Article

Blind Man, Mirror, and Fire: Aquinas, Avicenna, and Averroes on Thinking

Zhenyu Cai

Department of Philosophy and Religious Studies, Peking University, Beijing 100871, China; zc287@pku.edu.cn

Abstract: In Islamic tradition, the Falsafa school is well known for its naturalistic account of religion. When Falsafa’s theory of religion made its way to the Latin West, it was embraced and developed into the so-called “double truth theory” in Latin Averroism. However, this theory quickly lost its influence in the Latin tradition, primarily due to the critique by Thomas Aquinas. One of the key aspects of Aquinas’s critique is his criticism of the emanation theory of concepts and the doctrine of the unity of the intellect, which in turn undermines the foundation of Falsafa’s theory of religion, particularly their theory of natural prophecy. This paper aims to revisit the debate between Aquinas and Falsafa regarding the theory of intellect as the basis for natural prophecy, with a focus on highlighting Falsafa’s perspective. In particular, I examine how Aquinas’s arguments overlook the key insights that underpin Falsafa’s doctrine of the intellect.

Keywords: Aquinas; Avicenna; Averroes; emanation; the unity of intellect

1. Introduction

In Islamic tradition, the Falsafa school is well known for its naturalistic account of the nature of religion. According to this theory, religion is a system of opinions and behaviours established by the first legislator in a community. A religion is considered complete only when its legislator is a true prophet. But what defines a true prophet? According to Falsafa, a true prophet has a dual role: they can grasp the truth of the entire reality and establish the truth as a system of opinions and behaviours at the community level in a poetic way. In what sense is a legislator with this dual role considered a prophet? This is where Falsafa’s theory of natural prophecy comes into play.

Falsafa’s account of natural prophecy aims to provide a naturalistic explanation for the phenomenon of prophets in light of a specific theory of intellect. The general strategy can be outlined as follows: human intellectual cognition involves receiving intelligible forms from the agent intellect that exists independently in the celestial world and is often referred to as the giver of forms. It explains both the substantial changes in the sublunar world and how the human intellect acquires intelligible forms. Based on this theory of intellect, prophetic revelation is considered a limiting situation where the human intellect receives intelligible forms from the agent intellect. In this situation, certain gifted individuals can receive knowledge about the order of the entire reality from the agent intellect in a very short period of time and express this knowledge in their imagination.

Within Falsafa, there are different views on how the human intellect receives intelligible forms from the agent intellect. Al-Fārābī and Avicenna believe that every individual has the ability to receive these forms. This ability is called the material intellect. They developed the emanation theory of concepts to explain how the material intellect acquires concepts, suggesting that our concepts of external objects emanate from a separate agent intellect. In contrast to the views of Al-Fārābī and Avicenna, who believe that every human being has their own material intellect, Averroes argues that the material intellect is not possessed by individuals but is a distinct entity shared by all humans. This position is called the doctrine of the unity of the intellect. Both the emanation theory of concepts and the doctrine of
the unity of the intellect play foundational roles in Falsafa’s theory of religion, as the two different doctrines of intellect form the basis for the theory of natural prophecy. When Falsafa’s theory of religion entered the Latin West, it was accepted and developed into the “double truth theory” in Latin Averroism. However, this theory quickly lost its influence in the Latin tradition, primarily due to the critique by Thomas Aquinas. One of the key aspects of Aquinas’s critique is his criticism of the emanation theory of concepts and the doctrine of the unity of the intellect, which in turn undermines the foundation of Falsafa’s theory of religion, particularly its theory of natural prophecy.

This paper aims to revisit the debate between Aquinas and Falsafa regarding the theory of intellect as the foundation for natural prophecy by defending Falsafa’s perspective. In particular, I examine how Aquinas’s arguments miss the key insights that underpin Falsafa’s doctrine of the intellect.

I begin by presenting two arguments (the blind man and the mirror arguments) in Aquinas’s critique to outline his general argumentative strategy against Falsafa. I focus on how Aquinas reads Falsafa, particularly how his reading motivates his criticism. I then turn to Falsafa’s perspectives, first examining why Avicenna’s emanation theory of concepts is immune to the blind man argument and then considering how Averroes’s insight that there is an essential difference between thought and natural species motivates the unity of intellect. Following Avicenna, Averroes holds that, at the metaphysical level, different thoughts have a built-in shared content but that different individuals have no built-in shared content. Therefore, for Averroes, thought is not a natural species, and there is no problem with how a thought can be individualised as an instance under a natural species. He proposes the unity of intellect precisely to explain this peculiarity of thought.

If we place Avicenna and Averroes’s thought together, we find a picture that is different from Aquinas’ conception of thought. For Aquinas, thought is a property that can be attributed to an individual in the world. But for Falsafa, thought is metaphysically identical to the act of thinking and is not a categorical property of things in the world. Thought exists outside the world, as a mode of existence different from that of things in the external world. Especially for Averroes, the “I think” is not a state in which one owns their act of thinking but is an event in which the universal intellect thinks through me. I conclude with a remark on the wider cultural context of Islamic intellectual traditions, which may have fostered Averroes’s seemingly counterintuitive theory of intellect.

2. Aquinas’s Critique of Falsafa

Aquinas (2002) dealt with the theory of intellectual cognition in his Summa Theologiae, Part I, questions 84–89; he focuses on the intellectual cognition of material things (84–86), of oneself (87), of things higher than the rational soul (88), and of the separate soul (89). In question 84, Aquinas establishes the position that the intellect apprehends material things through the intelligible species. He then questions how the intellect acquires the intelligible species and proposes a criticism of the Platonic–Avicennian approach: the intelligible species come from separate forms or substances. Plato’s theory of forms represents the position of coming from separate forms, while Avicenna’s position represents coming from separate substances.

