CRITIQUE

An open access journal devoted to the critical study of
philosophy by undergraduates worldwide, published by
Durham University

Contents

Las Meninas, Diego Velázquez, 1656

LAS MENINAS

Las Meninas (translates as 'the ladies-in-waiting') is a 1656 oil painting by the Spanish master Diego Velázquez. Velázquez was born in Seville in 1599 and served as the palace chamberlain in the court of Philip IV of Spain from 1651 until his death in 1660. As part of his role, Velázquez painted royal portraits, and on the surface, this is what *Las Meninas* seems to be — a portrait of the King's Daughter, Margaret Theresa, positioned at the centre of the composition. However, a closer inspection reveals that Las Meninas is an extraordinary and mysterious painting. The composition is enigmatic and reflects upon the nature of the painter, the viewer, and the subject.

Velázquez included himself in the scene, standing just behind the young girls of the court, palette and brush in hand, looking directly at the viewer. Also present behind the young girls are two shadowy figures, believed to be Margaret Theresa's caretaker and a palace bodyguard. In the background, a figure stands waiting on the threshold of the room, also looking at the viewer. This is José Nieto Velázquez, the artist's brother. Left of this figure is a mirror that shows the reflections of King Philip IV and his wife, Queen Mariana, one of the more confounding details of the painting.

Some contend that the painting is thus from the viewpoint of the King and Queen. We see through their eyes as Velázquez paints them while members of their court observe the scene. The reflection in the mirror could therefore be a direct reflection of their faces or a reflection of their likenesses on the canvas Velázquez is painting. The work is stylistically and compositionally groundbreaking for its time.

Las Meninas is fundamentally a work concerning identity. The painting forces the viewer to ask questions such as "who am I meant to be?" and "who are you?". Viewers are inevitably drawn to speculate about the identities of the mysterious figures in the painting, the relationships between them, and the connection between them, the viewer, the art, and the artist.

Editor's Introduction

WHAT IS THE PURPOSE of an editor's introduction? Reading previous introductions, it seems there is a divide between those who see it as a procedural piece that simply gives some context to the journal and those who believe it should provide additional philosophical insight, the editors views on certain matters, or some neat conclusion that links all the articles.

I am much more inclined towards the former view. I believe that the purpose of the editor's introduction is to inform readers on the context of the journal, describe the process of putting it together, and to give thanks to those who deserve it. I do not think that a *Critique* reader is curious as to my views on Kant, psycho-physical harmony, or protest art. I believe that the papers and their authors speak for themselves.

Critique has had something of a tumultuous past, there was a widely publicised transphobia scandal involving an associate editor in 2018. More recently, many editions have been removed from the university website (hopefully temporarily) due to a complaint concerning past editorial practices.

I hope that the journal is now on steadier ground. In fact, I would argue that it is. Michaela Makusha's Identity edition of *Critique* is a fantastic example of how the journal is changing for the better. In another positive step, for this edition we accepted our first submission from outside the Anglosphere; a fantastic piece on Plato's *Phaedo* from a student at Fudan University in China. Furthermore, I am confident that the team of peer reviewers and editors put together for this edition will provide a strong base from which *Critique* can move forward.

I am confident that *Critique* has a positive, bright future and can continue to foster philosophical discussion in Durham and across the world. While much of this responsibility lies with myself and the other editors, the success or failure of an undergraduate philosophy journal ultimately depends on the involvement of undergraduates themselves. Thus, I strongly encourage anyone interested in philosophy, publishing, writing, etc., to get involved with future editions of *Critique* – either by joining the editorial board or by submitting a paper.

Finally, I want to thank all the peer reviewers who worked on this edition, Philip Goff for his insights, and most importantly, Jess Walsh for her invaluable contributions.

WOODY JEFFAY

'Immanent Form' and the Immortality of the Soul: Rethinking the Final Argument in Plato's *Phaedo*

ZIHAN GUO
Fudan University

INTRODUCTION

IN PLATO'S *PHAEDO* SOCRATES PRESENTS FOUR ARGUMENTS REGARDING THE IMMORTALITY OF THE SOUL: the 'cyclical argument,' the 'recollecting theory,' the 'kinship argument,' and the 'final argument.'[1] In the first three arguments, Socrates defines a form[2] as 'x-ness itself (τὶ αὐτό),' emphasising its unchanging nature in contrast with the ever-changing particulars. Forms reside in the invisible world and are accessible only through reason, while particulars, being the opposite, exist in the sensible world and are visible to us. Forms are ever-existing and immortal, while particulars are easy to be destroyed. Additionally, forms, as divine beings, rule over particulars (78d-80b).

However, in the 'final argument', Socrates introduces a different concept of 'x-ness,' one that exists within particulars, such as 'the tallness in Simmias' and 'the hotness in fire.' Previous scholars suggest that this 'x-ness in something (τὶ ἐν τινί),' as the name suggests, shares kinship with the 'pure' form, i.e. the 'x-ness itself' mentioned before. Therefore, they simply refer to it as the 'immanent form.'[3] However, some more recent scholars have reflected on it and argued that such a name is a mere arbitrary assertion. The 'x-ness in something' would simply be an alternative expression of characteristics of a particular ('one being x'), rather than suggesting that such characteristics have the ontological status of form.[4] In this essay, I take the opposite standpoint to such recent points of view. By reexamining the essential kinship

---

[1] The English translation of the text of *Phaedo* is from Cooper, J. M. and Hutchinson, D. S. ed. 1997, with slight modifications. The Greek texts are from Duke, E. A. et al. ed. 1995.

[2] The concept of 'form' in Plato's text is expressed through various Greek terms, such as εἶδος, ἰδέα, μορφή, etc.. Some scholars have suggested some slight differences of these terms when used to express the concept of 'form' (cf. Gallop ed. 2009, 93-94). However, I reckon that, when such terms are used to express certain 'x-ness itself', which is different from particulars, there are no apparent semantic differences. Therefore, in this essay, all these Greek terms are translated simply into 'form.'

[3] For the uses of this term (or similar expressions) among the scholars, cf. Frede 1978, 28; Bostock 1986, 179-180; and most recently Ebray 2023, 255-256; etc.. Vlastos, although he called 'x-ness in something' as 'immanent character,' insists that 'x-nesses in something' are entities that are just as forms. cf. Vlastos 1969, 299-300.

[4] Devereux 1994 gives the most representative account of such case. In that essay he gives detailed reasons for declining the 'immanent form,' especially by referring to other Platonic texts regarding the concept of form. Perl 1999 also reckons that 'x-ness in something' are merely 'appearances' of form and rejects the view that it is an entity. cf. Devereux 1994, 193-204 and Perl 1999, 303-304.

Critique, vol. MMXXIV, no. 1, pp. 1-12.

between 'x-ness itself' and 'x-ness in something', I will argue that 1) despite the distinctive differences between pure form and 'x-ness in something,' the latter shares one most crucial factor with the pure form, namely the unchanging identity, i.e. its resistance to change and opposition. Therefore, it is reasonable to call it the 'immanent form.' 2) It is based on the unchanging identity of 'immanent form' and its role on the particular that the 'final argument' is successful. Taking this view, I will then reconstruct Socrates' argument by differentiating two kinds of 'immanent forms', namely 'occasional' and 'essential,' considering their relationship with particulars. It is in the latter case that the soul's immortality can be proved.

NATURE OF 'IMMANENT FORM'

At the beginning of the 'final argument,' Socrates, after having defined the 'pure' form as 'x-ness itself', identifies what he considers to be the 'strongest argument (ἐρρωμενέστατος λόγος)' as the foundation for his discourse. He suggests that the strongest and safest answer why a particular manifests to be 'x' is 'x-ness itself,' i.e. the corresponding pure form[5]. For instance, the reason for a man's characteristic of being tall is the form 'tallness itself.' Based on this, Socrates argues that:

> ..., when the above had been accepted, and it was agreed that each of the Forms existed, and that other things acquired their name by **partaking of** them, he followed this up by asking: If you say these things are so, when you then say that Simmias is taller than Socrates but shorter than Phaedo, do you not mean that there is **in Simmias both tallness and shortness**?—I do. (102b)
>
> ..., ἐπεὶ αὐτῷ ταῦτα συνεχωρήθη, καὶ ὡμολογεῖτο εἶναί τι ἕκαστον τῶν εἰδῶν καὶ τούτων τἆλλα **μεταλαμβάνοντα** αὐτῶν τούτων τὴν ἐπωνυμίαν ἴσχειν, τὸ δὴ μετὰ ταῦτα ἠρώτα, εἰ δή, ἦ δ' ὅς, ταῦτα οὕτως λέγεις, ἆρ' οὐχ, ὅταν Σιμμίαν Σωκράτους φῇς μείζω εἶναι, Φαίδωνος δὲ ἐλάττω, λέγεις τότ' εἶναι **ἐν τῷ Σιμμίᾳ ἀμφότερα, καὶ μέγεθος καὶ σμικρότητα**; ἔγωγε.

This paragraph can be divided into two parts. In the first part, Socrates states that the pure form exists, and that each particular has its name by 'partaking of' (μεταλαμβάνον)[6] the form. Here, Socrates first uses the term 'partake' to illustrate the relationship between particulars and pure forms. The form 'x-ness itself' is the cause of a particular A's characteristic of being 'x,' because A acquires such characteristic by partaking of the very 'x-ness itself.' In the second part, Socrates further explains that, when Simmias is 'tall' and 'short' on different occasions, he must first have the 'tallness' and 'shortness' within him. In other words, a particular has the characteristic of being 'x', which indicates that there is an 'x-ness' within it. As aforementioned, this 'x-ness in it' has been referred to as 'immanent form.' Following this term, it could be concluded that a particular gets its 'immanent forms' and manifests the

---

[5] Socrates' method of giving answer reminds us of the concept of 'cause.' The concept of 'cause' (αἰτία) in ancient Greece, particularly in Plato and Aristotle's theory of causality, differs significantly from modern causality, as it does not strictly adhere to a linear causal law. Vlastos 1969 argues that the ancient Greek concept of 'cause' primarily addresses the question 'why something is such the case.' Depending on the context, it could signify causal, teleological, or essential connections. This essay suggests that Plato also uses the concept of 'cause' in this broad sense, referring to anything that can answer the question 'why something has such certain characteristic' as a 'cause.' cf. Vlastos 1969, 292-296.

[6] For the translation of the term, cf. Gallop ed. 2009, 192-193.

characteristics by partaking of the pure forms. The 'immanent form' emerges as a result of the partaking process.

However, what does 'immanent form' exactly mean? Or to be exact, should we catagorise the 'x-ness in something' as some type of 'form?' There are clear distinctions between the 'x-ness in something' and the 'x-ness itself' in the context of the *Phaedo,* which indicates that the 'x-ness in something' is not equivalent to the concept of pure form. Pure forms are not directly connected with specific particulars, whereas the 'x-ness in something,' as the name suggests, exists within particulars. Pure form is immortal and thus is not subject to generation or destruction. However, in Socrates' later argument, the 'x-ness in something' is likely to be destroyed when it encounters its opposite (102e). Pure forms rule particulars, whereas it seems that the 'x-ness in something' could be influenced by the particulars, since it exists within it (103a-e).

Although we acknowledge that 'x-ness in something' is not equivalent to pure form, it does not imply that, as the result of the partaking process, 'x-ness in something' shows no essential kinship with the pure form. In fact, it possesses the same essential feature as that of the pure form, namely the unchanging identity, which makes it reasonable to identify the 'x-ness in something' as some special kind of 'form' in the context of the 'final argument.' To clarify this essential commonality, it is necessary to revisit Plato's definition of the pure form in the 'kinship argument,' where, among several descriptions, he outlined the fundamental feature of the pure form.

In the 'kinship argument,' Plato gives several statements about pure forms: 1) they maintain their essence and do not accept any change; 2) they are immortal; 3) they are real yet invisible; 4) they rule particulars (78d-80b). However, it should be noted that these statements are not of the same importance in principle; on the contrary, the latter ones are derived from the first statement. Regarding the second, because pure forms retain their essence and remain unchanged, they are least likely to be compounds, which ensures that they are not divisible or subject to destruction (78c-78d). They would also not accept its being as others, which prevents generation or destruction. Regarding the third, because pure forms are inherently unchangeable and immortal, they are real and can only be comprehended through reason instead of sensory perception (78e). Finally, regarding the fourth, since immortal beings are divine, it is natural for divine beings to rule (80a)[7]. It is obvious that among these statements of pure form, the most fundamental one should be the first, i.e. its inherent unchanging identity[8]. Pure form, as 'x-ness itself,' are always keeping its essence of being 'x' (for example, 'bigness itself' are big in essence)[9]. It does not undergo any changes, nor does it accept the existence of 'non-itself' or the opposites within itself. Therefore, it 'always remains the same and in the same state' (78c), maintaining its complete identity permanently. It is because of this unchanging identity that pure forms fundamentally differ from particulars. The particulars, constantly changing and 'varying from one to another' (78d), are unable to always maintain their own identity.

---

[7] It is worth noting that regarding the last two statements, Simmias' 'harmony theory' suggests that the invisible may be the harmony of particular material rather than 'x-ness itself' (86A-D). Additionally, the analogy from Cebes of the weaver and the cloak suggests that a particular, instead of some 'x-ness itself,' may also 'rule' over another particulars (87C-E). They offer another evidence that 'invisibility' and 'ruling over particulars' are not essential distinctions between pure form and particulars.

[8] For the fundamental status of 'unchanging nature' to the pure form, see Ebray 2023, 3-5.

[9] This statement, featured as 'self-predication,' are analysed in detail in Nehamas 1979, 92-103.

Socrates' descriptions of form suggest that it is because of the 'unchanging identity' that pure form becomes what it is. In such cases, by examining the 'x-ness in something', we shall find that it also possesses this unchanging identity. Continuing the above example of 'tallness' and 'shortness' in Simmias, Socrates argues:

> *Now it seems to me that not only **tallness itself** is never willing to be tall and short at the same time, but also that the **tallness in us** will never admit the short or be overcome, but one of two things happens: **either it flees and retreats** whenever its opposite, the short, approaches, **or it is destroyed by its approach.** (102d-e)*
>
> ἐμοὶ γὰρ φαίνεται οὐ μόνον **αὐτὸ τὸ μέγεθος** οὐδέποτ' ἐθέλειν ἅμα μέγα καὶ σμικρὸν εἶναι, ἀλλὰ καὶ τὸ **ἐν ἡμῖν μέγεθος** οὐδέποτε προσδέχεσθαι τὸ σμικρὸν οὐδ' ἐθέλειν ὑπερέχεσθαι, ἀλλὰ δυοῖν τὸ ἕτερον, **ἢ φεύγειν καὶ ὑπεκχωρεῖν** ὅταν αὐτῷ προσίῃ τὸ ἐναντίον, τὸ σμικρόν, **ἢ προσελθόντος ἐκείνου ἀπολωλέναι·**

Although Simmias, as a particular, may simultaneously be both 'tall' and 'short' on different occasions, the inherent 'tallness in Simmias' would never admit 'shortness.' In other words, 'tallness' in nature always remains identical with itself and rejects both the alteration from itself and the approach of its opposites. So, when the opposite of 'tallness,' namely 'shortness,' approaches, because of the unchanging identity, there are two possibilities for 'tallness': it 'either flees and retreats or is destroyed' (102d). Plato does not clarify in which circumstances each of these possibilities would occur, as they represent two completely different destinies for an 'x-ness in something'[10]. However, it is evident that the 'x-ness in something' are totally unlike the changeable particulars. Particulars often contain different kinds of 'x-nesses in something' and hence manifest different characteristics on different occasions. By contrast, the 'x-ness in something' would never tolerate any change from itself at the approach of its opposite. Unlike the immortal pure forms, however, it would be possible for an 'x-ness in something' to be destroyed at the approach of its opposite. Nevertheless, as long as it exists and is within the particulars, it always remains what it is and does not undergo any change.

In such a case, although there are distinct differences between 'x-ness in something' and the pure form, 'x-ness in something' is consistent with Plato's fundamental statement of the pure forms, namely the unchanging identity, in its own way. Therefore, in this context, it is reasonable to regard 'x-ness in something' as a special sort of 'form' within a particular, i.e. an 'immanent form.'

Moreover, in defining the 'immanent form,' Socrates' 'strongest argument'[11], which argues that the reason for the particulars' characteristic of being 'x' is the corresponding pure form 'x-ness itself,' could be more clearly explained. Particulars firstly have the corresponding immanent 'x-ness' within them by partaking of 'x-ness itself', and by virtue of the immanent 'x-ness' they present the corresponding characteristics. Let us return to the above example where Simmias is taller than Socrates. Firstly, Simmias partakes of the pure form 'tallness itself,' then he has the immanent 'tallness' within him. The 'tallness' within him enables him

---

[10] 'Fleeing or retreating' might mean that the particular may be possible to reappear, while 'being destroyed' seems that it would cease to exist. Some scholars have argued that the option of 'fleeing and retreating' applies only to the case of the soul. However, Plato never directly presents such an argument. I believe he has left room in the text for various concrete circumstances. For relevant discussions, see Miura 2018, 195.

[11] See the beginning of these chapter.

to manifest the characteristic of being 'tall.' Socrates, on the contrary, by partaking of the pure form 'shortness itself,' has the 'shortness' within him. Thus, he manifests the characteristic of being short (102b-d). As can be seen, the partaking process could be described as a causal chain: pure forms — 'immanent forms' — characteristics of a particular[12]. In such a case, the pure form is always the decisive and fundamental cause of the characteristics of particulars, whereas the 'immanent form' is a result from the partaking process of the pure form within a particular. On the other hand, when it comes to a particular itself, the 'immanent form' could also be considered as the secondary cause of its characteristic, as it makes the particular manifest such corresponding characteristic. For example, Simmias' being tall is primarily because of the 'tallness itself', and secondarily because of the 'tallness in him'. As the secondary reason for a particular's characteristic of being 'x' is the 'x-ness within it', the changes and alterations of this 'x-ness within it' hence would impact on the particular[13].

As will be demonstrated below, Socrates' statement of the unchanging quality of the 'immanent form' and its role in serving as a secondary cause of the characteristic of a particular will play an essential part in his 'final argument' of the soul.

ABANDONING THE PURE FORM: DILEMMA OF THE 'KINSHIP ARGUMENT'

Having clarified the unchanging identity of 'immanent form,' Socrates takes it as a standpoint to carry on his final argument for the immortality of the soul. However, there is still one problem with this approach: if 'immanent form' fundamentally shares the same nature with pure form, why did Plato abandon the latter in favour of constructing a special sort of 'form' to prove the immortality of the soul?

In the 'kinship argument,' Socrates indeed attempts to prove the immortality of the soul through its similarities with pure forms. He argues that composite and compound things are by nature (φύσει) inclined to split up, while non-composite things are less likely to disintegrate. Uniform and unchangeable things are most likely to be non-composite, whereas changing things are composite. Therefore, pure forms are eternal and do not split up, whereas particulars will eventually disintegrate. According to Socrates, human can be referred to as a union of body and soul. The soul is 'akin to' (συγγενής) invisible forms and the body is more similar to a visible combination of materials. The soul also rules over the body, which is similar to that the form rules over particulars. When the soul is not confined to the senses of the body but get knowledge through reason by itself alone, it becomes 'more similar to' (ὁμοιότερον) and 'more akin to' (συγγενέστερον) the unchanging pure form. Thus, it draws as close as possible to the immortality (78c-80).

However, such a proof, accomplished through pure form, has an inherent defect: the soul, though it may be most akin to pure form, is not a form[14]. Let us return to the fundamental unchanging identity of pure forms. Should the soul be a form, it must always remain

---

[12] For a further elaboration of this divided structure, cf. Frede 1999, 120-124.

[13] Bostock also supports the perspective of taking 'immanent form' as cause for characters at particular's level. cf. Bostock 1986, 179-184.

[14] There are some scholars who argue that the soul is somehow pure form (cf. Fei 2019, 83-86). However, such an understanding fails to take into account the fundamental difference between the soul and the forms described below, which is discussed in detail in Burnet ed., 1911, 78-79 and Gallop ed. 2009, 140-141. In addition to Socrates' implication here that the soul is not a pure form, in the 'final argument' he directly compares the soul to particulars such as 'fire' and 'snow,' which further proves that Socrates thinks the soul are of the same type as particulars like fire and snow.

permanently what it is without accepting any change. According to Socrates' statement of the soul, it must always examine things through reason and be present with other pure forms. The truth, however, is that the soul is imprisoned by the body and is in constant danger of being compelled by the body to mistake the senses of the body as real. It might even remain in the sensible world after death, not wish to ascend into the invisible world, and ultimately reincarnate into a new body (81b-82c). The soul could be so easily affected by the body and even divert itself from its nature that it is in no way akin to the pure form. In terms of this, Socrates claims that the soul must philosophically 'practice death' in this world to remain as free as possible from the influence of the body and to maintain its identity, so that it would live with the pure forms after death (84a-b). However, even in this case, he could only use the term 'being similar to' (ὅμοιος) and 'being akin to' (συγγενής) to describe the soul's relation with the pure form as well. Nevertheless, the soul, however similar to pure form, is not a pure form. Since the soul likely to change its nature, it is more reasonable to classify it as a particular. Since the soul belongs to particulars, however special it is, it could never achieve eternal immortality as the pure forms. The immortality of the soul has not been proved yet.

The above problems in the 'kinship argument' are exposed in Cebes' question. Cebes compares the relationship between the soul and the body to that between a weaver and a cloak. The weaver, being stronger and more durable than the cloak he weaves, can wear and care for different cloaks throughout his life. Similarly, the soul is stronger than the body and can enter and care for different bodies many times. However, while the weaver is stronger than the cloak, he, as a particular, is still perishable (87b-d). Similarly, the soul may die at one point after tending to different bodies many times and then perish completely[15]. Cebes' rebuttal raises the problem: since the soul is a particular rather than a pure form, it cannot be truly eternal as the pure forms. It is ultimately finite. There is an unbridgeable gap between the pure form and the soul, therefore the proof of the immortality of the soul through pure form is fundamentally incapable of fulfilling its purpose.

