

Causation as Agency in Modal Meinongianism

Stelian Mădălin Mihai

Faculty of Philosophy, University of Bucharest, Bucharest, Romania

Email: stelianmadalinmihai@gmail.com

How to cite this paper: Mihai, S. M. (2024). Causation as Agency in Modal Meinongianism. *Open Journal of Philosophy*, 14, 805-819.

<https://doi.org/10.4236/ojpp.2024.144054>

Received: July 15, 2024

Accepted: October 9, 2024

Published: October 12, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper, I am going to explore an alternative explanation of causation in Graham Priest's modal meinongianism. Priest proposes an understanding of causation, which is either too confusing, or against the metaphysical core of modal meinongianism. In his proposals, causation is discussed in the context of defining purely fictional and abstract objects, by using a counterfactual approach. In this case, causation is understood as an existence-entailing relation. I will argue that such an account of causation proves ineffective. Instead, I will give another interpretation for causation, which is based on Peter Menzies and Huw Price's version of agency theory of causation. My interest in studying such a particular subject is to find an answer which is to be used in a further research. The end result of my work will be the fact that causation as agency can, at best, be useful in understanding intentional acts, but not suitable for explaining what it means for a particular object to causally interact with us.

Keywords

Possible Worlds, Impossible Worlds, Causation, Agency, Fictional, Abstract

1. Priest's Modal Meinongianism—A Better Principle of Characterization

In *Towards Non-Being* (2005), Graham Priest proposes a modal version of meinongianism. His purpose is to construct a better Principle of Characterization. Initially, the Principle of Characterization was devised by Alexius Meinong, in order to state that every object, irrespective of its ontological status, has all of its properties. This was called the “naïve version”, and the reasons for such a label were that it had had some unpleasant consequences, such as the possibility to construct any object whatsoever and assign existence to it. Even if Meinong's legacy was under heavy criticism, coming from Bertrand Russell and W.V.O. Quine, some thinkers considered that its spirit could be salvaged. In this context, there

were several attempts to resurrect Meinong's philosophical insight. Modal meinongianism could be considered one of the most recent attempts to improve the Principle of Characterization.

The Principle of Characterization allows every object to have all of its characterizing properties. Meinong's proposal implied that an object could be characterized in any way whatsoever, and that its ontological status is irrelevant in this endeavor. More precisely, one can characterize an object by choosing any given set of properties. For instance, one can choose the following set of properties $S = \{x \text{ is detective, } x \text{ lives in Baker Street, } x \text{ has unusual powers of observation and inference, } x \text{ exists}\}$, and construct an object. The conjunction of those properties is obviously referring to Sherlock Holmes. Now, the problem is that, according to Graham Priest, the Principle of Characterization "cannot be correct in full generality. If it were, not only could one run the ontological argument to prove the existence of God, and everything else, one could, in fact, prove everything" (Priest, 2005: pp. 7-8).

So, this could be the starting point. The Principle of Characterization, in its initial form, leads to some unsavory consequences. In this circumstances, Priest's proposal is an attempt to make it better. But why would he want to do such an improvement? Obviously, this is an interesting philosophical issue, but one might also believe that choosing such a particular subject implies a personal motivation as well. The desire to improve the Principle of Characterization is a means to an end, since Priest wanted to embrace a metaphysical perspective, called "noneism", and a reasonable way in which he had considered that such a perspective could be accepted was to have a consistent or satisfying version of the Principle of Characterization. In his own words, "I first became attracted to noneism when I found an approach to the characterization problem that I found plausible" (Priest, 2005: p. 8).

Noneism is proposed by Richard Routley, as an alternative metaphysical perspective about existence. The traditional perspective about existence has reached its final version in Quine's aphoristic expression "to be is to be the value of a variable", bound to a quantifier (Quine, 1948: p. 9). Such a perspective about existence is based on an assumption, which is "the prejudice in favor of the actual". Noneism is a rejection of such a prejudice. Routley talks openly about noneism in *Exploring Meinong's Jungle and Beyond* (2008). Here, noneism is largely defined by proposing seven theses. These theses are inspired from Meinong's theory of objects. Ultimately, noneism can be shortly defined as the metaphysical perspective, according to which not only that some objects do not exist, but also that it is possible to consistently talk about non-existent objects (Routley, 2018: pp. 3-4).

Now, even if noneism claims that we can talk about non-existent objects, it still needs to propose an explanation for the possibility of non-existent objects to bear properties. Routley himself considered that Meinong's Principle of Characterization was the proper solution to such a predicament. But nonetheless its initial form is not well equipped to accomplish this task, due to its unavoidable

shortcomings. In this context, if Priest wanted to accept noneism, he would only do it if he had a better version of the Principle of Characterization. And so, he proposed a better version. His novelty consists in saying that the Principle of Characterization “can hold unrestrictedly, provided only that its instances may hold, not at this world, but at others” (Priest, 2005: p. 8).

