Synthese (2024) 203:193
https://doi.org/10.1007/s11229-024-04617-9

ORIGINAL RESEARCH

The role of scenarios in paradoxes

Angelica Mezzadri[1]

Received: 30 January 2024 / Accepted: 30 April 2024

Abstract
This paper fills a gap in the existing metaphilosophical research on paradoxes by focusing on the role of scenarios. Typical philosophical paradoxes contain a scenario description whose contribution to paradoxes remains unexplored. I argue that scenarios are examples or instantiations of the abstract schema of paradoxes. As such, scenarios contribute to paradoxes on two levels. First, they make the argument more concrete, thus enhancing the dialectical force of paradoxes and facilitating their understanding, especially for non-experts. This function is external to the paradox itself, but has important practical implications for the use of paradoxes and their effect on philosophical debates, and it contributes to explaining why philosophical paradoxes are usually introduced by a scenario. Second, and more crucially, scenarios are essential to the epistemic dimension of paradoxes. By definition, paradoxes have two necessary components: the argumentative structure and the plausibility/implausibility of the premises/conclusion. By providing examples of the abstract schema, scenarios contribute to making the premises plausible. In particular, scenarios are the source of plausibility and justification for those premises that contain an empirically grounded assertion of existence. Examples of such paradoxes are the Sorites paradox, the Lottery paradox, and the Grue paradox. Contrary to the dialectical role, the epistemic function of scenarios is indispensable, as it connects paradoxes to the real world and underscores their significance in specific debates.

Keywords Paradox · Scenario · Plausibility · Justification

1 Introduction

Paradoxes have long captivated the minds of philosophers. Far from being mere intellectual curiosities, they have an important role to play: they raise serious problems by revealing tensions and contradictions between our pre-theoretical understanding of

✉ Angelica Mezzadri
angelica.mezzadri@unito.it

[1] Department of Philosophy and Education Sciences, University of Turin, Turin, Italy

Published online: 03 June 2024

phenomena and established philosophical theories. As a result, they trigger significant revolutions in philosophical debates, forcing philosophers to re-evaluate, refine, and sometimes completely reformulate their theories.

Because paradoxes are such a powerful tool, they have received a lot of attention from a metaphilosophical perspective. Questions about the definition of paradoxes have been discussed extensively in the literature with solid results. However, less attention has been paid to analyzing various aspects of how they work, such as the observation that philosophical paradoxes are usually presented through informal short case descriptions rather than formal arguments. The Lottery paradox, the Sorites paradox, Hilbert's hotel paradox, the Unexpected Examination paradox, and Zeno's paradoxes are just a few of the many paradoxes typically introduced by a scenario description. Despite the magnitude of the phenomenon, the role of scenarios in paradoxes has not thoroughly been investigated. To fill this gap, I will explore two research questions: (i) Do scenarios play an indispensable role in paradoxes? (ii) What specific contributions do they make to paradoxes?

First, I will test the elimination hypothesis, i.e. the idea that scenarios do not play an essential role in paradoxes and that they can be removed without consequences for the paradox itself. The analysis of the Sorites paradox will show that the elimination of the scenario affects the dialectical and epistemic strength of the paradox. This suggests that scenarios play a crucial role in paradoxes, and their contribution needs to be investigated in more detail. Based on the information gathered in the first part of the paper, the dual contribution of scenarios is then elaborated: first, they play a pragmatically relevant dialectical function, making the paradox more concrete and comprehensible; second, they play an essential epistemic function by lending plausibility to the premises. This hypothesis is then confirmed by the analysis of different cases, such as the Lottery and Grue paradoxes. Finally, I will test the limitations of these findings by comparing them with problematic cases such as Russell's paradox and paradox-like cases that do not fit the standard definition of paradoxes, such as Simpson's paradox.

2 Are scenarios essential to paradoxes?

This section examines whether scenarios can be removed from paradoxes, and if so, what consequences this removal might entail. Before turning to the investigation, it is important to clarify the definition of paradoxes that I will adopt, to examine the treatment of the role of scenarios in paradoxes in the existing literature, and to establish the basic terminology and common framework on which the investigation will be built.

2.1 Paradoxes, arguments, plausibility

According to the standard definition, a paradox is an unacceptable conclusion derived by apparently acceptable reasoning from apparently acceptable premises (Sainsbury, 2009). Paradoxes, then, are essentially arguments, where a conclusion is derived from

a set of premises. What distinguishes paradoxes from all other arguments is the plausibility of the premises, which clashes with the implausibility of the conclusion that supposedly follows from the premises, creating the typical paradoxical feeling. There are three ways to escape paradoxes: (i) accepting the conclusion despite its implausibility; (ii) rejecting the validity of the argument itself; or (iii) rejecting one or more premises of the argument despite their initial plausibility. According to Quine's (1966) famous taxonomy, paradoxes that are resolved by accepting their conclusions are veridical paradoxes: the reader's sense of confusion quickly dissipates upon further reflection on the argument. Second are paradoxes that are resolved by identifying a fallacy in the argument: these are falsidical paradoxes. Finally, paradoxes of the third kind are antinomies: in cases where the conclusion is false and the argument is valid, one of the premises must be abandoned. However, scholars often disagree on the path to take, which is one of the reasons why paradoxes in philosophy lead to great and heuristically productive debates. Paradoxes, however, exhibit varying degrees of complexity. While some are relatively easy to solve and bear minor consequences for philosophical theories, others (in particular antinomies) prove to be immensely challenging due to the individual plausibility and desirability of their premises, the validity of the argument, and the extreme implausibility or absurdity of the conclusion.