FROM 'IMMANENT FORM' TO THE IMMORTALITY OF THE SOUL

(1). METHODOLOGY

Cebes' question leads Socrates to rethink to method of proving the soul's immortality: since the soul belongs to a particular, it would likewise contain 'immanent forms' and hence represent different characteristics. If the soul is proved to be immortal, it must always manifest a certain characteristic of being alive. Hence, there must be some 'immanent form' opposed to 'death,' namely 'life,' serving as the cause of its being alive[16]. In other words, if Socrates is able to prove that there are some 'immanent form' of 'life' which is always within the soul, and that this 'immanent form' could always prevent the soul from dying, in such manner he could prove the soul's immortality. This approach should be considered more feasible considering that the soul belongs to particulars.

---

[15] Ebray 2023 argues that the soul's act of 'reweaving,' namely its exhaustion to care for the body is the cause for the future perish of the soul. However, regardless of this, the soul, just as the weaver, would still never reach eternity. cf. Ebray 2023, 254.

[16] In the autobiographical section preceding the 'final argument,' Socrates seems to have already noticed the importance of the question of 'cause' in arguing for the immortality of the soul; this emphasis on cause serves as a prelude to the 'final argument'. See Ebray 2023, 223.

(2). THE 'ESSENTIALLY IMMANENT FORM' AND 'BRINGER'

Taking the above approach, Socrates first distinguishes two types of 'immanent forms,' by differentiating two kinds of relationships between a particular and a certain 'immanent form,' namely 'the occasional' and 'the essential'[17]. Therefore, there are 'occasionally immanent form' as well as 'essentially immanent form.' It is necessary to analyse these two types respectively.

The 'occasionally immanent form,' as its name suggests, is the result that a particular *per accidens* partakes of a pure form. In this case the particular has this 'immanent form' and manifests corresponding characteristic in certain circumstances. For example, on the occasion where Simmias is taller than Socrates, he *per accidens* has the 'tallness' within him, which makes him tall. When the circumstances, in which the particular is situated, change, the particular may accordingly have an opposed 'immanent form.' When Simmias is compared to Phaedo, who is taller than him, he would then *per accidens* have the immanent 'smallness.' When this takes place, the 'tallness' in Simmias, as aforementioned, would retreat or be destroyed as 'smallness' approaches. However, as the 'tallness' within him would retreat or be destroyed, the identity of Simmias himself would never be affected, for 'tallness' is just *per accidens* within him. Simmias, whether taller or shorter than others, is always himself and maintains his identity, regardless of how the status of 'tallness' and 'smallness' change within him.

The 'essentially immanent form,' on the other hand, differs significantly from the former. While the 'occasionally immanent form' is irrelevant to the identity of a particular itself, the 'essentially immanent form' is tightly connected with the particular's identity. When a particular remains what it is, it always manifests certain characteristics, without which the particular cannot be referred to as itself. Since 'immanent form' accounts for the manifestation of characteristics, insofar as a particular is itself, it must necessarily have contained certain 'immanent forms.' For example, a particular 'fire' in nature always manifests the characteristic of being 'hot', therefore it must have the corresponding immanent form of 'hotness.' Only in such cases can it be called 'fire.' This also applies to 'snow' and 'coldness,' 'three'[18] and 'oddness,' and so forth.

The 'essentially immanent form' is so strongly linked with the identity of the particular itself that it brings about two implications for the particular. To begin with, when an 'essentially immanent form' meets its opposite, not only itself will retreat or be destroyed, but the particular, due to the essential connection, will not accept the opposite 'immanent form' either. It will likewise retreat or be destroyed. For example, 'coldness' is an 'essentially immanent form' of snow. If the opposite 'hotness' approaches and enters snow, the 'coldness' will retreat or be destroyed. Correspondingly, the particular 'snow', along with 'coldness,' will also retreat or be destroyed instead of containing 'hotness,' because having 'hotness' within it is never consistent with the identity of snow.

---

[17] For the discrimination of these terms, cf. Frede 2011, 147-149.

[18] The number in Plato's philosophy is quite complicated. As Aristotle puts famously, it should be an intermediate between visible particulars and forms. However, I reckon that here Socrates here simply takes 'three' as akin to particulars, which is common in ancient Greek mathematical thoughts, instead of introducing another dimension. So it is reasonable to put it here in parallel with 'snow' or 'fire.' cf. Mendell 2022, 160-164.

Furthermore, if a particular has an 'essentially immanent form,' it is the 'bringer' of this 'immanent form'[19]. Socrates does not imply how the 'bringing' process functions. He simply mentions that some particulars would somehow 'bring' its 'essentially immanent form' to other particulars when it occupies them. For example, 'three' would bring 'oddness' to something when occupying it (105a). Many scholars, following this description, suggest that the 'bringer' could simply transfer its 'essentially immanent form' into other particulars[20]. However, as aforementioned, the 'immanent form' would only be the result of the partaking process. In such case, a 'bringer' would be unable to directly bring the 'immanent form' to another. It would rather somehow 'stimulate' the particular it occupies to partake of the same pure form. Therefore, a better explanation of the 'bringing' process would be revised as follows: Particular A is the 'bringer' of the 'essentially immanent form' 'x-ness within it'. When A occupies another particular B, it will stimulate B to partake of the very 'x-ness itself'. In this way B could have the same 'x-ness within it'. For example, when fire, which essentially has 'hotness' within it, occupies a particular, it would stimulate the particular to partake of 'hotness itself' and have the corresponding 'hotness' within it. From a more general perspective, it would seem that fire 'brings' its 'hotness' to this particular. Hence the so-called 'bringing' process would be demonstrated as below:
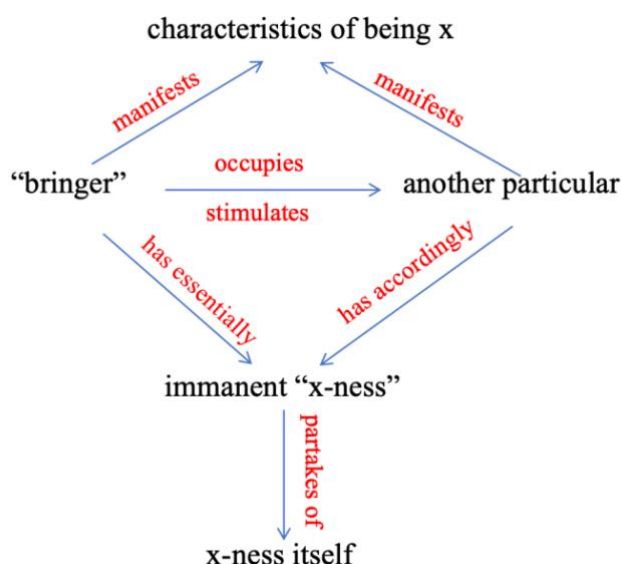
In this way a 'bringer' apprears to 'bring' its 'essentially immanent form' to another particular and makes it manifest the same characteristic. Moreover, Socrates seems to indicate that 'a particular having an 'essentially immanent form'' is the necessary and sufficient condition of 'it being its 'bringer.'' In such case, if a particular constantly serves as a 'bringer' of an 'immanent form,' it could also be identified to have it essentially[21].

On raising the concept of 'bringer,' Socrates gives a 'smarter answer' to explain the reason why a particular has the characteristic of 'being x,' beyond the aforementioned 'safest answer':

---

[19] I agree with David Sedley's judgement that a 'bringer' (he called 'essential bearer') must be a particular rather than form. See Sedley 2018, 213.

[20] cf. Ebray 2023, 260-266; Sedley 2018, 212-214; Frede 1978, 28-30, etc.

[21] This hypothesis seems to be taken for granted in the final proof of the soul's immortality below. cf. 105d-105e.

*...I say that beyond that safe answer, which I spoke of first, I see another safe answer. If you should ask me what, coming into a body, makes it hot, my reply would not be that **safe and ignorant one**, that it is heat, but our present argument provides a **smarter answer**, namely, fire, and if you ask me what, on coming into a body, makes it sick, I will not say sickness but fever...* (105b-c)

...λέγω δὴ παρ' ἣν τὸ πρῶτον ἔλεγον ἀπόκρισιν, τὴν ἀσφαλῆ ἐκείνην, ἐκ τῶν νῦν λεγομένων ἄλλην ὁρῶν ἀσφάλειαν. εἰ γὰρ ἔροιό με ᾧ ἂν τί ἐν τῷ σώματι ἐγγένηται θερμὸν ἔσται, οὐ **τὴν ἀσφαλῆ** σοι ἐρῶ ἀπόκρισιν ἐκείνην **τὴν ἀμαθῆ**, ὅτι ᾧ ἂν θερμότης, ἀλλὰ **κομψοτέραν** ἐκ τῶν νῦν, ὅτι ᾧ ἂν πῦρ· οὐδὲ ἂν ἔρῃ ᾧ ἂν σώματι τί ἐγγένηται νοσήσει, οὐκ ἐρῶ ὅτι ᾧ ἂν νόσος, ἀλλ' ᾧ ἂν πυρετός...

The 'safe and ignorant' answer, as aforementioned, is that the cause of something's characteristic of being 'x' is the pure form 'x-ness itself.' This, as the name indicates, is reliable but unable to inform us more beyond the form itself. However, on raising the concept of 'bringer,' Socrates is able to take an external particular as a cause to explain why something has certain characteristics. Since fire has the 'essentially immanent form' 'hotness,' it is the 'bringer' of 'hotness.' Therefore, when some particular is occupied by fire and stimulated to have the 'immanent form' 'hotness,' it is also reasonable to say that 'fire' is the cause of its characteristic of being hot. In such manner Socrates has established a causal relationship between two particulars from the 'essentially immanent form.'

(3). FINAL PROOF OF THE IMMORTALITY OF THE SOUL

Through introducing the concept of the 'bringer', Socrates begins to analyse the relationship between relative particulars and their respective 'immanent forms.' In this way Socrates begins his final proof of the immortality of the soul.

The final proof could be divided into two parts. In the first part Socrates identifies the soul as the 'bringer' of 'life' by concluding the relationship between the soul and the body. It is generally accepted that the soul 'brings' life to the body as it occupies it, for it is obvious that when the soul occupies the body, it is alive; however, when the soul leaves, the body dies and quickly dissolves. In this way, the soul appears to stimulate the body to have the 'immanent form' of 'life' and hence could be regarded as the 'bringer' of 'life.' Therefore, according to the aforementioned sufficient and necessary condition, 'life' is the 'essentially immanent form' of the soul. Since 'life' cannot contain its opposite, namely 'death,' the soul cannot accept 'death' either. So when death approaches and enters the soul, it would likewise be confronted with two destinies: it either flees and retreats, or is destroyed (105d-106a).

The first part, however, presents great danger to the soul, as it is still possible to be destroyed when death comes. Therefore, in addition to proving that the soul has the characteristic of being un-dead (ἀθάνατος), it is also necessary to prove that it is indestructible (ἀπώλευτος), i.e. it would not be destroyed when death approaches. Socrates makes the argument for this, again, by analysing the identity of the 'immanent form'. Firstly, the destinies of a particular and its 'essentially immanent form' are identified. If the 'essentially immanent form' is destroyed at the approach of its opposite, so will the corresponding particular. On the other hand, if some 'essentially immanent forms' would not be destroyed but retreat, the particulars would as well retreat insead of being destoryed. In the case of the soul and 'life,' when death approaches, 'life' would generally either retreat or be destroyed. However, if

something is destroyed, it ceases to exist. There would be no other way to go out of existence by 'dying' or accepting death[22]. As 'life' is opposed to 'death,' it could never allow itself to accept 'death,' because it, as an 'immanent form,' always remains its identity. Therefore, 'life' could not be destroyed. It can only flee or retreat as death approaches. The soul, having 'life' as its 'essentially immanent form', would be faced with the same destiny. When death comes, the soul would not be destroyed either. It would otherwise flee and retreat to the realm of the underworld.

CONCLUSION

Let us make an overview about Socrates' final argument of the immortality of the soul from the perspective of the 'immanent form':

1) The 'immanent form' cannot be regarded as simply 'characteristic' of a particular. It is reasonable to regard it as some special 'form,' for it shares the unchanging identity with pure forms. As a result of the partaking process, it serves as a secondary cause of the corresponding characteristics of particulars.

2) From the 'kinship argument', Socrates implies that the soul belongs to particular. Therefore, it is unlikely to prove its immortality by comparing it with pure forms. Cebes' question acutely points out of the problem.

3) Since the soul is a particular, it should have 'immanent forms' and hence manifest different characteristics. There are two kinds of 'immanent forms,' namely the 'occasionally immanent form' as well as 'essentially immanent form'. While the latter meets its opposite, the particular will retreat or be destroyed along with it. When a particular has an 'essentially immanent form', it becomes its 'bringer,' namely being able to stimulate other particulars when occupying them to partake of the same form and have the very 'immanent form'.

4) The soul is the 'bringer' of life, accordingly 'life' is its 'essentially immanent form.' When 'death' meets 'life,' since 'life' can accept neither death nor nonexistence, the soul would retreat along with 'life.' In such a case, the soul is both 'un-dead' and 'indestructible.' Therefore, it is immortal.

With the conclusions above the main purposes of this essay could be complete: I have not only clarified the righteousness of referring to 'x-ness in something' as 'immanent form', but also argued that 'immanent form' play an essential role on the 'final argument'. This interpretation of 'final argument' may serve as another perspective to understand this most complicated argument in *Phaedo.*

After the two questions are raised by Simmias and Cebes, all of the audiences are followed by a great deal of depression and grief. Socrates, on the other hand, remains calm and firm. He reminds his audience to take trust on reason (λόγος), and keep seeking truth through it:

> ***That we should not become misologues****, as people become misanthropes. There is no greater evil one can suffer than to hate reasonable discourse...* (89d)
> **μὴ γενώμεθα, ἦ δ' ὅς, μισόλογοι**, ὥσπερ οἱ μισάνθρωποι γιγνόμενοι· ὡς οὐκ ἔστιν...ὅτι ἄν τις μεῖζον τούτου κακὸν πάθοι ἢ λόγους μισήσας...

---

[22] I take this understanding from Dorothea Frede as convincing. cf. Frede, 1978: 301-302.

Socrates' eventual attempt to prove the immortality of the soul in his 'final argument' is his ultimate implementation of the principle of argumentation (λόγος). He has investigated into the nature of the soul and proved its immortality through dialectic discourses and arguments. Admittedly, not everyone agrees that the final argument is successful. For example, Plato's method of taking 'life' as an 'immanent form' of a particular has been criticised more generally by modern philosophers, who argue that 'being' could be at a more fundamental status than other characteristics[23]. However, it should be noted that, if we follow the context of *Phaedo* and Plato's presuppositions, the argument of the soul's immortality should be considered complete.

Moreover, in a deeper intellectual sense, the implementation of 'immanent form' in the 'final argument' is also an important expression of the 'spirit of λόγος' which Socrates carries out to the end of his philosophical life. By abandoning the 'ascent' to the 'pure forms,' Socrates returns to a discussion of the soul's own nature and takes a circuitous path that ultimately gives reliable proof of the soul's immortality and ensures the meaning of Socrates' 'philosophical death.' After all, only when the soul is proved immortal through reason, can the foundation for the meaning of his death be laid firmly, so would the meaning of philosophy be revealed.

REFERENCES

Bostock, D. (1986). *Plato's Phaedo*. Oxford, United Kingdom: Clarendon Press.

Burnet, J. (Ed.). (1911). *Plato's Phaedo*. Oxford, United Kingdom: Clarendon Press.

Devereux, D. T. (1994). "Separation and Immanence in Plato's Theory of Forms." C. C. W. Taylor (Ed.), *Oxford Studies in Ancient Philosophy*. Oxford, United Kingdom: Oxford University Press

Duke, E. A., Hickon, W. F., Nicoll, W. S. M., & Robinson, D. B. (Eds.). (1995). *Platonis Opera*. Vol. 1, Tetralogiae I-II. Oxford, United Kingdom: Oxford University Press.

Ebrey, D. (2023). *Plato's Phaedo: Forms, Death, and the Philosophical Life*. New York, USA: Cambridge University Press.

Fei, W. (2019). "Being and Life in Phaedo." *Philosophical Research,* no. 2, pp. 80-89.

Frede, D. (1978). "The Final Proof of the Immortality of the Soul in Plato's Phaedo." *Phronesis*, vol. 23, no. 1, pp. 27-41.

——— (1999). *Platons Phaidon: Der Traum von der unsterblichkeit der seele*. Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.

——— (2011). "Das Argument aus den essentiellen Eigenschaften (102a-107d)." J. Müller (Ed.), *Platon: Phaidon* (pp. 143-159). Berlin, Germany: Akademie Verlag.

---

[23] One of the famous rebuttals is from Kant, who argues that 'being' should not be a predicate. cf. Hintikka 1981: 128-130.

Gallop, D. (Ed.). (2009). *Plato Phaedo*. Oxford, United Kingdom: Oxford University Press.

Hintikka, J. (1981). "Kant on Existence, Predication, and the Ontological Argument." *Dialectica,* vol. 35, no. 1, pp. 127-146.

Mendell, H. (2022). "Betwixt and Between: Plato and the Objects of Mathematics." D. Ebrey & R. Kraut (Eds.), *The Cambridge Companion to Plato* (pp. 358-398). Cambridge, United Kingdom: Cambridge University Press.

Miura, T. (2018). "Immortality and Imperishability of the Soul in the Final Argument of Plato's Phaedo." G. Cornelli, T. Robinson, F. Bravo (Eds.), *Plato's Phaedo: Select Papers from the Eleventh Symposium Platonicum* (pp. 193-198). Baden-Baden, Germany: Academia Verlag.

Nehamas, A. (1979). "Self-Predication and Plato's Theory of Forms." *American Philosophical Quarterly*, vol. 16, no. 2, pp. 93-103.

Plato. (1997). *Complete Works*. (J. M. Cooper & D. S. Hutchinson, Eds.). Indianapolis, USA: Hackett Publishing Company.

Perl, E. D. (1999). "The Presence of the Paradigm: Immanence and Transcendence in Plato's Theory of Forms." *The Review of Metaphysics,* vol. 53, no. 2, pp. 339-362.

Sedley, D. (2018). "The *Phaedo*'s Final Proof of Immortality." G. Cornelli, T. Robinson, F. Bravo (Eds.), *Plato's Phaedo: Select Papers from the Eleventh Symposium Platonicum* (pp. 210-220). Baden-Baden, Germany: Academia Verlag.

Vlastos, G. (1969). "Reasons and Causes in the Phaedo." *The Philosophical Review*, vol. 78, no. 3, pp. 291-325.

Illusionism and Psycho-Physical Harmony

THOMAS MILLS
Durham University

PSYCHO-PHYSICAL HARMONY

The problem of psycho-physical harmony challenges us to explain why sensations and behaviour rationally match, such as why we aim to avoid pain. We may be able to produce evolutionary explanations for why dangerous stimuli are painful, but this relies on organisms already aiming to avoid painful experiences. We cannot explain avoidance of pain in terms of natural selection for fear of circularity. Goff invites us to consider Inverted Ian.[1] Ian's desires and aversions are inverted, so he wants to avoid eating burgers and desires to be hacked to pieces. However, Ian's world is different to ours: conscious beings there act to seek out their aversions and avoid their attractions. These factors combine such that Ian leads a life no different from yours or mine on the outside — but it's utterly bizarre! It is totally irrational to consistently avoid those things which you desire and seek out things that you are repulsed by, yet that's what Ian ends up doing. Given that Ian behaves exactly like we do, natural selection should have no reason to favour us over him, and some other explanation must be provided for why we are not like Ian. We must explain something broader than evolution: the fine-tuning of our universe to produce psycho-physical harmony.

Cutter and Crummett discuss various kinds of normative harmony — cases where a conscious state harmoniously provides reasons for our actions.[2] Our friend Ian is an instance of cognitive *dis*harmony, with his actions not rationally matching his desires. He takes action to seek out certain foods, even when his aversion to those foods justify the opposite action. We may also consider hedonic harmony, where pain sensations correspond with avoidance behaviour. Unpleasant sensations like pain seem to justify our avoiding them, but they do not appear to entail actual avoidance. As we saw above, natural selection can only associate danger with pain in a hedonically harmonic world, where organisms actually avoid unpleasant sensations, but provides no explanation for why this is the case. A third type is epistemic harmony, where our sensations rationally match the beliefs they bring about. For instance, my sensation of seeing a ball leads me to, and justifies, the belief that there is a spherical object before me, and not to some other unrelated belief.[3] These types of harmony apply to a large swathe of the experience-behaviour pairings in our day-to-day lives, so we really ought to have

---

[1] Goff, 2023, pp. 72-73
[2] Cutter & Crummett, forthcoming, pp. 6-8
[3] Ibid.

some explanation for why our experiences justify our behaviour despite not seeming to guarantee it.

We may also consider semantic harmony. Semantic harmony describes the connection between phenomenal experiences and verbal reports in terms of truth, rather than in terms of the rational match with behaviour emphasised in discussion of normative harmony. That is to say that our verbal reports regarding our experiences require a harmonic phenomenal state to be true. Saying that two colours you see are similar wouldn't be true if we had TV-static disordered vision, or if we experienced sounds when looking at objects.[4] In such cases where we have some other phenomenal states which made our reports false, a physically identical system would have no reason to make different reports, and hence would be consistently false in its reporting. But we take most of our verbal reports of phenomenal states to be truthful, so our phenomenal states and verbal reports are harmonic. This cannot be overcome merely by pointing to a common cause, since we are focused on the semantics of verbal reports.[5] We can imagine consistently co-occurring report-experience pairings which do not bear semantic harmony, such as saying 'my eye itches' whenever one is happy. Pointing to a common cause between phenomenal state and verbal report explains this co-occurrence but does not entail that the report has anything to do with the experience, as in this example. Again, our reports appear to be consistently truthful, so some explanation beyond common cause is required to make semantic harmony likely.

We may formulate the problem in Bayesian terms: given various hypotheses, which ones make harmony most likely? Cutter and Crummett suggest that theism, or a theism-adjacent view, would give very good chances for harmony appearing in our world, so psychophysical harmony can be taken as important evidence towards those theses.[6] If God wants people to live sensical lives, then He may create psycho-physical laws which bring about normative and semantic harmony. On the other hand, naturalistic atheism alone gives us no reason to expect harmony, so our world would seem terribly unlikely under such a doctrine. We may express this as P(harmony|theism) >> P(harmony|naturalistic atheism). Although for many theism has a lower prior probability than naturalistic atheism, harmony is so much more likely on theism that P(theism|harmony) > P(naturalistic atheism|harmony), and psycho-physical harmony should make us re-evaluate our views of atheism and theism.