Text 1:

But it is contrary to the nature (rationem) of sensible things for their forms to subsist without their matters, as Aristotle proves in a number of ways. Accordingly, Avicenna, after ruling out this view, posited not that the intelligible species of all sensible things subsist on their own without matter, but that they preexist immaterially in separated intellects. Such species are derived from the first of these separated intellects into the next one, and so on down to the final separated intellect, which he named the Agent Intellect. From this intellect, as he says, intelligible species emanate into our souls, and sensible forms into corporeal matter. (ST, I, q.84, a.4 (144))
This text clearly shows that Aquinas reads Avicenna’s theory of concept formation as the process through which the intelligible species emanate from the agent intellect. He also provides a concise theoretical reconstruction of Avicenna’s emanationist theory of conception.\textsuperscript{11} According to Aquinas, on the one hand, Avicenna agrees with Plato that the intelligible species must be acquired from forms external to the rational soul. But, on the other hand, Avicenna also accepts Aristotle’s powerful critique of Plato’s theory of forms, not believing that forms have an independent existence. To solve the difficulties surrounding the independent existence of forms, Avicenna proposes that forms essentially exist in the intellect. Thus, he replaces Platonic forms with forms in the intellect, which is a metaphysical revision that leads to his emanationist understanding of concept formation. The reconstruction of the Plato–Avicenna approach further paves the way for Aquinas’s general critique of this approach because Plato’s and Avicenna’s theories cannot explain why the cognitive nature of human intellect must depend on the body. Moreover, Aquinas raises a possible reply from the Avicennian perspective that emphasises the fact that Avicenna distributes a role to the senses in the process of intellectual cognition:

Text 2:

It might be said, however, in line with Avicenna, that the senses are necessary to the soul because they arouse it to turn toward the Agent Intelligence, from which it receives species. But this is not adequate. For if it belongs to the soul’s nature to understand through species emanating from the Agent Intelligence, then it would follow that the soul could sometimes turn toward the Agent Intelligence out of the inclination of its own nature, or even that, aroused by one sense, it turns toward the Agent Intelligence to receive the species of sensibles belonging to a sense that the person does not have. In this way, someone born blind could have knowledge of colours, which is clearly false. Accordingly, it must be said that the intelligible species by which our soul understands do not emanate from separate forms.\textsuperscript{12} (\textit{ST}, I, q.84, a.4 (145–6))

Aquinas notes that the senses play an important role in Avicenna’s doctrine of concept acquisition because Avicenna believes that grasping the intelligible requires the assistance of the senses, as sensory images can trigger the process of emanation.\textsuperscript{13} However, Aquinas points out a crucial difference between emanation and the role of the senses in Avicenna’s view: Emanation is a defining factor of intellectual cognition in human beings because it “belongs to the soul’s nature”, whereas the senses, as an occasional trigger, are not used to define human intellectual cognition.\textsuperscript{14} Therefore, for Avicenna, the definition of the rational soul does not rule out the conceivability of grasping the intelligible form from the agent intellect, even without a sensory occasion to trigger its emanation. In other words, for Avicenna, it seems possible to acquire the intelligible without the help of the senses. This possibility is a theoretical consequence that Aquinas considers counter-intuitive because, if it were possible, “someone born blind could have knowledge of colours, which is clearly false”.

Let us now turn to the mirror argument, which Aquinas presents in his well-known treatise \textit{De unitate intellectus contra Averroistas} as part of his systematic criticism of the unity of intellect:\textsuperscript{15}

Text 3:

He said that the understanding of that separate substance is my understanding or that person’s understanding, in so far as that possible intellect is joined to me or to you through phantasms which are in me and in you. He said that this is accomplished in the following way. Now the intelligible species, which becomes one with the possible intellect since it is its form and act, has two subjects: one, the phantasms themselves; the other, the possible intellect. So therefore the possible intellect is in contact with us through its form by means of the phantasms; and thus, as long as the possible intellect understands, this man understands.\textsuperscript{16} (\textit{DUI}, c. 3, par. 63 (49))
In this passage, Aquinas reconstructs Averroes’s so-called two-subject theory. Generally speaking, for Averroes, concept acquisition happens in the following way: first, the images acquired through the external senses need to be properly prepared by the cogitative power, and then the agent intellect sheds light on the images (as is the case in Avicenna’s theory), which finally triggers the process of abstraction in the potential intellect, although for Averroes the potential intellect is not in us but is separate and shared by all human beings. In this picture, the potential intellect is the bearer of the intelligibles, and the cogitative power is the so-called subject of truth insofar as the correspondence relation occurs between the intelligibles in the intellect and images in the cogitative power.17 According to Aquinas, Averroes posits the two-subject theory to save the phenomenon because Averroes believes that the thought of the potential intellect can somehow be attributed to the cogitative power insofar as the cogitative power triggers the potential intellect and serves as its subject of truth.

In DUI, Aquinas highlights two different aspects of UI as the targets of his criticisms: the two-subject theory and the thesis that the cogitative power triggers the actuality of the potential intellect. To criticise the two-subject theory, Aquinas presents three distinct arguments (the argument from the essential unity between soul and body, the mirror argument, and the colour-wall argument). In what follows, I will only focus on the mirror argument. This is not because it is the most powerful or cogent criticism, but because it is a relatively vivid one that clearly shows Aquinas’s general concern about UI and how his main argumentative strategy centres on the belief that human beings can think and own their own thoughts:18

Text 4:

Unless perhaps it be said that the possible intellect is in contact with phantasms as a mirror is in contact with the man whose appearance is reflected in the mirror. But such a contact clearly does not suffice for the contact of the act. For it is clear that the action of the mirror, which is to represent, cannot on this account be attributed to the man. Whence neither can the action of the possible intellect be attributed, on account of the above-mentioned joining, to this man who is Socrates, in such a way that this man would understand.19 (DUI, c. 3, par. 65 (50))

The mirror is like the potential intellect; the man facing the mirror is like the cogitative power. The proper condition between the mirror and man, which enables the image of the man to occur in the mirror, is like the light from the agent intellect. The image of the man that occurs in the mirror is like the intelligible. Given that, in the case of the mirror, the images in the mirror are not attributed to the man but to the mirror; likewise, in Averroes’s account of intellectual cognition, the intelligible should be attributed to the potential intellect and not to the cogitative power.20 Clearly, in the mirror case, the image in the mirror is possessed by the mirror and not by the man. Likewise, Averroes’s unity of intellect implies that the real possessor of the intelligibles is the potential intellect and not the cogitative power. Therefore, thoughts cannot be attributed to us in the sense that we own our thoughts; for Aquinas, this result is counterintuitive.