In this way, the Principle of Characterization receives a modal version. In order to understand how this modal version has been built, I will shortly present Priest’s semantics of intentionality.

2. Semantics of Intentionality—Possible and Impossible Worlds

Priest begins the presentation of this semantics by choosing a neutral stance about intentionality. Although he considers that intentionality is a fundamental feature of cognition, he does not explicitly take any side regarding the ontological status of the intentional objects. All he says is that ‘an intuitive understanding is quite sufficient’ (Priest, 2005: p. 5), given his purpose. His purpose is to build a formal device in order to be able to improve the Principle of Characterization. In this endeavor, intentionality is just the ability of thought to direct itself towards any object.

The formal system Priest builds so that he could construct a better version of the Principle of Characterization (CP) presupposes making use of the following elements: neutral quantification, a first-order predicate of existence and world semantics. In order to explain how such elements make up this formal system, it is mandatory to begin with a standard first-order formal language, L , which is enriched by adding modal operators (\Box , \Diamond), and a collection of intentional operators (Φ , Ψ , $X\dots$), such that, where t is a term and A is a well-formed formula, $t\Psi A$ reads “ t Ψ that A ”. If Ψ stands for an intentional act, such as expressing hope, then $t\Psi A$ could specifically mean that “ t hopes that A ”. In the class of intentional operators, a special place is occupied by the representational operator Φ , such that $t\Phi A$ reads “ t represents A as holding” (Ciprotti, 2014: p. 7).

Now, upon this basic syntax, Priest adds neutral quantification, a first-order predicate of existence and world semantics. In order to show the neutrality of quantification, Priest uses a different set of symbols for quantifiers (Π , Σ). The universally quantified formulas such as “ $\Pi x\varphi x$ ” read “All x s satisfy φ ”, and existentially quantified formulas such as “ $\Sigma x\varphi x$ ” read “Some x satisfies φ ”. If they are neutral, existence is not expressed by their usage. Existence is expressed by a first-order predicate, Ex . In this case, a statement like “ $\Sigma x(Ex\&\varphi x)$ ” is saying that something exists that satisfies $\varphi(x)$. Lastly, the introduction of world semantics implies that the evaluation of well-formed formulas depends both on the “assignments to variables in respect to interpretations and in respect to different worlds” (Ciprotti, 2014: pp. 7-8).

What makes this semantics of intentionality special is that it does not limit itself to the usage of only possible worlds, but it employs impossible worlds as well. In

what concerns the possible worlds, this semantics is a constant-domain modal S5. This means that “every object in the domain falls within the scope of Π and Σ , irrespective of the world” (Ciprotti, 2014: p. 8). Also, it is important to add that the existence of objects is not decided by the usage of quantifiers. An object exists in a given world only if it satisfies the first-order predicate of existence. Lastly, the final interpretation in this semantics is as follows: $M = \{P, @, I, D, \delta\}$, where “ P ” stands for possible worlds, “ $@$ ” is the actual world, and makes part of the set of possible worlds, “ I ” represents the set of impossible worlds ($P \cap I = \emptyset$), D is a non-empty set that stands for the general domain of L , and δ is the evaluation function, which assigns a denotation to every non-logical expression (Ciprotti, 2014: p. 8).

The introduction of impossible worlds serves a couple of purposes, such as solving some issues related to epistemic logic, but more importantly the introduction of impossible worlds is needed in order to be able to make a better CP. If intentionality is the ability of thought to direct itself towards any kind of object, and our speech might imply the usage of impossible objects, then in order to give a full account of intentionality it becomes necessary to include in this semantics impossible worlds as well. For instance, if I have the following intentional act, expressed in the sentence “I imagine that a round square exists”, this semantics must be able to logically represent it. Given its formal ingredients, such a statement might have the following logical interpretation:

$$@ \Vdash^+ a\Phi A(c_A). \text{ But } @ \Vdash^- A(c_A), \text{ which means that } @ R_{\Phi}^{\delta(a)} w, w \Vdash^+ A(c_A)$$

This roughly translates as follows: at the actual world, $@$, an agent (a) is having a representational act (Φ - imagines that), by which a round square object exists ($A(c_A)$). But such a set of properties does not hold at the actual world ($@ \Vdash^- A(c_A)$). Nevertheless, by dint of the accessibility relation, there is a non-actual world, w , where it is true that a round square exists. The problem that Priest identified at the Principle of Characterization, in its initial form, was that it could not have been accepted in full generality. Now, if impossible worlds are included in this semantics, it can be said in any case that ‘we just do not assume that an object characterized in a certain way has its characterizing properties at the actual world, only at the worlds which realize the way the agent represents things to be in the case at hand’ (Priest, 2005: p. 85). And so, a better version of CP has been achieved. Thus, every intentional act can be logically expressed and the unpleasant consequences of the initial CP are avoided.