OTHER ACCOUNTS OF HARMONY

Can physicalist accounts of mind explain normative or semantic harmony? One approach may be to make an identity between phenomenal and physical or functional states. If pain is identical to avoidance functions, then there should be no surprise that the experience of pain rationally matches our avoidance of it. However, functional identities are usually posited broadly — each phenomenal state is identical to *some* functional state, but individual pairings are not specified. This merely seems to shift the coincidence into the identities themselves, rather than their manifestations in experience and behaviour. How fortunate that avoidance-functional states are identical to phenomenal pain, rather than to some pleasant sensation! Since we have no reason

---

[4] Ibid., pp. 15-16
[5] Ibid.
[6] Ibid., p. 19

to expect the specific identities which would guarantee harmony, functional identities fail to satisfyingly explain psycho-physical harmony.[7]

We can make a similar argument against accounts which seek to ground the mental in the physical without identity:[8] why are these grounding relations harmonic, when it seems possible for a given physical system to ground some other phenomenal experience? These grounding laws seem no better than positing identities between phenomenal and functional states, and again harmony seems miraculous.

Another approach is to explain the normative roles of phenomenal states as a *result* of the functional states they are contingently paired with.[9] On this 'contingent normative roles explanation', we may say that sensations like pain are not intrinsically bad, but rather become bad via their association with certain functional states which natural selection links to danger. This explains the coincidence of normative harmony, since the badness of phenomenal states is not intrinsic to them, but reliant upon physical happenings, and it should not surprise us that physical happenings match rationally with their effects. Although this explain normative harmony, suggesting that phenomenal pain is not intrinsically bad seems unintuitive. Cutter and Crummett that phenomenal duplicates without physical bodies would still have normative phenomenal states.[10] For example, a disembodied Cartesian mind seeing a ball would still give the mind good reason to believe that there was a ball in front of it, whether or not any physical states exist which are linked to that phenomenal state.

Now, the plausibility of a Cartesian mind may be contested by physicalists, but even then, a contingent link between phenomenal states and their functional states seems to be silent on semantic harmony. The normative roles of phenomenal states being contingent would not lead us to expect reports about them to be true — certainly not reports which aren't related to normativity. We could expand upon this explanation in some other way to account for semantic harmony, but it is not obvious how we could do this while maintaining that phenomenal states really do exist, as their having contingent normative roles would require. Then this explanation does not do a very good job of explaining harmony in general, falling far behind the theistic hypothesis given by Cutter and Crummett.

Theism is not the only answer to the problem of normative harmony, but other promising approaches result in radical changes in our understanding of agency in the universe,[11] or affording phenomenal states much more power than we may initially expect.[12] None of these are promising approaches for anyone interested in defending a naturalistic physicalist picture of the world. Hypotheses such as God or pan-agentialism would lead us to expect psycho-physical harmony, while the obvious evolutionary explanations seem to fail to do any better than it being an incredible fluke. Then psycho-physical harmony gives us very good reason to support non-naturalistic understandings of the world, over the more mainstream naturalistic physicalism which many philosophers work under.

---

[7] Ibid., pp. 34-35; see also Mørch, 2018, pp. 300-302

[8] Ibid., p. 31

[9] Ibid., pp. 10-13

[10] Ibid.

[11] Goff, 2023

[12] Mørch, 2018

ILLUSIONISM ABOUT CONSCIOUSNESS

Given that simple physicalist explanations fail to explain harmony, I will outline a more radical approach which I believe goes a good distance in providing the sort of explanations required: illusionism. Illusionists deny that experiences really have phenomenal properties, and instead focus on quasi-phenomenal properties — physical properties which are misrepresented as phenomenal during introspection.[13] These quasi-phenomenal properties give the illusion of genuine phenomenal properties when we introspect, perhaps because we lack enough detail about the complicated processes happening in our brains at any time.[14]

It is important to be precise about what illusionists mean when they say that phenomenal properties do not exist. Phenomenal properties, often called qualia, are typically characterised as ineffable, qualitative, intrinsic properties of experiences which we may be directly, or infallibly, aware of. Dennett provides a host of thought experiments which highlight how confusing these concepts make our discussion of consciousness.[15] Considering the various characteristics listed above, Dennett suggests that none of them stand up to scrutiny and proposes that the idea of qualia be dropped altogether and replaced with complicated properties which could in theory be discussed objectively. Frankish develops this, arguing that any 'diet qualia', usually described as the 'what-it's-likeness' of an experience, are also problematic.[16] Philosophers either develop this vague concept into full-blown qualia, or reduce it to 'zero qualia' — the quasi-phenomenal properties which make us believe that there is a real qualitative aspect to our experiences.[17] Since illusionists reject the existence of qualia, they only accept quasi-phenomenal properties as belonging to conscious experiences.

Opponents of illusionism may suggest that illusionists do not believe in conscious experience at all, and thus illusionism can be ruled out immediately due to the obvious fact that we have conscious experiences. However, this all depends on what we mean by 'conscious experience'. If we maintain that it means qualia in all their glory, then illusionists do indeed deny its existence. But if we take this line of thinking, we seem much less justified in saying that our having conscious experiences is obvious — Dennett's arguments against qualia are enough to cast doubt on qualia as something we understand in an importantly privileged way. If we accept a weaker notion of conscious experience, as experiences which seem to have phenomenal properties, then illusionists would agree that we have such experiences, as a result of quasi-phenomenal properties.[18] I have glossed over much of the detail in Dennett and Frankish's arguments. In this section, I merely aim to make the illusionist position explicit and show that it is not blatantly wrong. In the Bayesian framework in which the psycho-physical harmony problem is presented, illusionism having some non-zero prior credence is good enough for it to be supported as a hypothesis, which I shall now begin to argue for.

---

[13] Frankish, 2016, pp. 14-15
[14] Frankish, 2016, p. 18
[15] Dennett, 1992
[16] Frankish, 2012
[17] Frankish, 2012, p. 675
[18] Frankish, 2016, p. 21

Semantic Harmony and Illusionism

Semantic harmony is reminiscent of the meta-problem of consciousness: the question of why we *believe* that consciousness is phenomenal, rather than reducible to physical processes. This is considered an 'easy' problem of consciousness — something that can be understood by the study of brains and their processes, without any appeal to consciousness itself. Given that this belief is not dependent on the actual status of phenomenal states, we arrive at the coincidence of their apparent truth, which Cutter and Crummett develop into the problem of semantic harmony.[19]

However, the meta-problem is also of special significance to illusionists. Frankish argues that the meta-problem is *the* problem of consciousness. That is to say that our reports and intuitions about consciousness are all that need explaining, since there is no phenomenal consciousness for our brains or judgements to be linked with.[20] We may expect many accounts of consciousness to have something to say about the meta-problem, whether or not they affirm the reality of phenomenal properties. But illusionists need say nothing more: like UFOs, once we explain the illusion, we have little justification for believing in genuine phenomenal properties.[21] There is plenty to be said on how the meta-problem is to be approached — this is the main project of illusionist accounts.[22] That said, tackling the problem of semantic harmony does not require an in-depth explanation of *how* illusions of consciousness come about. Only the assertation that phenomenal states are illusory is needed.

The illusionist response to semantic harmony is merely to say that no harmony exists. In denying the existence of phenomenal properties, illusionists deny the truth of our reports regarding phenomenal states. If we say 'I am in pain' and aim to pick out a phenomenal state, we are simply wrong. There can be no coincidental harmony if reports are systematically false, and thus there is no fine-tuning to explain.

Might we understand these reports in some other way? We may take them as theorists' fictions, as in Dennett's heterophenomenological method.[23] These reports can be taken to roughly track the physical goings-on of the brain and body: even if 'I am in pain' does not yield a truth about phenomenal states, it is generally correlated with bodily damage and certain brain states. Can we take *this* to be fortunately harmonious? I would suggest not. The verbal reports of pain produced by a physical system have as a causal factor that bodily damage or brain activity, so there is little room for coincidence. For another way of thinking about it, we could consider a philosophical zombie. We expect zombies' verbal reports to match their brains and bodies. But there is certainly no harmony here, since zombies have absolutely no phenomenal experience. Even if we take reports to have limited, fictitious truth, illusionism permits no semantic harmony.

---

[19] Cutter & Crummett, forthcoming, pp. 13-15
[20] Frankish, 2019, p. 85
[21] Ibid.
[22] Frankish, 2016, p. 37
[23] Dennett, 1993

NORMATIVE HARMONY AND ILLUSIONISM

Our desires, vision, and sensations justify our behaviour but seem unnecessary for action. Again, zombies are a good explanation. They behave just as we do, but they have no phenomenal experience to justify their behaviour. Comparing them with us, our consistent justification appears to be blind luck on top of a fully-functioning machine. Unlike with semantic harmony, illusionists cannot flatly deny that our experiences justify our actions. Illusions really do give us reason to believe things. Seeing a magic trick gives us *some* reason to believe in magic, even if this support becomes negligible upon having the trick explained to us. We might adapt the question of normative harmony to ask why our experiences *seem* to rationally match our behaviour. Why does pain *seem* unpleasant if bodies are wired to avoid damage anyway?

Illusionists assert that the way an experience seems to us is something describable through physiological language; understandable in terms of quasi-phenomenal properties which are misrepresented during our introspection. In varying how experiences seem to us, we would also have to vary the physiological system or relevant laws.[24] For example, to consider a world where bodily damage brings about hunger-like sensations, we would have to consider a world with different physiological laws governing how damage-states are introspectively represented, which is a physical process. However, this would undermine the problem of normative harmony. We ask why *our* world is harmonic — other worlds with different physiological laws being disharmonic need not concern us. Worlds with different physical laws are still useful to consider for more general fine-tuning problems of life and the universe. But psycho-physical harmony asks why *we* find harmony when other physically or functionally identical worlds may not; the only things we want to vary for this problem are phenomenal states. Without genuine phenomenal properties to vary, illusionism cannot permit a large probability space in which only a small region is harmonic. If we accept the illusionist view that apparently phenomenal properties are actually just misrepresentations of physical systems produced by physical processes, then it seems that any world with those same physiological laws would have those same misrepresentations.

Now we can work backwards to understand how these misrepresentations are associated with behaviour. The approach I favour bears similarity to the phenomenal powers view, which I will briefly outline. Mørch argues that the correspondence of pain with avoidance behaviour suggests that phenomenal pain is intrinsically powerful — it causes organisms to try to avoid it on the basis of its qualitative character alone.[25] This would rule out certain disharmony cases, such as consistently desiring to be in pain. However, a significant challenge to Mørch's approach is given by Cutter and Crummett: pain may motivate us to want to avoid it, but this does not make our physical behaviour harmonising with this desire any less coincidental.[26] Cutter and Crummett suggest that there is nothing inconceivable about phenomenal powers guaranteeing purely mental harmony without entailing actual avoidance

---

[24] I use the word 'physiological' instead of 'physical' to remain neutral on whether bodily processes can be micro-reduced to physics. If they can be, then physical laws would decide the processes of the body, and if not, then some non-reducible laws at the level of chemistry or biology would be in charge. These would be fine-tuned for life in the same way as the laws of physics are, but this should not affect our judgements of harmony *given* a life-producing universe, which is what the problem of psycho-physical harmony works under.

[25] Mørch, 2018, p. 302

[26] Cutter and Crummett, forthcoming, p. 28

behaviour. In other words, Mørch's view allows for what Goff describes as practical irrationality,[27] where we do not appropriately act on our desires. The phenomenal powers view does not say enough to explain normative harmony, since phenomenal powers explain our *attempts at avoiding* certain experiences, based on their phenomenal character, but not the successful manifestation of avoidance behaviour.[28]

I assert that an illusionist approach can provide such a link. As complex physiological properties, we have no reason to doubt the causal power of quasi-phenomenal properties on the body and brain. After all, these properties would affect and be affected by whatever physiological processes are happening in the brain, including the process of introspecting on them. This quasi-phenomenal power is not related to genuine phenomenal character, but instead expresses the physiological laws of our world — just as gravitation describes the power of masses to attract one another. Now, our introspective representation of quasi-phenomenal properties need not be a dead end: the representations themselves can bring about further brain processes, since illusionists understand introspection as a physical process. For instance, the misrepresentation of certain properties as being phenomenally painful could be causal factors of brain processes which lead to avoidance behaviour. By considering quasi-phenomenal properties, we can guarantee a link between behaviour and representations. From here, natural selection can favour those representations of quasi-phenomenal properties which cause beneficial behaviour, and we can arrive at the normatively harmonic world in which we live.

Is this a good explanation of normative harmony? Frankish highlights our ability to take illusions as powerful as a good reason to adopt illusionism over views which affirm the existence of non-physical phenomenal properties.[29] Taking illusions as intentional objects means we can make sense of their having causal power — not unlike how fictional stories can cause us to change our behaviour — without positing any psycho-physical law or intrinsic qualitative character to afford such causal power. Dennett understands talk of pain as a personal level of explanation, distinct from the physiological explanations we can afford from empirical research.[30] We are entitled to talk about pain as powerful — informing our actions — but we need not identify pain with or involve phenomenal properties in physiological causal accounts. These physiological explanations will be much more elaborate and will tell us everything we need to understand behaviour, while personal explanations are limited to discussing someone sensing pain and reacting to it. However, this limited vocabulary provides simplicity, as well as the ultimacy which Mørch advocates for as an explanatory virtue of the phenomenal powers account.[31] After all, if all we can say about pain's causal power is that it provokes avoidance behaviour, then we undoubtedly have an ultimate personal explanation why we avoid pain. This illusionist account works as a stronger explanation of psycho-physical harmony than phenomenal powers, providing a concise intentional account of harmony.

---

[27] Goff, 2018, p. 105

[28] If extended and combined with panpsychism, a phenomenal powers account may succeed in entailing normative harmony. This is discussed in the next section. However, taking phenomenal powers alone does not appear to guarantee psycho-physical harmony, as Cutter and Crummett point out.

[29] Frankish, 2016, pp. 27 and 29

[30] Dennett, 1986, p. 94

[31] Mørch, 2018, pp. 307-308

ILLUSIONISM AS A RESPONSE TO PSYCHO-PHYSICAL HARMONY

Illusionism is by no means an intuitive view of the mind. Chalmers suggests that Moorean arguments, which appeal to phenomenality being obvious to us, like the existence of our hands, are illusionism's strongest challenge.[32] The strength of the illusion of consciousness is something illusionists recognise. Phenomenality may *seem* obvious to us precisely because of these complex processes of the brain. However, this is not a particularly convincing response to firmly-held intuitions. Instead, pointing to the theoretical advantages of illusionism would be a better method of making the account appealing. One such advantage is a coincidence argument,[33] which is closely related to the problem of semantic harmony. By ruling out the truth of introspective reports, these reports do not seem luckily true, when considering meta-problem processes which would give us such intuitions regardless of actual consciousness.

As discussed above, physicalist grounding and identity accounts seem unsatisfactory explanations of psycho-physical harmony of both the semantic and normative varieties. Insofar as they have to posit some relation between physical properties or processes and genuine phenomenal properties, there seems to be coincidental harmony arising from this relation. Adopting an illusionist account, which removes any such relations, appears to be the simplest way for physicalists to explain harmony, despite worries about how unintuitive this move may be. If we are to maintain physicalism as a metaphysical approach, illusionism looks to be our best bet for giving physicalism a decent credence in the context of accounts of harmony.

Speaking of other accounts of harmony, how does illusionism compare to them? Psycho-physical laws designed by God are not an easy pill to swallow. Naturalistic atheism seems to be a more popular approach, both in philosophy and the sciences. Many would attribute theism a lower (but still significant) prior probability than atheism, which can then be augmented by theism giving us very good reason to expect harmony. Goff's pan-agentialism posits that the universe has a fundamental rationality, which combines with fundamental phenomenal consciousness to ensure physical processes rationally match phenomenal experience.[34] Again, panpsychism is a notable minority position in philosophy of mind more generally, so for many has a low prior probability.

We should also recognise that pan-agentialism is not the only approach for atheistic anti-physicalists: Mørch's phenomenal powers view is another possibility. The phenomenal powers view can be extended to explain our meta-problem intuitions,[35] and hence can resolve the problem of coincidence associated with semantic harmony. If powerful phenomenal properties are fundamental, then they also inform what we observe to be physical laws. This panpsychist phenomenal powers account would not permit a difference in phenomenal states without also changing the laws which constrain our behaviour,[36] as I have argued would be the case for illusionism. Then panpsychists have a choice between pan-agentialism and

---

[32] Chalmers, 2020, pp. 279-280
[33] Ibid., pp. 264 and 279
[34] Goff, 2023, chapter 3
[35] Mørch, 2020
[36] Ibid., p. 141

phenomenal powers as explanations of psycho-physical harmony.[37] I would suggest that this makes panpsychism a stronger approach than theism.

Returning to illusionism, the view looks to be in good company. While the denial of phenomenal character does not come easy to most people, other successful explanations of psycho-physical harmony are also far from our common understandings of the universe. Out of the explanations discussed here, illusionism is the only one compatible with physicalism and atheism and may be the most appealing choice for those accepting both of these broader metaphysical theses. Then I would suggest that physicalist atheists take the bold step to do away with conceptions of consciousness which would make our everyday experiences miraculous.

REFERENCES

Chalmers, D. (2020). "Debunking Arguments for Illusionism about Consciousness." *Journal of Consciousness Studies*, vol. 27, nos. 5-6, pp. 258-281.

Cutter, B. And Crummett, D. (forthcoming). "Psychophysical Harmony: A New Argument for Theism." *Oxford Studies in Philosophy of Religion*.

Dennett, D. (1986). *Content and Consciousness* (2nd ed.). London, United Kingdom: Routledge & Kegan Paul.

——— (1992). "Quining Qualia." A. J. Marcel & E. Bisiach (Eds.), *Consciousness in Contemporary Science* (pp. 42-77). Oxford, United Kingdom: Oxford University Press.

——— (1993). *Consciousness Explained*. London, United Kingdom: Penguin.

Frankish, K. (2012). "Quining Diet Qualia." *Consciousness and Cognition*, vol. 21, no. 2, pp. 667-676.

——— (2016). "Illusionism as a Theory of Consciousness." *Journal of Consciousness Studies*, vol. 23, nos. 11-12, pp. 11-39.

——— (2019). "The Meta-Problem is The Problem of Consciousness." *Journal of Consciousness Studies*, vol. 26, nos. 9-10, pp. 83-94.

Goff, P. (2018). "Conscious Thought and the Cognitive Fine-Tuning Problem." *The Philosophical Quarterly*, vol. 68, no. 270, pp. 98-122.

---

[37] This may depend on one's whether one finds phenomenal powers appealing from introspection alone. Goff (2020) presents a different account of the power of phenomenal states which separates their intrinsic character from non-phenomenal causal capacities. Here, pain does not entail avoidance, but 'pain+', the combination of phenomenal pain with this non-phenomenal aspect, would entail avoidance.

———— (2020). "Revelation, Consciousness+ and the Phenomenal Powers View." *Topoi*, vol. 39, pp. 1089-1092.

———— (2023). *Why? the Purpose of the Universe* (1st ed.). Oxford, United Kingdom: Oxford University Press.

Mørch, H. H. (2018). "The Evolutionary Argument for Phenomenal Powers." *Philosophical Perspectives*, vol. 31, no. 1, pp. 293-316.

———— (2020). "The Phenomenal Powers View and the Meta-Problem of Consciousness." *Journal of Consciousness Studies*, vol. 27, nos. 5-6, pp. 131-142.

Extending the ExM: Extended Emotional Dispositions

YONGPING DENG
University of St Andrews

1. INTRODUCTION

ACCORDING TO THE EXTENDED MIND THEORY BY CLARK AND CHALMERS, cognitive processes may exist outside of the body.[1] When the right circumstances are present, external entities can be integrated into an individual's cognitive process, becoming part of their extended mind. This approach gives cognition a central role, especially dispositional beliefs, which ignores the role of emotion in the philosophy of mind. While extending emotion seems counterintuitive, Colombetti and Roberts claim that proponents of ExM should also accept that all affective states (dispositional and occurrent) can extend beyond biological boundaries.[2] The former are long-term dispositions to undergo a certain kind of (occurrent) emotions, while the latter are occurrent states with event-like or temporal structure.

In this paper, I will focus on emotional dispositions and the ExM by C&C (1998) to argue against C&R's (2014) view that all affective states can be extended based on the extant ExM framework. Emotional dispositions are long-term tendencies to experience a specific type of (occurrent) emotions. I choose them for two reasons: i) As standing and intentionally directed states, they are most similar to the dispositional beliefs in the ExM; ii) C&R (2014) only use the ExM framework by C&C (1998) to analyse the extending of dispositional states, while their arguments for extending occurrent states are based on the further work by Clark.

I will first introduce the ExM with the famous Otto & Inga case. Then, I would look at the cognitivist view of the affective states, explaining two typical affective states (emotional dispositions and occurrent emotions) and their differences in more detail. Thirdly, I would evaluate the analyses of extending emotional dispositions by C&R (2014). I shall argue that the cases they proposed cannot all successfully show the material underpinnings of emotional dispositions extend beyond biological boundaries. In their analysis of emotional dispositions, there seems to be a gap between the extended affective state and the internal one before being extended, which I call '*translation problem*'. Lastly, I will focus on my proposal: to fully extend emotional dispositions in the sense of functional equivalence between a coupled affective system and an internal affective system, we need an additional assumption in the extant condition that the extended emotional disposition is *qualitatively identical* to the internal one before being extended.

---

[1] Clark & Chalmers, 1998
[2] Colombetti & Roberts, 2014

2. FROM EXTENDED COGNITION TO EXTENDED AFFECTIVITY

2.1 THE EXTENDED MIND THEORY (EXM)

In this section, I would like to clarify the Extended Mind Theory in this essay. There are two targets of the extended mind: i) standing mental states, such as dispositional beliefs; ii) fleeting or occurrent cognitive processes, such as calculating something.[3] This distinguishment is important because the arguments presented in the literature to support extending (i) and (ii) are different. The affective realm includes dispositional and occurrent phenomena, which require different ExM arguments to extend them.

The arguments for extending standing mental states, especially dispositional beliefs, are provided by C&C (1998) with the famous cases of Otto and Inga below. Basically, it is the idea that cognitive processes may exist outside of the body. In more detail, they argue that given the proper two-way interaction, any external entity linked with the human organism can create a coupled system that can be considered as a cognitive system in its own right.

