Counterfactual does not Entail Downward Causation

Dong Xin. Peking University

DOI: 10.47297/wspciWSP2515-469903.20200401

1. Background

There are numerous philosophers, called non-reductive physicalists(NRP), who agree with basic principles of Physicalism but insist that mental property cannot be reduced to physical property. The main goal of NRP is to save the causal power of mental property qua mental. However, throughout the 1990s, Jaegwon Kim developed the famous Supervenience Argument, proving that what purport to be NRP are ultimately untenable, since they cannot accommodate the causal efficacy of mental states and the irreducibility of mental at the same time.

Along this argument, there is a fatal step, that is, M caused M* by causing P*[①], called downward causation, which means M causing M* entails M causing P*. Here, P* is the supervenience base of M* while M has P as its supervenience base. Kim states that there is a general principle underlying this step: To cause a supervenient property to be instantiated, you must cause its base property to be instantiated[②]. Once this step is admitted, the following steps of Supervenience Argument seem to be well-reasoned, which is also reconstructed as

① Jaegwon Kim, *Mind in a Physical World: an essay on the mind-body problem and mental causation*, The MIT Press, 2000, p.42.

② Ibid.

Exclusion Argument. Based on this argument, the causal power of mental is eventually excluded by physical. What followed is the Exclusion Argument:

> (S) *Supervenience*: Mental properties supervene on physical properties. That is, if any system s instantiates a mental property M at t, there necessarily exists a physical property P such that s instantiates P at t, and necessarily anything instantiating P at any time instantiates M at that time.
>
> (DC) *Downward Causation*: If property A causes property B, then A must cause any base property of B instantiated on this occasion.
>
> (CCP) *Causal Completeness of Physics*: If a physical event has a cause that occurs at t, it has a (sufficient) physical cause that occurs at t.
>
> (NO) *Non-overdetermination*: No single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of causal overdetermination. And there is no systematic overdetermination in cases of mental causation.
>
> (I) *Irreducibility*: Mental properties are not identical with physical properties.

Conclusion: The causal power of M is excluded by the causal power of P. ①

Facing this exclusion argument, many NRP turn to the counterfactual theory of causation for help. However, they solve the Exclusion Argument in different approaches. What they all agree is that (S), (CCP), (NO) and (I) cannot be denied. But concerning with (DC), those NRP split into two groups. One group agrees with Kim that mental property has to commit downward causation in order to reserve causal efficacy and there is no problem with it (Lepore and Loewer 1987; Horgan 1989; Antony and Levine 1997; Loewer 2002, 2007). They use counterfactual theory to define causation in order to claim that mental property has causal power, no matter the effect of it is mental or physical. And their strategy to solve the exclusion problem is usually to prove that (DC), combing with (S) and (I) will not bring about violation of (NO) and (CCP). Therefore, these five premises together do not

① Zhong Lei: 'Can Counterfactuals Solve the Exclusion Problem?'. *Philosophy and Phenomenological Research*, Vol. LXXXIII, No.1, July 2011, p.138.

necessarily lead us to the conclusion.

The other group disagrees with Kim on (DC), announcing that it does not necessarily exist. As John Gibbons says, we do rule out systematic upward or downward causation and mental causation without downward causation is just one instance of more general pattern of intralevel causation without interlevel causation.① We can call this group the counterfactual autonomist solution. (Yablo 1992; Thomasson 1998; Marras 1998; Crisp and Warfield 2001; Gibbons 2006).

Let's explore autonomy solution further more. This theory claims that mental causes bring about mental effects but not physical effects, while physical causes bring about physical effects.② And it is compatible with the other four premises and therefore, can block the Exclusion Argument and save the causal power of mental property. Although some scholars use determination (see Thomasson 1998) or realization (see Schlosser 2009), instead of supervenience, to describe the relationship between M and P, they can still circumvent the Exclusion Argument by holding back the occurrence of downward causation.

I will not give pro and con analysis of these two solutions. Neither will I discuss the validity of Exclusion Argument. It is not the purpose of this paper. Here, I just accept all of them as our discussion background. And with it in mind, we can now move on to discuss Zhong Lei's paper "*Can counterfactuals solve the exclusion problem?*" and figure out if Zhong's argument is valid.

