Scientific Research Publishing

Open Journal of Philosophy, 2025, 15(3), 567-574
https://www.scirp.org/journal/ojpp
ISSN Online: 2163-9442
ISSN Print: 2163-9434

On the Ability Hypothesis Reply to the Knowledge Argument

Chi-Fai Lo

Department of Philosophy, School of Historical Studies, Birkbeck University of London, London, UK
Email: edcflo@gmail.com

How to cite this paper: Lo, C.-F. (2025). On the Ability Hypothesis Reply to the Knowledge Argument. *Open Journal of Philosophy, 15,* 567-574.
https://doi.org/10.4236/ojpp.2025.153034

Received: May 8, 2025
Accepted: July 28, 2025
Published: July 31, 2025

Abstract

This paper has critically assessed the Ability Hypothesis reply to the Knowledge Argument. By means of a modified version of Mary's Room, which is based upon an adaptation of Albert Einstein's thought experiment for Principle of Equivalence, the insufficiency of the Ability Hypothesis and its failed attempt to block the Knowledge Argument have been demonstrated. In addition, it has been revealed that a possible loophole due to ignorance exists in the Knowledge Argument for new scientific findings are yet to come to challenge our current concepts of physical reality.

Keywords

Ability Hypothesis, Knowledge Argument, Mary's Room, Thought Experiment

1. Introduction

In contemporary philosophy of mind there are two major theories about mind and body, namely dualism and physicalism. Dualism holds the view that mind and body are two fundamentally different kinds of substances; "mind is indivisible, extended or thinking, and incorporeal, whereas body is divisible, unextended or unthinking, and corporeal." (Blocker & Hannaford, 1974, pp. 134-135). That is, minds and bodies do not share any properties in common; for instance, unlike bodies, minds do not have physical properties like size, shape, colour and weight. The essence of a mind is to think, be conscious, and engage in other mental activities, whereas a body has physical properties like spatial extensions and being located in space. Most important of all, minds, unlike bodies, cannot disintegrate and die. Although mind and body are separate substances, there exists a two-way causal interaction between them, giving rise to the well-known mind-body problem (i.e. how an immaterial mind interacts with a physical body) which consti-

tutes a major challenge to dualism (Blocker & Hannaford, 1974: pp.134-135). On the contrary, physicalism claims that mind and body are not two distinct substances, both of which can be ultimately explained in terms of physical matter and processes (Blocker & Hannaford, 1974: p.139). Since all phenomena involving both mind and body are governed by physical laws, and can be apprehended through scientific understanding, especially through neuroscience and the study of the brain, the mind-body problem is resolved readily (Blocker & Hannaford, 1974: p. 139). Nevertheless, physicalism runs into the hard problem of explaining consciousness and subjective experience such as feeling pain, enjoying music or seeing beautiful flowers (Blocker & Hannaford, 1974: p. 144).

In 1982 Frank Jackson put forward the Knowledge Argument to challenge physicalism whether there exist aspects of human conscious subjective experience (qualia) that are beyond physical knowledge (Jackson, 1982). The argument is based upon a thought experiment commonly known as Mary's Room:

> "Mary is confined to a black-and-white room, is educated through black-and-white books and through lectures relayed on black-and-white television. In this way she learns everything there is to know about the physical nature of the world. She knows all the physical facts about us and our environment, in a wide sense of 'physical' which includes everything in completed physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this, including of course functional roles. If physicalism is true, she knows all there is to know. For to suppose otherwise is to suppose that there is more to know than every physical fact, and that is just what physicalism denies… It seems, however, that Mary does not know all there is to know. For when she is let out of the black-and-white room or given a colour television, she will learn what it is like to see something red, say. This is rightly described as learning—she will not say 'ho, hum.' Hence physicalism is false. This is the knowledge argument against physicalism." (Jackson, 1986).

In short, the Knowledge Argument can be stated as follows:

Premise 1: Mary, who has lived inside a black-and-white room for her whole life, has learned all of the physical descriptions about colour, but she has never seen any visible colours other than black and white.

