



---

# CRITIQUE OF RENÉ DESCARTES' CONCEIVABILITY ARGUMENT

Chenghao Li<sup>1</sup>

<sup>1</sup>Brigham Young University, USA

## Abstract

This paper first explains the difference between conceivability and metaphysical possibility from the perspective of possible worlds. Then, it introduces Saul Kripke's view on the metaphysical necessity of identity, contributing to a premise in Descartes' Conceivability Argument. Next, this paper interprets the Conceivability Argument. Then, this paper argues that even if "two things are distinct" is conceivable, we cannot conclude that their distinction is metaphysically possible, so conceivability does not entail metaphysical possibility. Thus, the Conceivability Argument is problematic. This paper also responds to a potential objection that, given humans have clear and distinct ideas, conceivability will entail metaphysical possibility.

## Keywords

Metaphysically, Descartes' Conceivability Argument, Philosophy, Substance Dualism

---

## I. Introduction

In *Meditations on First Philosophy*, René Descartes uses the Conceivability Argument to defend substance dualism, the idea that the mind and the body are distinct. The argument says that since it is conceivable that the mind exists without the body, it is metaphysically possible that the mind exists without the body; thus, the mind and the body are distinct. A problem appears when applying this argument to the way we conceive water and H<sub>2</sub>O. People can conceive that water is composed not of hydrogen and oxygen but something else. Thus, "water is not H<sub>2</sub>O" is conceivable. However, scientists have discovered that water is identical to H<sub>2</sub>O. Then, in every metaphysically possible world, "water" and "H<sub>2</sub>O" refer to the same thing, so "water is not H<sub>2</sub>O" is metaphysically impossible.

This paper first explains the difference between conceivability and metaphysical possibility from the perspective of possible worlds. Then, it introduces Saul Kripke's view on the metaphysical necessity of identity, contributing to a premise in Descartes' Conceivability Argument. Next, this paper interprets the Conceivability Argument. Then, this paper argues that even if "two things are distinct" is conceivable, we cannot conclude that their distinction is metaphysically possible, so conceivability does not entail metaphysical possibility. Thus, the Conceivability Argument is problematic. This paper also responds to a potential objection that, given humans have clear and distinct ideas, conceivability will entail metaphysical possibility.

## II. Conceivability and Metaphysical Possibility

This section introduces conceivability and metaphysical possibility from the perspective of possible worlds. There are two categories of possible worlds: conceivable and metaphysically possible. A conceivable world is a whole way that reality could be, which is logically consistent and can be fully described in terms of the properties and relations of the things within it. Each conceivable world is a self-contained and self-consistent system that may differ from the actual world in various respects, such as the laws of nature, the arrangement of matter and energy, or the values of specific properties. The world may be unreal, but we can imagine or conceive it. Metaphysically possible worlds are hypothetical scenarios that describe a reality consistent with fundamental laws of metaphysics, including causality, essence, and supervenience. What is metaphysically possible is constrained by things' real nature or identity.

For instance, a world containing a square with two sides is inconceivable and metaphysically impossible. A square must have four equal sides, and the phrase "a square with two sides" is conceptually incoherent

; namely, the meanings of “square” and “two sides” contradict each other, so no square with two sides exists. Some worlds are metaphysically possible but inconceivable. Consider a four-dimensional world: humans cannot conceive such a world, for human cognition is limited to three-dimensional space; however, they can depict and represent a four-dimensional space by mathematics, so such a world is metaphysically possible. Our real world is both conceivable and metaphysically possible. Similarly, we can conceive a world identical to our real world except that the sun is as large as the Earth; since such a world does not violate logic or metaphysical laws, it is also conceivable and metaphysically possible.

### III. Kripke’s Necessity of Identity

In *Identity and Necessity*, Kripke defends the necessity of identity that relates to metaphysically possible worlds (1971). For any objects  $x$  and  $y$ , if  $x$  and  $y$  are identical in the real world, they are also necessarily identical in all metaphysically possible worlds. To argue this, he starts from Joseph Butler’s view that “everything is what it is and not another thing.” From the perspective of possible worlds, Butler’s view means: “In all metaphysically possible worlds, any object is identical to itself.” For instance, in some metaphysically possible worlds, Richard Nixon is not the thirty-seventh president of the United States. Still, Richard Nixon cannot be Ronald Reagan, John Kennedy, or any other person in any metaphysically possible world.