While this picture works well in most contexts, doubts have been raised about the emphasis on the argumentative nature of paradoxes. In particular, (Lycan, 2010) argues that the standard definition misrepresents paradoxes as being intrinsically directional from premises to conclusion. Instead, paradoxes should be defined as inconsistent sets of propositions, where each one is highly plausible. This alternative perspective allows for different interpretations in identifying the conclusion of the paradox, with each proposition initially considered as an equally viable candidate for rejection. Doubts have also been raised about the interpretation of the term 'plausibility', which is used to describe the components of the paradox. Indeed, plausibility can be interpreted subjectively or objectively, and both interpretations raise concerns (Cowie, 2023).

Despite their differences, the standard and alternative definitions of paradoxes can be used interchangeably in many contexts. This is also true for the scope of this paper: my research questions and findings are compatible with all currently accepted definitions of paradoxes. For the sake of clarity, I will mainly use terms borrowed from the standard definition of paradoxes in what follows, but I could just as easily have used Lycan's alternative definition. This is possible because all alternative definitions identify two necessary components of paradoxes. First, paradoxes have an argumentative component, namely the logical connection between the premises and the conclusion of the paradox. Second, they have an epistemic component: the premises of the argument must be plausible, and the conclusion must be implausible or apparently false. Without the argumentative dimension, the paradox would be just a case description. Without the epistemic element of plausibility, the paradox would reveal an inconsistency but would lose its typical heuristic and transformative power.

Finally, the term "paradox" is used in some contexts in a broader sense, not only to denote contradictory sets of plausible propositions, but also those cases which, despite not revealing any inconsistency, create a sense of puzzlement in the reader. This usually occurs when the conclusion of a sound argument is unexpected and contradicts the reader's initial understanding of the case but is nevertheless undeniably true. Some

examples of this type of paradox are Simpson's paradox and Schrödinger's cat case. Just like standard paradoxes, these cases are usually introduced by a case description. Therefore, in the final section of this paper, the scope of the investigation is extended beyond standard paradoxes to include paradoxes in this broader sense. This extension offers the possibility of comparison with standard paradoxes and serves as further confirmation for the account under investigation.

Finally, it should be noted that none of the definitions of paradoxes explicitly mention scenarios, nor do they explicitly assign them a role in paradoxes. Therefore, while relying on the standard definition of paradoxes, I also need to examine how paradoxes are used in philosophical practice to investigate the role of scenarios. Ultimately, I will argue that scenarios play a relevant and non-eliminable role in paradoxes that is closely connected to their epistemic dimension. To this end, I will first present the opposite position, the elimination thesis, and show how the elimination of scenarios from paradoxes leads to a dialectical and epistemic loss for the paradox itself.

2.2 The elimination thesis

In this section, I examine whether scenarios can be eliminated from paradoxes without this having a major impact on them. If this is the case, it is unlikely that scenarios play a crucial role in paradoxes. On the contrary, if the elimination of scenarios reduces their effectiveness in any way, a more detailed analysis of their role is required.

To formulate the idea that scenarios may be dispensable for paradoxes, I draw inspiration from the ongoing debate on thought experiments, in which the role of scenarios occupies a central position. In particular, (Norton, 1991) advocates an eliminativist position for scenarios in scientific thought experiments. According to Norton, scientific thought experiments contain details that are irrelevant to the generality of the conclusion. Therefore, the same conclusions could be reached by reasoning without such specific details, making the scenario superfluous. This thesis is highly controversial in the debate on thought experiments, with the prevailing view being that a scenario is necessary for thought experiments to work. For the present study, however, I leave thought experiments aside and focus on the question of whether the elimination thesis applies to paradoxes.

At first, the eliminativist position seems to align with the standard definition of paradoxes. Indeed, it emphasizes the importance of the logical dimension of paradoxes and correctly assumes that the argumentative burden does not rest on scenarios. However, the eliminativist position dismisses the role of scenarios too soon, which raises some doubts. For example, it does not take into account the empirical data on how paradoxes are typically used and presented. Sainsbury (2009) and Clark (2012), among others, provide extensive lists of philosophical paradoxes. Each of these paradoxes is first informally presented and then analyzed. In both lists, all paradoxes are presented using a case description that corresponds to the best-known version of the paradox itself. Accepting this data, however, puts the eliminativist in a difficult position: the claim that scenarios are inessential to paradoxes despite their pervasiveness calls for an explanation of why they are practically always invoked in developing a paradox. On the contrary, by identifying the role of scenarios in paradoxes, I will also identify

the reasons behind their pervasiveness, which is an advantage over the eliminativist position.

These concerns suggest that the status of scenarios in paradoxes is not as straightforward as the eliminativist would like it to be. To address this issue, I will test the eliminativist thesis against the Sorites paradox.

2.3 The Sorites paradox

Consider the Sorites paradox of the heap:

> 10,000 grains suitably arranged make a heap. But, at no point can you convert a collection of grains that is a heap into one that is not, simply by removing a single grain. So it follows that a single grain makes a heap. For if we keep removing grains over and over again, say 9,999 times, at no point does it cease to be a heap. Yet we obviously know that a single grain is not a heap (Clark, 2012, p. 86).