Premise 2: When Mary steps out of her room or is given a colour television for the first time and sees something red, she learns what it is like to see something red.

Conclusion: Mary acquires some new knowledge that is concerned with subjective, qualitative properties of experiences (or qualia) rather than physical properties, so physicalism is false.

Jackson's conclusion apparently aims at disproving physicalism and favours dualism (Nida-Rumelin & Conaill, 2024). As it is difficult to deny that Mary does gain some new knowledge upon leaving her room or watching colour television, the Knowledge Argument undoubtedly inflicted a substantial impact upon the va-

lidity of physicalism, and has sparked a significant debate about the nature of consciousness and its relation to physicalism ever since.

For the past several decades physicalists have actively engaged in generating various responses to the Knowledge Argument, but none of these responses appears to be overwhelmingly convincing. Except the claim that the knowledge of what it is like to see red is not expressible in physical terms, the thought experiment does not provide any elaboration of what Mary learns from her first-person colour experience. Therefore, physicalists generally tend to focus on the nature of the new knowledge in hope of finding various holes in the thought experiment. One of the most influential objections to the Knowledge Argument is the Ability Hypothesis, which argues that Mary has learned an ability rather than knowledge of a fact (Nida-Rumelin & Conaill, 2024). In this paper I attempt to explain and assess critically the Ability Hypothesis reply to the Knowledge Argument. By means of a modified version of Mary's Room, which is based upon an adaptation of Albert Einstein's thought experiment for Principle of Equivalence, I argue against the Ability Hypothesis and conclude that it fails to block the Knowledge Argument. Besides, in the proposed thought experiment I reveal a possible loophole due to ignorance in *Premise 1* of the Knowledge Argument for new scientific findings are yet to come to challenge our current concepts of physical reality.

2. The Ability Hypothesis

The Ability Hypothesis, proposed by David Lewis and Laurence Nemirow to defend physicalism against the Knowledge Argument, claims that what Mary learns from her first-person colour experience (i.e. what it is like to see something red for the first time) is not any new physical knowledge (i.e. knowing that knowledge) for she has already learned all of the physical descriptions about colour (Nida-Rumelin & Conaill, 2024). Instead she gains certain new abilities (i.e. knowing how knowledge), namely the practical non-conceptual abilities to imagine what it is like to see red, to remember the experience of seeing red, and to recognise red when seeing red objects again (Lewis, 1983, 1988). In epistemology "knowing that knowledge" and "knowing how knowledge" are commonly contrasted as propositional knowledge and procedural knowledge, respectively. Whilst the former refers to explicit knowledge of facts or propositions, and can be expressed in declarative statements which can be true or false, the latter is procedural or skill-based knowledge that is often implicit, can be demonstrated through action, and learned through practice and experience, rather than through verbal expression. Their difference can be easily illustrated by the following two examples:

1) "I know that London is the capital of the United Kingdom." (knowing that knowledge).

2) "I know how to drive a car." (knowing how knowledge).

Undoubtedly, no one will deny that one can know all the physical facts about cars without knowing how to drive a car and needs to learn the skill of driving a car personally (i.e. a first-person experience). In short, the central idea of the Abil-

ity Hypothesis is that knowing what it is like to experience something is not about acquiring physical knowledge but about acquiring new abilities: (Nida-Rumelin & Conaill, 2024).

> "The Ability Hypothesis says that knowing what an experience is like just is the possession of these abilities to remember, imagine, and recognize. It isn't the possession of any kind of information, ordinary or peculiar… It isn't knowing-that. It's knowing-how." (Lewis, 1988).

For instance, when Peter, who knows all the technicalities, scales, and history of piano playing but has never played one, finds a piano and starts playing, he does not gain any new facts about music or the piano, but he does acquire new abilities about actually playing piano instead. Hence, on the one hand, the Knowledge Argument challenges physicalism's capability to explain qualia and supports dualism; on the other hand, the Ability Hypothesis criticises the misunderstanding of the nature of Mary's newly acquired knowledge in the thought experiment, and attempts to block the Knowledge Argument so as to defend physicalism by shifting the discussion from "knowing that knowledge" to "knowing how knowledge".

