On Referential Opacity in Spinoza's *Ethics*

KIAN MINTZ-WOO

UNIVERSITY OF AMSTERDAM

Abstract

In Spinoza's system, the identity of mental modes and extended modes is suggested, but a formal argument for its truth is difficult to extract. One prima facie difficulty for the claim that mental and extended modes are identical is that substitution of co-referential terms in contexts which are specific to thought or extension fails to preserve truth value. Della Rocca has answered this challenge by claiming that Spinoza relies upon referentially opaque contexts. In this essay, I defend this solution by analyzing what is required to establish that Spinoza recognizes referentially opaque contexts as part of his system. One objection that has been made to Della Rocca's account is that he only establishes the intelligibility, not the actuality, of such mode-identity. In this essay, I argue that the intelligibility of mode-identity is sufficient to establish the existence of opaque attribute-specific contexts.

1. Introduction

Spinoza's claim in the *Ethics* that the order and connection of ideas are the same as the order and connection of extended bodies has proved difficult to parse. Spinoza writes that an extended object (an object "under the attribute of extension") and its associated idea (the object "under the attribute of thought") "are one and the same thing" (2p7).[1] But Spinoza also maintains that objects under different attributes cannot have trans-attribute causal efficacy. Did Spinoza intend that mental objects are really the *same* as physical ones? If so, why can some contexts be truthfully applied to an extended object, but not to the same object when viewed as a thought? More subtly, are objects under these different attributes *intelligible* (or conceivable) as identical, or are they *actually* identical?Michael Della Rocca analyzes these issues by employing the distinction between referentially transparent and

---

1 For all quotations from the *Ethics*, I use Edwin Curley's translation (Spinoza, 1677) and his notation.

referentially opaque contexts (Della Rocca, 1993; Della Rocca, 1996). When contexts are referentially transparent, any way of referring to the same object within such a context will generate a sentence that has the same truth-value; the co-referring terms are exchangeable *salva veritate*. In referentially opaque contexts, exchanging coreferential terms can change the truth-value of the resultant sentence. Della Rocca argues that Spinoza is committed to the numerical identity of mental and physical objects, but could still maintain the causal independence of different attributes.

The key to this approach is that, although terms may be equal (i.e. refer to the same object), this does not mean that they can always be exchanged *salva veritate*. In referentially opaque contexts, the terms are not interchangeable. This can be used to examine modes under different attributes.

Della Rocca argues that Spinoza accepts an "explanatory barrier" between different attributes such that objects under one attribute cannot have causal relations (or explanatory relations) with the same object under a different attribute.[2] Spinoza writes that "[e]ach attribute of a substance must be conceived through itself" (1p10). So attributes are explanatorily isolated from one another. On the basis of this evidence Della Rocca concludes that Spinoza implicitly appeals to referentially opaque contexts.

In this essay, I address one of the responses leveled against Della Rocca's analysis.[3] The response is an attempt to accept Spinoza's claim about the connection between different contexts while denying that the numerical identity of mental and physical objects is required. The independence of the causal chains is accepted as an *intelligibility* or *explicability* claim, not a full-fledged ontological thesis. Garrett (2000) suggests that Della Rocca demonstrates the explicability of Spinoza's claiming the numerical identity of mental and physical objects while denying trans-attribute causality, but that this is not sufficient to establish the view that these objects are actually identical. In other words, Spinoza could claim that we can conceive of situations in which contexts are opaque without committing to the actual existence of such contexts. This would prevent the move towards referential opacity.

---

2 In this paper, I do not argue for the explanatory barrier, but take it as given. For defense of this barrier, cf. Della Rocca (1996).

3 Much of the critical response to both Della Rocca (1993) and Della Rocca (1996) has been positive. See, for instance, Schmaltz (2000) and Manning (1998).

I argue that Spinoza does require the stronger claim of the existence of referentially opaque contexts, and that Della Rocca has the resources to claim this. In particular, Della Rocca presents a counterexample to transparent contexts. The key is that he only needs the counterexample to be intelligible — i.e. that there is a conceivable situation which contradicts transparency — in order to demonstrate the actuality of referentially opaque contexts, as opposed to making substantive ontological claims.