If a better version of the Principle of Characterization has been achieved, then noneism can be accepted. If noneism is the general framework of Priest’s proposal, this might mean that noneism is the decisive factor in establishing the ontological categories that modal Meinongianism works with. In particular, when talking about the nature of non-actual worlds, Priest claims that ‘in the context of noneism, the obvious policy is to take all worlds other than the actual to be non-existent objects’ (Priest, 2005: p. 139). In the category of non-existent objects, Priest includes purely fictional objects and abstract objects as well (Priest, 2005: p. 135). For the purpose of this paper, in what follows I will be concerned only with

the intricacies of the distinction fictional-abstract.

3. Fictional and Abstract, as Two Different Modes of Non-Existence

The first useful consideration might be the distinction between fictional objects and purely fictional objects. Priest considers that some fictional objects, such as Napoleon in Tolstoy's *War and Peace* is an existent character. In this case, fictional objects which are real or have been existing are just fictional, whereas purely fictional objects are the objects that do not exist (Priest, 2005: p. 116). In this discussion, it will matter only the purely fictional objects, and for the sake of words efficiency, whenever I will use the expression 'fictional objects', I will be talking about purely fictional objects. Now, fictional objects do not exhaust the category of non-existent objects, because "Mathematical and other abstract objects can be taken to form another large class" (Priest, 2005: p. 124).

When talking about abstract objects, Priest begins with a very natural task, which is to determine their meaning. The starting point of such a resolve implies making a difference between fictional and abstract objects. Given noneism, "a major difference between standard purely fictional objects and abstract objects would seem to be in the *mode* of their existential status" (Priest, 2005: p. 136). If Zeus, for instance, does not exist, but it could have existed, an abstract object, like number 3 for instance, not only that it does not exist, but it necessarily does not exist. But such an account of abstract objects might be insufficient, since one can construct possible objects which have inconsistent properties. If such objects might be conceived, then an abstract object might have the possibility of existing, which is not desirable. So, Priest reaches another proposal, in which a counterfactual perspective is being use to define abstract objects. In this context, Priest claims that "an abstract object is one such that, *if it did exist it would still not causally interact with us*" (Priest, 2005: p. 136).

A fictional object, instead, is predisposed to having a different understanding, given the same counterfactual approach. This understanding can be inferred by the example that Priest gives when talking about Sherlock Holmes. Priest says that "Holmes, were he to have existed, would have entered into causal chains with us" (Priest, 2005: p. 137). From such a claim, it can be derived a general statement, which says that a fictional object is an object such that *if it did exist it would causally interact with us*. Both fictional and abstract objects are non-existent objects, which means that firstly they do not exist, and afterwards they could exist or cannot exist. In such a case, then maybe a logical representation could be reached. If existence is expressed by the first-order predicate of existence, *Ex*, then a fictional object is logically the following conjunction " $\sim Ex \& \Box \Diamond Ex$ ". The reason behind choosing " $\Box \Diamond Ex$ " to express the fact that an object might exist is to make sure that it is just possible, without actually entering the set of impossible worlds. If, for instance, Zeus is a fictional object, Zeus does not exist, and necessarily might have existed. In this case, it is guaranteed that there is at least one possible world which

includes a fictional object, without taking into account the actual world. So, fictional non-existence can be logically expressed as “ $\sim Ex \& \Box \Diamond Ex$ ”.

If a fictional object does not exist, but it might have existed, then it can causally interact with us. If an abstract object does not exist, and it necessarily cannot exist, it cannot causally interact with us. This might be the extended understanding of the counterfactual reading. If this is right, then this reading can be used to give some logical representations. Now, if the general definition of abstract objects, for instance, can be logically rephrased as ‘ $Ex \rightarrow \sim Cx$ ’, where ‘ Cx ’ stands for ‘causally interact with us’, the antecedent can be replaced with what it actually means for an abstract object to not exist. In this case, ‘ Ex ’ is to be replaced with ‘ $\sim Ex \& \Box \sim Ex$ ’. In consequence, the precise logical expression of a counterfactual representation of abstract non-existence is:

$$(NA): (\sim Ex \& \Box \sim Ex) \rightarrow \sim Cx;$$

The same reasoning can be applied in the case of fictional objects, and the end result be the following:

$$(NF): (\sim Ex \& \Box \Diamond Ex) \rightarrow Cx;$$

Such a logical representation of fictional and abstract non-existence goes against Priest’s proposal, since he devised the distinction between purely fictional objects and abstract objects as not exhaustive, meaning that some object can be at the same time purely fictional and abstract (Priest, 2005: p. 137). In my interpretation, I choose to make this distinction exhaustive, meaning that no fictional object can be abstract and no abstract object can be fictional. Maybe a methodological easiness is preferable in the context in acquiring knowledge about this topic. If this distinction can be understood more easily by taking it as exhaustive, then it might be justified to do just that. So, for the sake of understanding or simplicity, I will assume that the distinction fictional-abstract is exhaustive.