CASES OF OTTO AND INGA

> *Inga, a healthy adult, and Otto, who suffers from mild Alzheimer's, both heard of an interesting exhibition at the Museum of Modern Art (MoMA) and decided to see it. To ensure he does not lose important information, Otto carries a notebook with him no matter where he goes and consults it. Inga quickly remembered that the MoMA is on 53rd Street and headed there, while Otto consulted his notebook to find the MoMA's address before setting out.*

According to C&C, these two cases are entirely analogous: Otto's notebook serves the same purpose as Inga's memory; it stores information which can be accessed when needed and helps to guide Otto's intentional behaviour, just as Inga's memory does for her.[4] The entries in Otto's notebook meet the glue-and-trust criteria: though they may not always be consciously acknowledged, they are constant and can be accessed easily when needed.[5] They are also automatically endorsed.

They can justify their ExM by appealing to the parity principle, which introduces the functionalist credentials. Basically, the upshot is that the function rather than the location of a process matters when determining whether a process is cognitive. In more detail, if a process is considered cognitive when conducted internally (within the body), it should similarly be recognised as cognitive when conducted externally (in the external world). It is worth clarifying the functional equivalence here because their original passage is potentially misleading to think about the similarity between an external and an internal process. The required functional equivalence is not the one between the external process itself and the internal process, although certain expressions by Clark and Chambers might imply that the external element must independently possess functional equivalence to the internal component to be considered a part

---

[3] Wilson & Clark, 2009
[4] Clark & Chalmers, 1998
[5] Clark, 2010

of an extended mind. For instance, 'the entries in Otto's notebook play the same role as Inga's memory'.[6]

The proper interpretation is that, for a process to be considered as a component of the system of processes required for an individual to complete a cognitive task, it should fulfil the appropriate functional role and be causally integrated with other cognitive processes.[7] Thus, what is necessary is functional equivalence between a coupled system composed of inter-plus-outer processes and an entirely internal system. For example, Otto and his notebook work together to be functionally equivalent to Inga's internal memory system. In this way, his dispositional belief that the MoMA is on 53rd Street, or more precisely, his cognitive system implementing this belief, can thus be perceived as extended into the external world.

To end up, I will not respond to the criticisms of this ExM in this essay but based on its framework to discover the possibility of extending emotional dispositions. By introducing a novel application of ExM, this essay might contribute to the ongoing discussion on ExM, to some extent, providing clarity on its fundamental commitments.

2.2 THE AFFECTIVE STATES: COGNITIVIST VIEW

To avoid potential ambiguity brought by the word 'emotion', the affective states can be organised into two types: an occurrent state and a dispositional state.[8] For instance, if we learn that Lucy fears public speaking, we can interpret this information in two ways. On the one hand, we may conclude that Lucy is currently undergoing an episode of fear when faced with public speaking, accompanied by visible signs of anxiety and physiological changes like an increased heart rate. On the other hand, we might say that Lucy possesses a persistent disposition of fear regarding public speaking, leading to experiences and expressions of fear whenever confronted with such situations.

There are more complex categories under each type; for instance, the paper of C&R (2014) includes i) occurrent states: occurrent emotions and moods, and ii) dispositional states: emotional dispositions, sentiments, temperaments, and character traits. As I mentioned in the introduction, I will focus on the emotional dispositions in this essay. They are long-term tendencies to experience a certain type of (occurrent) emotions, which potentially lasts an individual's entire lifetime.[9] For example, Eve's dispositional resentment toward her parents. In contrast, occurrent emotions are occurrent states characterised by an event-like or temporal structure, typically lasting for a brief duration, usually a few seconds or minutes. For instance, Eve is currently undergoing an episode of resentment.

2.3 HYPOTHESIS OF EXTENDED AFFECTIVITY (HEA) & ANALYSES BY C&R

After looking at the ExM and the affective states, we can briefly propose the hypothesis of extended affectivity (HEA): that the physical realisers of *affective states* may exist beyond biological boundaries. According to C&R, HEA can apply to all affective states, including both dispositional and occurrent states.[10] Specifically, emotional dispositions can be said to have a

---

[6] Clark & Chalmers, 1998

[7] Menary, 2010

[8] Goldie, 2000

[9] Scarantino & de Sousa, 2021

[10] Colombetti & Roberts, 2014

physical basis shared across the human organism and the external entity. In more detail, they can be extended in the sense of functional equivalence between a coupled affective system and an entirely internal affective system.

To extend emotional dispositions, C&R try to see if it is possible to *directly* construct an extended emotional disposition case along the lines of Otto and his notebook case.[11] Because emotional dispositions are standing and intentionally directed states, they are most similar to dispositional beliefs in the ExM. The original case they use to extend emotional dispositions is Eve's dispositional resentment toward her parents.

CASE I: EVE AND HER DIARY

> *Eve often records her parents' actions that make her angry in her diary, such as 'my mother does not care about my feelings' and 'my father always criticises me and does not appreciate my achievements'. Even when she is not consciously exhibiting resentment, after reading these entries, she will. When she re-reads these entries, her resentment is reignited, even during periods of her life when her relationship with her parents is not a significant factor. Without the diary, it is possible that she could forget her negative experiences with her parents, fostering more positive memories and feelings towards them, which may help her avoid manifesting parent-directed resentment.*

C&R argues that if we endorse ExM, we should also acknowledge that Eve's diary is part of the coupled system that enables her to experience ongoing resentment towards her parents.[12] They suggest that the entries in her diary satisfy the criteria of glue-and-trust: they are consistently and readily available to her; she can access them effortlessly during her daily activities; and she does not question their contents but trusts them as she reads them. Also, the parity principle applies here: if an entirely internal system of Eve caused herself to experience resentment, we would readily recognise her as being in a state of dispositional resentment towards her parents.

3. EVALUATION OF THE ANALYSES BY C&R

In this section, I will evaluate the cases and arguments they use to extend emotional dispositions, demonstrating that the cases they proposed cannot all successfully show that the material underpinnings of emotional dispositions extend beyond biological boundaries.

Initially, I shall argue that case (i) does not demonstrate extended emotional dispositions but further reflects extended dispositional beliefs. In other words, what Eve put in her diary is merely sentences that record dispositional beliefs like Otto and his notebook, rather than something that genuinely records her resentment towards her parents. Suppose the activation of Eve's occurrent resentment requires her to treat specific externally recorded contents as accurate. In that case, it might be plausible to consider those externally recorded contents only as part of the foundation of dispositional belief. In this case, there is no more evidence to show that Eve indeed extends her resentment towards her parents into her diary.

---

[11] Colombetti & Roberts, 2014
[12] Ibid.

Just like Otto merely finds the sentence about the address in the notebook, it is highly possible that Eve merely extends the dispositional belief such as 'my mother does not care about my feelings'. Therefore, it is possible that an internalist approach to the relevant emotional phenomena is still maintained in this case.

Furthermore, someone might suggest using images rather than written texts to avoid the problem above because images reduce the inclination to *construe* the situation as endorsing a specific external proposition as true. For example, we can think about a new case (ii), in which there is no change in other things as in case (i), but Eve might draw a few images to record her parents' actions that make her angry. However, using images is still unsuccessful in extending emotional dispositions. Although this new case can avoid the problem that merely further reflects extended dispositional beliefs, it still cannot avoid what I call the 'translation problem'. It is the idea that there seems to be a *gap* brought by *translation* between the extended emotional depositions and the internal one before being extended. This problem might make them unable to function equivalently.

In detail, in the extant condition, there is no approach for us to record emotion itself directly. People simply rely on other objects to translate emotion (E), no matter the languages or images, and then extend these expressions that have been translated from emotion (TE). The translation process might lose some information about E; thus, TE is not equal to E, or in my view, TE cannot even count as a kind of emotion. Thus, TE needs to be translated back to E to be understood in the sense of emotion itself rather than those expressions. For example, in this new case, if these images trigger Eve's resentment, what she is first doing is still translating TE back to E. However, the second translation process might again lose information, after which we can get might not be E, but the emotion translated back from TE (ETE). Thus, there might be a gap between the extended emotional disposition and the internal one before being extended.

As I mentioned in section 2.1, for the final functional equivalence, the glue-and-trust criteria should be met by the entries in the external affective realisers. Namely, though they may not always be consciously acknowledged, they are constant and can be accessed easily when needed. They are also automatically endorsed. C&R argues that the entries in Eve's diary/image meet the glue-and-trust criteria.[13] I would say this is because the thing they actually extended here is those expressions of emotional dispositions (TE). However, the proper entries to be extended should be the emotion itself (E). If we merely extend those expressions rather than emotion itself, we can conclude that merely those expressions can be extended. An internalist approach to the relevant emotional phenomena is still being maintained in this case. As E and ETE might not be qualitatively identical, I would argue that the glue-and-trust criteria cannot be met in the extant condition, which means not all emotional dispositions can be successfully extended in the sense of functional equivalence between a coupled affective system and an entirely internal affective system.

However, they might suggest constructing cases in which the emotional disposition is completely subserved by non-representational artefacts, such as the resentment control ball in case (iii) below.

CASE III: EVE AND HER RESENTMENT CONTROL BALL

---

[13] Ibid.

Eve is disposed to be angry at her parents. However, she decides at some point to change this affective outlook. She uses a resentment control ball to regulate her resentment whenever she detects or predicts that her occurrent emotion will be anger at her parents. After a period of adjustment, the use of this resentment control ball enables Eve to undergo fewer occurrences of anger at her parents.

According to someone who thinks these non-representational artefacts can successfully extend the emotional dispositions, if one endorses ExM, one should say that the resentment control ball is part of the supervenience base of Eve's less resentment. Eve's resentment control ball meets the glue-and-trust criteria due to its stable availability, proficiency and uncritical deployment. The parity principle also applies: If the regulation of resentment took place entirely within Eve's body and led to a disposition for reduced resentment, we would readily interpret the process as a demonstration of Eve experiencing less resentment towards her parents.

I would argue that although the use of non-representational artefacts seems to avoid translation problems, C&R suggest it without explicitly noticing the significant translation problem rooted in cases (i) and (ii).[14] The use of non-representational artefacts is aimed to merely avoid the first problem I mentioned above: further reflecting extended dispositional beliefs rather than truly extended emotional dispositions. In their view, the results of using image (ii) and non-representational artefacts (iii) should be the same, which is a reduction of the inclination to construe the situations as endorsing specific external proposition content as true, as there is no candidate proposition. With this intention, they directly cut out the need for individuals to record or understand the emotion, which excludes the translation processes. However, the important point is that the proposition is one of the manifestations of those expressions translated from emotion (TE). In other words, they do not recognise the translation problem that there might be a problematic gap between E and ETE because of the unavoidable use of TE in the extant condition. This problematic gap is the essential point behind the problem that they focus on. Except for this point, case (iii) demonstrates a positive case that can support my proposal in the next section. As in the extant condition, we cannot really have a resentment control ball; thus, there needs to be an additional assumption to play the same role.

4. NEW PROPOSAL

As I argued in section 3, it is better to add an assumption to avoid translation problems in their arguments: the extended emotional disposition (ETE) is *qualitatively identical* to the internal one before being extended (E). I would demonstrate case (ii*) to see an example of this assumption. It differs from case (iii), where an imagined resentment control ball cut out the translation process, as it still includes the translation process but no translation problem with this additional assumption in the extant condition.

CASE II*: EVE AND HER DRAWINGS

Eve records her resentment through detailed drawings. These drawings become external representations (TE) of Eve's internal emotional dispositions (E). Nevertheless, the extending target is not the representations (TE) but her

---

[14] Ibid.

resentment itself (E). Assuming qualitative identity ensures Eve creates drawings that faithfully reflect her resentment, which means, for example, it can ensure Eve chooses colours and forms that perfectly record her internal resentment, and understands it without qualitative information loss, thus bridging the gap between E and ETE. In other words, this additional assumption makes Eve avoid the inherent translation problem brought by the ineffable core of her resentment, successfully extending her resentment.

Someone might argue that the parity principle itself does not require this assumption because it merely requires functional equivalence between a coupled affective system and an entirely internal affective system.[15] In more detail, this view might imply that the external processes themselves may not necessarily possess a phenomenological or embodied character. However, I would argue that it is a necessary assumption that could make the parity principle work in the context of emotional dispositions.

It is true that the parity principle does not explicitly require this to extend dispositional beliefs. However, I want to compare the affective and cognitive processes to see why this additional assumption is genuinely needed in the context of emotional dispositions. The content of cognitive processes does not need that kind of translation to be recorded or understood. In other words, we can directly record and understand the dispositional belief itself; thus, there is no gap between the extended dispositional belief and the internal one before being extended. For example, the sentence that the MoMA is on 53rd Street is constant before and after putting it in the notebook. It automatically satisfies the glue-and-trust criteria without other explicit explanations. The notebook's readiness for reliable deployment by Otto contributes to the coupled system's distinct dispositional profile, which characterises the dispositional belief. Then, the Otto-plus-notebook can be considered to function as equivalent to Inga's internal memory system. Thus, this is why, for cognition, no other requirements need to be met to be extended in the sense of functional equivalence.

However, suppose we want to fully extend all emotional dispositions in the sense of functional equivalence. In that case, we must add an assumption that ETE is qualitatively identical to E to revise the ExM framework to accommodate all emotional dispositions. The parity principle refers to functional equivalence that needs to be achieved between a coupled affective system and an entirely internal affective system. As I mentioned in section 3, this requires the entries in Eve's diary or image to meet the glue-and-trust criteria. These criteria implicitly indicate what in Eve's diary/image is always the one that has been put in the diary/image, which means the E should be the E itself. In an entirely internal affective system, the stored emotional dispositions do not experience the translation process to be recorded or understood, so the E is the E itself. By contrast, when we want to extend emotional dispositions, these criteria may be unable to be met.

It might be essentially because the emotional dispositions appear to contain some ineffable phenomenal core in affectivity that cannot be transferred or replicated,[16] which makes *themselves* cannot be directly and fully recorded or understood. Or worse, all extant mediums might not achieve the accuracy of expressing the remarkable depth and richness of emotional experiences. In other words, these emotional dispositions must experience the translation process, whether from E to TE or from TE back to ETE. However, due to the ineffable

---

[15] Menary, 2010
[16] Slaby, 2014

phenomenal core, the translation process might lose some information about E. Thus, what I call the translation problem occurs- there might be a gap between E and ETE. Although both are emotional dispositions for the same individual to record or understand, the extended emotional disposition (ETE) is *not qualitatively identical* to the internal one before being extended (E). Namely, the entries in external affectivity realisers cannot meet the glue-and-trust criteria in the extant condition. Therefore, we need to add the additional assumption to extend the emotional dispositions, in the sense of functional equivalence between a coupled affective system and an entirely internal affective system.

Lastly, I want to clarify the aim of this proposal by responding to some potential criticisms. Someone might argue that emotional experiences differ significantly between individuals, which makes it challenging to establish a universal standard for qualitative identity. However, this proposal does not aim to provide a single, universal definition of qualitative identity. Instead, what is important is to recognise the translation problem and make sure what is extended is E itself in each specific case. Therefore, with an ideal tool that considers the range of phenomenal and embodied characteristics of emotional dispositions, we can still honour individual variability of emotional experiences.

5. CONCLUSION

To conclude, I explore the possibility of extending emotional dispositions based on the ExM framework by Clark and Chalmers. The ExM successfully extends cognitive processes; however, emotional dispositions, characterized by an ineffable core, present an inherent '*translation problem*'. Specifically, I argue that most cases presented by Colombetti and Roberts cannot successfully demonstrate extended emotional dispositions in the sense of the functional equivalence required by the parity principle, due to this translation problem in the extant condition. Although one of their proposed tools could avoid it, they achieve this result without noticing this underlying translation problem, and this tool is unrealistic.

I propose to add an assumption *in the extant condition*: the extended emotional disposition (ETE) must be *qualitatively identical* to the internal one before being extended (E). This addition is significant to bridging the gap caused by translation, ensuring the glue-and-trust criteria can be met by external affectivity realisers. Then, we can successfully extend the emotional dispositions in the sense of functional equivalence between a coupled affective system and an entirely internal affective system. Additionally, it is important to emphasise that this additional assumption does not attempt to provide a standard definition of qualitative identity. Instead, the essential point is to make us recognise the translation problem and ensure that what is extended is E itself in each specific case.

Summarily, this essay might contribute to the ExM discussion by emphasising the unique problems of extending emotional dispositions and proposing an essential assumption to avoid them. This proposal seeks to refine the existing ExM framework, preserving the distinctive features of emotions, thus leading to a more nuanced understanding of extended emotional dispositions.

REFERENCES

Clark, A. (2010). "Coupling, constitution, and the cognitive kind: A reply to Adams and Aizawa." R. Menary (Ed.), *The Extended Mind* (pp. 81-99). Cambridge, Mass.: MIT Press.

Clark, A., & Chalmers, D. (1998). "The Extended Mind." *Analysis*, vol. 58, no. 1, pp. 7-19.

Colombetti, G., & Roberts, T. (2014). "Extending the extended mind: The case for extended affectivity." *Philosophical Studies*, vol. 172, no. 5, pp. 1243-1263.

Goldie, P. (2000). *Emotions: A Philosophical Exploration*. Oxford, United Kingdom: Oxford University Press.

Menary, R. (2010). *The Extended Mind*. Cambridge, Mass.: The MIT Press.

Scarantino, A., & de Sousa, R. (2021). *Emotion*. Stanford Encyclopedia of Philosophy.

Slaby, J. (2014). "Emotions and the Extended Mind." C. von Scheve & M. Salmela (Eds.), *Collective Emotions* (pp. 32-46). Oxford, United Kingdom: Oxford University Press.

Wilson, R. A., & Clark, A. (2009). "How to Situate Cognition: Letting Nature Take its Course." P. Robbins & M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition* (pp. 55-77). Cambridge, United Kingdom: Cambridge University Press.

Towards a Political Theory of Love and Revolution

GHAYAS OSSEIRAN
McGill University

'There is an extremely powerful force that, so far, science has not found a formal explanation to. It is a force that includes and governs all others, and is even behind any phenomenon operating in the universe and has not yet been identified by us.
This universal force is LOVE.'

Albert Einstein in a letter to his daughter Lieserl Einstein

'With the first link, the chain is forged. The first speech censured, the first thought forbidden, the first freedom denied, chains us all irrevocably.'

Judge Aaron Satie from *Star Trek: The Next Generation*

THE TRAJECTORY OF LOVE

IN HIS *REPUBLIC*, PLATO ENVISIONS HIS IDEAL KALLIPOLIS, a city whose civic and political structure was 'drawn by the painters who use the divine pattern'.[1] The ancient and medievals sought this divine nomoi, or Law, through rational speculation into a natural world organized with reason in Mind, by a God that is Nous and the Good. As Carlos Fraenkel puts it, what is best, 'whether the universe, an individual organism, a political community or a human life' is 'taken to be rationally ordered and hence divinely ordered' given the ancient's conception of God as Reason (Fraenkel 345). In turn, freedom was identified as rational autonomy in view of the rational organizational principles of nature and being. If the nature of God and the organizational principles of the cosmos were to differ, positive freedom would also reorient our motion in view of these new principles. While Kant distances his political philosophy from theocratic traditions and grounds our negative freedom in humanity's shared moral dignity, he maintains that our rational moral duties actualize our will's freedom. Aristotle supposes that Hesiod, was the first to consider love (erôs) to be a first principle of being, as did Parmenides and Empedocles.[2] Aristotle interprets however that such a principle of motion would make the appetites a dominant faculty of the soul.[3] In Harry Frankfurt's *Necessity, Volition and Love* he considers love to prompt human freedom and clarifies that duties of care belong neither to the

---

[1] Plato, 1943, 500e
[2] Aristotle & Reeve, 2016, 984b21-26
[3] Ibid., 984b23

cognitive nor the appetitive faculty for pleasure.[4] Rather, as Ibn Arabi would explore, love is a function of the heart or spirit.

      The aim of this paper is to explore the extent to which Kant's concept of political freedom lends itself to a divine Law of Love. In this first section, I will extract three principles of a divine Law of Love from Empedocles and Ibn Arabi's accounts of God as Love. Namely, incorporating a plurality of relative standpoints when constituting the general outlook of Kant's learned public, preserving our duties of care in the moral socio-political project and accounting for the immanence of love when considering the scope of a philosophy of history. In the second section, we'll consider how Kant's conflicted support for the French Revolution lends itself to the first extracted principle. In providing a post-colonial critique of his concept of political freedom by drawing on the works of Edward Saïd, we find Kant to fall practically short of the second principle. The completion of his concept will be explored through an exposition of Frankfurt's innovations on Kant's concept of freedom. In taking God to be Love as Empedocles, Frankfurt and Ibn Arabi do, the third principle places a relational condition on the nature of selfhood and expands the scope of political freedoms to animals and plant life capable of community and love. In the third section, I focus on Al-Farabi and Ibn Arabi's models of relational conceptions of selfhood to meet the third principle which remains, in part, absent from Kant's *Perpetual Peace*.

§1: GOD'S LOVE

Aristotle and Plato identify two names for God, Good (agathon, tagathon) and Nous. As a reference to the first principles of being, the first causes of the widest set of causal chains, the Good is taken by Plato to be the first and highest divine principle, seconded only by nous as the 'king for us of heaven and earth'[5] and the source of order in the cosmos.[6] Contrastingly, Aristotle considers Good and nous to be names of the same divine principle, holding neither priority nor posteriority from one another. While we can qualify the order of nature as good, the Good itself exists separately as a cause for the effect of the goodness of order.[7] In this respect, the Good for Aristotle is neither an idea or form of the good, nor a virtuous characteristic, but rather a final cause for-the-sake-of-which all good things move, the Good itself.[8] For Aristotle, this Good could materialize in as many ways as being can, whether as 'substance...quality, quantity or time, and also moving and being moved' .[9] Aristotle considers it a matter of consensus that the first principles are composed of contraries; what these contraries are, will differ amongst his counterparts. To some, they're odd and even, hot & cold, being one and being many, being and non-being, limited and unlimited, and some, like Empedocles, take the contraries to be 'love (philia) and strife'.[10] The introduction of Love as a motivating principle behind the order of the cosmos conditions three extracted principles. Firstly, equal and compassionate consideration for the multiplicitous parts of being from the unified perspective of God. Secondly, the engagement of love and strife in our moral and socio-

---

[4] Frankfurt, 2003
[5] Plato & Cornford, 1985, 28c6-8
[6] Menn, 1992, p. 546
[7] Aristotle & Reeve, 2016, 1075a11-15
[8] Ibid., 2016, 1218b7-12
[9] Aristotle & Ross, 1959, 1096a23
[10] Aristotle & Reeve, 2016, 1004b29-35

political spheres as dimensions of the good and bad themselves. Thirdly, the immanence of love as a living force within the cosmos requires a reconception of selfhood as relational.