In §2, I define my terms and delineate the objection that has been raised against numerical identity. In §3, I explain how referentially transparent and opaque contexts are necessary to address the issue raised in §2. In §4, I lay out Garrett's challenge. I focus on whether Spinoza can be read as endorsing the weaker claim that such contexts are *explicable* in an intra-attribute manner, without being actual. Finally, in §5, I examine three propositions. If the truth of the first two and the falsity of the third are shown, then certain contexts must be opaque. Della Rocca can respond to Garrett by showing that the intelligibility of opaque contexts is sufficient for their actuality. In this way, referentially opaque contexts are shown to be endorsed by Spinoza. I conclude in §6.

2. The charge against numerical identity

The numerical identity thesis holds that modes under different attributes refer to the same object. This term is related to philosophical issues about mind-body-identity, where it refers to the idea that my body is my mind. Spinoza appears to endorse this reading:

> The mind and the body are one and the same thing, which is conceived now under the attribute of thought, now under the attribute of extension. (3p2s)

This is a stronger claim than the claim that there is merely a parallel between modes under extension and modes under thought. The thesis of parallelism is articulated by Spinoza as follows: "The order and connexion of ideas is the same as the order and connexion of things" (2p7). This implies that modes in a causal chain under the attribute of thought directly correspond to a material causal chain, but leaves open the metaphysical status of the two chains. Numerical identity provides a simple answer: the parallel chains are the *same* chain, concerning the same modes. The only difference is that the modes are viewed first under one attribute, and now under another.

This is a striking claim, since contexts applied to modes under one attribute yield falsity when applied to modes under another attribute. This is due to the fact that different attributes are characterized by distinct qualities, e.g. objects under extension have width

and height. When a context can only truthfully be applied to a mode under a certain attribute I call it, following Garrett (2000), an "attribute-specific" context. Two examples of extension-specific contexts are "is larger than a breadbox" and "is red." These contexts are true only of objects under the attribute of extension, if we presume that thoughts lack size or color. Thought-specific contexts include "is a radical idea" and "contains within itself ideas about fairness."

Attribute-specific contexts offer a problem for numerical identity. How can it be that certain relations hold when a mode is viewed under one attribute and not another, and still pertain to the *same* mode? This objection can be strengthened by considering contexts about which Spinoza has firm commitments. One such commitment concerns causality, since Spinoza rejects trans-attribute causality.

> [...] so long as things are considered as modes of thinking, we have to explain the order of the whole of nature, or the connection of causes, through the attribute of Thought alone. And insofar as they are considered as modes of Extension, the order of the whole of nature has to be explained through the attribute of Extension alone. (2p7s)

So, when a certain attribute is being considered, all causal chains must be explicable with reference to modes under that attribute. Della Rocca (1996) calls this Spinoza's "explanatory barrier." This means that causal contexts are attribute-specific, in that the explanans and the explanandum must be both considered under the same attribute.

With this background in hand, it is not difficult to see where a puzzle about numerical identity emerges. Delahunty (1985) provides the following objection: If we take an attribute-specific context regarding causation, then it can be applied to an object under one attribute, but not to the same object under another attribute. On the numerical identity view, this context is applied twice to the exact same object, under different attributes. However, it is true when the object is viewed under one attribute, and false when it is viewed under another.

We can formalize this argument as follows. The thesis of numerical identity commits us to the identity of objects which correspond by the parallelism thesis. Let a and i be the same object under two attributes. In other words, the terms corefer as follows:

Mode under attribute of thought a = Mode under attribute of extension i.

But we choose a in such a way that it causes another mode under the attribute of thought. Let the mode which is the effect be b. Putting the modes together with a causal context

under the attribute of thought yields the following:

Mode under attribute of thought *a* causes mode under attribute of thought *b*.

Of course, using the identity in the first claim to substitute the coreferring term in the second claim yields:

Mode under attribute of extension *i* cause mode under attribute of thought *b*.

Given the proper choice of a mode *a* and *b*, the first two claims are true. The conclusion contradicts the "explanatory barrier"; it explains causation in a trans-attribute manner. Since the explanatory barrier is an important aspect of Spinoza's metaphysics, this acts as a *reductio* against numerical identity.