Beyond question, there are criticisms of the Ability Hypothesis from proponents of the Knowledge Argument, focusing on its inability to fully explain the subjective, qualitative aspects of experience; that is, the Ability Hypothesis fails to recognise that the new knowledge gained by Mary is not simply the acquisition of new abilities but also knowledge of what it is like to experience colour (Nida-Rumelin & Conaill, 2024). Particularly, in 1996 David Chalmers posed a striking argument against the Ability Hypothesis:

> "No doubt Mary does gain some abilities when she first experiences red, as she gains some abilities when she learns to ride a bicycle. But it certainly seems that she learns something else: some facts about the nature of experience. For all she knew before, the experience of red things might have been like this, or it might have been like that, or it might even have been like nothing at all. But she knows that it is like this. She has narrowed down the space of epistemic possibilities. No such new knowledge comes along when an omniscient mechanic learns to ride a bicycle." (Chalmers, 1996).

According to David Chalmers, Mary indeed acquires not only new abilities but also new knowledge from her first-person experience of colour, implying that the Ability Hypothesis is inadequate to explain "what it is like" and its nature (Chalmers, 1996). In addition, along this line of thought, Michael Tye suggested a hypothetical scenario in which, prior to seeing red, Mary's brain is "zapped" in such a way that her ability to form new memories is lost even though the overall functioning of her brain is kept unchanged (Tye, 2000). Thus, upon her first-person experience of seeing red, Mary is unable to remember it, imagine it, and recognise it. Then William Tye argues that if Mary still knows what it is like to experience colour while she is seeing red, then the "knowing how knowledge" is not

necessary for knowing what it is like to see red (Tye, 2000). Moreover, there are criticisms about equating "knowing how knowledge" with certain abilities in the Ability Hypothesis; (Nida-Rumelin & Conaill, 2024) for example, the current world record holder for the 100 m sprint, Usain Bolt, does know how to achieve his world record of 9.58 seconds, but he is no longer able to do so. Jason Stanley and Timothy Williamson even argue that "knowing how knowledge" is actually knowing *that such-and-such is a way to do so-and-so* (i.e. a form of "knowing that knowledge"), so it should not be equated with acquiring certain abilities (Stanley & Williamson, 2001).

3. A Thought Experiment

To further demonstrate the insufficiency of the Ability Hypothesis, we consider the following example which is adapted from a thought experiment devised by Albert Einstein for Principle of Equivalence at the beginning of 20th century. A group of ten students at a technical school are invited to participate an experiment of experiencing zero gravity; one of them is called Issac, who was born disabled and is unable to walk. They have learned Newtonian mechanics and Newton's Law of Universal Gravitation, but they have never experienced zero gravity. The students are inside a very big cabin and isolated from the outside world. Just like Galileo portrayed in popular science books, a student called Albert picks up two objects of identical shape but of different weights (marked 1 kg and 10 kg) from the floor and lets them drop together. The two objects fall and touch the floor simultaneously in the same way one would expect given his or her experiences on Earth. There is a balance on the floor, too. Albert confirms the weight of both objects with the balance. Then he steps on the balance and checks his weight. It reads 65 kg, just the same reading that he obtained from a health check two days ago. Hence, all the students infer that the cabin is situated on the Earth's surface.