Next, Kripke argues that if “any objects  $x$  and  $y$  are identical in the real world,” then “if  $x$  is identical to itself in all metaphysically possible worlds, then  $x$  is identical to  $y$  in all metaphysically possible worlds.” He proves this proposition with Leibniz’s Law and the Principle of Substitutivity. Leibniz’s Law says, “for any two things in the real world to be identical, they must share all their properties.” For instance, Hesperus and Prosperus are identical in the real world, for they are the same star: Venus. By Leibniz’s Law, Hesperus and Prosperus share all their properties, including slow rotations around the sun, no moons or rings, extremely hot temperatures, etc. The Principle of Substitutivity follows Leibniz’s Law, stating that if any two things,  $x$  and  $y$ , are identical in the real world, then if  $x$  has any property  $F$ , then  $y$  also has the property  $F$ . For instance, since Hesperus and Prosperus are identical in the real world, by knowing that Hesperus has a thick atmosphere, we can infer that Prosperus also has a thick atmosphere. In the proof, Kripke considers the property of “being identical to  $x$  in all metaphysically possible worlds.” Given  $x$  and  $y$  are identical in the real world, by the Principle of Substitutivity, if  $x$  is identical to  $x$  (itself) in all metaphysically possible worlds,  $x$  is identical to  $y$  in all metaphysically possible worlds. For instance, given Hesperus and Prosperus are identical in the real world, if “Hesperus is identical to Hesperus” is metaphysically necessary, then “Prosperus is identical to Hesperus” is also metaphysically necessary. Thus, by Leibniz’s Law and the Principle of Substitutivity, if “any objects  $x$  and  $y$  are identical in the real world,” then “if  $x$  is identical to itself in all metaphysically possible worlds, then  $x$  is identical to  $y$  in all metaphysically possible worlds.” From Butler’s view,  $x$  must be identical to itself in all metaphysically possible worlds. Thus, Kripke concludes that for any objects  $x$  and  $y$ , if  $x$  and  $y$  are identical in the real world, then  $x$  and  $y$  are identical in all metaphysically possible worlds, which reveals his identity of necessity. I construct Kripke’s argument on the necessity of identity as follows:

P1. In all metaphysically possible worlds, any object  $x$  is identical to itself. (By Joseph Butler’s view that “everything is what it is, and not another thing”)

P2. If  $x$  and  $y$  are identical in the real world, then if  $x$  is identical to itself in all metaphysically possible worlds, then  $x$  and  $y$  are identical in all metaphysically possible worlds. (By Leibniz’s Law and the Principle of Substitutivity with the property “being identical to  $x$  in all metaphysically possible worlds”)

C. If  $x$  and  $y$  are identical in the real world, then  $x$  and  $y$  are identical in all metaphysically possible worlds. (*Modus Ponens*, P1, P2)

An instantiation of conclusion C in the above argument is that if water and H<sub>2</sub>O are identical in the real world, then water and H<sub>2</sub>O are identical in all metaphysically possible worlds. Conclusion C can also be reinterpreted by conditional negation. The negation of C’s consequent is “ $x$  and  $y$  are distinct in some metaphysically possible worlds.” The negation of C’s antecedent is “ $x$  and  $y$  are distinct in the real world.” Then, C could be reinterpreted as “If  $x$  and  $y$  are distinct in some metaphysically possible worlds, then  $x$  and  $y$  are distinct in the real world.” If we call this reinterpretation C\*, then C\* states, “If it is metaphysically possible that two things are distinct, then they are distinct in the real world.” For instance, C\* implies that if it is metaphysically possible that the mind and the body are distinct, they are distinct in the real world. In Descartes’ Conceivability Argument, C\* is the third premise.

#### IV. Descartes' Conceivability Argument

Descartes defends substance dualism that the mind and the body are distinct. A body is a substance that fills space, is bounded by some shape, and is enclosed in a place. A mind is a thinking substance that doubts, understands, affirms, denies, wills, refuses, imagines, and senses. The body is extended, while the mind is not. To defend that the mind and the body are distinct, Descartes appeals to clear and distinct ideas. Clear ideas are those not subject to doubt, and distinct ideas are those fully understood as true. For instance, mathematical truths like " $1 + 1 = 2$ " are clear and distinct. The idea of pain in my body is clear but indistinct, for it could be falsely understood if I feel it while dreaming. Descartes argues that ideas of the mind and the body are clear and distinct, for he can conceive his mind as an unextended, thinking thing and his body as an extended, non-thinking thing. Also, we can conceive a mind without a body and a body without a mind (Descartes, 1993). For instance, we can conceive a powerful, unextended Artificial Intelligence that has cognitive states and understands humans' words; we can also conceive an extended philosophical zombie physically identical to a conscious human being but lacking consciousness entirely.