3. Abstractionism versus Emanationism

Although Aquinas considers Avicenna’s theory of concept acquisition to be emanationist, once we turn to Avicenna’s own writings, Aquinas’s reading is doubtful. Over the last two decades, scholars have debated whether Avicenna is an emanationist or an abstractionist with regard to concept acquisition. This debate arises from Avicenna’s seemingly different answers about how concepts are acquired. For example, in al-Nafs, we have the following:21

Text 5:

As for the intellectual faculty (al-quwwa al-‘aqliyya), when it reviews (atla ‘at) the particulars that are in the imagination, and the light of the active intellect (we discussed) sheds light upon [the particulars] in us, [the particulars] are transformed
In this text, Avicenna takes concept acquisition to be a process of preparation and reception in which the rational soul prepares images into a proper state, which allows them to be exposed to the illumination of the agent intellect. This process triggers the corresponding concepts to emanate in the intellect. However, in other texts, the role of the agent intellect seems completely absent:

Text 6:

So [the intellect] turns to (yuqibilu) these accidents, then extracts them (yanzi) as though it were (yuqashshiru) those accidents from it and throwing (yatrah) them away to one side until it arrives at the meaning (al- ma’na) in which they participate and by which they do not differ; and then [the intellect] acquires (yuhashilu) it and conceptualises (yatasawwaru) it. From the first moment that [the intellect] inspects (yuqashshiru) the mix in the imagination, it finds (yajidu) accidents and essential components, and of the accidents, [it finds] those that are necessary and those that are not. It separates (yufridu) the intentions from the mixed multiplicity in the imagination and takes its essence from [the mixed multiplicity]. (al-Burhan, 222 lines 8–11)

In this text, Avicenna vividly describes the process of concept acquisition as the extraction of abstract concepts (the intelligible form) from sensible images. Abstraction is depicted as an activity of peeling, as if images have a conceptual core wrapped in various sensible accidents that can be extracted by peeling off the sensible accidents. It is worth noting that Avicenna does not mention the role of the agent intellect. This omission might leave readers with the impression that the human intellect can abstract concepts from images on its own.

Scholars have developed various versions of abstractionism and emanationism to interpret textual complexities. According to the standard emanationist view, one acquires the concept of whiteness through the following mechanism: first, one receives the image of white through the external senses, and after complex processing by the internal senses, an event of emanation from the agent intellect is triggered through which the concept of white appears in one’s intellect (Black 2005, pp. 308–26; Davidson 1992, pp. 92–4; Lizzini 2010, pp. 223–42; Rahman 1958, p. 15; Taylor 2005, p. 180).

Note the explanatory relation between image preparation and emanation. Strictly speaking, preparation is sufficient but not necessary for emanation because Avicenna does not believe that it is always necessary to rely on the preparation of an image through the imagination every time a person thinks of a particular concept. Instead, relying on preparation usually happens only when one acquires the concept. When one’s intellect is powerful enough, the intellect can dispense with the help of imagination and directly trigger an event of emanation on its own (al-Nafs, 50, line 1).

On the contrary, according to the standard abstraction theory, after white images are prepared by the internal sensory system, the intellect then extracts the concept of white through its power of abstraction. The agent intellect still plays a specific explanatory role in this process: its illumination provides an environmental condition that enables the power of abstracting to function (Hasse 2001, pp. 39–72). Clearly, the most important difference between the theories of abstraction and emanation is that each provides a different explanation for the source of concepts. Another notable difference is that each also points to a deeper debate in light of the problem of concept acquisition: whether Avicenna is an empiricist or a rationalist.

In general, abstractionists tend to consider Avicenna an empiricist, while emanationists typically associate him with rationalism. One way to distinguish between the two is to examine whether Avicenna considers images essential for concept acquisition. From an empiricist perspective, concept acquisition and sensory experience have an intimate relationship. If we consider experience as the source of concepts, then the process of concept
acquisition cannot be understood without recourse to sensory experience. However, from the perspective of rationalism, Avicenna is not committed to such a strong relation between experience and concept acquisition. For example, one’s experience of F-things may help one acquire the F-concept, but this does not mean that concept acquisition without experience cannot be understood, as some concepts may originate from the intellect itself and not from experience. Having made these distinctions, I will argue that, regardless of whether Avicenna is an abstractionist/empiricist or an emanationist/rationalist, he has a reply to Aquinas’s blind man argument.

4. Does Avicenna Fall Prey to the Blind Man Argument?

According to Aquinas’ blind man argument, if the emanationist’s account of the intelligible form is correct, then a man born blind can have knowledge of colour. However, it is clear that a man born blind cannot have knowledge of colour. Therefore, the emanationist’s account is false. From the perspective of the abstractionist’s interpretation, Avicenna may have a quick reply to the blind man argument: the argument is both right and wrong. If we unpack the apparent intuition that it is inconceivable for a blind man to know what colour is, the meaning behind this intuition seems unclear because the concept—knowing what colour is—is ambiguous. One might offer at least three different interpretations of the intuition, each based on a different understanding of what it means to know what colour is.

I1: The essence of colour is nothing but the phenomenal colour that one directly experiences in one’s colour-experiences, so the real knowing of the what-ness of a colour is not a dry conceptual belief that one can hold based on one’s colour-concepts. Instead, the direct colour-experience brings one to the thing-itself—to literally see the colour itself.

I2: The essence of colour is what the intellect grasps by the concept of colour, so it is intellectual and not the qualia that one experiences. Therefore, having the conception of the essence of colour is sufficient for the knowledge of what colour is.

I3: The intellect grasps the essence of colour through the concept of colour but to have a real understanding of what a colour is, one needs to properly form an explanatory connection between one’s concepts and the relevant experiences.

I1 is the belief that the true cognition of colour only occurs through colour experience. Conversely, I2 and I3 involve the belief that grasping the essence of colour belongs to the intellect. The difference between I2 and I3 is that the former involves the belief that a conceptual grasp of colour through the intellect amounts to knowledge of what colour is, while the latter involves real conceptual knowledge of the essence of colours, which requires one to know how to apply concepts to explain their perception of colours.