2. The analysis of Zhong Lei's argument

Before I display Zhong's argument structure, I'd like to stress that who is his real opponent. In other word, I will point out whose theory is indeed threatened by Zhong's argument. Then, we can investigate whether the threat is successful.

In brief, Zhong claims that since those NRP admit supervenience as Kim stat-

① John Gibbons: 'Mental Causation without Downward Causation'. *The Philosophical Review*, Vol. 115, No.1, January 2006, p.89.

② Dwayne Moore: 'Counterfacutals, Autonomy and Downward Causation: Reply to Zhong'. *Philosophia*, Vol. 41, Issue 3, September 2013, p.834.

ed[①] and counterfactual theory, they have no choice but to admit downward causation. And according to the Exclusion Argument, if they admit downward causation, then they must accept that the causal power of mental property will be excluded by physical property. Therefore, Zhong states that counterfactual theory cannot help NRP to solve the exclusion problem as they expect it to.

As I have mentioned above, there are two solutions of Exclusion Argument. So, we must figure out which group of NRP Zhong is referring to in his paper. In my point of view, Zhong does not take the first group into account and his argument is ineffective for it because his attack of counterfactual theory depends on its entailment of downward causation. Confronted with his argument, scholars of this group can respond that they are fine with that counterfactual theory will entail downward causation. The bone of contention between them and Zhong should be whether the Exclusion Argument itself is persuasive and convincing, that is, whether downward causation will lead to the conclusion that the causal power of mental property is excluded. But it is obviously not what Zhong wants to testify but what he takes for granted.

In contrast, the second group of NRP shares with Zhong's assumption that in order to avoid the Exclusion Argument, we must abandon downward causation principle. As a result, it is obvious that if Zhong's argument succeeds, he can hit this group a fatal blow because the latter cannot accept counterfactual theory while remain deterring downward causation any more. Therefore, it is the second group of NRP, that is, counterfactual autonomist solution, not the first group that is threatened by Zhong. So, even if Zhong's argument is valid, it cannot directly lead to the conclusion that the counterfactual theory of causation cannot help the NRP to solve the exclusion problem, but only lead that it cannot help autonomist solution. Besides, as Zhong himself stresses in his paper, he don't plan to discuss the exclusion argument itself in his paper. The issue his paper is dealing with is not whether the exclusion problem can be solved, but whether it can be solved by the counter-

① Kim (2000) argued that there are varies kinds of physicalism and those different physicalisms hold different assumptions. However, there exist some assumption that once someone announces that he or she is an physicalist, he or she should accept it. He called this "minimal physicalism". And he said supervenience is the very assumption that should be accepted by all minimal physicalism.

factualist autonomy approach. ①

Now, since I have clarified the target of Zhong's argument, let's see how he achieves his goal.

Zhong points out that if someone accepts the counterfactual theory of causation and supervenience, he or she has to admit Downward Causation Principle. And according to the Exclusion Argument I have given out at the beginning of this paper, he or she must accept that mental is excluded by physical. As a result, the core of Zhong's argument is to prove that how counterfactual theory of causation and supervinience will necessarily entail downward causation. This proof follows as below:

> 1. M* is realized by P*, that is, any (~M*)–world is also a (~P*)–world, and any P*–world is also an M*–world.
> 2. A (~M& ~M*)–world is at the same time a (~M& ~P*)–world.
> 3. (~M&P*)–worlds are a subset of (~M&M*)–worlds.
> 4. If some (~M& ~M*)–world is closer to the actual world than any (~M&M*)–world , then (~M& ~P*)–world is closer to the actual world than any (~M&P*)–world.
> 5. So, if M* counterfactually depends upon M, then P* would also counterfactually depend upon M. ②

Conclusion: counterfactual theory of causation and supervenience together imply downward causation from M to P*.

In order to figure out the validity of his argument, we need to scrutinize whether Zhong uses the concepts of supervinience and counterfactual properly. In this section, I will examine the usage of supervinience first, since I, on the whole, agree with Zhong's employment of supervinience. In the next section, I will analyze the concept of counterfactual carefully and point out that there should be more steps in the argument in order to achieve Zhong's goal.

---

① Zhong Lei, 'Can Counterfactuals Solve the Exclusion Problem?'. *Philosophy and Phenomenological Research*, Vol. LXXXIII, No.1, July 2011, pp.140-141.

② Ibid, p.141.