If no fictional object can be abstract object and no abstract object can be fictional object, then (NA) and (NF) should be considered tautologies. Otherwise, one can build an interpretation, in which an impossible object might causally interact with us. So, if I am working with the assumption that a non-existent object can either be fictional, or abstract, then those logical representation of fictional non-existence and abstract non-existence must be proved to be true in all cases. And a proof might be devised in the following manner. Let us firstly take the case of fictional non-existence.

$$(NF): (\sim Ex \& \Box \Diamond Ex) \rightarrow Cx$$

The logical representation of fictional non-existence is governed by the logical operator of implication. For simplicity, (NF) might be rephrased as ‘ $A \rightarrow B$ ’. Now, $A \rightarrow B$ is false in just one case, which is when the antecedent is true and the consequent is false. There are several ways in which a material implication can be transformed into a tautology, since, by default, it is not. One of them is to show that an equivalent formula is a tautology. So, the general strategy of reasoning is to find a logically equivalent formula, to prove it a tautology and to deduce that the initial formula is also a tautology, given the relation of equivalence.

So, (NF) can shortly be written as ' $A \rightarrow B$ '. To reach a logically equivalent formula, one can use contraposition. By contraposition, ' $A \rightarrow B$ ' becomes ' $\sim B \rightarrow \sim A$ '. Returning to the actual implication that defines fictional non-existence, given the usage of contraposition, (NF) becomes:

$$(NF'): \sim Cx \rightarrow \sim(\sim Ex \& \Box \neg Ex)$$

And, in order to make it always true, the easiest way is to assume that " $\sim Cx$ " is always false. If " $\sim Cx$ " is always false, then " Cx " is always true. So, (NF') is a tautology if " Cx " is always true. This means that (NF) is a tautology for the same reason. Ultimately, in order to make sure that fictional non-existence is logically well represented as a tautology, " Cx " must always be true.

The same line of reasoning can be applied for (NA). By contraposition, (NA) becomes:

$$(NA'): Cx \rightarrow \sim(\sim Ex \& \Box \sim Ex)$$

In order to make sure that it is a tautology, then " Cx " must always be false. If it is assumed that " Cx " is always false, then (NA) also becomes a tautology, given the same justification. So, abstract non-existence can logically be represented as a tautology if " Cx " is always false.

In the end, the main difference between fictional non-existence and abstract non-existence is that, in the case of fictional objects they must always have the possibility of causally interact with us, whereas abstract object must never have the possibility of causally interact with us. But what is the meaning of this relation of causality? According to Priest, causation is 'an existence-entailing relation' (Priest, 2005: p. 141). And this leads us to shortly presenting the aforementioned relation.

4. Causation as Existence-Entailing

The distinction between predicates that entails existence and predicates that do not follows from assuming noneism. If some objects do not exist, then it is quite fitting to propose a distinction between those two types of predicates. In this case, there are predicates which entail existence, such as "hug", and if a hugs b , then both a and b must exist. Also, intentional predicates are existence-entailing. If a worships b , then a must exist, but b may or may not exist. Furthermore, there are some non-intentional predicates that do not entail existence, such is identity. An object is identical with itself, irrespective of its existence. For other predicates, it could be debatable if they are existence-entailing or not. One example that stands as a justification for such an observation could be a case in which a experiences b , and b is just an illusion. In this case, Priest does not provide an answer, but only a question – "does it follow that b exists?" (Priest, 2005: p. 60).

Lastly, Priest mentions that it is not his purpose to further debate or discuss this topic. What he adds to this discussion is that the following constraint, "*if* ($q_1 \dots q_i \dots q_n$) $\in \delta^*$ ($P, @$) then $q_i \in \delta^*$ ($E, @$)" – applies only at the actual world. So, if an object has existence-entailing properties, that object must be part of the domain of the actual world. If this is so, how could one understand the following claim: As a matter of fact, it seems to me that existence-entailing is world-

invariant, at least at possible worlds. Thus, for example, if a hits b at such a world, w , then a and b exist at w (Priest, 2005: p. 60). It seems a bit strange, because existence-entailing was supposed to apply strictly to the actual world, but if in a possible world a predicate implies a physical interaction between two objects, then those objects must exist in this possible world. Given such considerations, this distinction between predicates that entail existence and predicates that do not entail existence is somehow unclear or unfinished.