If we unpack the apparent intuition that someone born blind cannot have the knowledge of colour from the perspective of I1 or I3, emanationists would not deny that a blind person cannot know what colour is. To clarify this point, we need to understand Avicenna’s theory of sensation and perception. Avicenna refers to the sensible forms received by the external senses as the near sensible and the objects represented by these sensible forms as the far sensible (al-Nafs, 66, lines 6–14). The soul, through the external senses, can only feel the near sensible forms and cannot know whether they represent external objects.
To perceive external objects, cooperation among different faculties is required: first, various forms perceived by the external senses must be unified in the common sense, and then the estimative power or intellect is required to determine these forms as possessing external existence. In short, we can distinguish between sensation and perception in Avicenna’s theory of the external and internal senses, with the former being acquaintance without intentionality and the latter being cognition with intentionality.24

In this framework, a blind person lacks both the sensation and perception of colour: even if one can independently acquire the concept of a colour without one’s sensory system, one cannot identify the perceptual objects that can be explained by the colour concept. Therefore, neither I1 nor I3 can establish knowledge of colour. Thus, Avicenna does not deny that a blind person does not know colour given either interpretation of the intuition.

The controversy lies in I2. Emanationists accept that I2 is a real possibility for a blind person. The question then is whether the blind man can have knowledge of colour in the sense that he can have colour concepts. It is not clear whether this belief is self-evident. It is also possible that people do not share a self-evident intuition concerning the case of the blind man, just as people may not have a self-evident intuition about the necessity of images or imagination for intellectual cognition. In such cases, empiricists might judge that the blind man does not have colour concepts, as this aligns with their position regarding concept formation. However, rationalists might judge the opposite. It is important to note that the belief that people born blind cannot understand colour is not a self-evident thesis in contemporary cognitive science. Instead, it is typically considered an opinion of empiricist philosophers, and several recent studies have challenged this belief.25 Considering that this seemingly intuitive belief has been critically examined through empirical science, it is difficult to admit that the claim that the blind man cannot have colour concepts is self-evident.

According to Avicenna’s emanationist theory, it is possible for a blind person to acquire the concept of colour in the sense of I2 because of his rationalistic stance on the relation between concept acquisition and sensibles: the two are not intimately related. Avicenna can provide independent support for his rationalistic position on the basis of at least three other commitments. First, Avicenna believes that the definition of a conception does not necessarily include reference to the sensibles, which means that a conception is at least open in its essence to whether it depends on the sensibles for acquisition. Second, he believes that it is empirically true that the intellect can operate independently of perceptual systems, which means that an intimate relation between the sensibles and conception does not exist (al-Nafs, 221 line 15–223 line 10). Third, the openness of a conception to its relation to the sensibles further guarantees the possibility of the divine intellection of the sensibles. The divine intellect is precisely the kind of intellect whose conception does not depend on sensation or perception.26

5. Revisiting the Unity of Intellect

Averroes presents his most detailed discussion of the unity of intellect in his Long Commentary on the De Anima of Aristotle, Book 3, Chapter 5, where he considers three different explanations for the ontological status of the material intellect. These include Alexander of Aphrodias’ view that the material intellect is the disposition of the body, Ibn Bajja’s view that the material intellect is the disposition in the imaginative power, and Themistius’s view that the material intellect is a separate substance. Averroes rejects the first two positions and attempts to work out his own view based on a reworking of Themistius’s position. He first raises several problems with Themistius’s position, and then reformulates the problem as a concise dilemma. According to Averroes, the key to building a proper view on the material intellect is to resolve a dilemma about the intelligible. If the intelligible is the same in you and me, then it seems that everyone would be thinking at the same time. However, if the intelligible in you and the intelligible in me are individualised, then it seems that intellectual cognition and teaching would be impossible. Towards the end of Chapter 5, Averroes provides his ultimate answer to this dilemma.27
Text 7:

(7A) For if the thing understood in me and in you were one in every way, it would happen that when I would know some intelligible, you would also know it, and many other impossible things [would also follow]. If we assert it to be many, then it would happen that the thing understood in me and in you would be one in species and two in individual [number]. In this way the thing understood will have a thing understood and so it proceeds into infinity.

(7B) Thus, it will be impossible for a student to learn from a teacher unless the knowledge which is in the teacher is a power generating and creating knowledge which is in the student, in the way in which one fire generates another fire similar to it in species, which is impossible. That what is known is the same in the teacher and the student in this way caused Plato to believe that learning is recollection.

(7C) Since, then, we asserted that the intelligible thing which is in me and in you is many in subject insofar as it is true, namely, the forms of imagination, and one in the subject in virtue of which it is an existing intellect (namely, the material [intellect]), those questions are completely resolved.

First, Averroes restates the dilemma in Text 7A, without discussing the first horn in detail because he explained it earlier. He quickly moves to the second horn and presents two short arguments to explain why the individualisation of the intelligible is problematic. The first argument suggests that the individualised intelligible leads to an infinite regress of the intelligible. The traditional interpretation proposes that the infinite regress arises from the fact that the individualised intelligible, strictly speaking, is not truly the intelligible (because the true intelligible should be in some way common to you and me). Therefore, it triggers a further effort to acquire the real intelligible, which, in turn, leads to an individualised intelligible and so on to an infinite regress (Averroes 2011, pp. 328–9, note 114; Taylor 2004, pp. 125–6). Stephen Ogden has recently argued that the infinite regress, as understood by the traditional interpretation, is merely a potential regress (Ogden 2021, pp. 441–5).

However, in my view, the traditional interpretation does not involve a potential regress. The target of the first horn is that the intelligible form in you and in me is one in every way. However, in the second horn, the intelligible is still in you and in me but not one in every way. Rather, it is in the way that our intellectual cognition aims to grasp something common to you and me, but to explain why people have different thoughts, the common object is then taken as an individualised form in a given subject. If this is the case, then the problem is as follows: how could an individualised object become something common? In order to grasp the common object, one’s intellect still needs to grasp a common intelligible that begins with the individual intelligible. However, the intellect’s effort will still end up grasping an individualised intelligible, which further requires a move towards commonality, which leads to an infinite regress.