First comes first. In his paper, he uses a formula to summarize supervenience, that is:

> M* is realized by P*, that is, any (~M*)−world is also a (~P*)−world, and any P*−world is also an M*−world.

It may seem unfamiliar with those definitions we usually use in the first sight, but it actually conforms to the ordinary definition of it in two aspects.

On the one hand, Zhong's formula reveals an assumption he holds throughout his paper. That is, multiple realization is a specific kind of supervenience and the concept of supervenience and realization can be used alternatively. Zhong emphasizes that he will assume (reasonably) that mental properties are multiply realized by physical properties.[①] And he proves that the fact that M* is multiply realized by P* and Q* is compatible with the formula that any (~ M*)-world is also a (~ P*)-world, and any P*-world is also an M*-world.

Christensen and Kallestrup once argued that the notion the realization would raise challenges to the key step of Zhong's argument, that is, if M* counterfactually depends upon M, then P* would also counterfactually depend upon M. They claimed that whether you understand a physical realizer as core or total, it would make Zhong's argument invalid. Let's take $P_{total} = P_{core} + C$(C represents all the relevant physical background conditions, including properties concerning pertinent laws of nature). If you understand the physical realizer as core, there would be no P^*_{core} that is necessarily sufficient for the instantiation of M* because the physical laws may differ. If you understand it as total, the Downward Causation Principle will be problematic since it announces that M can cause laws of nature, which is absurd.

However, Zhong has defended himself recently, defining the physical realizer as primary, that is, M* is realized by $P^*_{primary}$---necessarily (in a nomological sense) any object that instantiates $P^*_{primary}$ also instantiates M*.[②] According to this definition, Zhong's argument can avoid the challenge about the necessity of realizer and

① Ibid, p.134.

② Zhong Lei: 'Why the Counterfactualist Should Still Worry About Downward Causation'. Erkenntnis, Vol 80, No.1, Feb 2015, p.169.

the causal problem of laws of nature as well. I agree with Zhong's defense that the usage of realization will not get his argument into hot water.

On the other hand, the formula he gives not only exhibits the nature of realization, but also shows the feature of one kind of supervenience. Kim once offered two classic definitions of supervenience. One of them concerns the relationship between the indiscernibility of two set of properties and reaches a simple slogan, that is, a set of properties A supervenes upon another set B iff "there cannot be an A-difference without a B-difference"[①]. The other one concerns the correlation of the existence of two set of properties and reaches the form we are all familiar with, that is, A weakly(or strongly) supervenes on B iff necessarily, if anything x has some property F in A, then there is at least one property G in B such that x has G, and (necessarily) everything that has G has F. [②]

Here, I don't want to discuss which definition of supervenience is better or more appropriate. It is not this paper's job. And I won't appeal to the rejection of superveniece to help the counterfactual autonomist solution out of the trap, as Dwayne did[③]. I will just figure out how Zhong understand the principle of supervenience and apply it.

In my view, Zhong's usage conforms to the strong version of superveniece, since he talks about the relationship of M and P across all possible worlds. The reason why he prefers to it may be that he wants to avoid superfluous trouble when he functions counterfactual on supervenience, for the discussion of the former is not limited to only one world. And the reason why he employs this form of definition rather than the ordinary one may be that he wants to unify the forms of counterfactual and supervenience in order to make his argument more obvious and convincing.

Either way, the formula itself is qualified to be a definition of supervenience and realization at the same time. And this specific definition will not bring any

① http://plato.stanford.edu/entries/supervenience/

② Jaegwon Kim: 'Concepts of Supervenience'. *Philosophy and Phenomenological Research*, Vol.45, December 1984, pp.163-165.

③ Moore mainly suggests that we can reject supervenience and then defend for counterfactualist autonomist, since Zhong's argument presupposes that the Superveniece Principle is true.

trouble to Zhong's argument. In the next section, I will discuss whether the usage of counterfactual will make Zhong's argument problematic.

3. The analysis of counterfactual

3.1 Assumption about counterfactual causation

After the analysis of superveniece, we should now move on to the other important concept in Zhong's argument – counterfactual. In this part, I want to reveal that Zhong actually has an implicit assumption when he gives out the key argument. And in the second part, I will try to demonstrate that this assumption is not appropriate and as a result, we need a renewed version of Zhong's argument. In the third part, I will use two methods to prove that the renewed version of argument is invalid. And I suggest that in order to defend his argument, Zhong should discuss more about his original assumption.