The Conceivability Argument starts that, given clear and distinct ideas of the mind and the body, "the mind and the body are distinct" is conceivable, and if "two things are distinct" is conceivable, "the two things are distinct" is metaphysically possible; thus, "the mind and the body are distinct" is metaphysically possible. By Kripke's necessity of identity, if it is metaphysically possible that two things are distinct, they are distinct in the real world. Since "the mind and the body are distinct" is metaphysically possible, the mind and the body are distinct in the real world. The Conceivability Argument is constructed as follows:

- P1. "The mind and the body are distinct" is conceivable. (We can conceive a mind without a body and a body without a mind)
- P2. If "two things are distinct" is conceivable, then "the two things are distinct" is metaphysically possible.
- C1. "The mind and the body are distinct" is metaphysically possible. (*Modus Ponens*, P1, P2)
- P3. If it is metaphysically possible that two things are distinct, they are distinct in the real world. (This is C\* in Kripke's argument on the necessity of identity)
- C2. The mind and the body are distinct in the real world. (*Modus Ponens*, C1, P3)

The Conceivability Argument has been both praised and criticized over the years. The argument brings to light the fundamental question of the relationship between the mind and the body. However, some philosophers have criticized the argument, for it does not convincingly explain how the mind and the body interact. In the next section, I will show that the Conceivability Argument is problematic in its second premise (P2).

#### V. Against Descartes' Conceivability Argument

This section argues that given "water is not H<sub>2</sub>O" is conceivable but metaphysically impossible, the second premise (P2) in the Conceivability Argument fails, making this argument unsound.

The Conceivability Argument presumes P2 that "if two conceivable things are distinct, then it is metaphysically possible that they are distinct." From the perspective of possible worlds, P2 indicates that all conceivable worlds are metaphysically possible, so conceivability entails metaphysical possibility. To reject P2, I will consider the example of water and H<sub>2</sub>O. Water is a colorless, odorless, and tasteless liquid essential for most life forms on Earth. Before discovering that water is composed of hydrogen and oxygen atoms, humans had various theories about the nature of water. Some ancient Greeks believed water, earth, air, and fire were among the four essential elements. Other ideas suggested that water was a single substance that could transform into different forms, such as ice or steam. In the eighteenth century, the phlogiston theory proposed that water was a compound of hydrogen and phlogiston, an imaginary substance thought to be released during combustion. It was not until the nineteenth century that scientists could determine the accurate composition of water through experiments like electrolysis. Thus, like ancient Greeks and phlogiston theorists, humans can conceive that water is made not from Hydrogen and Oxygen but from something else, so "water is not H<sub>2</sub>O" is conceivable. However, scientists have discovered that water is composed of hydrogen and oxygen atoms with the chemical formula of H<sub>2</sub>O, so water is identical to H<sub>2</sub>O. Even though "water" and "H<sub>2</sub>O" are two concepts, they refer to the same thing with the same properties and cannot be separated. By Kripke's necessity of identity, since water and H<sub>2</sub>O are identical in the real world, they are identical in all metaphysically possible worlds, so "water is not H<sub>2</sub>O" is metaphysically impossible. Then, a world where water is composed not of H<sub>2</sub>O molecules but of other substances (like hydrogen and phlogiston) is conceivable but metaphysically impossible. I use the following argument to summarize this:

P1. If  $x$  and  $y$  are identical in the real world, then  $x$  and  $y$  are identical in all metaphysically possible worlds. (Conclusion C in Kripke's argument on the necessity of identity)

P2. Water and H<sub>2</sub>O are identical in the real world.

C1. Water and H<sub>2</sub>O are identical in all metaphysically possible worlds. (*Modus Ponens*, P1, P2)

C2. "Water is not H<sub>2</sub>O" is metaphysically impossible. (C1)

P3. "Water is not H<sub>2</sub>O" is conceivable. (We can conceive water is composed of hydrogen and phlogiston instead of H<sub>2</sub>O)

C3. "Water is not H<sub>2</sub>O" is conceivable but metaphysically impossible. (*Conjunction*, C2, P3)

Given C3 in the above argument, "Water is not H<sub>2</sub>O" makes conceivability no longer entail metaphysical possibility, which denies the second premise in the Conceivability Argument and makes it unsound. My argument against the Conceivability Argument could be constructed as follows:

P1. Assume the second premise in the Conceivability Argument is true; if "two things are distinct" is conceivable, then "the two things are distinct" is metaphysically possible. (*Assumption for Indirect Proof*)

C1. If "water and H<sub>2</sub>O are distinct" is conceivable, "water and H<sub>2</sub>O are distinct" is metaphysically possible. (*Universal Instantiation*, P1)

P2. "Water is not H<sub>2</sub>O" is conceivable but metaphysically impossible.