3. Opacity in Spinoza

Della Rocca's solution relies upon the claim that certain contexts are referentially opaque in Spinoza's system. I begin by sketching referential opacity and then move on to the use of this concept in defense of Spinoza's commitment to the numerical identity of mental and physical objects.

Referentially opaque contexts are contexts where exchanging coreferential terms can change the truth-value. This is most easily illustrated with an example. Consider the following identity:

(A) Dwayne Michael Carter, Jr. is Lil' Wayne.

This claim is true; i.e. the two terms corefer — they both refer to the rapper from New Orleans. Furthermore, we can find a context that is true for the second term, but falsified by the first:

(B) Jessica knows Lil' Wayne made the top selling album of 2008.

Since Jessica follows online music charts obsessively, she knows that this is so; i.e., (B) is true. However, since she is not aware of Lil' Wayne's given name, exchanging the coreferring terms in the second sentence yields falsity:

(C) Jessica knows Dwayne Michael Carter, Jr. made the top selling album of 2008.

This shows that this epistemic context is an opaque context, i.e. exchanging coreferential terms fails to preserve truth-value.

We can define opacity more schematically. A context f is referentially opaque if the following condition is satisfied: a term a and a term b which are coextensive (so $a=b$), and a context f such that the truth of $f(a)$ does not entail the truth of $f(b)$. If there are no such terms, then f is a referentially transparent context.

Newlands (forthcoming) suggests that reading Spinoza while assuming referential transparency offers certain challenges. Since Spinoza takes two things to causally depend on each other if and only if one conceptually contains the other, Newlands claims that Spinoza could be read as denying the explanatory barrier. For instance, if you begin by assuming referential transparency in causal contexts, then you can use *modus tollens* to derive that no explanatory barrier exists. Instead of beginning with the barrier and concluding with opacity, he suggests it is possible to conclude that there is no barrier from referential transparency.

But I take this reading to be difficult to maintain. Assuming referential transparency requires that immanent causal contexts are transparent, which I will argue contradicts 2p6dem.

Consider Della Rocca's suggestion that referentially opaque contexts are behind the problem of substituting coreferential terms and the explanatory barrier expounded in §2. That opaque contexts are necessary for causality is established by the claim that G-d only causes modes insofar as he is viewed under the appropriate attribute:

> [T]he modes of each attribute involve the concept of their own attribute, but not of another one; and so (by IA4)[4] they have G[-]d for their cause only insofar as he is considered under the attribute of which they are modes, and not insofar as he is considered under any other. (2p6dem)

With Della Rocca's suggestion in mind, the quote implies that causal contexts which include G-d must be opaque. Why? For a mode under the attribute of extension to be caused by G-d implies that, in this instance, G-d must be conceived under the attribute of extension (and similarly for other attributes).

However, if we consider G-d under a different attribute, then the same context makes a false statement. I.e. by 1p18, substance (G-d) considered under the attribute of thought is no longer the cause of the extended mode. This is true even despite the fact that the two

---

4   "The knowledge of an effect depends on, and involves, the knowledge of its cause."

terms— G-d under the attribute of thought and under the attribute of extension — are coreferential.

The use of referentially opaque contexts addresses the charge in §2 by making these causal claims sensitive to the attribute under which G-d is viewed. The demonstration of 2p6 suggests that, at least for G-d, referentially opaque contexts are operative. A further issue concerns other modes, particularly finite modes for which causality is transitive. If it can be established that Spinoza also views the causal contexts of finite modes as referentially opaque, then this solves the problem. Recall Della Rocca's "explanatory barrier," from §2. When we consider modes under the attribute of thought, Spinoza maintains that they must be exclusively explained through and caused by ideas (similarly for extended modes).

The explanatory barrier appears to be the perfect solution. Since the causality of mental modes "ha[s] to [be] explain[ed] … through the attribute of Thought alone" and the causality of extended modes ha[s] "to be explained through the attribute of Extension alone," causal relations between finite mental and extended modes cannot be explained in a trans-attribute manner (2p7s). But the next inference — that Spinoza takes causal contexts for finite modes to be referentially opaque — is open to attack. Finite causal contexts may only be *intelligibly* opaque, and not *actually* opaque. I explore this objection in the next section.