In other words, in the second horn, the infinite regress is caused by an irreducible gap between the end of the act of conception (grasping the common intelligible) and the actual but unacceptable result of that act (the individualised intelligible). The regress is not potential because our intellect must constantly and actually strive to move from the individual intelligible to the common intelligible, given that the end of the act of conception is to grasp the common intelligible. If grasping the intelligible amounts to an infinite process of grasping the individualised intelligible, then it means that the end of the act of conception has never been realised; that is, there has never been any conception. However, since we do have real conceptions, there is no actual infinite process of striving to achieve conception.

After Averroes’s infinite regress argument, he criticises the individualisation of the intelligible form from another perspective, which centres on the possibility of teaching. If the individualisation of the intelligible form were possible, it would lead to an unacceptable consequence: the activity of teaching would be impossible. From the perspective of the impossibility of teaching, the discussion of individualisation of the intelligible form
contains a new argument, one different from the argument of infinite regress. Let us call it the impossibility argument of knowledge transmission.

Averroes introduced a concise analogy to illustrate the impossibility argument of knowledge transmission. Consider how fire causes something to burn. When an object comes near a fire, it can become hot and eventually be consumed by the flames. In this case, the fire can cause the object to receive its own form. Now consider the process of knowledge transmission, for example, when a teacher, who knows the definition of a circle, imparts it to a student. In this case, the teacher first possesses some intelligible form, and through their teaching, they cause in some way a new concept of a circle in the student’s mind.

The question that Averroes has us consider is whether the transmission of knowledge and the transmission of the form of fire follow the same pattern. His answer is negative. In the case of fire, the burning heat in the fire and the burning heat produced in the object do not share a common content. But in the process of knowledge transmission, the intelligible form that the teacher possesses and the intelligible form that the student acquires essentially share a common content. When a student acquires the concept of a circle under the teacher’s instruction, what the student grasps is precisely what the teacher grasps. Therefore, we cannot confuse an intelligible form with things like fire. Were one to confuse the intelligible form with, say, fire, the intelligible form would lose the ability of sharing the same content with other intelligible forms. However, this ability is precisely the prerequisite for imparting knowledge. Based on this insight, we can reconstruct Averroes’ argument as follows:

The impossibility argument of knowledge transmission

P1: If the intelligible form can be individualized, then the intelligible form must become a member of a species of natural things.

P2: If the intelligible form is a member of a species of natural things, then different intelligible forms cannot share the same content.

P3: If different intelligible forms cannot share the same content, then teaching is impossible.

P4: Teaching is possible.

C1: Different intelligible forms can share the same content.

C2: The intelligible form is not a member of a species of natural things.

C3: The intelligible form cannot be individualized.

Averroes explains P1 at the beginning of Text 7: If we believe that an intelligible form can be attributed to an individual, then an individual can possess an individual intelligible form just as they possess an individual colour or quantity. Furthermore, we can take the intelligible form itself as a species under which there are many individualised intelligible forms that constitute the members of the species. P3 is evident given Averroes’s illustration of the activity of teaching. P4 can be understood as a premise established on the basis of experience. It seems that the most problematic premise is P2, as it is unclear why being members of a species of natural things amounts to the rejection of sharing the same content. For example, there are different pictures of Averroes. Some might be made of paper, while others might be digital. However, all these pictures are made of some kind of physical natural thing. Clearly, all these pictures share a content, which is that they all represent Averroes. But it is hard to deny that these pictures are natural things. Therefore, P3 is not only unclear but also seems easily falsified.

Averroes is not concerned with any possible form of representation but rather with the feature of the intelligible form conceived of as sharing the same content. In other words, the key to understanding P2 lies in how Averroes understands how different intelligible forms might share the same content. Although Averroes disagrees with Avicenna on many issues, he does follow Avicenna in understanding the intelligible form as the conception of quiddity.31 Therefore, it may be helpful to return to Avicenna’s metaphysical analysis of
the intelligible form. Avicenna is well known for initiating a trend in Arabic philosophy that analyses the intelligible form as the occurrence of quiddity in the mind.

Text 8:

The one form in the intellect is related to the many, and it is on this consideration (al-i’tibār) a universal, it is one meaning in the intellect, whose relation to any given animal does not differ; that is, the form of any of them is present to the imagination immediately; the intellect then extracts its meaning (ma’nahū) abstracted from accidents, [then] the form itself occurs (ḥaṣala) in the intellect. This form is the one which occurs by abstracting animality from any individual image, taken either from an external existent or from something that plays the role of an external existent even if it itself does not exist externally but [is something] the imagination invents. (al-Ilahiyyat, p. 156, lines 10–18)

In my view, Aquinas rightly points out that Avicenna’s understanding of the ontological status of intelligible forms is a correction of the theory of Platonic forms. From Avicenna’s perspective, to explain why different people in different places at different times can provide the same answer for what a thing is or, to put it more technically, conceive the same quiddity, Plato believes that the quiddity of things is a kind of entity that can be separated from the human intellect and subsist on its own. But this position leads to many absurdities. To solve these problems, Avicenna proposes that quiddity itself is not an independent entity but either exists in things or in the mind. However, as far as its existence in things is concerned, quiddity is individualised, and only when it exists in the mind can it maintain its sameness in different conceptions.32 We might find a clue in Avicenna’s doctrine of the triple distinction of quiddity for understanding why sharing the same content is a feature that cannot be instantiated by individualised things in the world.

Strictly speaking, according to Avicenna and Averroes, the idea that we share the same content in intelligible forms means that the same essence exists in different intelligible forms as their common content. The intelligible horse in my mind makes horseness available to me. However, this does not mean that the form makes me stand in a cognitive relation to an abstract object outside my mind that exists independently. Rather, horseness is available to me precisely because it exists in my mind: Horseness literally exists in my mind and becomes an aspect of the intelligible form horse. This aspect is not individualised in my conception of horse because it is also common to any other possible conception of horse, whether in my mind or in others’ minds. In this sense, although horseness as the common content cannot exist independently as a Platonic form, it is still independent in the sense that it has what might be called inter-intellectual subjectivity. By this phrase, I mean that the same horseness as the common content, though inseparable from any intellect, always exists and is open to different intellects as an aspect shared by any possible intellect.33

In light of the inter-intellectual subjectivity of the quiddity as content, we can draw a sharp distinction between quiddity in the mind and quiddity in things. When horseness is in the external world, it is individualised as different horses. The individualisation of horseness in the external world not only means the rejection of the so-called universal in things but, more importantly, also that different individualised horses are not connected through horseness as a common aspect cognitively open to different horses. Were one horse in the world to instantiate the inter-intellectual subjectivity, it would not be a horse in the world but a concept. In this sense, the inter-intellectual subjectivity sets a clear boundary between the mind and the external world, which might be Averroes’s reason for supporting P2.

