlds, this is not sufficient to construct the kind of closeness relation among pairs of worlds necessary for trans-world causation. Consequently, it is not sufficient to show that the possibility of trans-world causation follows from Lewis' theory of causation either.

Finally, one might want to defend García-Ramírez's proposal by arguing that embracing option (d) does not require finding a way of composing member and RCS-relational similarities into one closeness relation; that it is enough to state that in some contexts, when comparing pairs of worlds, we care more about relational comparison than about member comparison, and at other times it is the other way around. Context decides this, and it is not fixed at all which respect would trump the other in which situations. So, it might be unfair of me to ask for a principled way to determine trumping of respects when the idea of option (d) should be that there aren't such principles. Furthermore, the high context sensitivity of trans-world counterfactuals and of trans-world closeness is inherited from the high context sensitivity of intra-world counterfactuals and intra-world closeness. When evaluating ordinary counterfactuals we only consider, certainly, the content of individual worlds, but even then there are many relevant respects of comparison. And which one of these trumps the other is a matter that varies from context to context. So, the truth-value of ordinary counterfactual type claims is not something fixed once and for all, it changes from token to token. Just as in ordinary comparisons of pairs of objects contextual factors decide which respect —member comparison or relational comparison— is the most relevant, there is no fixed principled way to do this when comparing pairs of possible worlds.

The point is well taken; however it does nothing to further García-Ramírez's goal. Embracing a contextualist solution cannot save us from the dilemma of choosing between member and relational similarity. As I have shown, both alternatives have vastly undesirable consequences: a closeness relation based only on member similarity makes all relevant trans-world counterfactuals trivially false, while a closeness relation based only on RCS-relational similarity makes most intra-world counterfactuals trivially false. Therefore, it is no good to say that we can choose sometimes one, and other times the other when both horns of the dilemma have such undesirable consequences.

In the end, all of our attempts at developing García-Ramírez's sketch into a full proposal seem to have resulted in triviality or unnaturalness. Taking member similarity to be a necessary condition for closeness results in undesirable triviality, making trans-world causation impossible. Taking relational similarity to be a sufficient condition for closeness makes most intra-world causation impossible, which is even worse. And finally, we have also seen that there is no way of weighing or comparing both respects of similarity to get a single closeness relation that is not trivial and entailed by Lewis' theory. Consequently, contrary to what García-Ramírez claims, we can conclude that Lewis' counterfactual theory of causation does not actually entail the possibility of trans-world causation. Without such potentially undesirable consequence, his reductio argument is unsound, and fails to compel us to abandon Lewis' theory.[4]

[4] I would like to thank Alessandro Torza, Eduardo García-Ramírez and the anonymous referees in charge of reviewing previous versions of the manuscript

## REFERENCES

García-Ramírez, E., 2012, "Trans-World Causation?", *The Philosophical Quarterly*, vol. 62, no. 246, pp. 71–83.

Konikowska, B., 1997, "A Logic for Reasoning about Relative Similarity", *Studia Logica*, vol. 58, no. 1, pp. 185–226.

Lee Bowie, G., 1979, "The Similarity Approach to Counterfactuals: Some Problems", *Noûs*, vol. 13, no. 4, pp. 477–498.

Lewis, D., 1986, *On the Plurality of Worlds*, Blackwell, Oxford.

———, 1973, "Causation", *The Journal of Philosophy*, vol. 70, no. 17, pp. 556–567.

Morreau, M., 2010, "It Simply Does Not Add Up: Trouble with Overall Similarity", *The Journal of Philosophy*, vol. 107, no. 9, pp. 469–490.

*Received: June 4, 2012; revised: August 16, 2013; accepted: February 5, 2014.*

for their helpful and encouraging comments. This research was supported by the "Epistemología y metafísica de la modalidad" PAPIIT IA 400412 research project.