4. Is intelligibility sufficient?

In his review of Della Rocca (1996), Garrett questions whether Della Rocca's argumentation is sufficient to guarantee that such opaque contexts occur for finite modes. Perhaps these arguments only show that they *conceivably* occur. It is worth quoting Garrett at length:

> How can Spinoza support the claim that attribute-specific contexts are referentially opaque? Here again Della Rocca sees the explanatory barrier between the attributes at the basis of Spinoza's thought […] Della Rocca argues [that] it is intelligible or conceivable that "mode of extension A is extended" and "mode of extension A = mode of thought 1" are true while "mode of thought 1 is extended" is false. This, he concludes, would show that "whether it is true to say that a thing is extended depends on how that thing is described," and hence (since a parallel argument applies to thought) it shows that attribute-specific contexts are referentially opaque (146). In fact, however, what follows directly seems at most to be that it is *intelligible or conceivable* that "whether it is true to say that a thing is extended depends on how that thing is described." Whether *actual* referential opacity follows from this or not depends on the logical behavior of the "it is intelligible that…" operator. (Garrett, 2000, p 226, italics his)

The three following propositions are meant to demonstrate the falsity of referential transparency (i.e. show that referential opacity holds in some contexts). Let us label these propositions in turn:

(1) Mode of extension *A* is extended.

(2) Mode of extension *A* = mode of thought 1.

(3) Mode of thought 1 is extended.

I shall call the intelligibility (or conceivability) of (1) and (2) together with the unintelligibility of (3) Della Rocca's "intelligibility counterexample." When Della Rocca presents the intelligibility counterexample, he thereby shows that it is intelligible that this attribute-specific context is opaque. Why is this? If all attribute-specific contexts are non-opaque, i.e. transparent, then it is inconceivable that (1) and (2) are true while (3) is false. In other words, this acts as a counter-example to transparency. Then Garrett's question is as follows: Is this sufficient to guarantee the actuality of such opaque contexts?[5]

In other words, Garrett wonders if Spinoza's metaphysical commitments, under Della Rocca's interpretation, are underdetermined. The consistency of this account with Spinoza's claims is not sufficient to show that Spinoza endorses this account.

This is a serious objection, but it mischaracterizes Della Rocca's position. His claims are stronger. He does not merely *assert* the intelligibility of referentially opaque contexts relating to the three propositions. Instead, Della Rocca argues that the negation of (1) and the truth of (3) are both *inconceivable* and *unintelligible*. Furthermore, the intelligibility counterexample demonstrates that transparency does not (always) accurately describe such contexts. In the following sections, I argue that these stronger claims are sufficient to ground the actuality of referentially opaque contexts.

5. A counter-example to transparency

In this section, I argue that Garrett's objection (cf. §4) shows that Garrett does not recognize the strength of Della Rocca's claims (Della Rocca, 1996, pp.144-155). In particular, when we consider the intelligibility counterexample, it is not simply that the truth of the first

---

5 In this context, "possibility" could also be used instead of "actuality", but such usage might engender confusion. Since Spinoza's necessitarianism prevents any modal conception of possibility (e.g. possible world semantics), the sense of "possibility" intended — a metaphysical modality, as opposed to event modality —might be difficult to define.

two is intelligible along with the falsity of the third, but also that the falsity of (1) and the truth of (3) are *unintelligible*.

Let us examine this more carefully. Consider what it would mean for proposition (1), "Mode of extension *A* is extended", to be false. If it was false, then a mode under the attribute of extension would not be extended. But this is contradictory for Spinoza: being under the attribute of extension implies being extended. So not only is it intelligible that (1) is true; its negation is unintelligible.

Now consider (3): "Mode of thought 1 is extended." By 2p6, we have it that the modes under an attribute are caused by G-d only insofar as he is considered under that attribute.[6] Since the mode of thought 1 is extended, it is caused by the unique, existing substance (i.e. G-d) insofar as the substance is considered as extended. But since a mental mode 1 cannot be caused by anything extended, 1 cannot be caused by the extended substance. Thus, the truth of (3) is likewise ruled out.