Frankfurt's account of God as Love demonstrates the first principle. He begins by reasoning that God is the only lover who isn't subject to the 'profoundly distressing anxieties and sorrows' that come with loving a person who may be harmed, lost, or taken away from us in the course of their lives.[11] While the wounds that love may importantly teach us caution in how much or trustingly we love, and how many people we allow in our hearts, God is not subject to these restrictions. For Frankfurt, this uninhibited and inexhaustible love 'desires an altogether unconstructed plenitude in which every possible object of love is included'. Following Aquinas' observations, Frankfurt recognizes God's love of Being, of the living world as opposed to non-being. God's love, he continues, does not discriminate between one beloved in existence from the next given the shared grounds existence looks to God from, and has as its central tenant the 'limitless expansion of the varieties of existence'.[12] This claim, seemingly innocuous, sheds some light on the love and preservation of multiplicity; of cultures, for example, political views, identities, sovereign nations, when striving for unity in a public forum. In contrast, and as we'll explore in the second section, the Platonic efforts at homogenizing the character of his Republic under the guardianship of the ruling class have been used to justify the colonial conquest of cultures and peoples deemed less capable of wielding their freedom responsibly. The love of Being for Frankfurt leaves only one purpose for the Universe, to simply exist, as 'many', in harmony, rather than harmonizing the 'many' by becoming an amorphous 'one'.[13]

Our deontological and utilitarian moral duties are often abandoned in the name of the love of friends and family and the strife of losing them. The second principle that Empedocles engenders is that love and strife are not at odds with morality, or supplanters of it, but unanimous with the good and bad. He considers Love as the reason behind the order of the cosmos, why the good is Good, and the cause of unity between the elements in the universe, while strife is the cause of their division.[14] Alternatively, some conceived of these first principles of the good, orderly, and loving organizational force in the cosmos, as mathematical and unfeeling. Plato followed the Pythagoreans in thinking of numbers as substance. The Pythagoreans and Plato had a mathematical understanding of the One. This numerical reduction of the Good has largely been criticized as the depersonalization and mechanization of the Good.[15] With love being the motivation behind the Good, in civic life, Empedocles takes love and strife to be the moral good and bad themselves.[16] Love fits contraries together in precise ratios, it evens out the cosmos and harmonizes its multiplicitous parts.[17] In addition, Empedocles names the world soul being conscious of itself, the collective organism in which humanity is only a part of in the sublunar world, the Sphairos. This singular life force finds humanity committing itself to unity as one animal and requires love from all of us to try to fit together in harmony. In the Sphairos love helps us approximate Oneness and promotes the synthesis of seemingly contrary fragments of the natural and civic world.[18] As opposed to the

---

[11] Frankfurt, 2003, p. 173
[12] Ibid.
[13] Ibid.
[14] Aristotle & Reeve, 2016, 985a22-29
[15] Ibid., 985b21-986a2
[16] Ibid., 2016, 985a6
[17] Empedocles & Leonard. 1973. p. 22
[18] Ibid., p. 29

Platonic effort at emulating or enforcing an artificial unity in the polis, love as a driving force for unity allows civic societies to harmonize organically and consciously.

The third principle requires us to revisit the nature of God as immanent and selfhood as relational. If Empedocles is correct in taking the nature of the One or First Cause to be Love, the love he considers to be an element or force that brings matter together ought to participate in the Unity of God as Love. Love is the Good, 'a starting point both as a mover (since it brings together) and as matter, since it is part of the mixture'.[19] In this respect, it is both transcendent as a motivating organizational principle of the cosmos, and an immanent force Aristotle denotes as an 'element' that aggregates Empedocle's elements of fire, water, earth, and wind (or matter as such) together and disaggregates them in strife.[20] For Ibn Arabi, the nature of God's love and immanence is much more comprehensive. Let's consider the analogy between Love, the Beloved, and the Lover on one hand, and God's Essence and Light, the transcendent God as they're loved apart from Creation, and the loving Worshiper on the other. Ibn Arabi employs a unique metaphysical and theological notion of God that can be quite revealing, namely the Unity of Being or Existence. Although he never used the term himself, Ibn Arabi's students compiled and named his account of the nature of the God-Cosmos-Human relation the Unity of Existence or wahdat al-wujûd. God for Arabi is both transcendent and immanent (and a third mode that would require too wide a diversion). For Arabi it was not sufficient to claim the unity of God within itself, 'there is no God but God', it would be more proper to supplement the claim with God's Names, that there is 'no Reality but the Real' or 'there is no Love but the Loving'.[21] For Ibn Arabi, there is no existence apart from God, including humans and animals, trees and rivers. A believer, in this respect, holds 'two visions' (nazar) of the Real. The first respects God as the originating Essence that is independent and outside of Creation neither corporeal nor temporal, but incomprehensible and eternal. The second sight reflects and experiences God as the spirit permeating both the matter and immaterial substrates of the cosmos.[22] The first vision of the divine Law deals with the transcendence of God; if you see a book and say it is not God, you thereby qualify God, limit Them, imply a duality between creation and Creator, and thus imply a reality distinct from the singular Reality of God. In the second vision, the immanence of God suggests that the book is God, now we have restricted God to this particular book, we have qualified Them in time and place knowing full well God is the incorporeal, eternal, and necessary Being on which material Creation depends. A healthy faith respects both paradoxical visions, that God is both outside of time and place, and in time and place, all of creation and none of creation, both the Ever-Living and the Sustainer.[23] Now that we've cleared that concept, let's review the analogy. The notion of a Unity of Being allows us to claim that 'there is no Love but the Loving'. From the Real's perspective, there is no real distinction between the Love, the Beloved, and the Lover. 'Seeking to unite' with the Beloved, is a process Ibn Arabi describes as developing one's capacity and authenticity of love, one's akhlaq, or volitional character on the path of the divine Law. At this perspective or degree of clarity, he holds that the Lover's act of loving God, is really the self-disclosure of God's spirit, given that Love is the sustaining force that animates the cosmos and this act of Love. This loving is also an expression of God's Loving Essence, Being-itself, and the mediating force

---

[19] Aristotle & Reeve, 2016, 1071b1-5
[20] Ibid., 985a22-29
[21] Chittick, 1993, p. 300
[22] Ibid., p. 107
[23] Ibid., p106

that holds the cosmos together as Empedocles suggested. The recipient, source, and medium of expression of this Love is the Beloved, as they stand apart from the corporeal world, in-themselves, the 'luminous' Reality. All Love is borrowed from God, received from God, expressed through God, and expressed towards God. As 'Arabi puts it, 'nothing comes to You from anyone; nothing comes from You to anyone; everything comes from You to You, You are everything, and that is all' for 'God is, and nothing is with Him'.[24]

§2: KANT'S AGENTS & SPECTATORS OF HISTORY

Love is not always peaceful. Sometimes love is delivering a harsh truth to a dear friend. Sometimes, love is revolution; a love of freedom in the face of oppression, a love of life in resistance to the onslaught of death. Love is also deposing a regime corrupt to the core with a bloodthirst for power over the life and death of another. Love acts for all our sakes; to liberate the persecuted beloved from tyranny, to free the hearts of the tyrant from the poison of their absolute power. On July 14th, 1789, starved and destitute, a group of French revolutionaries stormed the Bastille Prison as a symbolic catalyst for the deposition of the feudal system, its nobility, and bourgeois elites. With restrained excitement, 19th-century philosophers stood in awe at the unfolding monument of organized agency and sacrifice in the name of their natural right to freedom. Notably, both Hegel and Marx adorned Kant with the title of 'the philosopher of the French Revolution'. Although reluctant at the sight of civil unrest, Kant certainly admired the 'grandeur of soul' in the deposition of the old martial nobility. He considered it a leap in human history towards the highest purpose of our natures, 'a cosmopolitan existence'.[25] Not only was Kant inspired by the revolution, but the revolution was notably inspired by his anti-authoritarian ideals. Theremin once reported to Kant that Sieyès, an influential French political theorist and a member of the Constituent Assembly commissioned to draft the new French constitution, was adamant on introducing Kant's philosophy into the French conscience as a complement to the revolution.[26] In this section, I'll argue that Kant's discussion of revolution offers an avenue for the first principle of a divine Law of Love, incorporating marginalized voices in historical discourse, a direct challenge to the cultural logic of imperialism that Edward Saïd notes as characteristic in 19th and 20th century thought. Although Kant falls short of the second principle of preserving our duties of care by bolstering the spectator of history's impartial view when discussing war and peace, Frankfurt's innovations capture the volitional necessities of love in our moral sphere.

In the 20th century, a significant force in the cultural and intellectual activities of Western 'publics' at once glorified the universal character of Western enlightened thought as well as dismissed non-western thought as infantile, pre-historic, unsophisticated, or in some cases, subhuman. Edward Said collects a great litany of the dominant cultural and intellectual traditions that emerged from 18th and 19th-century French and English imperialism. On the one hand, the conqueror was glorified. In his comments on the British education in India derived from the ideologies of Macaulay and Bentinck, Charles Trevelyan viewed the role of the colonizer as 'Platonic Guardians'. 'In a Platonic sense', the role of the enlightened British settlers was to 'awaken the colonial subjects to a memory of their innate character, corrupted

---

[24] Addas, 2000, p. 79
[25] Kant & Reiss, 1991.
[26] Arendt & Beiner, 1982, p. 44

as it had become...through the feudalistic character of Oriental society'.[27] As Isaiah Berlin notes in his *Two Concepts of Liberty*, the Platonic and paternalistic co-option of citizens' positive freedom grants states the ability to 'ignore the actual wishes' of the empirical selves it reigns over, and coerce, 'bully, oppress, [or] torture' them on behalf of their 'real', innate and rational self.[28] Said poses that there was an implicit mission of the conquest of Africa, the Orient and Asia, to universalize cultural discourses in a globalizing world, according to Western ideals that were self-professedly transcendent of non-western cultures.[29] In his reflections on the colonization of Congo and Egypt Leroy-Beaulieu notes that 'la formation des sociétés humaines, pas plus que la formation des hommes, ne doit être abandonée' to the local communities who would instead rule 'au hasard'. The goal of colonization he considers is to 'mettre une société nouvelle dans les meilleures conditions de prosperité et de progrès'.[30] Cultural representations of British and French identity up till the mid-19th century carried with them the confirmation of European power and superiority understood in the antithesis of non-European barbarism. Basil Davidson's survey of writings of travelers, writers, and commanders visiting Africa until the mid-20th century notes a common character of these writings, 'a single domination attitude' that was convinced in their representation of indigenous populations as 'primeval', remnants of 'humanity as it had been before history began'. Convinced also in their election as vicegerents of God sent to this continent to liberate the local population from themselves.[31] The firm and unmoving belief in the English imperial destiny as a mandate of God's will is echoed in Ruskin's 1870 Slade Lectures at Oxford University. Supported by the rise in ethnographic literature in linguistics, racial theory, and history, Ruskin praises the British man as a race of 'the best northern blood', civilized, superior, and charged with the liberation of the primitive and 'subject races' (108). The superiority of 'British blood', culture, 'a religion of pure mercy', advances in natural science, pre-empted Britain rule, and naturalized the subservience of the colonized whose land is 'fruitful waste ground' without a British master to 'plow and sow for', for 'little pay', to 'behave kindly and righteously for' and to 'bring up their children to love' the British master otherwise suffer the consequence of their hatred.[32]

It becomes clear how the nature of colonialism and the cultural logic of imperialism alienates the crucial voices of non-western knowledge production in international discourse; thereby violating our first extracted principle of a divine Law of Love. The greatest political freedom for Kant is the freedom 'to make public use of one's reason at every point';[33] if violated, it is also the condition for an immanent revolution. In turn, the impartiality of one's thoughts along with what Kant names 'enlarged thought' or enlarged mentality is predicated on the first extracted principle from our divine Law of Love, taking other's potential views into account when investigating a topic. Even if the actual judgment of others is amiss in our empty room, Kant maintains the importance of considering the possible judgments of others 'by putting ourselves in the place of any other man'.[34] The importance of this freedom of public and diversified thought is clarified as the condition of rational speculation and the moral law

---

[27] Said, 2014. p. 109
[28] Berlin, 1966, p. 24
[29] Said, 2014, p. 108
[30] Ibid., p. 107
[31] Ibid., p. 100
[32] Ibid., p. 104
[33] Kant & Beck, 1963, pp. 4-5
[34] Kant & Meredith, 1973, §40

itself when it takes into account Kant's position that reason is not intended 'to isolate itself but to get in community with others' and ought to maintain internal coherence in its rational or moral judgments.[35] For Kant, the frequency and quality of our speculative reason would be impaired if taken out of the cosmopolitan community. So much so that if a person's freedom to communicate their thoughts publicly and idiosyncratically is taken away, Kant considers their 'freedom to think' to have been confiscated ([DA] Kant 131-137) despite noting that thinking is a solitary activity. This freedom of public thought importantly presupposes that citizens are allowed to reach their own conclusions without force of coercion or indoctrination. Kant takes our freedom to be empowered by the rights to free information, thought, speech, and publicizing one's views. He denotes this a liberty of the 'scholar' or 'world-citizen' as opposed to the layperson. While a soldier may have the duty of obeying their superior's command as a layperson or citizen of a nation, the soldier may very well critique the command and 'lay them before the public for judgement' as a scholar who belongs to a larger community, 'a society of world citizens'.[36]

Not only is the 'public' co-opted by an implicit cultural logic of imperialism, but colonialism also impairs the volitional capacities of colonized peoples by introducing this logic as an objective mode of meaning-making. Kant's *sensus communis* is a 'sense common to all' and a mental capacity that situates our private judgments in community. It is representative of the 'collective reason of humanity' which functions as an apriori account of a collection of general or inter-subjective standpoints and is unique to humans given that it depends on communication and sociability.[37] It serves as a counterbalance against our private judgments when being critical of our own thoughts and desires. It also serves as a determinant for the coherence and internal consistency of our internal lives. When racial, cultural or reproductive hierarchies pose as this common sense, this can infringe on the negative freedoms of the colonized for self-determination. In addition, this coercive force in marginalized voice's capacities for meaning-making becomes an infringement of our first principle in our divine Law of Love as a result of its effacement of the relative standpoints that make up the whole of the global community.

As a spectator, Kant held a conflicted view of the French Revolution, yet one that nevertheless lends itself to our first principle of inclusion of marginalized voices in historical discourse. On one hand, he held great admiration for these events; he considered the revolution's efforts to be far 'too important' and 'interwoven with the interest of humanity' to be forgotten.[38] On the other hand, Kant condemned its organizers.[39] From the detached yet sympathetic vantage point of a world spectator, or Weltbetrachter, of the revolution, Kant is at once removed from the emotionally ambient climate of French revolutionaries who have experienced the tyranny of the bourgeoisie firsthand, and are directly engaged, with their lives on the line, in the dismantlement of the old regime and the hope for a new one. These actors have no choice but to be engaged in the unfolding events and to assume the particular perspective of a sufferance that cannot be avoided by receding into one's mind. Even as a disengaged spectator, however, Kant remarked a 'wishful participating that borders closely on enthusiasm' for the revolution. Even if citizens of a polis had 'no lack of welfare to complain

---

[35] Kant & Pluhar, 1996, p. 897
[36] Arendt & Beiner,1982, p. 45
[37] Kant, 2004, p. 30
[38] Kant & Beck, 1963, pp. 143-148
[39] Arendt & Beiner, 1982, p. 45

of', it's the fate of every ruler that denies their people the freedom of self-determination and co-legislating their affairs with the state, to be deposed by revolution according to Kant.[40] Nevertheless, at all other times, and when incompatible with morality, the people's pursuit of 'their rights by revolution' remains unjust.[41] In Kant's Perpetual Peace, he clarifies that violent revolution is at times a necessary injustice against the unlawful order of a community. Its positive constitutional effects can neither be justly overturned to its historic conditions, nor does its failure exempt the rebels from punishment under the old and presiding regime.[42] While the spectator may condemn the agents of revolution by appealing to a principled judgment, Kant also recognizes that the actor is subject to an entirely different principle of action that might necessitate revolution against an unjust state that denies them fundamental rights and conditions of justice. If the judgment of the spectator were to be acted on, Arendt determines Kant would have been a 'criminal'.[43] Kant grounds his maxims of action in revolution when necessary and the advancement of peace rather than war in all cases. Despite the role that wars may play in Kant's 'hidden design of nature', namely tempering humanity for cosmopolitan peace, such speculative judgment may never ground an actor's maxim for action in the advancement of war.[44] Alternatively, while spectators may be removed from the necessity of revolution, this does not supplant revolutionary maxims for action against tyranny.

In *Perpetual Peace*, Kant develops arguments for a philosophy of history marked by an overarching force of progress. In contrast to Mendelssohn who argues that only individuals progress while humanity's moral character stagnates across history,[45] Kant poses that human history has shown itself capable of progress, and hope for better futures has proven instrumental for a 'right-thinking' persons advancing 'the common good'.[46] History is thus motivated by a steady advance of the self-perpetuating Good, while evil remains by nature self-destructive rather than self-advancing as Machiavelli would take it.[47] He attributes this perpetual progress to 'providence' or 'Nature.' This providence is perceptible to world-spectators but not to actors on the stage of history, and in turn the 'public' ought to consist of the general perspective of the spectator. I argue that Kant's characterization of this spirit of history is inappropriately grounded in the view of the disinterested spectator. Such a view disenfranchises marginalized voices who have no choice but to be agents in war, colonial subjugation, or social hierarchies from the 'public' forum, and thus falls short of the second extracted principle which grounds our moral and socio-political systems in the agential and lived experiences of love as the good and strife as a collective moral failing. A public of spectators discludes 'actors' in history who have no choice but to take on the particular view of agents swept up in an 'other' 's history, and marginalized communities from circles of knowledge production through implicit racial, reproductive or religious hierarchies informed by the unreflected cultural logic of imperialism. Such a public also enforces a social imaginary, or sensus communis, that has been co-opted by the historically privileged voices of disinterested actors who enjoy the benefits of their successful Empires.

---

[40] Ibid., p. 47
[41] Kant & Reiss, 1991, p. 184
[42] Kant & Beck, 1963, p. 120
[43] Arendt & Beiner, 1982, p. 54
[44] Ibid., p. 53
[45] Kant & Reiss, 1991, p. 88
[46] Ibid., p. 89
[47] Arendt & Beiner, 1982, p. 51

In contrast to Kant's evacuation of the 'public' from lived and particular experiences of agents of history who must instead assume the impartial and disinterested role of the 'world spectator' it would be a valuable epistemic practice to investigate one's biases when it comes to the accidental properties of members of a public forum, their sex, gender, race. Rather than merely abstracting from them as innocuous to the conclusion of our critical investigation where an enlarged mentality is the goal of deliberation, we could fail to include the perspective of the marginalized groups. Failure to investigate these biases could result in the lack of representation of these groups in public and international forums, lack of sufficient imagination or empathy to include their anticipated views in their absence, or discluding their views due to the private biases that call into question the credibility or relevance of the marginalized group's perspective when taking an account of the views that make up a public realm founded on cultural values deemed superior to those of the marginalized community. What seems to be missing from Kant's view of political freedom, from the west-centric 'public' of disinterested spectators which marginalizes foreign traditions from academic circles, and the enlargement of Kantian mentality to empathetically incorporate the particular views of agents as cognizant of the hidden design of Nature, is an account on the volitional necessities of love and care.

The second extracted principle from our divine Law of Love prompted the consideration, promotion, and facilitation of citizens acting on their duties of care to be of moral and political significance. Conversely, instances of strife, war, and divisions amongst peoples, classes, and nations ought to be combatted on our political stages. In qualifying love as a motivating force in our will's actualization of its freedom, Frankfurt innovates Kant's concept of freedom in line with our second principle. Kant defines the autonomy of the will as 'the property the will has of being a law to itself'.[48] While the will may develop its volitional laws in view of its base desires, wealth or even power, our internal harmony rests on employing our freedom to live in line with the essential natures of our being, and in effect, in personal or civic harmony with the same organizational principles which constituted the natural world we are a part of. Kant identifies two categories of voluntary yet authoritative necessities on our free will, the necessities of ambition and prudence, and the necessities of duties. Kant takes this duty to be respect and obedience to the moral Law. Frankfurt on the other hand investigates the duties of love.[49] For Kant a will that is moved solely by their own personal interest is inevitably heteronomous, which leads Kant to preclude personal interests from 'the essential nature of a person's will'.[50] Such a person would be motivated entirely by circumstances outside themselves and would not be said to be autonomous,[51] nor anything but a passive agent.[52] A fate Frankfurt attributes equally to Kant's conception of the selfless and pure will, as well as agents who deny themselves the necessities of love, or passive agents of love whose care is conditional on the benefits they derive from their beloved.[53] This passive love is as Kant warns, motivated solely by their own personal interests. Where Frankfurt and Kant disagree is in Kant's assumption that a will moved by contingencies, wills necessarily on the basis of what is 'adventitious' rather than what is essential to their own nature.[54] Contrastingly, Frankfurt

---

[48] Frankfurt, 2003, p. 131
[49] Ibid., p. 129
[50] Ibid., p. 131
[51] Ibid., p. 132
[52] Ibid., p. 133
[53] Ibid.
[54] Ibid.

assures Kant that a person's interests can be contingent and nevertheless belong to the essential and individuated nature of their will, namely when the will is moved by the contingencies of unconditional and active love.[55]

An autonomous will for Kant is a 'pure will.' Such a will is devoid of personal interests, 'empirical motives, preferences and desires'.[56] Kant's pure will is an undifferentiated and amorphous identity in Frankfurt's view and impinges on an ideal Frankfurt, and our first extracted principle, hold at high esteem, individuality. Departing from Frankfurt's characterization, a world populated by only pure, formless wills devoid of substance, would look like a dystopian kind of social unity, one where synchronicity and unanimity are gained in exchange for the loss of authentic wildness of a culturally and individually multiplicitous society seeking unity through co-existence and care rather than homogeneity and order. A brutal characterization might I add, that Kant would categorically deny in assuring that a pure selfless will facilitates seeing the 'dear self' everywhere rather than efface selfhood altogether.[57] Frankfurt takes Kant to be preserving the Platonic tradition of generalizing the essential nature of humans, not as merely rational, but also as being reductively subject to 'austerely impersonal dictates of the moral law'.[58] In contrast, Frankfurt proposed that a human's essential nature is unique to every individual's volitional character, and that morality ought to be carried out authentically through duties of care rather than impersonal duties towards an authoritative moral law. A character is constrained not merely by the mandates of the moral law, but the deeply individuating duties of Love we hold as well.