One way to understand such a distinction is to use some examples and see more specifically how it works. For instance, “if Santa Claus thinks about Pegasus at w , then Santa Claus exists at w though Pegasus need not; if Holmes kicks Moriarty at w , then both exist at w ; and this holds whenever w is a possible world. What happens at impossible worlds, of course, is another story; nonexistent things, for instance, may think or kick at some impossible world” (Berto, 2011: p. 16). Given such an interpretation, it is still difficult to grasp the meaning of existence-entailing, because another distinction is used. More precisely, existence-entailing behaves differently at different types of worlds. In possible worlds, an object may have existence-entailing properties, which will lead to acknowledging that it exists. At impossible worlds, it is possible to have non-existent objects that possess existence-entailing properties. I suspect that the actual difference is between the way a non-existent object does not exist. In any possible world, an object non-existence is contingent, meaning that it could have been part of the actual world. In an impossible world, an object is necessarily non-existent, meaning that it could never have been part of the actual world. Furthermore, if it is taken into account the difference between fictional and abstract objects, it could be said that fictional objects possess existence-entailing properties, whereas abstract objects (being necessarily non-existent) might also possess existence-entailing properties, but only at impossible worlds, where everything can hold.

Maybe the distinction between predicates that entail existence and predicates that do not is not so clear after all. And the lack of clarity comes from not having a precise criterion of demarcation. As Francesco Berto says, ‘The distinction between properties that entail existence and properties that do not is, no doubt, merely intuitive. I wouldn’t be able to indicate a general criterion to distinguish between the two kinds, and I guess people have different and fuzzy views on this’ (Berto, 2012: p. 150). So, how can one use the idea of existence-entailing to explain another difficult concept, such as causation? At first sight, if causation is existence-entailing, and to exist is to be concrete (Priest, 2005: p. 136), one way of understanding this is to claim that possible objects need to be concrete, since causation is a relation between two existing things, or concrete things. More precisely, if Santa Clause exists at a possible world, he must be concrete, if it has the capacity of causally interact with us, “us” being the actual world. In this case, causation implies an interaction between objects of the same ontological nature, which is being concrete or material. In the case of impossible objects, because they cannot ever exist, they will never have the capacity to causally interact with us. This means

that they can never be concrete or material. But, as it has been implied, necessarily non-existent objects might have existence-entailing properties.

If an impossible object, which is necessarily a non-existent object, has existence-entailing properties, does not this mean that it exists? So, an impossible object is both existent and non-existent. This might be a strange conclusion to arrive at, since the metaphysical framework of Priest's proposal is noneism. If noneism claims that some objects do not exist, then it is difficult to accept that some objects that do not exist. Therefore, in this interpretation, if causation is understood as existence-entailing, then it leads to infringing the metaphysical core of modal meinongianism, since the distinction between existent and non-existent blurs or fades away. Overall, if the notion of existence-entailing is used to explain causation, it can be claimed that either this attempt is compromised from the beginning, since it implies explaining a difficult concept by using another difficult concept, or, in a given interpretation, causation as existence-entailing has the unwanted consequence of breaking the metaphysical core of modal meinongianism.

Another particular reason to avoid this path of conceiving causation as existence-entailing might very well come from the interpretation I have proposed for the distinction between fictional and abstract objects. In my interpretation, it is assumed that such a distinction is exhaustive, meaning that no fictional object can ever be abstract and no abstract object can ever be fictional. Let us mention again the conclusion of the previous section:

(NF): $(\sim Ex \& \Box \Diamond Ex) \rightarrow Cx$, Cx always true;

(NA): $(\sim Ex \& \Box \sim Ex) \rightarrow \sim Cx$, Cx always false;

In the case of fictional objects, if ' Cx ' must always be true, ' $(\sim Ex \& \Box \Diamond Ex)$ ' also must always be true, since this is required for considering that (NF) is a tautology. Given the meaning of the antecedent, a fictional object does not exist at the actual world, but it necessarily makes part of a random possible world. So, in a possible world, it exists. If it exists, then it is concrete. If it is concrete, it can causally interact with us. Moreover, the possible world in which that objects exist must also exist or be concrete, since there is this pretense of ontological sameness between a world and its denizens. Therefore, possible worlds are concrete. In what concerns the abstract objects, they do not exist at the actual world and cannot exist in any possible world. They exist at an impossible world, but they can never causally interact with us. But if they exist at an impossible world, they are concrete, since to exist is to be concrete. If this is so, an impossible object is both abstract and concrete, which might be the case, given a certain understanding of impossibility. Nevertheless, if an abstract object can be concrete, then the class of abstract objects becomes more spacious or changes its extension. If this is the case, then the given logical understanding of abstract non-existence has to be reviewed. Or, I do not want to change it, since the entire interpretation depends on it.