Very soon the students hear the announcement: "Gravity is now being switched off.", and find themselves approaching to a weightless situation steadily. Albert finds the balance under his feet showing 0 kg and can now lift the two weights upward with his fingers effortlessly. All of them, including Issac, can move up and down as well as sideways freely like flying birds; they also chase one another inside the cabin. After a short while, there is another announcement: "Gravity is now being switched on.", and all the students settle down on the floor gradually. Then a door is opened, and they find that the cabin is aboard a huge spaceship which accelerates uniformly in deep space (namely being far away from all significant masses and their gravitational attraction) at exactly 9.81 $m \cdot s^{-2}$, i.e. the acceleration due to gravity on the Earth's surface. The experimenters inform them that switching off (or on) the gravity, in fact, refers to stopping (or resuming) the acceleration. Hence, the students realise that they cannot tell whether the cabin is situated on the Earth's surface or moving at a uniform acceleration of 9.81 $m \cdot s^{-2}$ in deep space, from their vantage point as observers inside the cabin. Such indistinguishability of these two different situations is a manifestation of Principle of Equiva-

lence, which makes a revolutionary impact on our concept of space-time and gravitation as well as lays foundation for General Theory of Relativity (Einstein, 1923).

Now comes the big question: "What do we learn from this thought experiment?" First, prior to opening the door of the cabin, there is a consensus among the ten students that they learn what it is like to experience zero gravity for the first time although, like the physicists in the 19th century or before, they think they have already learned all physical facts about gravity, which are derived from Newtonian mechanics and Newton's Law of Universal Gravitation. Perhaps, reiterating the same argument, proponents of the Ability Hypothesis may claim that the students acquire the abilities to imagine what it is like to be weightless, to remember the experience of being weightless, and to recognise weightlessness when experiencing zero gravity again, rather than knowledge of a fact. This interpretation, however, neglects that the students enjoy the excitement of being able to move around freely like flying birds in the zero gravity environment.

Second, in addition to the first-person experience of zero gravity, Issac learns what it is like to be free from his disability and feelings of inferiority for the very first time. It is inevitable that he also learns what it is like to be handicapped by his disability again after switching on gravity. Obviously, none of his schoolmates is able to apprehend such an experience of Issac because it is beyond their first-person experience of zero gravity, and they were all born healthy. By analogy with Thomas Nagel's arguments in his essay "What is it like to be a bat?":

> "Our own experience provides the basic material for our imagination, whose range is therefore limited. It will not help to try to imagine that one has webbing on one's arms, which enables one to fly around at dusk and dawn, catching insects in one's mouth; that one has very poor vision, and perceives the surrounding world by a system of reflected high-frequency sound signals; and that one spends the day hanging upside down by one's feet in an attic." (Nagel, 1974).

which claims that the nature of subjective experience cannot be fully captured or understood through the objective scientific methods and the existence of an intrinsic "what it is like" aspect to consciousness is fundamentally inaccessible to a third person perspective, one may argue that, despite extensive physical knowledge about the physiology and psychology of a disabled person (i.e. the figurative bat), there is something about Issac's subjective experience that escapes the understanding of his healthy schoolmates (and anyone who was born healthy). Besides, it is too far-fetched to repeat the Ability Hypothesis to challenge against Issac's conscious subjective experience, namely enjoying his new-found freedom in the zero-gravity environment and feeling emotional pain due to the loss of his short-lived freedom after switching on gravity. Undoubtedly this emotional turmoil experienced by Issac for the very first time since birth cannot be interpreted as acquiring some kind of abilities. In particular, the emotional pain experienced by Issac is a central part of what philosophers and neuroscientists refer to as qualia;

it is introspectively accessible, qualitatively rich, and known from the first-person point of view (Chalmers, 1996; Block, 1995; Singer et al., 2004). As a result, the claims of the Knowledge Argument are being upheld.

Finally, after knowing that the cabin is aboard a spaceship travelling at a uniform acceleration of 9.81 $m \cdot s^{-2}$ in deep space, the students realise that they learn a new piece of physical knowledge about gravity, namely Principle of Equivalence and General Theory of Relativity. In other words, being unaware of Einstein's theory of gravity, the students, just like the physicists in the 19th century or before, have erroneously thought that all physical facts about gravity are known to them. This concern poses a serious problem on the validity of *Premise 1* of the Knowledge Argument because it took more than two hundred years to discover the insufficiency of Newton's Theory of Gravitation and to formulate Theory of General Relativity based upon Principle of Equivalence.