In contrast to Kant, Frankfurt's account of the volitional necessities of love captures an important feature of our first extracted principle: the importance of individuality in preserving the various perspectives making up the whole of society. In describing this shift, Frankfurt notes that the essence of a triangle can be captured by the necessary characteristics of its interior angles adding up to 180 degrees. Plato and Frankfurt's reading of Kant identify a human's essence to be similarly generalizable as a rational and moral being. Frankfurt, on the other hand, takes a human's essential nature to be invariably personal and unique to their volitional character. The essence of a person's will doesn't follow from their genus as a 'human being or as a biological organism', but rather from the characteristics of their will. While a triangle, much like the Platonic forms or mathematical objects holds an essence as an 'a-priori matter of definitional or conceptual necessity' a person on the other hand, is subject to change, corporeality, and transformation (138). In the discussion of different kinds of substances of λ3 in Aristotle's Metaphysics, he clarifies that while Plato takes souls to be individuated irrespective of matter Aristotle firmly stands by the notion that souls are individuated by the body. Aristotle's view is that despite a rational soul being capable of intrinsic individuation, ultimately a soul can't exist outside of matter.[59] Frankfurt stresses that this individuated volitional character is not the collection of impulsive drives for motion or one's desires, but rather, the character one develops in ordering these desires, in the self-directed effort at creating a set of underlying principles, values, and things we care about to differing degrees which individuates the currents of our essential natures.[60] This process, unlike the interior force of our

---

[55] Ibid., p. 135
[56] Ibid., p. 132
[57] Ibid., p. 134
[58] Ibid., p. 132
[59] Aristotle & Reeve, 2016, 1070a4-27
[60] Frankfurt, 2003, p. 113

base desires which Aristotle seeks to overcome with reason, is very much reflective, intentional, and guided just as much by the object of one's care as it is by reason.[61] It's fruitful to note that while reason is dethroned as the supreme authority in our internal lives this does not however mean that Frankfurt seeks to subordinate reason to affect. Rather he places rational speculation on the same stage as one's impulses, desires, moral and loving duties when ordering them in accordance with our mandates of care, just as Kant orders them in obedience to moral duty.

§3: LOVE OF GOD

Our understanding of freedom developed significantly when we considered God to be Love in addition to Good and nous. A second way divine Love may change our understanding of freedom is if it gives us a reason to revisit our conception of selfhood. In this section, I explore the third extracted principle of a divine Law of Love, namely its requirement of an immanent conception of God and a relational view of the self rather than an individualistic one. If this turns out to be the case freedom would have to include the organization and personalization of our volitional character to serve not the self heteronomously as Kant might object to Frankfurt, or the community selflessly as Frankfurt takes issue with Kant, but rather the self as an indivisible part of the community and the community as a constitutive extension of the self. I will consider two models for relational selfhood. The first is Al-Farabi's model of a transcendent God and a weak relational selfhood evident in his *Principles of the Views of the Citizens of the Virtuous City*. For the medieval Islamic philosopher every human being, including the ruler, is incomplete without the community that sustains them, the artisans that weave the person's clothes, the merchant that supplies them with produce they had no time to grow on their own, the soldiers that keep them safe while they slumber and the rulers that order the city according to a common set of principles codified into a divine Law.[62] I consider his view weakly relational because although he argues that human life is invariably interdependent and co-sustained and that human beings share the same beginnings and ends for the sake of which we move, he does not require an ontological relationality of selfhood. Alternatively, Ibn Arabi's model of the spiritual ecosystem living beings are a part of in their participation in God's self-disclosure in creation, requires a strong relational view; whereby humans are not only socially entwined and share a common destiny, but are also co-constitutive given that they share a common life force.

Love for Frankfurt creates a unique symbiosis, or entanglement between the beloved and the lover. Betraying our duties of love towards our beloved is also a betrayal of one's self, one's volitional character shaped by their care for the beloved. This self-betrayal certainly entails for Frankfurt, 'a rupture in his inner cohesion or unity', a rift in their will. Reminiscent of Al-Farabi's condition for Happiness to have unity within one's self, betraying one's love fractures our volitional unity in what would amount to psychic distress, and would mobilize, Frankfurt adds, the betraying person towards 'self-preservation'.[63] If we extend the notion of selfhood in relationality with the beloved as well, this would mobilize the lover towards preserving the sanctity of the unity of the beloved's volitional character by making amends or avoiding harming them. Frankfurt argues that love is authoritative most irreducibly as an

---

[61] Frankfurt, 2003, p. 114
[62] Fārābī & Walzer, 1985. p. 229
[63] Frankfurt, 2003, p. 139

outcome of the instinctive need to 'protect the unity of the self'.[64] A unity of the self that is threatened by self-betrayal, and in turn our basic need for self-respect. Love then makes it so that harming the beloved is harm to the self, and I add that harm to the self certainly causes psychic or even existential distress in the lover. The question then remains, whether love expands the scope of selfhood to also ontologically include the beloved or our communities of love. This concept of selfhood is not solely relational as Aristotle and Al-Farabi take it, that the self is better off living in a community and dependent on sociability to exercise their virtues or to live a Good life. Rather it invites a rather modern view of selfhood, whereby the beloved, or the community, constitutes a part of the indivisible seed of the self.

Frankfurt proposed a curious characteristic of the nature of self and divine love. The self's love for one's self and God's love for a self, cannot by nature 'require the lover to jeopardize or sacrifice his own true interests'.[65] By true interests, I take Frankfurt to mean the interests of the essential nature of their wills. What if one's self-interests are at odds with another's, how then can God preserve the conflicting and opposing interests of this pair? There are two ways in which the essential natures of their wills can be accorded by God's love. The first is offered by Al-Farabi, who suggests that the true purpose of the individual can only be understood in the context of the purpose of the world they inhabit. God need only serve the true purpose of the world with Love, in order to capture the true interests of two living beings who are living in accord with their true natures. To the same end, the second solution is found in Ibn Arabi's concept of God the immanent spirit of history which offers a relational view of the Self, in which an individual spirit or rûh is not only indistinguishable from the spirit, which is housed in the friend we conflict with, but from the collective Spirit of human history as well.

Consider the following passage from Al-Farabi's work on Aristotle's philosophy:

> '[I]f the human being is a part of the world, and if we wish to know his end, activity, advantage and rank, first we have to know the end of the whole world so that it will become clear to us what the end of the human being is and that the human being is a necessary part of the world because through his end the ultimate end of the world is attained. If, therefore, we wish to know the thing for which we must strive. We need to know the end of the human being and the human perfection for which we must strive. For this reason, we must know the end of the world in its entirety; and we cannot know this without knowing all the parts of the world and their principles: what it is, how it is, from what it is, and for what it is. And this [we need to know] concerning the whole world and concerning each of the parts of which it is composed.'

Al-Farabi makes synonymous knowledge of the ends towards which a human being must strive with knowledge of the end of 'the whole world'. Presumably, this is a conclusion derived from Aristotle's argument for the necessity of a First Cause of being due to the incoherence of an infinite regress of material, formal, efficient, and final causes.[66] From this view, we find that all participants in Creation owe the elements that constitute their bodies, the intelligent design that these bodies were formed in, and the motion of their lives in time and in this corner of the cosmos, to the same Mover. All of existence originated from the same Cause and is moving towards the same Final Cause. Now, Fraenkel notes that understanding this universal dependence, and 'grasping the world's purpose' through consistent education in

---

[64] Ibid.

[65] Ibid., p. 168

[66] Aristotle & Reeve, 2016

metaphysics and the natural sciences is a condition for Al-Farabi's concept of freedom. In order to be self-directed and a participant in the sovereignty of the Virtuous City, an inhabitant must be familiar with the principles that shape the world and for the sake of which the world moves, in order to situate themselves in and act in accordance with the perfection of the world as a whole.[67] In the previous section, we found Aristotle considers a few contenders for this end, the Good, nous, and Love as he suggests in his critique of Empedocles and Anaxagoras' accounts. Al-Farabi would attempt to resolve Frankfurt's contentious proposition by arguing that if the two friends were knowledgeable of the purpose they share in the ends for-the-sake-of-which the natural world moves, their true interests wouldn't conflict. Right knowledge of their essential natures then, brings these two friends into accord.

Alternatively, Ibn Arabi proposes that the two friends not only share in the first principles that motivated their existence, that organize nature, or the final ends they move towards, but also in the immanent and loving Spirit they house. For Ibn Arabi, God is the Spirit of the cosmos, which extends beyond its borders, and acts within the cosmos as the life force that animates the bodies (jism) of living beings and gives form to the matter of the cosmos which 'Arabi takes to be God's body.[68] Ibn Arabi considers Adam and his children to have been made as 'a miniature of the world', their bodies contain an imprint or the form of the cosmos, a collection of all universals and forms, a familiarity at least, with the forms of 'rivers, mountains, and stars' and other existents. The spirit of human beings contains the form of the divine and gathers together the 'truths of the divine Names' and in the ideal humans are actualized and balanced with one another.[69] Explicitly, the 99 names of God are contenders for the qualities gathered in the fitra, or the original and universally shared human disposition. In the context of our discussion, we can consider 'the Good', 'Nous', and 'Love' to be Names as well. The spirit differs from the soul and body in that the spirit as the 'breath of God' already embodies the virtues of the fitra, while the soul has to contend with the desires of the body and actively orient its life towards these Names or principles.[70] The rûh or spirit of a living being grants humans 'knowledge and awareness, luminosity, desire, power, speech, generosity, justice, and so on'.[71] All our internal faculties in addition, our senses, 'imagination, memory', intellect ('aql) are all modes of the same internal spirit.[72] The goal of the divine Law's path is to actualize these potential forms in experience and action in view of the 'governing spirit' in the internal world (al-batin) and in the external world (al-zahir), rather than the matter, or base things as Al-Farabi denotes them, of bodies. The true essence of our conflicting friends is thus but the embodied self-disclosures of the same Spirit of God. In this regard, while my interests may be at odds with another's, our true essences will be in accord with the natural design of human history. Both selves from the Spirit's perspective are the same Self, and their true interests will always be preserved in line with the totality of the Spirit rather than its individuated, locally and temporally situated, self-disclosure.

Al-Farabi attributes the relationality of the conflicting friends to the rational apprehension of the Principles of Being, and thus to be self-directed is a feature reserved to human rational souls. Animal souls and vegetative souls in contrast to human ones, don't share

---

[67] Carlos, 2010, p. 353
[68] Chittick, 1993, p. 92
[69] Hakim, 1993, p. 281
[70] Chittick, 2004, p. 99
[71] Ibid., 93
[72] Ibid., 97

in God's quality as nous, and are thus precluded from any meaningful discussion of freedom, cognition of the unity of being, or the capacity to voluntarily move in view of the divine Law. Curiously enough, Ibn Arabi considers plants and inanimate objects to be better equipped to receive the self-disclosures of the divine Law despite their inability to communicate it.[73] This living fabric which encompasses all of creation derives this life from 'the divine self-disclosure' (al-tajallî al-ilahî) of God's Spirit in all existing things. These self-disclosures of God in every corner of creation are witnessed passively by all existing things except for 'angels, mankind and the jinn' who know God only by what is taught to them (Quran 18:64). Contrastingly, inanimate objects and 'plants' don't need to acquire this knowledge, these existents experience the self-disclosures of God 'from behind the veil of the unseen'.[74] While it is common for Muslim thinkers to believe all beings come from God, and to God they must return, Ibn Arabi's concept of Unity of Existence affords him a more intimate vision of affective unity that all rational or irrational participants in the cosmos can experience on their way from the First to the Final Cause. The sensuous experience of this unity of Spirit, becomes not only the condition for self-direction and ideal political legitimacy, but it also generates the belief that the fates of all living and existing beings, rather than just humans, are indivisibly entangled.

It is troublesome enough to conceive of the self's ontological relationality with their loved one, to be not only constituted by our duties of love towards them, but invariably composed of the same spirit. What perhaps requires a more thorough and precise investigation is the relationality humans hold with animal and vegetative souls in our World Soul, or Sphyros, natural providence for Kant, or even in the Spirit of History of Ibn Arabi's 73d chapter of his Futûhat Makkiyā. Love of an immanent Spirit is certainly cause to reconsider how separable and isolated our love is from the common Breath all living beings under the Sun share. Frankfurt carries forward Descarte's tradition of human exceptionalism in the necessities of care on the will. He remarks that while 'subhuman animal species' are subject to desires and preferences, it doesn't entail a capacity for Happiness outside cycles of 'the pleasures of gratification and to the pains of frustration'.[75] Rather, he submits that non-human animals lack the 'psychic complexity..to care about anything'.[76] Traditionally non-human animals are regarded as inescapably propelled by the impulsive drives of their appetites, taken astorm by their passions, and incapable of both reason and liberation from causal determinacy for the ancients. If care truly is neither solely a cognitive attraction to benefit nor an affective attachment to pleasure, one need to look no further than any non-human Mother nursing her cubs and protecting them from harm to attribute them the capacity for care and love that doesn't simply rise from the biological need for self-preservation. Whether the enduring desire to care for their younglings is an exercise of their 'own volitional activity rather than...[desire's] own inherent momentum' requires further investigation in animal behaviour. While this discussion might seem tangential it serves a very pertinent role in a more complete discussion of freedom. If reason nor rational morality are the sole avenues towards which a living being may express their freedom in a world organized solely by nous and the Good, it would be valuable to ask which living beings this expanded view of organizational principles of nature includes. Love

---

[73] Chittick, 1993, p. 100
[74] Ibid.
[75] Frankfurt, 2003, p. 157
[76] Ibid., p. 158

can liberate a will and bring them in harmony with the natural world, in turn, all living beings capable of love ought to also be included in our discussions of political freedoms and rights secured by the state. The humanism of the 18th century championed by 20th-century idealists might have resulted in the establishment of universal and inalienable human rights that follow from one's basic human dignity. An expansion of the conditions necessary to liberate oneself to their dignity would also find us including communities of non-human animals and the natural ecosystem among the right bearers of our community under the Sun.

REFERENCES

Addas, C. (200). *Ibn 'arabī, the Voyage of No Return*. Cambridge, United Kingdom: Islamic Texts Society.

Arendt, H., & Beiner. R. (1992). *Lectures on Kant's Political Philosophy*. Chicago, USA: University of Chicago Press.

Aristotle, & Ackrill, J. L. (1987). *A New Aristotle Reader*. Oxford, United Kingdom: Clarendon Press.

Aristotle, & Reeve, C. D. C. (2016). *Metaphysics*. Indianapolis, USA: Hackett Publishing Company.

Aristotle, & Ross, W. D. (1959). *The Nichomachean Ethics*. Oxford, United Kingdom Oxford University Press.

Berlin, I. (1966). *Two Concepts of Liberty; an Inaugural Lecture Delivered before the University of Oxford on 31 October 1958*. Oxford, United Kingdom: Clarendon Press.

Chittick, W. C. (1993). "Two Chapters from the Futûhât Al-Makkiyya." S. Hirteinstein (Ed.), *Muhyiddin Ibn 'Arabi: A Commemorative Volume* (pp. 264-290). Elements Books.

——— (2004). *Imaginal Worlds: Ibn al-'Arabī and the Problem of Religious Diversity*. Suhail Academy, 2004.

Empedocles, & Leonard. W. E. (1973). *The Fragments of Empedocles*. Open Court.

Fārābī, and Walzer. R. (1985). *Al-Farabi on the Perfect State: Abū Naṣr al-Fārābī's Mabādi' Ārā' Ahl al-Madīna al-Fāḍila: A Revised Text with Introduction, Translation, and Commentary*. Oxford, United Kingdom: Clarendon Press.

Fraenkel, C. (2010). "Theocracy and Autonomy in Medieval Islamic and Jewish Philosophy." *Political Theory*, vol. 38, no. 3, pp. 340-366.

Frankfurt, H. (2003). *Necessity, Volition, and Love*. Cambridge, United Kingdom: Cambridge University Press.

Hakim, S., & Tiernan. M. (1993). "Knowledge of God in Ibn 'Arabi." S. Hirteinstein (Ed.), *Muhyiddin Ibn 'Arabi: A Commemorative Volume* (pp. 264-290). Elements Books.

Kant, I., & Meredith. J. C. (1973). *The Critique of Judgement*. Oxford, United Kingdom: Clarendon Press.

Kant, I., & Pluhar, W. S. (1996). *Critique of Pure Reason*. Hackett Pub.

Kant, I., and Beck, L. W. (1963). *On History*. Bobbs-Merrill.

Kant, I. (1991). Political Writings. Cambridge, United Kingdom: Cambridge University Press.

——— (2004). *Introduction to the Metaphysics of Morals*. Kessinger Pub.

Menn, S. (1992). "Aristotle and Plato on God as Nous and as the Good." *The Review of Metaphysics*, vol. 43, no. 3, pp. 543-573.

Plato. (1943). *Plato's The Republic*. New York: Books, Inc.

Plato, & Cornford, F. M. (1985). *Plato's Timaeus*. London, United Kingdom: Macmillan.

Said, E. W. (2014). *Culture and Imperialism*. Vintage Digital.

Arresting Images: A Brief Look Into Evaluating Protest Art

LARA THAIN
University of St Andrews

INTRODUCTION

AESTHETICS IS A BROAD DISCIPLINE that delves into philosophical questions relating to art and its meaning. One specific area of art that has failed to garner the attention it deserves is protest art. Protest art is uniquely positioned as it can enact real-world change by communicating messages, building community, and subverting the status quo. This functionality of protest art sets it apart from most categories of art and is why it is overdue to be examined on its own. This article will first focus on categorizing what protest art *is* and then on how to evaluate it. I will do this by setting criteria for protest art and examining specific examples of successful protest art; I will focus on visual protest art such as murals, posters, and paintings. I will show that any piece of protest art can be broadly categorized as 'art', however, just because it is a successful piece of 'art' does not mean the piece is a successful piece of protest art. There are different categories of protest art and sub-categories within those categories and they should be judged according to the correct category when it comes to aesthetics. For example, protest art encapsulates performance, songs, visual art, and movies which are different categories of protest art and should be judged respectively. This article will examine the value of protest art and demonstrate why it is important for a successful protest movement due to its potential to effectively communicate and engage audiences. I will also respond to the objection that protest art is not simply art being used to protest. Protest art is overdue to be examined on its own due to its importance in broader protest movements and this article will sketch foundational ideas in an area that has not received as much attention in aesthetics.

WHAT IS PROTEST ART?

WHY CONTEXTUALISM:

When judging art, one must judge it in categories. One should not compare a cubist painting against an impressionist painting to see if it is successful because the category of cubism is essential in understanding the nature of the painting itself. This essay uses Kendall Walton's *Categories of Art* as foundational, creating a framework on how to judge protest art based on Walton's claim that:

If a work's aesthetic properties are those that are to be found in it when it is perceived correctly, and the correct way to perceive it is determined partly by historical facts about the artist's intention and/or his society, no examination of the work itself, however thorough, will by itself reveal those properties.[1]

---

[1] Walton, 1970, p. 363

In his essay, Walton also gives us a framework for judging art in 'categories'. The criteria for what makes a successful piece of protest art is different from what makes a successful artwork in other categories, and thus they cannot be compared. Straightforwardly, just as 'art' is an umbrella for many different categories of art, 'protest art' is also an umbrella term. Many different mediums and styles of art can fall into the category of protest art such as films, posters, songs, etc. The variety within protest art is what causes it to become hard to define in certain moments. Protest art is a larger category of art than many art movements and is not exclusive like other forms of art. For example, a piece of pop art could also be protest art, but a piece of pop art could not also be impressionist.

Two theories that often juxtapose each other in the philosophy of art are formalism and contextualism; Formalists believe that you should be able to judge the value of art based on the piece's directly perceptible properties alone while contextualists believe that historical background and knowledge of the piece is necessary to be able to judge it correctly.[2] If you were to judge pieces from a formalist perspective, you would judge them using the wrong criteria. It is vital to use a contextualist view to judge protest art because the aesthetic value goes beyond that of just the actual artwork itself, so it is necessary to judge art in categories. Carlson writes in *Appreciation and the Natural Environment* that, 'Works of different particular types have different kinds of boundaries, have different foci of aesthetic significance, and perhaps most important demand different acts of aspection. In knowing the type we know what and how to appreciate'.[3] A contextualist view allows us to categorize art correctly and thus be able to appreciate it appropriately. All art must be appreciated in its context, this is why how it is curated, and talked about and the history behind it all contribute to the audience's experience of the piece. Protest art is not made for museums, meaning the curation of it is especially vital. Protest artists have to carefully consider how they display their art so that it is shown to the public. For example, Diego Rivera's piece discussed later was effective in part due to the fact it was displayed in an institution that symbolized the epitome of capitalism, which he was protesting. A contextualist view considers this as well as the history and intentions of the artist.

One reason that a contextualist view is not only beneficial but also necessary for protest art is due to the fact non-aesthetic variables influence a viewer's experience of protest art. For example, seeing a Greek statue displayed in Greece allows people to connect it with the history associated with the piece in a more influential manner than if they saw the statue displayed in a different country- yet the form of the statue remains the same no matter what country it is displayed. This is where a formalist perspective falls short as it does not consider how broader non-aesthetic aspects influence the viewer's experience of the piece. We must also consider how non-aesthetic aspects affect art from a moral standpoint. The history of an artwork matters beyond just aesthetics because it also affects the rights of the art itself. This is important for protest art because it is specifically made to be displayed to the public. A contextualist view protects protest art from being used in ways it wasn't intended and displayed correctly. Most categories of art remain in that category of art even if no one were to ever view the piece. For example, if I were to sketch a flower in a notebook and then put it in a cupboard never to be seen, that would not stop the sketch from ceasing to be a sketch. This is where protest art is unique; if you did not display the piece, it would cease to be protest art. Protest art is reliant on the audience to be categorized as protest art.