In his article, Zhong first illustrates Lewis' idea about counterfactual as followed:

> According to the standard analysis of counterfactual conditionals, the conditional "if A had not been instantiated, then B would not have been instantiated" is understood as "some world where neither A nor B is instantiated (i.e., (~A&~B)- world) is closer to the actual world than any world where A is not instantiated but B is instantiated (i.e., (~A&B)-world)" (Lewis 1973).[①]

Then, combining this conditional with the definition of counterfactual dependence, we can rewrite Lewis' definition into the following formula:

(i) A□→B(□→means counterfactual dependence in Lewis' sense) is true iff either (1) there are no possible A-world, or (2) some (A&B-

① Zhong Lei: 'Can Counterfactuals Solve the Exclusion Problem?'. *Philosophy and Phenomenological Research*, Vol. LXXXIII, No.1, July 2011, p.134.

world) is closer to the actual world than any (A& ~ B−world).

(ii) ~ A□→ ~ B is true iff either (1) there are no possible ~ A− world, or (2) some (~ A& ~ B−world) is closer to the actual world than any (~ A&B−world).

According to this formula, we should first revise the inference from step 4 to 5. Based on (ii), step 5 should be adapted to "if ~ M* counterfactually depends upon ~ M, then ~ P* would also counterfactually depends upon ~ M" so that it can match the inference from step 4. As (i) said, if Zhong really means "if M* counterfactually depends upon M, then P* would also counterfactually depend upon M", he should prove that if some (M&M*)-world is closer to the actual world than any (M& ~ M*) -world , then (M&P*) -world is closer to the actual world than any (M& ~ P*)-world. I believe it is a writing mistake. It indeed would not influence the whole argument. Now, we have a revised step 5 in form language :(~ M□→ ~ M*)→(~ M□→ ~ P*).

However, can we just skip from revised step 5 to the conclusion Zhong gives? In fact, there is an implicit assumption underlying the argument. That is, in Zhong's mind, A counterfactually causing B means that if A had not occurred, then B would not had occurred. That is, A causes B iff ~ A□→ ~ B. If this assumption is correct, then inferring from revised step 5, we can safely arrive Zhong's conclusion that if M causes M*, then M causes P*.

3.2 A renewed version

However, whether this assumption is correct worth discussing further. As I have mentioned above, since Zhong takes ~ A□→ ~ B as the sufficient and necessary condition of A causes B, he smoothly derives 'M causes P*' from 'M causes M*' through the following steps that he does not give out explicitly: M cause M* →(~ M□→ ~ M*)→(~ M□→ ~ P*)→M cause P*. But can we take ~ A□→ ~ B as the sufficient and necessary condition of A causes B?

In fact, David Lewis distinguished counterfactual dependence and causation early in 1973. He proposed that let c and e be two distinct possible particular events. Then e depends causally on c iff the family O(e), ~ O(e) depends counterfactually on the family O(c), ~O(c)… The dependence consists in the truth of two counterfactuals: O(c)□→O(e) and ~O(c) □→~O(e).[①] That is, we can briefly conclude the distinction between counterfactual dependence and causation:

Counterfactual Dependence: e counterfactually depends on c iff O(c) □→O(e).

Causation: e is caused by c iff O(c)□→O(e) and ~O(c) □→~O(e).

Nevertheless, many philosophers ignore the first conditional of causation since Lewis took it for granted and focused his attention on ~O(c) □→~O(e). But Lewis did not mean that the first conditional should be ignored under all circumstance. He just claimed that if c and e are actual events, then it is the first counterfactual that is automatically true. Then e depends causally on c iff, if c had not been, e never had existed.[②] That is to say, if we want to ignore the first conditional, we must prove that both of the causal relatum should actually happen.

Therefore, if Zhong wants to keep the assumption and skip from "~M□→~P*" to "M cause P*", he has to prove that M and P* are actual. However, we can conceive that in the actual world, M causes M* and they are realized by Q and Q*, respectively. In this case, M and P* are not both actual. That is to say, the case that M causes M* cannot guarantees that M and P* are both actual. As a result, when we examine the causal relation between M and P*, we cannot ignore the first conditional.