All in all, given the assumption that I am working with, another reason has been provided for avoiding the attempt to understand causation as existence-entailing. In what follows, I will propose another understanding of causation, which might

be more suitable for the given interpretation of the distinction between fictional and abstract objects.

5. Causation as Agency

In presenting this topic, I will be using Peter Menzies and Huw Price's version of agency theory of causation. In their paper, "Causation as a Secondary Quality" (1993), causation is defined as follows: "an event A is a cause of a distinct event B just in case bringing about the occurrence of A would be an effective means by which a free agent could bring about the occurrence of B " (Menzies & Price, 1993: p. 187). In this definition, a means-end relation is being used, which is to be understood as "agent probabilities". Menzies and Huw explain that agent probabilities work as conditional probabilities, "assessed from an agent's perspective under the supposition that the antecedent condition is realized *ab initio*, as a free act of the agent concerned" (Menzies & Price, 1993: p. 190).

The formal notation for conditional probabilities is " $P_A(B)$ ", which shows the probability that B would happen, if an agent chose to realize A . In other words, agent probability, devised as conditional probability, rests on the agent's rational capacities, which are employed in deciding that A , as a means, is sufficient or effective in accomplishing B , as a goal. In consequence, the ability of decision-making is detrimental to defining and understanding agent probabilities (Menzies & Price, 1993: p. 190). This might become more intelligible if, for instance, a concrete situation is presented.

Let us suppose that an agent, Mike, finds himself in the context of making a decision about his broken car. He understands the problem, but he is not sure whether to fix it himself or go to a professional service. Nevertheless, the problem presents itself as a conditional. More precisely, Mike understands that if the car stayed broken, he would not be able to use it or, in other words, he considers what would happen under different choices. Mike's purpose is to have a functional car, and for such a purpose he has two means, namely to fix it himself or to have it fixed by a mechanic. In this case, Mike might judge the efficiency of both means and ultimately decides that having the care fixed by a mechanic is the most efficient way in which he can have his car functional again. If A is the choice to fix it himself, and A' is the choice to have it fixed by a mechanic, then in this case Mike evaluates A' as being more effective or having a greater probability of achieving his purpose. So, because $P_{A'}(B) > P_A(B)$, then $P_{A'}(B)$. In terms of causation, the effect B , which is having the car functional, is achieved by cause A' , which is taking the car to a mechanic to have it fixed. In this scenario, it can be observed that causation is linked to Mike's ability to make a choice or to intervene and decide. It could also be pointed out that causation is not something abstract or objective, but rather something subjective, which depends on Mike's perspective or agency.

The notion of causation in itself is to be understood as a relation between agents and their experience. This means that causation is extrinsic. Being as such, causation is "to be explained in terms of the way in which an agent's producing,

manipulating, or *wiggling* one event affects the probability of another event” (Menzies & Price, 1993: p. 190). As an analogy, causation is like color, since color, as a secondary quality, exists only in relation to an observer, as his experience. But if causation is a secondary quality, this does not mean that it should rest on an infallible account of experience. In other words, such as it happens in the case of colors, when sometimes red is not red, for instance, one can be mistaken about the actual cause of an effect. If sometimes red is not red, due to some extraordinary conditions, then an action might have a different cause behind it, or it might have been accomplished by accident, meaning that the real cause was not actually devised by a particular agent, in a given context (Menzies & Price, 1993: p. 193).

The agency theory of causation, envisioned by Menzies and Price, faces several objections. One of them is the circularity objection, which happens to be the most frequently put forward objection. More precisely, as Menzies and Price shows, if “*A* is a cause of a distinct event *B* just in case *bringing about A* would be an effective means by which a free agent could *bring about B*” (Menzies & Price, 1993: p. 193). And they further go on, by explaining why this is a vicious circle. The reason behind such a fallacy would be the double usage of the term ‘bring about’. Even worse, if a careful analysis is employed in the case of the terms ‘means’ and ‘free agent’, it would seem that those terms as well imply the notion of causality. If the purpose is to define causality, and in the definition of causality the notion of causality is already presupposed, or, in other words, it is not possible to define the terms “bring about”, “means” and “free agent” without already using a notion of causation, then the definition is strongly or heavily circular (Ahmed, 2007: pp. 123-124). Their reply on this matter works as an argument by analogy, saying that, if in the case of colors, the notion of ‘looks red’ is not circular, then also the notion of “bring about” is not circular. The actual justification for this is saying that “the dispositional explanation of the concept *red* need not fall into the trap of circularity” (Menzies & Price, 1993: p. 194). And because of that, neither the notion of causation as agency needs to fall in the trap of circularity.