In summary, this thought experiment which is a modified version of Mary's Room has demonstrated the insufficiency of the Ability Hypothesis and its failure to block the Knowledge Argument. It has also shown a possible loophole due to ignorance in *Premise 1* of the Knowledge Argument for new scientific findings, especially in neurosciences, are yet to come to challenge our current concepts of physical reality.

4. Conclusion

In this paper I have attempted to explain and assess critically the Ability Hypothesis reply to the Knowledge Argument. A modified version of Mary's Room, which is based upon an adaptation of Albert Einstein's thought experiment for Principle of Equivalence, has been proposed to demonstrate the insufficiency of the Ability Hypothesis and to conclude that it fails to block the Knowledge Argument. Like those criticisms raised by proponents of the Knowledge Argument such as David Chalmers, I have argued that Issac acquires not only new abilities but also new knowledge from his first-person experience of weightlessness. More specifically, what Issac learns from his unique conscious subjective experience is what it is like to be a disabled person experiencing weightlessness first-hand for the very first time, what it means to be acquainted with or to experience freedom from disability and feelings of inferiority in the zero-gravity situation, as well as what feeling emotional pain due to the loss of his short-lived freedom after switching on gravity is like. Likewise, I have revealed a possible loophole due to ignorance in *Premise 1* of the Knowledge Argument; the students' knowledge of gravity is incomplete and limited by Newton's Law of Universal Gravitation. In fact, *Premise 1* is prone to having loopholes; for example, if one revises the Knowledge Argument by stipulating that Mary is blind from birth, then inconsistency will appear in *Premise 1* according to Gilbert Harman's argument:

> "The person blind from birth fails to know what it is like to see something red because he or she does not fully understand what it is for something to be red, that is, because he or she does not have the full concept of something's

being red. So… the person blind from birth does not know all the functional facts." (Harman, 1990).

All in all, the proposed thought experiment has succeeded in not only demonstrating the failed attempt of the Ability Hypothesis to block the Knowledge Argument but also revealing a possible loophole due to ignorance in *Premise 1* of the Knowledge Argument.

Acknowledgements

The author thanks Dr. Sarah Patterson for her useful comments and suggestions.

Conflicts of Interest

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

References

Block, N. (1995). On a Confusion about a Function of Consciousness. *Behavioral and Brain Sciences, 18,* 227-247. https://doi.org/10.1017/s0140525x00038188

Blocker, H. G., & Hannaford, W. (1974). *Introduction to Philosophy* (pp.134-135). D. Van Nostrand Company.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory* (p. 144). Oxford University Press.

Einstein, A. (1923). *The Principle of Relativity.* Dover.

Harman, G. (1990). The Intrinsic Quality of Experience. *Philosophical Perspectives, 4,* 31-52. https://doi.org/10.2307/2214186

Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly, 32,* 127-136. https://doi.org/10.2307/2960077

Jackson, F. (1986). What Mary Didn't Know. *The Journal of Philosophy, 83,* 291-295. https://doi.org/10.2307/2026143

Lewis, D. (1983). Postscript to Mad Pain and Martian Pain. In D. Lewis (Ed.), *Philosophical Papers* (Volume 1, pp. 130-132). Oxford University Press.

Lewis, D. (1988). What Experience Teaches. *Proceedings of the Russellian Society, 13,* 29-37.

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review, 83,* 435-450. https://doi.org/10.2307/2183914

Nida-Rumelin, M., & Conaill, D. O. (2024). Qualia: The Knowledge Argument. In E. N. Zalta, & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy*. Spring.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but Not Sensory Components of Pain. *Science, 303,* 1157-1162. https://doi.org/10.1126/science.1093535

Stanley, J., & Willlamson, T. (2001). Knowing How. *Journal of Philosophy, 98,* 411-444. https://doi.org/10.2307/2678403

Tye, M. (2000). Knowing What It Is Like: The Ability Hypothesis and the Knowledge Argument. In G. Pryer, & F. Siebelt (Eds.), *Reality and Human Supervenience: Essays on the Philosophy of David Lewis* (pp. 223-238). Rowman & Littlefield.