Someone may raise a question: since the first conditional is indispensable, do we have a chance to ignore the second conditional? According to Lewis, if c and e do not actually occur, then the second counterfactual is automatically true: so e depends causally on c iff the first counterfactual holds. However, as I mentioned above, we can conceive an actual world, in which M occurs and P* not occurs. As a result, in an actual world that M causes M*, if we want to examine whether M cause P*, we need to prove both of the conditionals. That is, M□→P* and ~M□

① Lewis: 'Causation'. *Journal of Philosophy*, Vol. 70, No. 17, 1973, p.563.

② Ibid.

→ ~ P*. Since Zhong has successfully proved that (~M□→~M*)→(~M□→~P*), he also need to prove that (M□→M*)→(M□→P*), in order to get to the conclusion that M causes P*.

To sum up, we need a renewed version of Zhong's key argument as followed:

1'. M* is realized by P*, that is, any (~M*)−world is also a (~P*)−world, and any P*−world is also an M*−world.

2'. A (~M& ~M*)−world is at the same time a (~M& ~P*)−world.

3'. (~M&P*)−worlds are a subset of (~M&M*)−worlds.

4'. A (M&~M*)−world is at the same time a (M&~P*)−world.

5'. (M&P*)−worlds are a subset of (M&M*)−worlds.

From 2' and 3', we have:

6'. If some (~M& ~M*)−world is closer to the actual world than any (~M&M*)−world , then (~M& ~P*)−world is closer to the actual world than any (~M&P*)−world.

7'. So, if ~M* counterfactually depends upon ~M, then ~P* would also counterfactually depend upon ~M.

From 4' and 5', we have:

8'. If some (M&M*)−world is closer to the actual world than any (M& ~M*)−world , then (M&P*)−world is closer to the actual world than any (M&~P*)−world.

9'. So, if M* counterfactually depends upon M, then P* would also counterfactually depend upon M.

10'. A causes B iff B counterfactually depends on A and ~B counterfactually depends on ~A.

From 7', 9' and 10', we have:

Conclusion: Counterfactual theory of causation and supervenience together imply downward causation from M to P*.

In the next part, I will try to figure out that whether we can successfully deduce step 8' and 9' from step 4' and 5'.

3.3 Deduction problem

We should begin with the step 1', 4' and 5', since the whole argument depends on them. I will symbolize them slightly in order to make the relationship they indicate more explicit.

1'. M* is supervenieced by P*, that is, any (~ M*)-world is also a (~ P*)-world, and any P*-world is also an M*-world.

4'. A (M& ~ M*)-world = a (M& ~ P*)-world.

5'. (M&P*)-worlds ⊂ (M&M*)-worlds.

As described in step 1', the relationship between (~ M*)-world and (~ P*)-world should be the same as that between P*-world and M*-world. So, if a (~ M& ~ M*) -world = a (~ M& ~ P*) -world, then a(~ M&P*) -world = a (~ M&M*)-world. And if (~ M&P*)-worlds ⊂ (~ M&M*)-worlds, then (~ M& ~ M*)-worlds ⊂ (~ M& ~ P*)-worlds. We have to choose between these two situation and not adopt "=" and "⊂ " at the same time as Zhong does.

Then, which relationship should we choose? What does Zhong mean when he said any (~ M*)-world is also a (~ P*)-world? Obviously, to complete this sentence, we get this: any (~ M*)-world is also a (~ P*)-world, but not vice versa. Because if M* is multiply realized by P* and Q*, we can have some (~ P*)-world being also a M*-world. For this reason, I think we should not identify (~ M*)-world with (~ P*)-world as well as P*-world with Q*-world. Therefore, equation is out of the game.

Now, we have (M&P*)-worlds ⊂ (M&M*)-worlds and (M& ~ M*)-worlds ⊂ (M& ~ P*)-worlds. Then, combining these two premises with the antecedent of step 8', we find it hard to deduce that (M&P*)-world is closer to the actual world than any (M& ~ P*)-world since it is possible for a (M& ~ P*)-world to be closer to the actual world than all (M&P*)-worlds. Consequently, we fail to achieve M□→P* and M causes P* as well.