Prima facie, the scenario seems to contribute to the construction of the paradox, as it serves to generate the inconsistency arising from the use of vague terms like “heap”. A more thorough examination of the Sorites paradox’s mechanics becomes necessary to assess the elimination thesis. First, it is worth noting that the heap scenario can indeed be dispensed with. The same paradox can be constructed using different scenarios and vague terms. Consider terms such as "tall," "yellow," "forest," and "bald." A person with one hair on their head remains bald, just as a bucket of yellow paint retains its yellow hue when a drop of red paint is added. Cutting down a single tree does not change a forest into a non-forest. In each of these scenarios, we encounter perplexing conclusions, where we are asked to accept that a person with a lot of hair on their head is bald, that a bucket of mostly red paint is yellow, or that a lone tree constitutes a forest. As all these scenarios yield the Sorites paradox, the heap scenario itself is not indispensable. Still, by replacing the heap scenario with alternative scenarios, we are not obtaining an abstract and scenario-free version of the paradox. Thus, while individual scenarios may not be essential to the paradox, the presence of some kind of scenario seems to be required.

However, the elimination thesis does not pertain to any specific scenario; rather, it posits that scenarios as such can be eliminated from paradoxes. To achieve this, the paradox must be presented as an argument. Such an argument may be reached by a process of abstraction and elimination of unnecessary details. In the case of the Sorites paradox, comparing different alternative scenarios aids in discerning eliminable elements. For example, heaps of sand, forests, and bald people are interchangeable and non-essential to the paradox. Moreover, comparing different Sorites scenarios can help identify the common elements, thus revealing the necessary components of the paradox. By pinpointing these essential elements, one can create a more abstract version of the paradox that transcends irrelevant details. Roughly, the abstract structure of the Sorites paradox is composed of something that is described by a vague term. Subsequently, a principle is required, dictating that anything sufficiently similar to the

original entity falls under the same vague term. Lastly, the result of the repeated application of the principle must not fit the vague term to create a paradoxical conclusion.

More precisely, three elements are necessary to build an instance of the Sorites paradox. First, we need a predicate for which it is possible to create a series where the members are ordered with respect to a dimension that decides the application of the predicate. For example, for the predicate "heap" members are ordered by the number of grains. For the predicate "tall" members are ordered by height. Second, neighboring members of the series must differ only slightly from one another and be almost impossible to distinguish to human perception. As a result, if the predicate applies to a member of the series, it also applies to its neighbor. For example, a person with a height of 180 cm is only slightly different from a person who is 181 cm. Therefore, if the 181 cm person is tall, the 180 cm individual is too. Lastly, the predicate must unambiguously apply to the first member of the series while unambiguously not applying to the last. For instance, a collection of 100,000 sand grains constitutes a heap, whereas a single grain does not (Hyde and Ruffman, 2018). This presentation of the Sorites paradox is more abstract than the one based on the heap scenario. It does not refer to particulars such as heaps, colors, forests, or tall people. Rather, it speaks of predicates and series of members to which the predicate applies. For this reason, the abstract version of the paradox meets the requirements of the elimination thesis. It contains no detail, and it represents the abstract argumentative structure of the paradox.

The elimination thesis posits that the scenario, comprising particulars and details, holds no relevance to the paradox, which can be conveyed equally effectively through argumentation alone. However, this does not seem to hold for the Sorites paradox. There is a substantial difference between the original and the abstract rendering of the paradox. The former is immediately comprehensible, and highly persuasive, and its conclusion strikes readers as highly implausible, if not outright false. The paradoxical nature of the scenario resonates intuitively and instantaneously with proficient English speakers. The abstract version of the paradox lacks the same persuasive impact as the original version. Nothing about it strikes the reader as immediately puzzling or problematic. Instead, most readers would likely struggle to grasp the paradox, which would become easier to understand by reintroducing the scenario as an example.

This speaks to the different dialectical power that the two versions have and significantly undermines the elimination thesis. Interestingly, a similar argument was raised for thought experiments. Specifically, (Brown, 1992) challenges Norton's elimination thesis by arguing that although the particulars invoked by scenarios may be irrelevant to deriving the conclusion, they remain relevant to understanding it, and they play a vital role in making the conclusion comprehensible and persuasive to the reader. For this reason, it is very common to present thought experiments using scenarios, rather than their more abstract or formal version. In this sense, scenarios are explanatory examples of the abstract structure of the paradox: by providing a concrete example of how the paradox interacts with elements of our world, language, and cognitive life, they make it easier to grasp, especially for laypeople.

Moreover, by providing examples of vague predicates to which the paradox applies, the Sorites scenarios provide plausibility to the premise of the argument concerning

the existence of such vague predicates. Indeed, the abstract formulation of the Sorties paradox points out the inconsistency of vague predicates. However, the revealed inconsistency in itself does not elevate the argument to the status of a paradox: the inconsistency would be uninteresting and inconsequential if such vague predicates were not in use in English. Since vague predicates are numerous and frequent, the inconsistency is a genuine paradox with urgent and real consequences for our use and understanding of these terms. By providing examples of such vague predicates, the Sorites scenarios effectively argue for the plausibility of the existence premise and thus contribute fundamentally to the epistemic dimension of the paradox.