One might have a more global concern about the argument as a whole: even if one accepts the soundness of the fire argument, the argument establishes only that the intelligible cannot be individualised. There is still a gap from the non-individualisation of the intelligible form to positing a single intellect for all human beings. What is Averroes’s demonstration for his move from the rejection of the individualised thesis to his doctrine of the unity of the intellect?
In my view, there is indeed no demonstration, but to clarify this point, we must return to Averroes’s general argumentative strategy for the unity of the intellect. One common misunderstanding of Averroes’s strategy is that he accepts that we can rightly attribute thoughts to an individual. Therefore, although he produces a few arguments that step towards the unity of the intellect, he still posits the two-subject theory to save the phenomenon that individuals think. Unfortunately, he is not aware that his attempts are doomed to failure given the large gap between his theory of intellect and our intuition that thoughts which occur in me are naturally mine. However, the infinite regress argument and the fire argument are clearly targeted at the individualisation of the intelligible form, which serves as a precondition of attributing thought to an individual. If Averroes, on the one hand, tries to produce an argument against the attribution of thoughts to the individual, how could he, on the other hand, save the phenomenon? A more plausible proposal might be that Averroes is not interested in saving the phenomenon. On the contrary, his aim is to show that the seemingly natural intuition does not hold up to reflection and should be rectified.

The rectification proceeds in three steps. First, he needs to show why the intelligible form cannot be individualised. Second, if the intelligible form cannot be individualised, he needs a theory of intellect that can offer the best explanation for the non-individualisation of the intelligible forms. And, finally, he needs an error theory to explain what the seemingly natural intuition really means. In the three-steps reading, the unity of the intellect is posited for the second step, not through demonstration but through inference from the best explanation. The two-subject theory is for the third step, which does not aim to save the intuition but clarifies the meaning of “an individual thinks” in light of his new theory of intellect that accommodates the non-individualisation of the intelligible forms. For Averroes, “an individual thinks” precisely means that the universal intellect thinks through an individual. Therefore, it is unfair of Aquinas to criticise Averroes by appealing to the intuition that Averroes tries to rectify. By presupposing the intuition as evidently true, Aquinas misses the field of debate.

6. Concluding Remarks

Through a detailed analysis of Avicenna and Averroes on the intellect in the context of Aquinas’s critique, I have shown that Falsafa’s positions cannot be easily refuted by Aquinas’s arguments. However, my aim has not been to defend Falsafa’s theory of intellect from Aquinas’s critical insight. Falsafa’s theories are still problematic in several respects. For example, in my reconstruction of Averroes’s argumentative strategy for the unity of intellect, I suggest that the inter-intellectual subjectivity of the intelligible form might serve as Averroes’s reason for adopting P2. However, the move from inter-intellectual subjectivity to P2 is not cogent. One way to bridge the gap is to add a hidden or implicit premise that all natural species are in the external world. Given that all natural species are in the external world, the intelligible forms must be in the external world in some sense if they are taken to be natural species. If that is the case, the forms then lose their inter-intellectual subjectivity, and we arrive at P2. However, the belief that all natural species are in the external world is itself problematic and is a view that Avicenna does not seem to adopt.

My defence of Falsafa, therefore, is limited to showing the deep theoretical divergences between Falsafa and Aquinas, which are likely concealed by their common Aristotelian terminology and their specific arguments and counter-cases. For instance, scholars might assume that philosophers working in the Aristotelian tradition are empiricists about concept formation. However, the deep divergence between Avicenna and Aquinas is precisely about the fundamentally different intuitions of rationalists and empiricists. In the case of Averroes, the divergence seems more counter-intuitive at first glance, given that Aquinas appeals to the idea that we are the primary owners of our thoughts, an intuition that seems to be common sense. In addition to the theoretical analysis of why Averroes thinks this seemingly common sense is problematic, I also hope to add a final remark on a possible
religious condition that might make Averroes’s embrace of the attribution of thoughts to the universal intellect seem more understandable.

It is worth starting from the well-known *Hadith al-Nawâfil*:

**Text 9:**

Allâh said “I will declare war against him who shows hostility to a friend of Mine. And the most beloved things with which My servant comes nearer to Me, is what I have enjoined upon him; and My servant keeps on coming closer to Me through performing *nawâfil* (praying or doing extra deeds besides what is obligatory) till I love him, then I become his hearing with which he hears, and his sight with which he sees, and his hand with which he grips, and his leg with which he walks; and if he asks Me, Twill give him, and if he asks My Protection (Refuge), I will protect him (i.e., give him My Refuge); and I do not hesitate to do anything as I hesitate to take the soul of the believer, for he hates death, and I hate to disappoint him.” (al-Bukhari 1997, pp. 275–76, revised based on Khan’s translation)

*Hadith al-Nawâfil* is commonly believed to be one of the essential *hadith* that shapes the Islamic conception of the nearness of Allâh. Especially within the Sufi tradition, the perplexing claim that “I become his hearing with which he hears” is read as an indication of the divine presence, a fundamental doctrine of Sufism, which holds that believers can somehow witness the divine by breaking their own self-centred consciousness. For example, al-Junayd, one of the central figures in the formative period of Sufism, has proposed a clear explanation of how a believer can see or hear through Allâh by being in the state of *fanî* (the breaking or the annihilation of the self):

**Text 10:**

They are distinguished by their knowledge of truth before Allâh when Allâh creates in them the faculty of true knowledge of Himself. This faculty emanates from Allâh and must be attributed to Him and *not to the person in whom it is endowed*. The possession of this faculty marks the fullness of endeavours before God. (Kader 2014, p. 162)