Someone may object that the proof above is weak, so let's turn to validity proofs in Lewis-counterfactuals system for help and try to construct a countermodel. Our purpose is to certify that (M□→M*)→(M□→P*). Given that M* is multiply realized, we assume that M* is realized by P* or Q* only, then we get (M□→(P*∨Q*))→(M□→P*) instead. Now, we can investigate whether this formula valid in Lewis's system.

An LC-model is a three-tuple $\langle W, \leq, V\rangle$, where:

i) W is a nonempty set

ii) V is a function that assigns either 0 or 1 to each wff relative to each member of W

iii) $\leq$ is a three−place relation over W("nearness relation"; read "$x \leq_w y$" as "x is at least as near to/similar to w as is y".)

iv) V and $\leq$ satisfy the following conditions:

C1: for any w, $\leq_w$ is strongly connected

C2: for any w, $\leq_w$ is transitive

C3: for any x, y, if $y \leq_w x$ then $x=y$ ("base")

v) For all wffs, φ,ψ and for all $w \in W$:

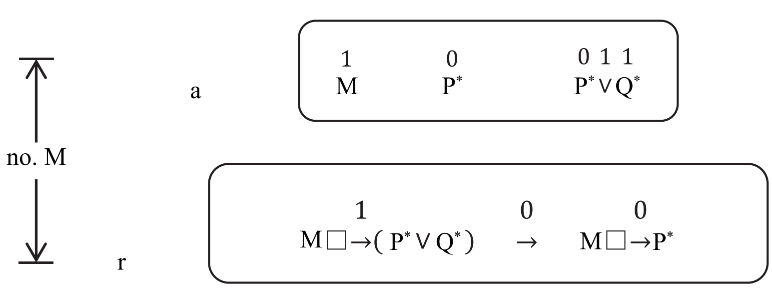
a) $V(\sim\varphi,w)=1$ iff $V(\varphi,w)=0$

b) $V(\varphi\rightarrow\psi,w)=1$ iff either $V(\varphi,w)=0$ or $V(\psi,w)=1$

c) $V(\Box\varphi,w)=1$ iff for any v, $V(\varphi,v)=1$

d) $V(\varphi\Box\!\!\rightarrow\psi,w)=1$ iff either φ is true at no worlds, or: there is some world, x, such that $V(\varphi,x)=1$ and for all y, if $y \leq_w x$ then $V(\varphi\rightarrow\psi,y)=1$ ①

In this particular case, this formula is invalid, as the following countermodel shows:



We first suppose for reduction that $(M\Box\!\!\rightarrow(P^*\vee Q^*))\rightarrow(M\Box\!\!\rightarrow P^*)$ is false at some world r. Therefore, $M\Box\!\!\rightarrow(P^*\vee Q^*)$ is true at r, but $(M\Box\!\!\rightarrow P^*)$ is false at r. I begin with the false subjunctive: $(M\Box\!\!\rightarrow P^*)$. This forces the existence of a nearest

① Ted Sider: *Logic for Philosophy*, Oxford University Press, 2010, p.220.

M-world, named a, in which P* is false. Since M is true there, this rules out the true $M\Box\!\!\rightarrow(P^*\vee Q^*)$ in r being vacuously true and rules out the case that M is not true in any world. Then, given $V(M\Box\!\!\rightarrow(P^*\vee Q^*),r)=1$, for every x, if x is a nearest-to-r M-world, then $V(P^*\vee Q^*,x)=1$. It is easy to find that world a is a nearest-to-r M-world. Therefore, $V(P^*\vee Q^*,a)=1$. Since P* is false at a, Q* must be true there. Here's the official model:

$W=\{r,a\}$

$\leq_r=\{\langle r,a\rangle,\langle r,r\rangle,\langle a,a\rangle\}$

$V(M,a)=V(Q^*,a)=1$; all other atomics 0 everywhere else.

Therefore, I have proved that $(M\Box\!\!\rightarrow(P^*\vee Q^*))\rightarrow(M\Box\!\!\rightarrow P^*)$ is invalid in Lewis' system.

As a result, either we use Zhong's argument structure or Lewis' system, we cannot prove that $(M\Box\!\!\rightarrow M^*)\rightarrow(M\Box\!\!\rightarrow P^*)$. In another word, we cannot deduce step 8' and 9' and cannot harshly conclude that supervenience and counterfactual theory will entail downward causation.