Since the negation of (1) and the truth of (3) are unintelligible, the "intelligibility" of (2) (i.e. mode identity) is the only claim of those involved in the intelligibility counterexample that remains to be considered. But before turning to the rest of the argument, we should address an important question: What does "intelligibility" mean in this context?

Intelligibility is actually as weak as it appears. Its negation, for Della Rocca, is contradiction. We see this in his claim that the falsity of (1) is unintelligible since "for Spinoza, it follows from the notion of mode of extension that each such mode is extended" (Della Rocca 1996, p.146). In other words, the truth of this statement is related to analyticity; unintelligibility may take the form of *contradicting analytic statements*. So all that is required for the intelligibility of a statement is that it is not contradictory in this sense. In particular, intelligibility does not relate to whether the terms of the statement refer to extant objects.

Furthermore, mere intelligibility does not establish actuality: "If Spinoza were justified in holding that it is *intelligible* for a mental thing to be identical with a physical thing, then [this would] not help establish that mental things are actually identical with physical

---

6 If these modes are finite, they are caused by other finite modes in G-d, while they are caused by infinite modes in the case that they are infinite modes. However, I bracket this distinction, as it is immaterial for the present argument.

things" (Della Rocca, 1996, pp.145-146, italics his). So there is a clear distinction between intelligibility and actuality.

If intelligibility is such a weak concept, isn't Garrett correct in claiming that it does not establish the actuality of attribute-specific contexts? The answer is that Della Rocca establishes the intelligibility of mode identity as a precursor to the actuality of attribute-specific contexts. If he can establish the *intelligibility* of mode identity, then the opacity of attribute-specific contexts follows (Della Rocca, 1996, pp.144-145). Why is this? Because the intelligibility of (1), (2) and not-(3) shows that "…is extended" is an opaque context. In other words, if it can be established that (2) is intelligible — not "actual" or "possible," as these are loaded words in Spinoza's system — then "is extended" *is* an opaque context. The intelligibility of the same context with two coreferential terms having different truth-values does not show that it is intelligible that the context is opaque; it shows that the context *is* opaque. If no attribute-specific contexts were opaque (i.e. all were referentially transparent), then it would not even be conceivable that different coreferential terms could yield differing truth-values.

Here is Della Rocca on the subject:

> To see why, for Spinoza, the identity of a mode of extension and a mode of thought is intelligible, let's assume that the opposite is true. In particular, let's assume that it follows from the concept of an extended thing and from the concept of a thinking thing that no thinking thing is identical with an extended thing. If this is the case, then we could offer the following explanation for the non-identity of mode of extension *A* and mode of thought 1: Mode of extension *A* is not identical with mode of thought 1 *because* mode of extension *A* is extended (and mode of thought 1 is thinking). (Della Rocca, 1996, p.147)

This explanation crosses the explanatory barrier: It explains the fact of *A*'s non-identity with 1 by appealing to trans-attribute facts. Thus, for Spinoza, the unintelligibility of (2) cannot be maintained.

But an immediate objection presents itself: this argument may be accepted as valid, but its assumptions may be denied. In particular, there is no need to assume that unintelligibility of mode identity requires the concepts to be explanatory in this sense. It may be that the modes are non-identical, but this cannot be *explained* by appealing to these trans-attribute facts. In other words, it could be a contingent fact that there is no extended mode such that it is identical with a mode of thought. It is not by the definitions of "extended" and "thinking" that there are no identical modes, but it simply *happens* that there do not

exist such identical modes. The intelligibility counterexample fails to take the leap from conceptual possibility to actuality. I take Garrett's objection to lie along these lines.

To respond to this objection, consider what it means. The claim is that there are no *actual* counterexamples; i.e. no referents for the coreferential terms that are used in the intelligibility counterexample. But the conclusion that the intelligibility counterexample is meant to produce is not about existence. The conclusion is that some contexts are opaque. It implicitly relies upon a conditional:

(4) If all attribute-specific contexts are transparent, then it is inconceivable that exchanging coreferential terms in attribute-specific contexts can change truth-value.

It is correct to say that Della Rocca cannot appeal to the existence of an object for which the terms (an extended mode and a thinking mode) are coreferential. To do so would be to beg the question in favor of Spinoza's view.