In al-Junayd’s interpretation of *fanî*, one essential moment is when the individual is aware of the fact that their capacity to know Allah is “from Allah and must be attributed to Him and not to the person in whom it is endowed”. To put it simply, for al-Junayd, breaking the self is not a mysteriously conscious state in which one can literally get rid of their first-person perspective; rather, it is a state that might be termed as *the reverse of subject*, in which one is aware of the fact that the operation of their capacity to know is ultimately attributed to the divine principle of the world. Therefore, *the reverse* means that one can break out their everyday belief that their thoughts are primarily their own and arrive at a new state of consciousness in which they know that they are just a place through which the divine thinks. Al-Junayd’s conception of *fanî* marks out a theme that is repeated again and again in the later development of Sufi traditions and culminates theoretically in Ibn al-’Arabi’s idea of the disclosure of the absolute.

We can see a similarity between Averroes’s thoughtful struggle in his theory of intellect and the Sufi view of *fanî*: both try to reach the reverse of the subject. By pointing out this similarity, I do not intend to claim that Averroes is a Sufi or that he is trying to develop an insight from the Sufi tradition with his Aristotelian philosophical resources. Instead, I intend to draw attention to the notable fact that, given that the Sufi movement had a widespread influence throughout the Islamic world during Averroes’s time—and that Muslim Spain would later witness Ibn Arabi’s age—Averroes was living in a religious world in which our commonplace beliefs about thought attribution would be problematic. This religious condition might be strong evidence for Averroes’s doubt about what we take to be intuitive. In other words, one lesson that we can learn from Averroes is that, if we examine the problem of the individualisation of thought from a historically and religiously different perspective, we can understand that the belief that one’s thoughts are primarily
attributed to oneself is not always accepted as naturally true and cannot serve as the basis for determining whether a theory of cognition is correct. Therefore, although we may not agree with Averroes’s solution and specific argument, we can follow his overall strategy for examining whether seemingly intuitive beliefs have theoretical problems.

Funding: This research was funded by [The National Social Science Fund of China] grant number [23CZJ021].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1. Although *Falsafa* usually refers to philosophers who engage in philosophy in the Islamic intellectual tradition that treats Greek philosophy as paradigmatic (e.g., al-Fārābī, Avicenna, and Averroes), I use the term in this article to refer to Avicenna and Averroes. The naturalistic account of religion refers to the approach taken by the *Falsafa* school, which seeks to explain the nature of religion using the best available theoretical framework of its era while staying within the bounds of rationality.

2. For a general sketch of Al-Fārābī’s Theory of Prophecy, see (Walzer 1957).

3. For an overview of Al-Fārābī and Avicenna’s theory of intellect, see (Taylor 2006; Gutas 2012b).

4. For an overview of Averroes’s theory of intellect, see (Davidson 1992, chp. 6–8).

5. For a general account of Latin Averroism, see (Marenbon 2007).

6. For an overview of Aquinas’ critique of Averroes’ theory of intellect, see (Taylor 1999).

7. Texts 1 and 2 are from (Aquinas 2002). For the Latin text, see (Aquinas 1889), henceforth, *ST*. The page number of the English text is in brackets.

8. There is an ongoing debate concerning whether, for Aquinas, intellectual cognitions of material things must be mediated through the intellectual awareness of the relevant intelligible species. An affirmative answer may lead one to the so-called representationist’s reading, whereas a negative answer may lead one to a direct realist reading. In this article, I leave the question open. For a detailed examination of the debate, see (Baltuta 2013).

9. *ST*, I, q.84, a.4 (144): “Respondeo dicendum quod quidam posuerunt species intelligibiles nostri intellectus procedere ab aliquibus formis vel substantiis separatis. Et hoc dupliciter. Plato enim, sicut dictum est, posuit formas rerum sensibilium per se sine materia subsistentes; sicut formam hominis, quam nominabat per se hominem, et formam vel ideam equi, quam nominabat per se equum, et sic de aliiis.”

10. *ST*, I, q.84, a.4 (144): “Sed quia contra rationem rerum sensibilium est quod eorum formae subsistent ab separatis substantiis, ut Aristoteles multipliciter probat; idea Avicenna, hac positione remota, posuit omnium rerum sensibilium intelligibiles species, non quidem per se subsistere ab aliqua materia, sed praecox etiam immaterialiter in intellectibus separatis; a quorum primo derivatur huismodi species in sequentem, et sic de aliiis usque ad ultimum intellectum separatum, quem nominat intellectum agentem; a quo, ut ipse dicit, effluent species intelligibiles in animas nostras, et formae sensibiles in materiam corporalem.”

11. Note that Avicenna himself did distinguish between intelligible species and intelligible forms. Therefore, in Text 1, Aquinas’s reconstruction of Avicenna’s doctrine may not be historically accurate. In the next section, I will discuss how Aquinas’s emanationist reading of Avicenna’s theory of concept formation is currently the topic of a heated scholarly debate.

12. *ST*, I, q.84, a.4 (145–6): “Si autem dicatur, secundum Avicennam, quod sensus sunt animae necessarii, quia per eos excitatur ut convertat se ad intelligentiam agentem, a qua recipit species; hoc enim suum habet; quia si in natura animae est ut intelligat per species ab intelligentia agente effluxas, sequetur quod quandoque anima positio se convertere ad intelligentiam agentem ex inclinatione suae naturae, vel etiam excitata per alium sensum, ut convertat se ad intelligentiam agentem ad recipiendum species sensibilium quorum sensum aliquis non habet. Et sic caecus natus possit habere scientiam de coloribus, quod est manifeste falsum. Unde dicendum est quod species intelligibiles quibus anima nostra intelligit, non effluent a formis separatis.” Note that, in this text, when Aquinas points out that, for Avicenna, the human intellect receives the species of the sensibles (species sensibilium) from the agent intellect, the species of the sensibles refer to the intelligible form of the sensible. This is because, for Avicenna, only the intelligible forms emanate from the agent intellect.