If this reasoning is sound or not, it is not my interest to judge. In this particular case, all that matters is that this theory of causation makes appeal to agency and has a metaphysical perspective about causation, which makes it a secondary quality. Since it employs agency, causation is understood in terms of intervention or decision-making, based on a conditional probabilistic reasoning, and a kind of perspectivism, by which an agent has subjective fallible experiences. If causation as agency is circular or not, this is the topic of another debate. In order to provide a reasoning for using it, even if this theory might be considered too problematic to give a proper account of causation, I could make appeal to a counterfactual construction or justification. In this case, I could say that were this theory to be less problematic, it would be a fitting tool or instrument to give a better interpretation of causality, not in general, but just in the case of Priest’s definitions of fictional and abstract objects.

6. Is It Really Fitting?

According to Priest, causation is an existence-entailing relation. Since the relation of existence-entailing is in itself confusing or insufficiently clear, it can be reasoned that this is a good reason to try to find another understanding of causation. Evidently, this is not a good reason to accept another explanation, or, in particular, the agency theory of causation. In order to see if the aforementioned theory might be a fitting replacement of causation as existence-entailing, another argument or justification is needed.

Given the assumption that I am working with, which is that the distinction between fictional and abstract objects is exhaustive, the non-existence of fictional and abstract objects has been logically represented as follows:

(NF): $(\sim Ex \& \Box \Diamond Ex) \rightarrow Cx$, Cx always true;

(NA): $(\sim Ex \& \Box \sim Ex) \rightarrow \sim Cx$, Cx always false;

This logical representation of Priest's understanding of fictional and abstract objects points out the need to understand " Cx ", or what it means to causally interact with us. If causation as existence-entailing have been considered inadequate, then let us see if causation as agency could prove better. In order to seek an answer, maybe a good way of searching implies to consider again the path that leads to constructing a modal version of Meinong's proposal. As I have mentioned, Priest wanted to accept noneism, but this perspective needed an explanation regarding the properties of non-existent objects. The CP seemed a sound explanation, but unfortunately not in initial form, since it had had some unacceptable repercussions. So, Priest proposed a modal version of CP. The modal version of CP says that 'an object characterized by a representation has the characterizing properties, not necessarily in the actual world, but in the worlds (partially) described by the relevant representation' (Priest, 2005: p. 84).

The essential term in the modal version of CP is "representation". As I have mentioned, the semantics of intentionality includes a special intentional operator, which is the representational operator. The general form of the operator is the following: "...represents...as holding [in the matter at hand]" (Priest, 2005: pp. 84-85). The special mentioning, "in the matter at hand", outlines the subjective element of each representation, since it depends on the agent that represents something. In this way, as Priest says, "if I read a novel I represent the world of the novel to myself in a certain ways" (Priest, 2005: p. 85). Since Priest's proposal is a semantics of intentionality, I will introduce another assumption, which is that every intentional act is a representational act. If intentionality is the ability of thought to direct itself towards any object, then the way in which this directing happens is by representation.

In the agency theory of causation, causation was understood as "an event A is a cause of a distinct event B just in case bringing about the occurrence of A would be an effective means by which a free agent could bring about the occurrence of B " (Menzies & Price, 1993: p. 187). This understanding of causation implied a means-end relation. Let us replace A with "a set of properties", and B with "a

representational act". In this case, the definition becomes: "a *set of properties* is a cause of a distinct *representational act* just in case bringing about the occurrence of a *set of properties* would be an effective means by which a free agent could bring about the occurrence of a distinct *representational act*". In other words, an agent cannot have a representational act without firstly having a set of properties. In the sentence 'I imagine that a round square exists', the set of properties {*round, square, exists*} comes metaphysically first. This understanding of representational acts comes from an example that Meinong used in order to prove that negative existential statements can be true. His example is about a perpetual movement machine, which does not exist. In order to say that it does not exist, its set of properties must firstly be established. The conclusion of the argument is that the set of properties (*Sosein*) comes metaphysically first or before any ontological determinations (Perszyk, 1993: pp. 49-50). In this way, before intentional acts can be formed, there must be a set of characterizing properties, which, given agency theory of causation, works as a cause for intentional acts or representational acts.

So, the cause is the agency, which makes up a particular set of properties, whereas the effect might be the unique representational act an agent is having at a given moment. This fits, since Priest says that "if I read a novel I represent the world of the novel to myself in a certain ways" (Priest, 2005: p. 85). In this case, I might represent the same novel in a totally different way. The generic cause is to have a set of properties, whereas the particular causes are the precise sets of properties. This can be said of the effects too. The general effect is to have a representational act, but the particular effects are the precise or singular representational acts. The link between a cause (set of properties) and an effect (representational act) is not something objective, since human subjectivity is behind it. So, in simple terms, it could be argued that causation in this case is a secondary quality, since it depends on the agent's perception or cognition.