But what does referential transparency mean? Given a transparent context, if a term which makes the context true is replaced with any other coreferential term, then the statement retains its truth-value. But referential opacity does not generate ontological commitments. In particular, the possibility of referential opacity does not lend itself to the existence of the referential objects. There is no restriction on transparent or opaque contexts that limit themselves to existent objects.

This may appear to be counterintuitive in light of the fact that co-referentiality requires referentiality. However, if one considers an opaque context where the referent is no longer actual, then the opacity of the context does not change.

Consider the aforementioned example of Lil' Wayne (§3). If we change the example in such a way that the coreferring terms no longer refer, I suggest that this does not change the opacity of the contexts. Suppose that Lil' Wayne dies. This does not affect the truth of (A), since it still holds intuitively that:

(A) Dwayne Michael Carter, Jr. is Lil' Wayne.

This is true even given that the referent of both terms is no longer extant. Suppose that Jessica (who knows that Lil' Wayne was behind the top selling album of 2008) never learns

his given name after Lil' Wayne's death. Then (B) still holds and (C) is still false:

(B) Jessica knows Lil' Wayne made the top selling album of 2008.

(C) Jessica knows Dwayne Michael Carter, Jr. made the top selling album of 2008.

By the definition of opacity, the falsity of (C) together with the truth of (A) and (B) show that "Jessica knows that … made the top selling album of 2008" is an opaque context. This works despite the reference failure of the two terms in (A).

The conditions of opacity and transparency do not include ontological claims about the objects to which the context can be applied. Thus, the intelligibility of such identities (and the corresponding change in truth-values) is all that is required.

By showing that this is intelligible, Della Rocca can appeal to the converse of (4) ("If it is conceivable that exchanging coreferential terms in attribute-specific contexts can change truth-values, then not all attribute-specific contexts are transparent") and conclude that there are attribute-specific opaque contexts.

So even though Della Rocca only demonstrates the intelligibility of an extended mode being identical to a thinking mode, this is (in the context of the intelligibility counterexample) sufficient to show that attribute-specific contexts are opaque. Since the consequent of (4) is violated by the intelligibility counterexample, it follows that some of the attribute-specific contexts are opaque. *Contra* Garrett, a stronger conclusion than the mere intelligibility of opacity can be gleaned from this argument. The intelligibility counterexample demonstrates, on Spinoza's terms, the existence of referentially opaque attribute-specific contexts.

6. Conclusion

The role that referential opacity plays in maintaining consistency in Spinoza's metaphysics cannot be understated. The difficulty in establishing the existence of such contexts underlines the need for careful examination of Spinoza's argumentation. Della Rocca's intelligibility counterexample, predicated on the intelligibility of mode identity, and leading to the production of referentially opaque contexts, ultimately lays the groundwork for the heavy lifting needed to make sense of actual mode identity.

This essay was concerned with the smaller aspect of this argument. Garrett's objection illustrates the need for Della Rocca's implicit conditional (4) about referential transparency. I

take the implicit conditional to be feasible, but further work is needed to determine whether it is analytic with respect to the definition of the constitutive terms or in need of its own defense.

References

Delahunty, R. J., 1985. *Spinoza*. London: Routledge and Kegan Paul.

Della Rocca, M., 1993. Spinoza's argument for the identity theory. *Philosophical Review*, 102(2), pp. 183-213.

-----, 1996. *Representation and the Mind Body Problem in Spinoza*. New York: Oxford University Press.

Garrett, D., 2000. Review: Representation and the mind body problem in Spinoza. *Philosophy and Phenomenological Research*, 61(1), pp. 223-6.

Manning, R., 1998. Review: Representation and the mind body problem in Spinoza. *Philosophical Review*, 107(4), pp. 603-6.

Newlands, S., Forthcoming. Another kind of Spinozistic monism. *Nous*.

Schmaltz, T., 2000. Review: Representation and the mind body problem in Spinoza. *Mind*, 109(435), pp. 580-3.

Spinoza, B., 1677. *Ethics*. Translated by Edwin Curley 1996. 2nd ed. London: Penguin Books, London.