13. See Text 3 in Section 2.

14. Aquinas insightfully notes that the emphasis of the role of the senses is not essential for Avicenna’s theory of human intellect. One significant clue that justifies Aquinas’ observation is Avicenna’s well-known analogy of a horse and rider. According to Avicenna, in the process of concept acquisition, the relation between the rational soul and the senses is like that of a rider and a
horse crossing a river. Once the rider successfully crosses the river, they can abandon the horse and continue their journey on their own; likewise, when the intellect first acquires the intelligible form, the intellect itself becomes more powerful, such that it can trigger the re-awareness of the same form solely through its own will without relying on the senses. For the horse–rider analogy, see (Avicenna 1985, p. 374, ll. 8–14).

Texts 3 and 4 are from (Aquinas 1968). For the Latin text, see (Aquinas 1976, pp. 289–314), henceforth, DUI, followed by the chapter and passage number in the Latin text and the page number of the English text in brackets.

DUI, c. 3, par. 63 (49): “Dixit quod intelligere illius substantiae separatae est intelligere mei vel illius, in quantum intellectus ille possibilis copulatur mihi vel tibi per phantasmatas quae sunt in me et in te. Quod sic fieri dicebat. Species enim intelligibilis, quae fit unum cum intellectu possibili, cum sit forma et actus eius, habet duo subjecta: unum ipsa phantasmatas, alius intellectum possibilum. Sic ergo intellectus possibilis continuaturo nobiscum per formam suam mediantibus phantasmatibus; et sic, dum intellectus possibilis intelligit, hic homo intelligit.”

For a detailed reconstruction of the two-subject theory, see (Davidson 1992, pp. 289–90).

For a detailed reconstruction of Aquinas’s different arguments against UI, see (Wu 2017). It is worth noting that, in addition to the theoretical criticisms, Aquinas also argues that UI is not a faithful reading of Aristotle himself. For the purpose of this paper, I will solely focus on the theoretical dimension of Aquinas’s critique. Whether Averroès’s theory of intellect is a faithful interpretation of Aristotle’s relevant doctrines, or whether Averroès indeed aims to present Aristotle faithfully, remains an open question.

DUI, c. 3, par. 65 (50): “Nisi forte dicatur quod intellectus possibilis continuaturo phantasmatibus, sicut speculum continuat tur homini cuius species resultat in speculo. Talis autem continuatio manifestum est quod non sufficit ad continuationem actus; manifestum est enim quod actio speculi, quaee est repraesentare, non propter hoc potest attribui homini: unde nec actio intellectus possibilis propter praedictam copulationem posset attribui huic homini qui est Socrates, ut hic homo intelligeret.”

One might point out that, in the mirror case, we do somehow attribute the images in the mirror to the man because, if one were asked to whom the image belongs, one could say that the image belongs to the man. Nonetheless, we should make a distinction between what the image is about and what possesses the image; however, in both cases, the term “belongs to” might be used in our natural language.

Text 5 is from (Avicenna 1959), henceforth al-Nafs; Text 6 is from (Avicenna 1956), henceforth al-Burūhān; and Text 8 is cited from (Avicenna 2005), henceforth al-Ilāhiyyāt. Texts 5 and 6 are my own translation; Text 8 is revised based on Marmura’s translation.

For a recent overview of the debate, see (Ogden 2020, pp. 2–7).

For the empiricist’s reading, see (Gutas 2012a, pp. 391–436). For the rationalist’s reading, see (Zarepour 2020, pp. 819–33; Kaukua 2020, pp. 215–40).

For Avicenna’s account of perception, see (Avicenna 2013, p. 175, ll. 3–12, p. 176, ll. 1–6). For a detailed study of Avicenna’s theory of sensation, see (Black 2014, pp. 185–214); for perception and intentionality, see (Kaukua 2014, pp. 215–42).

That people born blind cannot understand colour is not a self-evident thesis in contemporary cognitive science. See (Kim et al. 2021; Striem-Amit et al. 2018).

For Avicenna’s account of the divine intellect, see (Avicenna 2005, p. 285, ll. 8–13).

Averroès, Long Commentary on De Anima = Averrois Cordubensis Commentarium Magnum in Aristotelis De Anima Libros (Averroes 1953), henceforth LCDA. For a detailed reconstruction of Averroès’s discussion of different commentator’s views on the material intellect, see (Davidson 1992, pp. 282–95).

The English translation is from (Averroes 2011, pp. 328–29).

A more detailed reconstruction of the infinite regress argument might run in the following way: Suppose the intelligible form can be individualised. If the intelligible form can be individualised, then the intelligible form must become a member of a natural species. If the intelligible form is a member of a natural species, then the intelligible form will have another intelligible form, and so on infinitely, which will in turn make the intellectual cognition impossible. However, one does have actual intellectual cognitions, so the intelligible form cannot proceed into infinity. Therefore, the intellectual cognition is possible, and the intelligible form cannot be individualised.

According to Ogden’s new account of the infinite regress, if one has an individualised F in their mind, then there must be a universal and intelligible F* in their mind through which F occurs as an individualised form. Since F* itself is intelligible, one needs a further F**; thus, it proceeds to infinity. However, it is unclear why Averroès would accept the quasi-Platonic principle that an individualised concept subsists in the mind through the participation of its universal and intelligible species.

For Averroès’s account of the apprehension of essence, see (Weimer 2019).

For a detailed reconstruction of Avicenna’s quiddity–existence distinction, see (Bertolucci 2012, pp. 257–88); for Averroès’s critical reworking of the quiddity–existence distinction, see (Menn 2011, pp. 51–96); and for a study of how Avicenna’s quiddity–existence distinction is received in the post-Avicennian traditions, see (Benevich 2017, pp. 203–58).

It’s important to note that the concept of intellectual subjectivity in the realm of the intelligible does not align with the realist perspective on universals. This stance is distinct in that it does not subscribe to universals as inherent in individual natural entities, nor does it adhere to Platonic forms.

For a general introduction to Hadith al-Nawūfī and its reception in the spiritual traditions of Islam, see (Ebstein 2018, pp. 271–89).
For Ibn al-ʿArabī's account of al-Nawāʾil, see (Chittick 1989, pp. 325–31).

References


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.