These considerations suggest that for certain paradoxes, such as the Sorites paradox, the removal of the scenario has implications for the epistemic force of the paradox and its crucial ability to generate a sense of paradoxicality in the reader. The Sorites scenarios connect the paradox to a context, transforming it from a potential paradox about potential predicates into an actual paradox about actual vague predicates.

3 The role of scenarios in paradoxes

The information gained from the comparative analysis of the Sorites paradox in its abstract and its scenario version can be used to develop a more comprehensive understanding of the function of scenarios in paradoxes. The last section has shown that removing the heap scenario has an impact on the effectiveness of the paradox. This suggests that scenarios play a relevant role in paradoxes, which is worth further investigation. Starting from the idea that scenarios are essentially examples, or possible instantiations, of the paradoxical argument, I argue that they contribute to paradoxes in two different ways.

The first function of scenarios is trivially related to the fact that they are examples. As in many other contexts, examples clarify and explain, in the case of paradoxes scenarios illustrate the abstract structure and make paradoxes more concrete. This in turn makes paradoxes easier to understand, especially for non-experts. In addition, scenarios serve as examples to illustrate the connection between the paradox and certain debates and provide a concrete context in which the paradox occurs. For example, the Sorites scenarios show examples of the kind of vague predicates that generate the paradox, highlighting the relevance of the paradox for English speakers.

Illustrating, clarifying, contextualizing: all these functions are related to the way paradoxes are presented, communicated, and received by the audience, and they fall under the umbrella of the dialectical role of scenarios. By reinforcing their dialectical power, scenarios have practical implications for the way paradoxes are used and received by both laypeople and experts. This might partially answer the original question: If scenarios are inessential to paradoxes, why are paradoxes typically presented using scenarios? One of the reasons why this happens is that scenarios are usually more effective than abstract arguments in getting the message across, especially to an audience of non-experts, as they are specifically designed and conceived to maximize their dialectical appeal. However, the dialectical function does not fully explain the theoretical cost of removing the Sorites scenarios from the paradox. In particular, without the examples that the Sorites scenarios provide, the premise concerning

the existence of the vague predicates that give rise to the paradox seems to lose its plausibility and grip. This is an indication that although the scenarios are essentially examples of the abstract schema, they play a crucial role in paradoxes related to the plausibility and justification of their premises.

According to standard definitions of paradoxes, the abstract structure of a paradox reveals an inconsistency. However, inconsistencies only rise to the status of paradoxes if they are generated from a set of plausible propositions that initially appear to be true. Basically, not every randomly generated inconsistency is interesting, but only those that appear to be realized or realizable. The plausibility of the propositions that form the paradoxical sets depends on various sources, such as common sense, background theory, or reality, but the analysis of the role of scenarios reveals a pattern. The ability of paradoxes to provoke a sense of puzzlement in the reader is related to the fact that the inconsistency they reveal is usually hidden and not obvious in everyday life. The scenarios described by paradoxes are often cases where the normally hidden inconsistency becomes obvious. The paradoxical argument then contains an assertion of existence that relates to this particular case or scenario. Take the Sorites paradox as an example: the argument contains a premise that asserts the existence of vague predicates that give rise to the paradox. These existence assertions often have an empirical basis: they are plausible and justified because they refer to a possible or actual state of affairs. The plausibility of the Sorites paradox, for example, is based on the existence of vague predicates in English. By providing examples of such predicates, the Sorites scenarios link the abstract assertion of existence to concrete cases, thus demonstrating the plausibility of the premise itself.

In general, scenarios play an essential epistemic role in paradoxes, since they provide plausibility and justification to one of the premises of the paradox, namely the one that expresses an assertion of existence with an empirical basis. While the dialectical role of scenarios has important practical consequences for paradoxes, it lies outside the paradox itself and is therefore not essential for its existence and definition. In contrast, the epistemic function of scenarios is fundamental, since it contributes to a necessary component of paradoxes, namely the plausibility of the premises. This account of the role of scenarios in paradoxes takes the Sorites paradox as a model and can therefore seem very specific, maybe applying only to a limited number of paradoxes. However, this is not the case. As far as the dialectical function is concerned, it can be assumed to apply to any paradox that contains a scenario. Moreover, I argue that scenarios play an epistemic role in all those paradoxes that contain an empirically grounded assertion of existence. Examples of such paradoxes, besides the Sorites paradox, are the Lottery paradox, the Grue paradox, Zeno's arrow paradox, Newcomb's paradox, the Unexpected Examination paradox, and the preface paradox. So which paradoxes are left out? Essentially those paradoxes that either have no existence assertion as a premise or that have a non-empirically grounded existence assertion. To further validate this account and examine its scope of application, the next sections will test it against two famous philosophical paradoxes: the Lottery paradox and the Grue paradox.

3.1 The Lottery paradox

Consider the paradox of the Lottery:

> Suppose there will be only one winning ticket in a fair lottery with a thousand tickets. If you have just one ticket, say number 192, it is reasonable to believe that it won't win. But the same will be true of every other ticket, although you know that one of the tickets will win. Then, taken together, your set of beliefs will be inconsistent (Clark, 2012, p. 126).