---

[2] Dickie, 2001
[3] Carlson, 1979, p. 267

Another reason that a contextualist view is necessary is due to the fact protest art is inherently functional. There is the view that to be art a piece must inherently be 'useless'. Peter Lamarque writes in *The Uselessness of Art,* 'But in the case of art, 'uselessness' is used to refer to a kind of value' and argues that valuing art for their usefulness, 'seems to distort altogether their true worth as art'.[4] Art can be useless, but it does not *need* to be it. For protest art, it has to have a function to be protest art and the fact that it is useful does not negate its value as art itself but the opposite. The usefulness of protest art allows people to aesthetically experience it to a fuller extent as they are engaging not only with just a single artwork but also with the movement that the art is for. This dissertation will use the meaning of movement generally as people in a specific time-period working towards a societal change. The usefulness of protest art causes it to be a piece in a larger movement and it can never be taken out of context, this does not negate its value as art itself but instead simply changes the audience's aesthetic experience.

COMMUNICATES A MESSAGE

Protest art is made with a message in mind against an existing institution; this paper uses institution broadly to mean physical or social organizations or movements. Successful protest art can communicate its message in a succinct and easily comprehensible way, but at its core, the function of protest art is communicating against an institution. If a piece is not made with this function in mind, it is not truly protest art. Other categories of art also have the purpose of communicating a message, however, they are distinctive as they are promoting or supporting an institution. This can be seen in propaganda and religious art. However, another distinction between religious and political art and protest art is that protest art *must* be displayed publicly to be categorized as such while the others do not have to be.

A contextualist view stops artwork from 'accidentally' becoming protest art. If the artist did not intend for their piece to serve the function of protesting an institution, then it cannot truly be protest art as protest art is directly connected to a protest movement. A protest is organized intentionally to communicate a message and thus the art that supports a protest is made with a thought-out message, so we must judge it accordingly. The intentions of the artist are essential when categorizing it due to protest art's function of communicating; We cannot judge a piece of protest art without knowing the message that the piece is trying to communicate. Art can function as a speech act for the artist, so just as the intentions of an orator would be examined, the intentions of an artist should be examined when they create a piece of protest art. This is important for protest art as the function of protest art is communicating a message, so it is explicitly a speech act. Daisy Dixon explores the idea that art can not only be a speech act but also be used as counter-speech that potentially breaks down a piece metaphysically saying that 'an artwork can illocute in a *derivative* sense'.[5] If art can illocute then the intentions of the artist are critical in categorizing protest art as protest art and not in another category. The intentions of the artist while making protest art also help us differentiate protest art and artists simply protesting. This also means knowledge about the institution that the artist is protesting is necessary as well. A successful piece of protest art would communicate

---

[4] Lamarque, 2010, p. 206
[5] Dixon, 2022, p. 412

its message without insight from the artist, but that does not disqualify the piece as protest art itself.

DISPLAY MATTERS:

Communicating a message involves both the communicator and the audience. Therefore, for a successful piece of protest art, the artist will have to consider who exactly will see their piece and how it will affect them. All protest art has the intention of communicating a message, but not every piece is successful in communicating. However, all protest art must be displayed in an intentional way that is available for the public to view. If it is not displayed in a public manner, this means a message cannot be communicated to an audience and it cannot be categorized as protest art. This is because the art will have failed to fulfill function and it could be categorized as another type of art, but it cannot be protest art. This means that how protest art is displayed matters. A piece of protest art must be displayed in a public manner: a poster on a lamppost, a sign at a protest, or a billboard. However, this leaves museums in a difficult situation with regard to protest art. On one hand, it gives validation and larger recognition to a piece and thus a larger movement, but also many museums uphold some of the institutions that protest art is trying to change. Consequently, art that is made specifically for museums and galleries is not truly protest art. There are often barriers such as admission costs and locations that cause art in museums to not truly be 'public'.

Some protest art may be hung in museums, but this is after it has served its function. For example, many pieces of protest art have historical and cultural significance and remain in museums. These pieces of protest art are historical and dormant pieces and thus are in their own sub-category of protest art. Historical pieces of protest art were active pieces at some point and thus they can still be considered 'protest art'. Historical factors influence an audience's perception of categories of art. This is what causes historical protest art to remain protest art, despite one of the primary criteria of protest art being the ability to communicate, and historical pieces no longer having this ability. People still perceive these pieces as protest art, despite losing their function; they can comprehend their historical significance and thus these pieces move into the sub-category of historical protest art instead of simply not being protest art. This means that essentially, once a piece is protest art, it will always be protest art, however unlike other categories of art, the piece will move into the sub-category of historical protest art once the piece loses its function. One of the challenges that protest art faces is its timeliness. Artists need to be able to make protest art that is relevant to current issues which is increasingly difficult in our modern society as technology has allowed information to change at rapid rates. Protest art is time-specific, just as some art must be displayed in a certain manner, protest art needs to be made in a certain time frame for it to be relevant as protest art.

One may argue that a piece of protest art will never lose its function, and perhaps that is a possibility with some pieces, however, this depends on the institution that is being protested. For example, the Suffragist movement created much protest art supporting votes for women. Now, women have the right to vote in the UK, so these pieces no longer need to have the purpose of communicating this message. Unlike many other categories of art, how protest art is categorized is highly reliant on the political atmosphere at the time. All pieces of art are products of the time and place that they are made in, but the atmosphere in which protest art is made is highly influential to the piece as it directly speaks to a specific climate or institution. This is why a contextualist view must be taken with protest art. Protest art is timely in the sense

it deals with human issues and thus is ephemeral. The functionality of the piece is dependent on the time-period it is in; analogously, perfume loses its smell after some time and once it loses its smell that does not stop it from being perfume itself but it has lost the aspect that inherently makes it perfume- the functionality is essential for its nature. This is the same with protest art, once it is no longer relevant to the current time, it only is protest art because it once was and is now only useful from a historical perspective.

There could be some very specific scenarios in which a piece of protest art goes from being a historical piece of protest art that was used in museums to one that is active again. For example, if a regressive policy were to happen or if an issue that was once resolved became prevalent again, then pieces of protest art that had become historical could potentially become active again. For example, in the 1920s-1930s the N.A.A.C.P. would hang a banner that read 'A Man was Lynched Yesterday' which would become a recognizable piece in American history. In 2015, Dread Scott created his own version of the banner with the same infamous typesetting which read, 'A Man was Lynched by Police Yesterday'.[6] By making the two banners indistinguishable at first glance, Scott makes a statement by saying that the first banner isn't a piece of historical protest art despite it being treated like so. Ideally, once a piece of protest art is historical it would remain in that sub-category but it is not necessary that it does.

NOT MADE FOR PROFIT

Unlike other forms of art, protest art cannot be made for profit. This is because the act of creating a piece of protest art is, in itself, an act of protest. If an artist were to create a piece of protest art for profit, the act of creating the artwork would no longer be a protest as the main intention is for profit and not protest. When taking a contextualist view of art, intentions are important and the extrinsic motivation of money precludes protest art because it dilutes the internal motivations of the artist. For example, if someone were to become a parent only because they were receiving the money to do it, you would not consider that person the child's mother or father and instead just their caregiver; the motivation of money devalues the very nature of the thing and the same can be said with protest art. Money is supposed to act as an indicator of how people value art- with more expensive art as highly valued. Unfortunately, it is not as simple as this; some artists such as Jeff Koons and Damien Hirst have made large profits despite intense criticism, so we can see that profit is not an accurate measurement of aesthetic value. Therefore, money should not be taken into consideration when evaluating art generally because there is no correlation between aesthetic value and profit, but especially not with protest art as it indicates to the audience that the motivations of the artist are not truly for protest.

This does not mean there are no borderline cases, which are up for interpretation. If a piece of protest art was commissioned by someone else and used as a protest, is it protest art? This would depend on the payment of the piece and the intentions of the artist. If the artist supported the cause that they were commissioned to make for, this would be reflected in the aesthetics of the piece, and if the artist completely disagreed with the cause they most likely wouldn't accept the commission. Being paid for the supplies and time they spend on the piece is not inherently problematic in itself; it is when the artist is making a piece for profit first and foremost and they do not support the cause, that the piece is not truly protest art because the

---

[6] Rogers, 2016

intentions of the artist are not there. Instead, the piece would be something along the lines of a menu- informative, made with intentions, and maybe even pretty, but not art itself. In the case where an artist somehow got tricked into making a piece or a piece was used without permission, then these artworks would not be protest art because the intentions of the artist are not truly to make protest art.

<div align="center">HOW TO JUDGE PROTEST ART?</div>

In the first section of my dissertation, I have outlined criteria for what categorizes art as protest art. Using Walton's categories of art, I show that a contextualist view is necessary to correctly categorize and judge protest art. Three main criteria are delved into as necessary for protest art including communicating a message, displaying the piece publicly, and not being made for profit. Just because something qualifies as protest art does not make it good art or even good protest art, which means different criteria for judging what is *good* protest art are needed. The main qualifications that I will use to judge good protest art are: how well it communicates its message, how aesthetically pleasing it is, and how it subverts the status quo. Another important qualification for 'good' protest art is how effective it is as a piece of protest; however, I will argue that how effective a piece of protest art is directly correlated with how well it fulfills the aforementioned criteria e.g. a piece of protest art that is effective is also aesthetically pleasing, subverting and has a clear message. If the piece fulfills the first three criteria, it should subsequently fulfill the fourth. This section of the essay will dive into the specifics of the three criteria and then examine examples of protest art that meet them and thus also fulfill the fourth.

COMMUNICATES MESSAGE WELL

Protest art is functional as it intends to communicate a message with the audience, however, just because the intentions are there does not mean the message is communicated effectively. A piece of 'good' protest art not only has an intended message but also has a message that the audience can understand easily and has an impact on them. For example, if the piece has a message but a too deep understanding of a niche concept is required to understand it so only a handful of people understand the piece, it is not as effective as a piece of protest art, though it could still be judged as protest art just not as 'good' as other pieces. For example, jokes often require cultural and language knowledge that does not translate and due to the nature of a joke, if it needs to be explained, it no longer remains funny. The same can be said with protest art, if too thorough of an explanation is needed then it loses some of its value. However, like all things, there is a nuance. For example, there could be specific pieces of protest art where the artist intends for the message to be slightly opaque or understood only by a specific group of people. This could be the case with political protest art where the artist hopes to avoid retribution by the government but still wants to communicate the message to an audience. One could argue that if this were the case, it would not be protest art, but this is where the distinction between protest art and protest by artist matters-which I delve into later.

Protest art also must be judged in its own criteria separate from other types of art by communicating a message, protest art often employs slogans and writing. In many categories of art, the use of writing on the actual artwork would be detrimental to the aesthetic experience of the piece, however, many times successful protest art uses words to improve its

understandability. This is normally used as a way to clarify the intentions of the images and allow exact information to be spread. For example, one of the most effective types of visual protest art is posters or billboards. The criteria for what makes a visually appealing poster is quite different from other pieces of visual art because of the functional purpose of posters. Slogans and writing help to clarify meanings and thus can often make pieces better at communicating and thus better pieces of protest art. The slogans and writing on the piece also cannot be meaningless- more effective protest posters have slogans that can be remembered easily and stick in the minds of the audience.

AESTHETICALLY PLEASING

Pieces of protest art that are more aesthetically pleasing become more effective for an audience. This is because people are more likely to engage with aesthetically pleasing pieces- and one of the most important functions protest art has to do is engage with an audience. So naturally, it follows that aesthetically pleasing protest art is more successful. How do we judge something as aesthetically pleasing? This is a question that is at the forefront of aesthetics and there is no one criteria that has been made that can satisfy all people. In my argument, I will not layout specific criteria for what makes art aesthetically pleasing but instead make some general comments about what makes aesthetically successful protest art.

There are different categories of protest art and each piece should be judged in their respective genre. For example, a piece of protest art could also be categorized as pop art. If we were to judge how aesthetically pleasing that piece was, we would judge against pop art, comparing: do the colours work as well, is it eye-catching, does it use space effectively? The fact that it is a piece of pop art does not hamper the fact that it is also protest art if it fulfils the criteria. If the piece of pop art is aesthetically pleasing that allows it to become a more successful piece of protest art; it engages audiences effectively and thus improves how well its message is communicated. However, what works aesthetically for pop art does not work for other types of art and thus other categories of protest art. What works for a poster will not work for a mural due to differing aesthetic criteria.

There is the worry that something is so aesthetically pleasing in its genre that people become distracted and do not appreciate the fact that it is also protest art. However, if this happens then the artist fails to fulfil one of the criteria of good protest art which is that it must communicate well. A good piece of protest art does not negate its message for aesthetics to the point that the audience becomes distracted. A good piece of protest art perfectly intertwines protest and any other category it may be in together; So, the piece can become aesthetically pleasing while simultaneously still maintaining the message it hopes to communicate as a piece of protest. A successful piece of protest art combines elements so it can become more successful in both categories e.g. a piece of protest pop art blends both protest and pop elements to become more successful in both categories.

SUBVERTS THE STATUS QUO

Good pieces of protest art catch the eye of the audience and then maintain their attention for long enough that the audience can engage with the piece, which is essential as the main function of protest art is to communicate a message. One way that pieces of protest art can effectively engage audiences is by subverting expectations. One of the problems that protest art has to

overcome is captivating the attention of people quickly and spreading information. For example, a poster is posted on the street and has to communicate its message rapidly enough that people can understand when they walk by or be engaging enough that people are willing to stop to look at it further. One way that protest art can engage people quickly is by having some form of shock value; People are more likely to look twice at something that subverts the status quo and engage with something controversial. Protest art already is going against a current institution, so the message is often already subverting the status quo- so it is up to the artist to create a piece that reflects their message properly. This can include coming up with a catchphrase or image that stands out to the public. By subverting the status quo, pieces are more to also more likely to increase discussion with viewers thus causing more people to engage with the piece.

One group that effectively used subverting the status quo in their protesting was the Black Panthers. Artists such as Emory Douglas were especially prolific in creating protest art for the movement which realigned black masculinity with symbols that were traditionally associated with white power and masculinity; the Black Panthers very effectively used subverting the status quo to their advantage, but movements have to carefully consider nuances when they subvert the status quo. By realigning with white masculine symbols, the Black Panthers excluded other groups such as women by fostering symbols of misogyny.[7] It was, however, a measure that very effectively grabbed the attention of many people and brought awareness to the movement.

EXAMPLES OF PROTEST ART
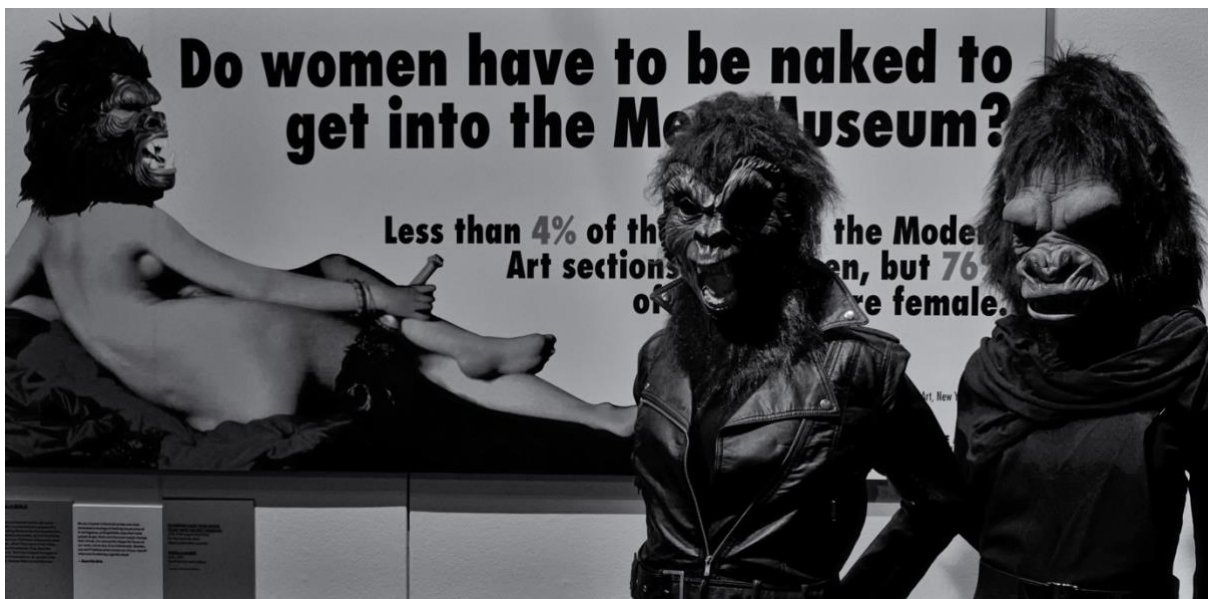
Members of the Guerilla Girls pictured with their work *Do Women Have to be Naked to get into the Met. Museum?* (1989).[8]

The Guerilla Girls were a feminist group in the art world that made posters such as the one above.[9] It is categorized as a piece of protest art because it communicates a message to the

---

[7] Doss, 2007

[8] Licensing: Eric Huybrechts, Guerrilla Girls - V&A Museum, London, CC BY-SA 2.0

[9] McQuiston, 2019

audience, is intended to be a piece of protest art, is displayed in a public manner, and was not made for profit. It can also be judged to be a successful piece of protest art because it subverts expectations, is aesthetically pleasing, and communicates its message effectively. When the viewer looks at the piece, their eye is first drawn to the woman with the guerrilla head, which subverts the expectations of the audience. Next, your eye is drawn to the slogan at the top of the poster, which communicates the main message. In an aesthetic and engaging way, the viewer learns about the protested institution with pink used to highlight keywords. Below, you can see who made the poster which gives meaning to the guerrilla mask. The piece is aesthetically pleasing due to its variable properties- the highly intense colours are eye-catching and help to focus the viewer toward the slogans. As the piece fulfils these categories, it is also able to fulfil its intended purpose, which is to enact change. This piece is highly successful in subverting expectations as it catches the eye of a viewer because of the intense colours and intrigues due to the subversive elements from the guerilla mask and informs the viewer simultaneously. It is hard to quantify something such as impact in a quantitative way, but certain indicators can be looked at. For example, with the Guerrilla Girls, issues of representation in museums have not been resolved, however, a huge amount of awareness was spread about women in art with their provocative imagery and slogans. They are also looked at by others as a reference when making protest art, which indicates that they were successful.

Another artist who successfully made many pieces of protest art was Keith Haring, whose pieces were often displayed in public places such as in transit terminals or through murals. Much of his work was drawn in the NYC subway system for viewers to see, thus allowing many different kinds of people to view it due to its very public nature (Pih, 2019). Its public display is not the only reason that his work constittes successful protest art. Art can communicate a message without words allowing it to be accessible to people without a language barrier. His anti-nuclear work protests the abuse of power using images that are recurring throughout Haring's work, allowing the viewer to more quickly understand his work. For example, Haring often protested for nuclear disarmament using the image of an atom and used an image of a barking dog to represent those in power. Haring didn't use specific slogans in his subway art as he hoped to provoke people into creating interpretations for themselves. Haring also created a performance out of creating his art. By displaying his pieces in the subway, they were technically graffiti, another reason why he had to use simple and recurring symbols to create his message, but he would also often attract crowds when he was working on these pieces. By creating a performance out of creating his art, he risked ticketing or arrest, making the act of creating his art an act of civil disobedience in itself. This adds a layer of subversion as it is not commonplace to see an artist arrested for their work, nearly creating a spectacle out of the piece. Along with all of this, Haring's imagery and symbols themselves are eye-catching and aesthetically pleasing to the viewers. Haring's work is often criticized for being over-commercialized, however, it gives weight to how aesthetically pleasing audiences find his work. The public nature of his work, aesthetically pleasing properties and iconic symbolism in Haring's work culminate in pieces that are highly successful pieces of protest art.

Recreation of the mural *Man at the Crossroads* (1933) called *Man, Controller of the Universe* (1934) by Diego Rivera

Diego Rivera is famous for his public murals protesting different causes such as anti-capitalist and anti-religious views, but perhaps his most notorious is *Man at the Crossroads* because of the reaction it caused and subsequent destruction. The mural was commissioned by the Rockefeller family to be painted in Rockefeller Center in New York City, however, the communist imagery that Rivera added as a form of protest to the mural led to it being destroyed.[10] However, it is perhaps because of this destruction that the mural was even more effective protest art. The piece was simple to understand, displayed in a public manner and aesthetically in line with Rivera's other works, making the piece already successful protest art. It was because of this success that it was called to be destroyed, but by doing this, the Rockefeller family only gave merit to Rivera's piece as a form of protest. The artwork was highly effective protest art and this can be seen in the reaction to its destruction- if it was ineffective, the piece would never have been destroyed. The protest was also specifically effective because of its location; the mural went against everything that Rockefeller Center represented. In response, many artists protested the destruction of the mural which drew further attention to the issues that Rivera was protesting, showing the effectiveness of the piece. *Man at the Crossroads* was an effective piece of protest art not simply for its pure visual aesthetics, but because of the other elements that influenced the aesthetic experience of the viewer: where it was located, how it was made, how it was talked about, and the history surrounding the piece. *Man at the Crossroads* shows us a contextualist view is necessary when judging protest art, as without all of these elements the mural would not be nearly as effective as it was.

WHAT IS THE VALUE OF PROTEST ART?

Protest Art as an Effective Means of Protest:
Questions about the purpose of art have been around for millennia and philosophers have tried to tackle it for just as long, going back to Plato. This is as art as is an effective means of communication allowing protest art to be an effective way to engage and inform audiences about a protest movement. Linguistic messages are often effective and used to support protest art, but knowledge of the specific language is needed which is exclusionary to some of the

---

[10] Lee, 1999

audience. Just as people read language, people 'read' art in much the same way. Aesthetic cognitivism, which believes good art can spread understanding, gives us a framework to start looking at how effective protest art can be in communicating with an audience.[11] Aesthetic cognitivism believes that knowledge can be acquired through art and that engaging with it is not just an aesthetic encounter. For example, you could see a painting about rising sea levels that causes you to have an emotional reaction that deepens your understanding of the severity of climate change; the fact the painting was hung in a boat could move you even more deeply. This dissertation assumes that it is possible to spread knowledge through art and that it is vital for protest art, as its specific function is to communicate a message through the piece. Engaging with an artwork in a thoughtful manner allows us to have emotional insights and with successful protest art, should allow us to deepen our knowledge around a specific cause. Many pieces of protest art communicate knowledge in obvious ways through slogans and writing, but some pieces can effectively communicate ideas to the audience just through symbols.