C2. The second premise in the Conceivability Argument is false. (*Indirect Proof*, C1 and P2 conflict with each other)

P3. If any premise in the Conceivability Argument is false, this argument is unsound. (*The basic rule of logical deduction*)

C3. The Conceivability Argument is unsound. (*Modus Ponens*, C2, P3)

One clarification is that my argument against the Conceivability Argument merely shows that the mind and the body may not be distinct in the real world. Still, it does not exclude the possibility that the mind and the body are separate substances. In other words, it does not deny Cartesian substance dualism but rather argues the Conceivability Argument is flawed.

## VI. Objection and Response

Descartes may respond to my argument that conceivability will entail metaphysical possibility if humans have clear and distinct ideas in their conception. Since people have clear and distinct ideas of the mind and the body, they will not conceive that the mind and the body are indistinct. Therefore, the Conceivability Argument still holds.

I have two responses to this objection. First, Descartes argues clear and distinct ideas are ideas not subject to doubt and fully understood as true. Given this account, I use illusionism and physicalism to respond to Descartes' objection. According to Keith Frankish, illusionism holds that phenomenal consciousness and minds are illusory, and that experiences do not possess any physical or non-physical qualitative properties (Frankish, 2016). Physicalism holds that minds and mental states are reducible or identical to physical states and processes in the brain. Then, illusionism and physicalism suggest that the mind is either illusory or physically reducible. Illusionism rejects substance dualism since the mind does not exist; physicalism also denies it, for the mind could be eventually reduced to the body. From the illusionists' and physicalists' accounts, they have clear and distinct ideas of the mind and the body but still deny substance dualism's claim that the mind and the body are distinct. Even though we assume that if humans have clear and distinct ideas in their conception, conceivability will entail metaphysical possibility, humans may still not have clear and distinct ideas of the mind and the body. The mind-body problem is still an active area of debate and research in the philosophy of mind. Various schools of philosophical thought and theories attempt to address the issue, including physicalism, non-reductivism, illusionism, etc., but no consensus has been reached yet.

My second response focuses on the limitation of Descartes' "clear and distinct ideas." For instance, Descartes claims that pure mathematics is clear and distinct (1993). However, Gödel's First Incompleteness Theorem states that no consistent axiomatic system for arithmetic is complete; namely, an axiomatic system for

arithmetic has some unprovable true proposition. Pure mathematics is an axiomatic system for arithmetic, and mathematical propositions like “ $1 + 1 = 2$ ” relies on Peano Axioms in arithmetic. By Gödel’s First Incompleteness Theorem, we cannot have clear and distinct ideas of all true propositions in pure mathematics, for some are unprovable. Thus, “clear and distinct ideas” are limited in that certain true ideas are not clear or distinct. Similarly, the ideas of the mind and the body may encounter the same problem as those unprovable true propositions in pure mathematics. Given this limitation, we may be unsure that the ideas of the mind and the body are clear and distinct, even though they seem so to Descartes. Therefore, the Conceivability Argument is still problematic.

## VII. Conclusion

Starting from Kripke’s necessity of identity, this paper introduces conceivability and metaphysical possibility from the perspective of possible worlds and constructs Descartes’ Conceivability Argument. Then, this paper argues that “water is not  $H_2O$ ” is conceivable but metaphysically impossible, which denies the second premise in the Conceivability Argument and makes it unsound. This paper also responds to a potential objection to Descartes’ “clear and distinct ideas,” one from philosophical theories of mind and the other from the limitation of “clear and distinct ideas.” Further critiques of Descartes’ Conceivability Argument may focus on the ontological gap between conceivability and metaphysical possibility, the zombie argument proposed by David Chalmers, and the dualistic ontology that the mind and the body interact with each other.

## Works Citation

Descartes, R. (1993). *Meditations on First Philosophy*. Hackett Publishing Company.

Frankish, K. (2016). Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies*, 23(11-12), 11-39.

Kripke, S. & Munitz, M. K. (Ed.) (1971). *Identity and Individuation*. New York University Press.