The paradox deals with the rationality of belief, specifically targeting two rules: the acceptance rule, which allows belief in propositions with a high probability, and the agglomeration rule, which licenses belief in the conjunctive proposition $p\&q$ for the subject that rationally believes that p and rationally believes that q. In the paradox, the two rules are tested against a lottery of 1000 tickets, only one of which is a winning ticket. Suppose the acceptance rule licenses belief in propositions that have a probability of 0.99 or higher. In this situation, it is rational to believe that each ticket is a losing ticket. Applying the agglomeration rule to these beliefs leads to the justified belief that every ticket is a losing one. However, since the lottery has one winning ticket (and the probability of this proposition is equal to 1), it is rationally required to believe one of the tickets is a winning ticket (Sorensen, 2022).

To access the abstract structure of the paradox, the lottery scenario must be removed. The paradox is then left with two rules, the acceptance rule and the agglomeration rule, both of which serve as premises of the argument. According to these rules, it is rationally acceptable to believe in propositions that have a high probability, and it is rational to believe in the proposition $p\&q$ if you have the rational belief that p and the rational belief that q. However, there are cases in which the probability of proposition p is so high that it is rational to believe that p and the same is the case for q, but the belief in $p\&q$ appears irrational. The existence of such probability distributions shows that the two rules are incompatible.

This reconstruction of the paradox qualifies as an abstract version and fulfills the requirements of the eliminativist since it does not contain any concrete scenario or detail. However, it is not as effective on a dialectical level as the original version. While the lottery scenario exemplifies the abstract argument and makes it more concrete and generally easier to understand, the abstract version is less immediate and may appear more controversial to the non-expert audience. In particular, the claim that there are cases that show the inconsistency of the rules of agglomeration and acceptance is not in itself very convincing. The lottery scenario, which provides a concrete example of such probability distribution, improves the comprehensibility of the paradox and generally makes it more accessible to a non-expert audience. Moreover, the reference to the lottery scenario enhances the sense of puzzlement created by the paradox. Indeed, when thinking of real-life lotteries, people do not usually think of buying a lottery ticket as irrational, and this is also true of the lottery depicted in the paradox. However, the paradox suggests that it is in fact irrational to do so, since according to the agglomeration and the acceptance rules, one should believe that there is no winning ticket, and if one believes that no ticket will win, then it is irrational to buy a ticket.

However, the contribution of the lottery scenario to the lottery paradox is not merely dialectical, but also epistemic. It contributes to making its premises plausible. In particular, the argument contains multiple premises, whose justification and plausibility derive from different sources. For example, the rules of agglomeration and acceptance are justified by their connection to relevant background theory on the rationality of belief, as well as appearing generally plausible and intuitive to common sense. Consider though the premise asserting the existence of cases in which the probability of a proposition p is so high that it is rational to believe that p, the same is true of q, but believing $p\&q$ seems irrational. This proposition is an empirically grounded assertion of existence. It states that there is at least a probability distribution or scenario that, unlike typical cases, can reveal and make visible a hidden inconsistency between two well-established rules. The justification and plausibility of this kind of existence claim is not based on background theory or even common sense, but rather on its connection with real or possible cases. In the paradox this connection is established by the lottery scenario, which is a case where the hidden inconsistency emerges, and thus demonstrates the plausibility of the existence premise. The justificatory force of the lottery scenario rests on the fact that the lottery described in the paradox is possible and resembles real lotteries. In this sense, the lottery scenario shows that cases that uncover the inconsistency of the agglomeration and acceptance rules are realistic and possible, thus making the paradox itself a pressing matter.

To sum up, the lottery scenario provides a possible instantiation of the abstract argument, contributes to the dialectical function of the paradox, and evokes the typical feeling of puzzlement in the reader. Moreover, it justifies and gives plausibility to the existence claim contained in the argument, thus contributing significantly to the epistemic success of the paradox.

3.2 The Grue paradox

Consider now the Grue paradox.

> If we generalize on the basis of emeralds we have examined we can reach the conclusion that all emeralds are green. Now define a new adjective, 'grue': x is grue if it is green and examined (by now), or blue and unexamined (by now). If we generalize on the basis of previously examined emeralds it seems we can also reach the contrary conclusion that all emeralds are grue (Clark, 2012, p. 83).

In this paradox, the scenario introduces the definition of the predicate "grue", whose stipulated meaning is that something is grue either if it is green and has been examined at time t or if it is blue and has not been examined at time t. To investigate the role of the scenario in the paradox, it is useful to reconstruct the argument.

The paradox of "grue" deals with the concept of confirmation. Specifically, it considers a very intuitive principle for confirmation, according to which a generalization is confirmed (at least to some degree) by any of its instances. For example, the proposition "All emeralds are green" is confirmed by the instance "This emerald is green". The principle of confirmation is taken as a premise of the argument, together with the stipulated definition of "grue". As all examined emeralds are green, they count as grue

according to the definition of the predicate. Therefore, the principle of confirmation assures that the observed data of green emeralds counts as confirmation of the hypothesis that emeralds are grue. However, this is absurd. Indeed, this would also amount to confirmation of the fact that all the non-examined emeralds are blue, which strikes the reader as false and in conflict with the observed data. It is unacceptable that observing green emeralds confirms that yet unobserved emeralds are blue (Sainsbury, 2009).