4. Conclusion

I have tried to prove that combining the concept of supervenience and counterfactual theory of causation will not surely imply that there necessarily exists downward causation between M and P*. However, I am not saying there is no downward causation at all in all cases. That is another topic, which needs much more research. In this paper, what I want to argue is that we can hold supervenience and counterfactual theory of causation without worrying about the occurrence of downward causation. If Zhong insists that downward causation will occur, it is his burden to prove that we don't need the conditional $M\Box\!\!\rightarrow P^*$ to examine the causal relation between M and P*.

I have mentioned that autonomist solution is the one who is threatened by Zhong. Now, since we have removed the threat, we can save autonomist solution for a while and regard it as successful in preserving mental causation and irreducibility because it can be constructed as followed:

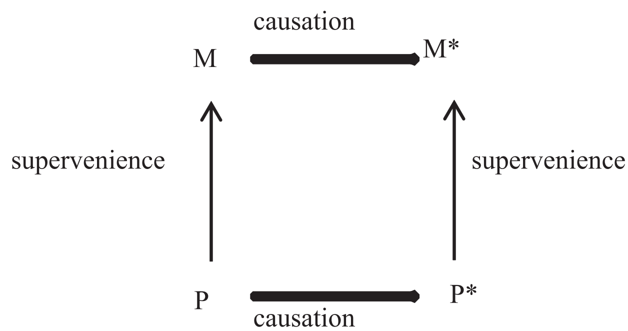


Fig. 1

Fig. 1 As the figure shows, autonomist solution saves the autonomy of mental property while preventing it from floating out of the scope of physics by using supervenience to correlate mental property with physical property. Besides, it does not violate the causal closure of physics since nothing outside physics will break into the causal chain of physics. In addition, there is no overdetermination as well because causal relation exists only in horizontal direction, not vertical. And thanks to the elimination of downward causation, we won't worry that P* has two sufficient causes, M and P.

Therefore, if we successfully reject Zhong's argument, we can save the counterfactual theory of causation as one of the solutions, which can block the exclusion argument. In other word, using counterfactual theory will not definitely lead us to downward causation, which is a premise of exclusion argument. As far as we can see, it seems that autonomist solution, whose assumptions contain supervenience and counterfactual theory, sounds reasonable and valid.

Works Cited

Amie Thomasson: 'A Nonreductivist Solution to Mental Causation'. *Philosophical Studies*, Vol.89, 1998, pp.181-195.

Brian P. McLaughlin: 'Varieties of Supervenience'. In Elias E. Savellos & U. Yalcin (eds.), *Supervenience: New Essays*, Cambridge University Press, pp.16-59.

David Lewis: 'Causation'. *Journal of Philosophy*, Vol. 70, No. 17, 1973, pp.556-567.

Dwayne Moore：‘Counterfacutals, Autonomy and Downward Causation：Reply to Zhong’. *Philosophia*, Vol. 41, Issue 3, September 2013, pp.831-839.

Jaegwon Kim：‘Concepts of Supervenience’. *Philosophy and Phenomenological Research* Vol.45, December 1984, pp.153-176.

----：*Mind in a Physical World：an essay on the mind–body problem and mental causation*, The MIT Press, 2000.

John Gibbons：‘Mental Causation without Downward Causation’. *The Philosophical Review*, Vol. 115, No.1, January 2006, pp.79-103.

Jonas Christensen and Jesper Kallestrup：‘Counterfactuals and downward causation’. *Analysis*, Vol. 72, Num. 3, July 2012, pp.513-517.

Markus E. Schlosser：‘Non-reductive Physicalism Mental Causation and the Nature of Actions’. In H. Leitgeb & A. Hieke（eds.）, *Reduction：Between the Mind and the Brain*, Ontos, 2009, pp.73-89.

Ted Sider：*Logic for Philosophy*, Oxford University Press, 2010.

Terence E. Horgan：‘From supervenience to superdupervenience：meeting the demands of a material world’. *Mind*. Vol.102, No.408, October 1993, pp.555-86.

Zhong Lei：‘Can Counterfactuals Solve the Exclusion Problem?’. *Philosophy and Phenomenological Research*, Vol. LXXXIII, No.1, July 2011, pp.129-147.

----‘Why the Counterfactualist Should Still Worry About Downward Causation’. Erkenntnis, Vol 80, No.1, Feb 2015, pp.159-171.