Christer Themptander's with his piece *We Never Forget Wounded Knee* (2005).[12]

In Christer Themptander's poster *We Never Forget Wounded Knee* (2005), there are no words or slogans, but there still is a clear message to be understood by the audience. A Lokota Man is seen behind an American flag, where the stripes act as bars for the man. Even without any knowledge about the Wounded Knee Massacre, the audience is still able to gain knowledge from the piece; it is reasonable to assume that someone could learn from this image that the American Government has unjustly oppressed Native Americans. The transference of knowledge is essential for good protest art and is part of the aesthetic experience of this piece. Just enough information is given by the piece to pique the interest of the viewer and entice them further; if the viewer has properly engaged with the piece, they will most likely find out the title of the poster which will further explain what the artist is protesting. This piece was specifically made hoping to spread information about the Wounded Knee Massacre and its ramifications, making aesthetic cognitivism at the core of the piece and the aesthetic value of it coming secondary.

---

[11] Graham, 1996
[12] Licensing: Skinersolen, Christer Themptander, CC BY-SA 4.0

Protest art is a particularly effective form of protest due to its position of being able to communicate with people quickly and without barriers. Many anti-cognitivists would argue that knowledge is only propositional and thus something like art cannot communicate, however, if you define knowledge in a broader sense then art can communicate information with an audience. Defining knowledge while excluding arts fails to account for the fact that under the right circumstances, even visual art can have propositional content. In *We Never Forget Wounded Knee*, no specific information about the massacre is disseminated to the audience, however, knowledge about the American Government's attitudes towards Native Americans is given to the audience and this broader knowledge is just as important as specific facts. Under the right circumstances, art can have propositional content as well as emotionally inform the audience. In this way, protest art functions as a speech act as the artwork is an act done by an artist asserting their views and informing the audience. The exhibition of an artwork is an act performed by the artist and thus is an indirect speech act similar to gesturing or facial expressions.

Protest art is especially effective in comparison to other forms of protest because of its function of educating the audience in a manner that is likely to engage people. Images can be processed quickly by people and thus protest art is effective at catching people's eyes when otherwise they might not have engaged in a protest. The aesthetic elements of the piece allow people to remain engaged; for example, there are many informative books and speeches around specific protests, however, people often will not remain engaged due to the high time commitment or becoming bored. As visual art can be processed quickly by the viewer, it does not risk boredom by the audience in the same way. As it does not rely solely on language, it also does not risk language barriers which excludes some people from being able to engage with the piece. This culminates in pieces that can engage with a large audience and are informative to them. One may argue if the piece can be processed so quickly then it would become forgettable to the audience, however, if it is truly a successful arresting image then it will stick in the viewer's mind despite the fact it is quick to process. Protest art is not the only effective form of protest, however, due to its ability to engage with large audiences and create change it is essential for creating successful protest movements and other acts of protest cannot replace it.

PROTEST ART V. ART USED AS PROTEST

One could argue that protest art is not a category of art in itself, but art that is used to protest. If this were true and protest art was not its own category but instead just art done for a protest, it implies that the piece is first and foremost art and it just happens to be used for protest. This could be the case with some pieces of art, but for protest art, the fact it was made for protest is intrinsically entwined with the fact that is art; it is not first an artwork and then a piece of protest, both aspects are equally important in practice and one piece of the art's identity cannot be removed without metaphysically changing the nature of the piece completely.

There is a difference between protest art and artists protesting. A contextualist view is necessary to differentiate these two because the difference often depends on how the artist identifies their work at the time. For protest art, it is essential that the artist identifies it as a sort of activism that is part of a larger movement, so when an artist makes protest art they should not only identify as an artist but also as an activist at the time. It is possible to have a protest of just one, but in that scenario, their protest has to reach a large audience. For example, if I were

to protest as just one person against another person, it would not truly be a protest and instead be giving criticisms. If a piece happens to have some protest aspects to it, but the artist was not intending the piece to be activist, then the work would not be a piece of protest art. This is the same with any of the criteria listed in the first section of this essay, if the piece fails to qualify as protest art, but still has some elements then the piece is an artist protesting but not protest art. However, if an artwork fails to meet the public display criteria then it could also simply fall into the category of political art. This is an important distinction because it allows us to judge art according to the correct category that it falls into; a piece of protest art should not be judged against an artist protesting because the intentions of the pieces are completely different. It also allows us to not categorize every protest made by an artist as art. Someone can identify as an artist but that does not make every action they perform art. For example, someone could professionally be an artist but as a form of protest, they go into a museum and graffiti a painting. The graffitied painting would not be a form of protest art in itself. This is not to say that you cannot make a new piece using materials of other pieces, which is the case in many mixed media pieces and collages, however in this instance as a form of protest by graffitiing on the piece, the artist fails to create a new piece and thus does not create protest art. The line where a new piece is created through the destruction of a previous piece is blurry, but it often lies when enough of the piece has been changed that the intentions of the new piece are more prominent than the original piece. This matters for protest art as it is not unheard of for people to destroy or add to art pieces as a form of protest. Commonly when this occurs, the artist fails to create a piece that is original and thus not a piece of protest; instead, it is simply an artist protesting. Another example of artists protesting, and not protest art, is when many artists boycotted commissions by the Rockefellers after what happened with Rivera's mural. This is an example of artists protesting because despite it having to do with protest art and being done by artists, the action of protesting alone is not art.

Conclusion

In this essay, I outlined what protest art is and how we should judge it. I have also said that for artwork to be categorized as protest art it has to be created with intention, communicate a message to an audience, be displayed publicly, and not be made solely for profit. I then set criteria to help define what properties 'successful' protest art has. These properties include effective communication, aesthetically pleasing properties, and subversion of the status quo. Then examples of protest art are given to show how they exemplify these properties. I then show why protest art is such an effective form of protesting due to its unique positioning. Finally, I respond to one objection that could arise from this paper. This dissertation outlines what protest art is and how to judge it due to the importance that it plays in our society. It can communicate, create community, and enact visible change and is a vital element to successful protest movements. Thus, a clear understanding of what protest art is and what makes it good is essential as it helps us understand how to communicate and make change with art. A thorough examination of protest art is long overdue because few other mediums silently play such a vital role in our society and can make such important changes.

References

Barthes, R. (1977). *Image, Music, Text*. Translated by S. Heath. London, United Kingdom: Fontana Press.

Bell, C. (1913). "Post-Impressionism and Aesthetics." *The Burlington Magazine for Connoisseurs*, vol. 22, no. 118, pp. 226-230.

Carlson, A. (1979). "Appreciation and the Natural Environment." *The Journal of Aesthetics and Art Criticism*, vol. 37, no. 3, pp. 267-275.

Christensen, A. P., Cardillo, E. R., & Chatterjee, A. (2023). "Can art promote understanding? A review of the psychology and neuroscience of aesthetic cognitivism." *Psychology of Aesthetics, Creativity, and the Arts*.

Davies, S. (1990). "Functional and Procedural Definitions of Art." *Journal of Aesthetic Education*, vol. 24, no. 2, pp. 99-106.

Dickie, G. (2001). *Art and Value: Themes in the Philosophy of Art*. Malden, Mass.: Blackwell.

Dixon, D. (2022). "Artistic (Counter) Speech." *The Journal of Aesthetics and Art Criticism*, vol. 80, no. 4, pp. 409-419.

Doss, E. (1999). "'Revolutionary Art Is A Tool For Liberation': Emory Douglas And Protest Aesthetics At The Black Panther." *New Political Science*, vol. 21, no. 2, pp. 245-259.

Graham, G. (1996). "Aesthetic Cognitivism and the Literary Arts." *Journal of Aesthetic Education*, vol. 30, no. 1, pp. 1-17.

Haring, K., Pih, D., Dujardin, P., Hemmes, T., Hans-Jürgen Lechtreck, Tate Gallery

Liverpool, Palais Des Beaux-Arts (Brussels, Belgium and Museum Folkwang Essen (2019). *Keith Haring*. Liverpool: Tate Liverpool; London.

Kieran, M. (2010). "The Vice of Snobbery: Aesthetic Knowledge, Justification and Virtue in Art Appreciation." *The Philosophical Quarterly (1950-)*, vol. 60, no. 239, pp. 243-263.

Lamarque, P. (2010). "The Uselessness of Art." *The Journal of Aesthetics and Art Criticism*, vol. 68, no. 3, pp. 205-214.

Lee, A. W. (1999). *Painting on the Left: Deigo Rivera, Radical Politics and San Francisco's Public Murals*. California, USA: University of California Press.

Mcquiston, L. (2019). *PROTEST!: A History of Social and Political Protest Graphics*. London, United Kingdom: Frances Lincoln Pub Ltd.

Rogers, A. (2016). "Does This Flag Make You Flinch?." New York Times.

Tepper, S. J. (2011). *Not Here, Not Now, Not That! Protest over Art and Culture in America*. Chicago, USA: University Of Chicago Press.

Tormey, A. (1973). "Aesthetic Rights." *The Journal of Aesthetics and Art Criticism*, vol. 32, no. 2, pp. 163-170.

Walton, K.L. (1970). "Categories of Art." *The Philosophical Review*, vol. 79, no. 3, pp. 334-367.

REVIEWS

THE CRISIS OF NARRATION
BYUNG-CHUL HAN
Cambridge, Polity Press, 2024, 100pp.

THE OTHER DAY, I CAME ACROSS AN ARTICLE IN *THE TIMES* BY WILLIAM HAGUE HEADLINED: 'How to stop technology leading to anarchy'.[i] The piece is premised on Jonathan Haidt's book, *The Anxious Generation*, which, to summarise simply, argues that the technological revolution has caused more bad than good. As Haidt argues — and Hague agrees — society has been 'over-protected in the real world and under-protected in the virtual world'.[ii] This has led to what he names the mental illness epidemic amongst the youth, specifically those that fall under Generation Z.

Similar to the narrative many books about technological crises follow, Haidt suggests that while technology has greatly benefitted society — from machines that mitigate the effects of climate change to the time-saving efficiency of Artificial Intelligence — social media has permanently damaged the younger generation. The evidence is there. The UK alone has mourned the deaths of schoolchildren who have taken their lives from what social media has made possible: online hate; triggering algorithms; grotesque images no child should ever see.

This is just one example of a sphere technology has plagued; another article *The Times* published a day later discussed the impact the digital landscape has on our memories. Laura Freeman argued that the digitalisation of once tangible forms — such as, the substitution of eBooks for books — is 'wearing at the fabric of our memories'.[iii] Freeman wrote: 'I can't see paintings I've glimpsed on Instagram the way I can see in my mind's eye paintings I've stood before in the flesh. You need the anchors around them: the paintings either side, the gallery benches, the postcards.'[iv]

Freeman's analysis is a helpful introduction to Byung-Chul Han's newest work, *The Crisis of Narration.* In a similar vein, the Korean-German philosopher argues that technology has invaded every aspect of our life. As the title suggests, Han does not focus on Generation Z or the destruction of tangible items, but another epidemic society is facing due to the rise of technology, that is, one's ability to narrate.

In Han's previous works — his most famous being *The Burnout Society,* originally published in 2010 — he captures the truth of our information-overloaded, phone-fixated, AI-generated society. In this new book, he reveals the impact this movement has on storytelling. That is, the pitch we sell other people. The reason why we should be liked and reshared. Thus, storytelling is now mostly 'a matter of commercialism and consumption' as social media has become a sale of the self.[v] Quality doesn't seem to matter — the important thing is to simply sell whatever you have to offer and hope it gains traction.

This line of reasoning is highly compelling and pinpoints the obvious damage the digital space has on today's society. In an era where one cannot travel to and fro without following a map conveniently packaged on an app; can find a detailed answer to their question within a few taps; or even satisfy their hunger with door-to-door food apps, whilst technology has made life notably easier, a cost occurs along the line. Han identifies that this simplification has turned

even the joys of storytelling into a transaction. Ultimately, we have commoditised the very thing that made us wholly human: the ability to convey stories with fervour.

His work parallels particularly with Mark Fisher's argument in *Capital Realism: Is There No Alternative?*.[vi] He argues that a lack of alternatives to the capitalist system, which he considers deeply flawed, has meant it is easier to imagine the end of the world than it is to imagine the end of capitalism. Fisher's work draws on one aspect of contemporary capitalism, that is, mental health phenomena such as depression, addiction, and attention deficit hyperactivity disorder (ADHD). He suggests that although these states can be influenced by genetics, they should also be understood as a by-product of the capitalist push to consume digitalised goods. Essentially, the pressure to stay in-the-loop and digest more material than the brain can process has exacerbated these conditions, as Jonathan Haidt also identified in his book.

I agree with Fisher's point. Ultimately, it is capitalism that has manufactured the once-unnatural — now-natural urge to reach for our phones without a second thought. This has not only impacted our attention span — of which the average is now at 8.25 seconds — but has given storytelling a competitive edge as users fight over how long they can keep watchers engaged. In essence, technology has naturalised all the tendencies that were once deemed abnormal.

Han's preface opens: 'Everyone is talking about narratives'.[vii] By 'narratives', Han means the stories that frame our character profiles. The stories we vividly remember from childhood, the stories we tell to impress others, or ones we neatly package into bullet points on our CVs. He argues that narratives provide our anchor in being and furnish life with 'meaning, support and orientation';[viii] characteristics that the digital age has altered. Instead, narratives are now 'contingent, exchangeable and modifiable'.[ix] In essence, the narratives we communicate in today's age are not absolute; they are subject to our own creation.

Han's argument is perfectly summarised in the first few pages:

> 'No amount of storytelling could recreate the fire around which humans gather to tell each other stories. That fire has long since burnt out. It has been replaced by the digital screen, which separates people as individual consumers. Consumers are lonely. They do not form a community. Nor can the 'stories' shared on social media fill the narrative vacuum. They are merely forms of pornographic self-presentation or self-promotion. Posting, liking, and sharing content are consumerist practices that intensify the narrative crisis.'[x]

Storytelling used to unite us around the campfire; it connected the youth to the past, helped us imagine our future and reminisce on the past. It invited laughter, belonging, sadness of what has ended, perhaps a fear of what lies ahead. For Han, narrating saturates our lives with meaning because he impactfully writes: 'only with narration does life elevate above its sheer facticity, above its nakedness'.[xi] There is no denying the power of narration; it is what gives humans their innate desire to sit around the fire.

Yet, this campfire — a metaphor for the pre-technological takeover when the youth rolled around in mud than played games on their iPads — has been replaced by a digital screen, Han suggests. 'The fire has long since burnt out', and thus, the ability to forge connections without the very source that creates this experience is impossible. So, as a substitute — or an antidote — we have become our own storytellers to an audience who are also their own storytellers.

However, Han has identified a deeper issue here. As society continues to feed into this digital narrative, we become immersed within the same narrative codes. Essentially, we are the observers of other's narratives and the protagonist of our own narrative. The problem here is these streams have merged. As every narrative has entered the virtual and digital realm, accessible by anyone and everyone within the tap of a screen, this realisation — the blending of once distinct mediums — is what gives rise to this crisis of narration. This, as Stuart Jeffries writes, is 'the existential tragicomedy of modern humanity'.[xii]

Yet, this competition is a lonely, soul-destroying one. Competitive individualism characterises today's society. It is interesting the way Han describes the evolution of the narrative — it is a 'pornographic self-presentation or self-promotion'. That is, a highly abused medium in which humans fall victim to the desire to present and promote oneself. Han goes further to suggest that the more we engage with these digital narratives by 'posting, liking and sharing' content, the more this 'narrative crisis' is intensified.[xiii]

This is an interesting start to his work, given this is all packaged into its first pages. In the first chapter, Han even draws on the decline of the novel, suggesting that 'the ultimate decline of the narration comes not with the novel but with the rise of information under capitalism'.[xiv] He too thinks modern novels — for example, #BookTok on TikTok — mark the decline of the narrative, suggesting that it has become bourgeois individualism personified. This, too, is an insightful analysis of the sheer scope of information devices store and instantaneously deliver when we request it. This information, scarily enough, has permeated the novel — an outlet we thought was immune to the technological touch.

In one section, Han even applies the death of the story to GP surgeries. In a place where personalised accounts of illnesses are necessary, this procedure has also become technologised. Instead, patients are referred to online GPs, asked to tick reductive boxes that match with their symptoms. Whilst medical care is made more efficient with the help of technology, this efficiency erodes all meaning. Patients are reduced to words on a screen waiting to be seen and excused rather than providing meaningful, rich accounts of their condition. As Jeffries summarises: 'We are reduced to identikit puppets with identikit symptoms',[xv] that is, homogenous beings with no narratives to define ourselves by.

In another analogy, he writes: 'On the internet [...] the dream bird cannot build a nest. The information seekers drive him away. In today's state of hyperactivity, where boredom is not allowed to emerge, we never reach the state of deep mental relaxation. The information society is an age of *heightened mental tension* because the essence of information is surprise and the stimulus it provides. The tsunami of information means that our perceptual apparatus is permanently stimulated. It can no longer enter contemplation. The tsunami of information fragments our attention. It prevents the contemplative lingering that is essential to narrating and careful listening.'[xvi]

The fundamentals in his argument are clear: in the digital era, humans are no more than manifestations of technology, but merely living data sets. This he says, 'hides behind the illusion of freedom and communication'.[xvii] We think we are free to write, to speak, to publish our narratives, but this is a mere hallucination. Instead, by contributing the facilitation of narratives, we are rather feeding into this blind sightedness. In essence, we are the creators of our own demise.

To explain the evolution of the narrative, Han uses many binaries which are oftentimes jarring. From contrasting homosapiens to Phono *sapiens* and storytelling to storyselling, these dichotomies can be highly reductive given there is a middle ground. Arguably, there is a current

push to revert our natural state whilst utilising the technology around us. For example, take meditation playlists on YouTube or Spotify. By utilising the power of technology, this outlet simultaneously utilises the power of the voice — transmitting one from the digital realm into a state of presence. Yet, I wonder if Han would say that this natural state can ever be reached when technology must be used to reach it? Can we ever return to the pure form of narration without completely wiping out technology? Is there no in-between? If not, is Han suggesting that the fire will never return?

In section six, Han takes his argument to an extreme. He first claims that the digital narrative 'reflects the smooth consumer world that is the opposite of the world of shocks. [...] His art is intentionally relaxed and disarming. What he wants above all is *to be liked*.'[xviii] Shock is necessary to feel, think and create raw narratives. This could be the shock of figuring out how to walk, being challenged in a meeting, or acclimatising to new environments; all of which shape our own permanent narratives. He has a point here — it is the shocking events that we recall. The strange and unexpected things that etch our memories over the mundane, expected events. Han writes: 'Etymologically, a screen [*Schirm*] is a protective barrier. A screen *bans* reality, which becomes an image, thus screening us from it. We perceive reality almost exclusively via digital screens. [...] On a smartphone screen, reality is so attenuated that it can no longer create any shock experiences. *Shocks give way to likes*.'[xix]

Here, he suggests that the screen suppresses shock, inviting comforting sensations of non-threatening episodes that do not surprise us. We can filter what we see, who we get notifications from, who we decide to share material with. A digital world that is friendly, comforting, and safe; one we escape to when reality gets too hard to deal with, or, too shocking for the system to process. Thus, shocking encounters are no longer the norm, disturbing encounters — anything that makes us feel uncomfortable — is banned. Understandably so, radical books and art, information about the real grotesqueness of current events — and everything in-between — has become desensitised.

Han remains dedicated to his reasoning. To him, homosapiens — the entirety of the human species — have degenerated into Phono *sapiens*. That is, shaped, controlled and glued to a technological escape of one's choice, so much so we are no longer human, but merely constructs of the very outlet we have grown to depend upon. Later, Han insists that the birth of Phono *sapiens* 'announces the end of the human being as someone with fate and history' in virtue of sacrificing oneself to social media.[xx] In his eyes, we have become non-beings that cannot distinguish the real from the digital experience.

And what makes the book even more enjoyable is that it is no more than 200 pages long. His ambition, he claimed in an interview for *El Pais*, is to make his work accessible and understandable, not brimmed with complex language, non-stop sentences, and incomprehensible dialogue typical of many philosophical books. Rather, his style is simple and digestible; an 'easy read'. If only today's digital landscape was this simple.

HOLLY DOWNES
Durham University

---

[i] William Hague, (2024, April 1), *How to stop technology leading to anarchy*, Retrieved from The Times: https://www.thetimes.co.uk/article/how-to-stop-technology-leading-to-anarchy-kss9jxzxh.

[ii] Ibid.

[iii] Laura Freeman, (2024, April 1), *Digital world will leave us clutching thin air*, Retrieved from The Times: https://www.thetimes.co.uk/article/digital-world-will-leave-us-clutching-thin-air-59ls900ks.

[iv] Ibid.

[v] Byung-Chul Han (2024), *The Crisis of Narration*, Cambridge: Polity Press, page 14.

[vi] Mark Fisher (2009), *Capital Realism: Is There No Alternative?*, Winchester: Zero Books.

[vii] Byung-Chul Han, op. cit., p. vii.

[viii] Ibid. page vii.

[ix] Ibid. page viii.

[x] Ibid. page ix.

[xi] Ibid. page 27.

[xii] Stuart Jeffries (2024, February 18), *The Crisis of Narration by Byung-Chul Han review — how big tech altered the narrative*, Retrieved from The Guardian: https://www.theguardian.com/books/2024/feb/18/the-crisis-of-narration-byung-chul-han-review#:~:text=feeds%20the%20machine.-,I%20found%20Han%20most%20relatable%20when%20he%20reflects%20on%20the,nor%20the%20patience%20to%20listen.

[xiii] Byung-Chul Han, op. cit., page ix.

[xiv] Ibid. page 5.

[xv] Stuart Jeffries, op. cit.

[xvi] Byung-Chul Han, op. cit., page 6.

[xvii] Ibid. page 7.

[xviii] Ibid. page 48.

[xix] Ibid. page 46.

[xx] Ibid. page 20.

EDITORS

EDITOR
Woody Jeffay, Durham University

ASSOCIATE EDITOR
Jess Walsh, Durham University

PEER REVIEWERS

Manon B. Danielsen, Durham University

Anastasia Khairova, Vrije Universiteit Amsterdam

Nathan Michel, McGill University

Neil Williams, University of Liverpool

Felix Porée, King's College London

Justin Lee, St Andrews University

Faye Perry, St Andrews University