The paradox arises from the conflict between the application of the naïve principle of confirmation and the introduction of a time-dependent criterion. While the observed objects meeting the criterion are confirmed, the application of the confirmation principle leads to the counterintuitive conclusion that all non-examined objects possess a property or condition inconsistent with observed data, challenging the reliability of the confirmation principle. The scenario contributes to the paradox by providing a possible and plausible instance of a time-dependent predicate, namely "grue". Despite the seemingly peculiar nature of the definition of grue, it aligns with accepted norms for defining predicates in English, and therefore it qualifies as a plausible predicate. This aspect is pivotal to the workings of the paradox since only a plausible predicate raises the contradiction to the status of an actual paradox and represents a concern for the naïve principle of confirmation. The "grue" scenario provides an instantiation of the schema that meets the requirement of possibility and plausibility, and thus it enhances the epistemic dimension of the argument, elevates the inconsistency to the status of paradox, and makes it a pressing problem for the philosophical debate on confirmation theory.

Finally, the introduction of an example of the abstract schema of the paradox makes the argument easier to understand for the reader. In this sense, the use of "grue" to exemplify the argument enhances the dialectical power of the paradox. Therefore, the scenario plays both an essential epistemic function, as it makes the premises of the paradox plausible, and an important dialectical function, as it exemplifies and clarifies the argument.

4 Testing the boundaries: limitations and applicability

Up to this point, it seems that scenarios play an important role in paradoxes. They are essentially examples of the abstract schema and thus contribute to the paradox in two ways. First, they increase the dialectical force of paradoxes, making them less abstract and easier to understand and they frame the argument by providing a context in which the paradox occurs. Second, scenarios contribute to the epistemic dimension of paradoxes by making their premises plausible. In particular, they are an indispensable source of justification for the empirically grounded assertions of existence contained in paradoxes such as the Lottery and the Sorites. But what happens to paradoxes that contain no such premise? And what about paradoxes that do not conform to the standard definition? This section extends the investigation by considering problematic cases.

4.1 Russell's paradox

Consider this presentation of Russell's paradox:

> Define "R" as the set of all sets that are not members of themselves. If R is a member of itself, by definition it is not a member of itself. Conversely, if R is not a member of itself, by definition it is a member of itself (Irvine & Deutsch, 2021).

Russell's paradox is different from previous examples: it is more technical, more abstract, and generally less accessible to non-expert audiences. However, it may still be possible to identify a scenario within the paradox and to differentiate between a scenario and an abstract version of the paradox, allowing for the specific contributions of the scenario to be tested. The working hypothesis is that the scenario in Russell's paradox consists of the description of R. This is not a typical example of a scenario. Usually, we think of scenarios as something richer in detail, and less abstract. Thus, qualifying the description of Russell's paradox as a scenario would mean stretching the notion of scenario a bit. However, exploring the impact of removing any reference to R or sets is worthwhile.

Formally, Russell's paradox can be expressed as $\forall x\ (x \in R \leftrightarrow \sim (x \in x))$, where R is the set of all sets that are not members of themselves, and x ranges over sets. To eliminate references to the scenario, one should express the paradox roughly as asserting that it is contradictory to postulate the existence of an object that stands in a relationship with itself if and only if it does not stand in that relationship with itself. This version of the paradox conforms to the elimination thesis and successfully captures the argument of the paradox without making any reference to sets or R. The elimination of the details concerning sets seems to have a significant impact on the paradox itself. In particular, the removal of the scenario impacts the dialectical force of the paradox. While the argument successfully identifies a contradiction, it may result in an assertion that is very abstract and hard to understand, especially for the non-expert audience. The scenario exemplifies the abstract schema and shows a context, set theory, in which the paradox occurs. As such, it makes the paradox more effective and immediately convincing, at least for experts of that specific context.

The importance of context can also be explored through a comparison with the Barber paradox. The Barber paradox shows that it is impossible to have a barber who shaves all and only the men of a village who don't shave themselves. Indeed, if the barber shaves himself, by definition he doesn't shave himself. On the contrary, if he doesn't shave himself, by definition he does. Therefore, the abstract schema of the Barber paradox is essentially the same as Russell's paradox: it is contradictory to postulate the existence of an object that stands in a relationship with itself if and only if it does not stand in that relationship with itself. Since the two paradoxes can be reduced to the same abstract argument through the elimination of their scenarios, they are indistinguishable by the elimination thesis. However, this conflicts with the way we treat them: they have very different subject matters and, consequently, very different impacts on the philosophical debate.

On one hand, Russell's paradox reveals that a set composed of all and only the sets that are not members of themselves cannot exist. This conclusion contradicts

our naïve understanding of set theory, which suggests that for any condition you can formulate, there is a class whose members meet the condition. Since this principle is deeply ingrained in our thinking about sets and classes, it is not easily given up, and the conclusion of Russell's paradox appears highly implausible. Russell's paradox revolutionized set theory by showing how our naïve understanding of set theory is inconsistent (Quine, 1966). On the other hand, the Barber paradox faces a more nuanced judgment. Some regard it as a pseudo-paradox, finding the existence of the barber immediately implausible. Others consider it a paradox, albeit one that engenders only a moderate sense of paradoxicality in the reader, as the solution is very straightforward: the barber cannot exist (Irvine & Deutsch, 2021). Either way, the two paradoxes have the same abstract structure, but very different impacts on the philosophical debate: this difference seems to be determined by the object of the paradox, which is specified by the scenario. Thus, by providing an instance of the abstract schema, scenarios make the argument interesting and relevant for a specific debate.