On the other hand, what does not actually fit is the pragmatic approach of causation as agency, since it is presupposed that the agent is rational or makes decisions based on a conditional probabilistic reasoning. Causation as agency is rather applicable in concrete, real situations, in which one has to make a decision. So, it is more about the external reality, then the workings of the mind. In this case, the actual purpose of the agency theory of causation does not apply to a representational act. Even so, in order to make this analysis complete or as complete as possible, I will try to understand the definitions of fictional and abstract objects by taking into account causation as agency.

Given (NA), in the particular case of an existent round square, it can be asked: what does it mean for this object to never causally interact with us? It cannot mean to be concrete or have the same ontological nature as us, since causation as existence-entailing has been ruled out. If 'us' means the actual world, 'us' can be taken as our concrete reality or the entire domain of objects, or as concrete, particular instances, in which an agent has a representational act. If the purpose is to see if agency theory works in this case, then it might be more plausible to define "us" as

particular or concrete representational acts. In such an interpretation, the existent round square, as an impossible object, never causally interacts with the agent that is having the representational act about it. If there is not conceivable an object without properties, it can be said that every object is a set of properties. If the set of properties was the cause of the representational act, the understanding of the definitions of fictional and abstract objects, in terms of agency causation, is happening at another level. Maybe a schematic representation might bring more clarity:

$C \approx E$, means that a cause (C) leads to (\approx) and effect (E).

$Ra =_{df.} \Phi(Ax)$.

Any representational act (Ra) is a way of using (Φ) a given set of properties (Ax). In this case, Ax is a cause to any representational act, because causation is the intervention of an agent in making a thing leading to another. So, the intervention of an agent leads to constructing Ax , in order to have a representational act about an object.

Now, Ax is the object of the representational act. (NF) and (NA) show how some of these objects might be understood. If an object has a set of properties that makes it only possible, then it is a fictional object. If an object has inconsistent properties, then it is impossible or abstract. In this case, " Cx " is actually a relation between the agent and Ax , not between Ax and Ra . In particular, a fictional object always causally interacts with us, whereas an impossible object never interacts with us. If interaction is one way of saying that there is a relation between an agent and Ax , then causation as agency is not a preferable alternative interpretation, since it rules out the connection to impossible objects. In other words, an agent can never construct an Ax , that implies inconsistent or contradictory properties.

7. Conclusion

Causation as agency, as an attempt to propose a better interpretation for causation in Priest's definitions of fictional and abstract objects, does not really work. Causation as agency, if applied, it would rule out impossible objects, which affects the generality of the Principle of Characterization. At best, causation as agency can explain the general workings of any intentional act, if it is assumed that every intentional act is a representational act.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Ahmed, A. (2007). Agency and Causation. In H. Price, & R. Corry (Eds.), *Causation, Physics and the Constitution of Reality: Russell's Republic Revisited*. Oxford University Press. <https://doi.org/10.1093/oso/9780199278183.003.0006>
- Berto, F. (2011). Modal Meinongianism and Fiction: The Best of Three Worlds. *Philosophical Studies*, 152, 313-353. <https://doi.org/10.1007/s11098-009-9479-2>

-
- Berto, F. (2012). *Existence as a Real Property: The Ontology of Meinongianism*. Synthese Library, Springer. <https://doi.org/10.1007/978-94-007-4207-9>
- Ciprotti, N. (2014). On Modal Meinongianism. In M. Antonelli, & M. David (Eds.), *Logical, Ontological, and Historical Contributions to the Philosophy of Alexius Meinong* (pp. 1-36). De Gruyter. <https://doi.org/10.1515/9783110351187.1>
- Menzies, P., & Price, H. (1993). Causation as a Secondary Quality. *British Journal for the Philosophy of Science*, 44, 187-203. <https://doi.org/10.1093/bjps/44.2.187>
- Perszyk, K. J. (1993). *Nonexistent Objects: Meinong and Contemporary Philosophy* (Nihoff International Philosophy Series 49). Kluwer. <https://doi.org/10.1007/978-94-015-8214-8>
- Priest, G. (2005). *Towards Non-Being: The Logic and Metaphysics of Intentionality*. Oxford University Press.
- Quine, W. V. (1948). On What There Is. *The Review of Metaphysics*, 2, 21-38. <http://www.jstor.org/stable/20123117>
- Routley, R. (2018). *Exploring Meinong's Jungle and Beyond*. Springer Nature Switzerland AG. <https://doi.org/10.1007/978-3-319-78793-0>