Finally, consider the epistemic dimension of Russell's paradox. The argument clearly contains an assertion of existence that needs justification. In fact, the paradox essentially asserts that it is contradictory to postulate the existence of an object that stands in a relationship with itself if and only if it does not stand in that relationship with itself. However, the abstract schema per se fails to provide plausibility to the premise concerning the existence of the object described by the paradox, and the reader has no reason to initially assess the existence of the contradictory object as plausible. Up to this point, the situation seems to be very similar to that of the Sorites paradox. The set scenario provides an instance of the abstract schema that makes the inconsistency interesting and relevant for set theory. However, there is a fundamental difference between the Sorites case and Russell's case. While in the Sorites case the heap scenario directly gives plausibility to the empirically grounded existence premise by providing a possible instantiation of the abstract schema, the existence premise of Russell's paradox does not seem to be empirically grounded. Rather, the plausibility of R and the justification of the existence premise derive from a theoretical source, namely the unrestricted comprehension principle, according to which for any well-defined property there is a set of all and only the objects that have that property. This would mean that the set scenario of Russel's paradox does not impact its epistemic success: it would indicate the object and the context of the existence premise, but it would not be responsible for its plausibility. This would correspond to a partial counterexample to the account this paper presents, which applies in its epistemic part only to the paradoxes that contain an empirically grounded existence claim.

However, a further examination of the justification of the principle of unrestricted comprehension reveals an empirical root, since naïve set theory is connected to our ordinary experiences with collections and sets. In this sense, the existence premise concerning R would receive plausibility and justification from the unrestricted comprehension principle, but ultimately and more fundamentally from the empirical facts from which naïve set theory emerges. Therefore, it is reasonable to think that the R-scenario contributes to the plausibility of the existence assertion because, by describing a set, it recalls our pre-theoretical experience with sets and collections, according to which we find the existence of a set like R plausible.

Therefore, despite the fact that Russell's paradox differs from the Sorites, Grue, and Lottery paradoxes because of its scenario, which is abstract and minimal rather than concrete and rich in detail, it nevertheless seems consistent with my analysis. The description of R makes the existence of an object that stands in a relation to itself if and only if it does not stand in that relation to itself seem plausible because it is theoretically supported by the unrestricted comprehension principle and, more fundamentally, because it is empirically grounded in our pre-theoretic understanding of sets and collections.

4.2 Simpson's paradox

Thus far, my analysis has mainly focused on a strict notion of paradox. However, the term "paradox" is used in a broader sense in many philosophical contexts, usually to indicate cases that do not fit the standard definition, but still produce in the reader the sense of puzzlement that typically characterizes paradoxes. This can lead to confusion and disagreement about the correct way in which these borderline cases should be classified. For example, Quine's taxonomy of paradoxes would classify many of these cases, which share the same psychological and sociological features of paradoxes but do not reveal inconsistencies, as veridical paradoxes. Typically, veridical paradoxes initially confuse the reader, but then show that the unexpected conclusion is actually true. In a stricter context, however, they would only be considered pseudo-paradoxes. Simpson's paradox, Montague's paradox, and the Monty Hall problem are some famous examples of such cases. Since these cases are very similar to paradoxes and contain scenarios, it is worth investigating whether my thesis about the role of scenarios in paradoxes generalizes to such paradox-like cases.

Consider the example of Simpson's Paradox from Clark (2012, 233). The example compares the survival rates for a particular surgery in two different hospitals. For patients with initially good conditions, the survival rate is better at hospital A:

Good condition	Survived	Not survived	Total	Survival rate %
Hospital A	490	10	500	98
Hospital B	810	90	900	90

Similarly, patients with worse initial conditions also have a higher survival rate at hospital A:

Bad condition	Survived	Not survived	Total	Survival rate %
Hospital A	260	240	500	52

Bad condition	Survived	Not survived	Total	Survival rate %
Hospital B	30	70	100	30

Considering that both for patients with good initial conditions and patients with bad initial conditions hospital A has better survival rates, one would think that hospital A would have the overall best survival rates. However, the opposite is the case, as hospital B has the best overall survival rate:

Overall	Survived	Not survived	Total	Survival rate %
Hospital A	750	250	1000	75
Hospital B	840	160	1000	84

The conclusion drawn from Simpson's paradox is puzzling and seems unacceptable, yet it is statistically valid. Indeed, Simpon's paradox is a statistical phenomenon where the correlation between two variables in a population is reversed when examining subgroups within the same population. The underlying mathematics of this phenomenon is sound, and no inherent inconsistency is revealed. Indeed, the surprising phenomenon can be explained by the differing proportions of patients with favorable prior conditions treated at the two hospitals. Patients with favorable prior conditions tend to have better outcomes, and both hospitals A and B reflect this pattern. The overall success rate of hospital B appears higher due to 90% of its patients having favorable prior conditions, compared to only 50% at hospital A. Balancing the proportion of patients with prior good conditions in both hospitals would cancel this effect.

Simpson's paradox, therefore, qualifies as a paradox only in an extended sense: it produces a sense of puzzlement in the reader, and its conclusion appears unacceptable, even though it remains statistically acceptable and unproblematic from an argumentative point of view. The question then arises as to what renders the conclusion seemingly unacceptable or implausible to readers and whether scenarios play a role as they do in traditional paradoxes. Unlike standard paradoxes, cases like Simpson's paradox are not inherently inconsistent. Disengaging the scenario from the case, specifically any reference to hospitals and surgeries, leaves us with a statistical phenomenon demonstrating the potential for an association between two variables in a population to emerge, vanish, or reverse when the population is divided into subgroups. This may be surprising, yet not inherently perplexing or paradoxical. The paradoxical appearance of Simpson's paradox emerges only when scenarios are introduced. Indeed, scenarios illustrate the statistical phenomenon and elucidate how it manifests in the real world. For instance, Simpson's paradox is frequently illustrated through scenarios related to hospitals, surveys, or admission rates. When presented with a tangible instantiation of the abstract statistical phenomenon, readers tend to formulate

conclusions that they perceive as plausible. In the context of the hospital scenario, readers expect that hospital A, displaying superior success rates for patients with both good and poor prior conditions, also has the best overall success rate. However, the case's conclusion contradicts this expectation, leading readers to deem it implausible and paradoxical. In this sense, scenarios contribute to the epistemic dimension of the case and play an ineliminable role in making the case look like a paradox. Moreover, they surely contribute to the dialectical dimension of these cases as well. By exemplifying the statistical phenomenon, the various scenarios of Simpson's paradox make it easier to understand for the non-expert reader.

Therefore, a brief analysis of Simpson's paradox seems to suggest that scenarios play in paradox-like cases a role that is quite similar to what they play in standard paradoxes: namely they instantiate the argument and thus contribute to the epistemic and dialectic dimension of the cases. They give plausibility to the premises and make the conclusion appear implausible, and generally contribute to making the case easier to understand by providing concrete examples.

5 Conclusion

This paper fills a gap in the existing metaphilosophical research on paradoxes. Notably, not much attention has been given to the fact that paradoxes are typically presented via a scenario description. To understand the role of scenarios in paradoxes I have first investigated whether they can be eliminated from paradoxes. Case analysis has shown that scenarios cannot be removed from paradoxes without impacting the effectiveness of the paradox. Having clarified that scenarios are essential to paradoxes, I have investigated their contributions, with the aim of better understanding the functioning of paradoxes and, at the same time, explaining the observed use of scenarios in paradoxes. Fundamentally, scenarios provide an example or instantiation of the abstract schema of the paradox. This brings scenarios to contribute to paradoxes on two levels. First, they improve the dialectical power of paradoxes, facilitating their comprehension, particularly for non-experts, by rendering paradoxes more concrete and contextualized. Second, and more crucially, scenarios are essential for the epistemic dimension of paradoxes. Indeed, paradoxes become significant only when the inconsistencies they uncover appear realizable. By providing examples of the abstract schema, scenarios give the necessary element of plausibility to the premises, thus making the paradox relevant to the philosophical investigation and creating the characteristic sense of puzzlement in the reader. In particular, scenarios are the source of plausibility and justification for those premises that contain an empirically grounded existence claim. Contrary to the dialectical role, the epistemic function of scenarios is indispensable, as it connects paradoxes to the real world and underscores their significance in specific debates. Interestingly, case analysis suggests that scenarios play this role not only in standard paradoxes but also in paradox-like cases such as Simpson's paradox. Future work could expand the investigation into these cases, as well as include other scenarios, like thought experiments. Moreover, it would be interesting to expand the investigation on the function of scenarios in paradoxes and paradox-like

cases from the dimension of justification to the dimension of discovery, where they could also play a role.

Acknowledgements I would like to thank Joachim Horvath, Vincenzo Crupi, Daniel Dohrn and Thomas Grundmann for critical and helpful comments on earlier versions of this paper. I also thank the anonymous reviewers who made useful suggestions for improvement.

Funding Open access funding provided by Università degli Studi di Torino within the CRUI-CARE Agreement.

Declarations

Conflict of interest The author has no competing interests to declare that are relevant to the content of this article.

References

Cowie, C. (2023). What are paradoxes? *Journal of the American Philosophical Association, 9*(1), 154–171. https://doi.org/10.1017/apa.2021.48

Hyde, Dominic, and Diana Raffman. 2018. "Sorites Paradox." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2018. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2018/entries/sorites-paradox/.

Irvine, Andrew David, and Harry Deutsch. 2021. "Russell's Paradox." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2021. Metaphysics Research Lab, Stanford University. Retrieved May 4, 2023, from. https://plato.stanford.edu/archives/spr2021/entries/russell-paradox/.

James Robert, Brown. (1992). Why empiricism won't work. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association., 1992*(2), 271–79. https://doi.org/10.1086/psaprocbienmeetp.1992.2.192841

Lycan, W. G. (2010). What, exactly, is a paradox? *Analysis, 70*(4), 615–622. https://doi.org/10.1093/analys/anq069

Michael, Clark. (2012). *Paradoxes from A to Z* (3rd ed.). Routledge.

Norton, J. D. (1991). *Thought experiments in Einstein's work*. Rowman & Littlefield.

Quine, W. V. 1966. *The Ways of Paradox and Other Essays*. New York: Random House.

Sainsbury, R. M. (2009). *Paradoxes* (3rd ed.). Cambridge University Press.

Sorensen, Roy. 2022. "Epistemic Paradoxes." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2022. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2022/entries/epistemic-paradoxes/.