



Title	Two Ways To Reject Kripke's Modal Argument : Absurdity, and The Primitiveness of the Person
Author(s)	Malik, Luke
Citation	メタフュシカ. 2008, 39, p. 59-72
Version Type	VoR
URL	https://doi.org/10.18910/8896
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

Two Ways To Reject Kripke's Modal Argument: Absurdity, and The Primitiveness of the Person

Luke Malik

In this paper we consider Kripke's Modal Argument (KMA) against the identity theorist (the philosopher of mind that identifies the mental with the physical). We offer two objections against this argument. First, we argue that KMA leads to a form of dualism, which we dub "KMA dualism". We argue KMA dualism leads to the practical irrelevance of the mental with respect to the physical, and we argue that this leads to absurdity, which may be used to reject KMA. Second, we argue that if the person or self is taken as primitive, as P. F. Strawson suggested, the mental and the physical cannot be separated. Since the separation of the mental and the physical is the starting point of KMA, this stops KMA getting started. So, the first objection attacks KMA indirectly and the second objection attacks KMA face on. Both put the blame on the premise of KMA for the problems we identify.

The paper consists of four sections. In section 1 we show that if KMA is correct, then every statement of identity of the type that states that there is an identity between a mental token or type and a physical token or type is false. In section 2 we show that if that is the case, then KMA leads to the practical irrelevance of mental tokens and types. This leads, in section 3 to our first objection—the objection from absurdity. And in the last section, section 4, we state the objection from the primitiveness of the person or self. Both objections aim to put the blame on Kripke's starting premise: the thought that it is imaginable or conceivable that the mental is separable from the physical. Before we get going, we should say that the use of the term "mind," "mental," and related word forms will be restricted. These terms will only refer to what David Chalmers calls "type-II phenomenon."¹ Following Chalmers "type-II phenomenon are those of subjective experience."²

Above we said we would show that every statement of identity of the type that states that there is an

¹ Chalmers, D. J., *The Problems of Consciousness*, at <http://consc.net/papers/montreal.html> extracted 12th April 2008.

² Ibid. p. 2

identity between a mental token or type and a physical token or type is false. This we will now show.

1. 1.

To begin with we need to introduce the reader to a few key terms and themes. First, we explain some of the vocabulary used below. A “designator” is a term that designates something, which is often a name or a description. A “possible world” is an imagined world which is as much like *this actual* world as possible, usually with some minimal difference we are asked to imagine it possessing, if this is possible at all. And an “identity statement” is a statement composed of a designator D_i and another designator D_n placed either side of the identity sign, and it looks something like this ‘ $D_i = D_n$.’ A “rigid designator” is a designator that designates the very same entity in all possible worlds or designates the very same entity in all possible worlds in which that entity exists and nothing in worlds in which it does not exist. The first kind of rigidity is called “obstinate” and the second type of rigidity “persistent.”³ We do not need to choose one of these alternatives because it remains unclear as to whether choosing one or the other of these alternatives carries any substantive, non-terminological, consequence.⁴ However, we will talk as if we accept the obstinate sense of rigidity (but nothing depends on this). A “non-rigid designator” (sometimes referred to as a “flaccid” or an “accidental” designator) is a designator that designates different entities in different worlds.

So consider a designator ‘A’. Take it that this designates something *George Orwell*. The designator is said to designate rigidly when the designator ‘A’ picks out *George Orwell* in all possible worlds. And the designator will be non-rigid when the designator ‘A’ does not pick out *George Orwell* in all possible worlds, but, say, picks out *the writer of 1984* (since it may have been another author in another world).

Now, consider a designator ‘B’ which designates something *Eric Arthur Blair*. Likewise, the designator is said to designate rigidly when the designator ‘B’ picks out *Eric Arthur Blair* in all possible worlds. And, again, likewise, it is non-rigid when the designator ‘B’ does not pick out *Eric Arthur Blair* in all possible worlds, but, say, picks out *the writer of Homage to Catalonia* (since it may have been another author in another world).

Kripke contends that if any two *rigid* designators ‘A’ and ‘B’ are placed either side of the identity sign, and ‘ $A = B$ ’ constitutes a true identity statement, then what they refer to must co-occur in all possible worlds. So, if ‘A’ always picks out *George Orwell* and ‘B’ always picks out *Eric Arthur Blair*, and ‘ $A = B$ ’ is a true statement, then *any* occurrence of *George Orwell* should be co-occurrent with an occurrence of *Eric Arthur Blair* in all possible worlds.

In cases where we can imagine the one without the other in some possible world, we will be imagining a lack of co-occurrence. We can infer either it is not true that the designators in question constitute a true statement of identity, or it is not true that both the designators in question are rigid designators. It follows,

³ Salmon, N., *Reference and Essence*, (Oxford: Basil Blackwell, 1982)

⁴ Stanley, J., “Names and Rigid Designation”, in Hale and Wright, (ed.), (Oxford: Blackwell Press, 1997) pp. 555-585

if the designators in question are rigid designators, we can infer that the statement of identity is false. It is necessarily false at that because identities are either true in all worlds or in none, according to Kripke. Of course, if the designators in question are not rigid, it could turn out that the statement in question is a contingent statement of identity. However, a contingent statement of identity does not state a contingent case of identity, because identities between objects are not contingent, rather a contingent statement of identity attributes contingent properties to the objects involved in the identity claim.⁵ With these concepts in hand we are now in a position to elaborate KMA.

1. 2.

So take it that 'M' is a term that rigidly designates a mental type, say, *pain*. And take it that 'P' is a term that rigidly designates a physical type, say, *the firing of C-fibres*. With the theorist of mind who wants to identify *that* mental type with *that* physical type, we may want to argue that an identity relationship exists captured in a true identity statement of the following kind 'M = P'. However, again, with the identity theorist we may also admit that it is imaginable that it is not the case that if there is an instance of that physical type, then there is an instance of that mental type in some possible world, which amounts to an imagined lack of co-occurrence of the physical and the mental in some possible world. Now, this is the key premise, this is the fulcrum of Kripke's attack, so we state it again: *We can imagine or conceive that the physical and the mental are separable*. Now, in terms of what we have just said above, either the identity statement is false or one of the two terms in question does not refer to what it picks out rigidly. That is, it is not true that (a) both the terms constitute an identity statement of the kind 'M = P' and (b) the terms refer to what they pick out rigidly. Since it is true that each of the two terms, respectively, do refer to what they pick out rigidly, the identity claim is false.

Furthermore, Kripke argues that if 'M' designates a mental type, such as *pain*, then it always designates it rigidly. So, as said, if 'P' rigidly designates its type, here *the firing of C-fibres*, no contingent identity statement can follow. However, it is conceivable that 'P' may designate a physical type non-rigidly. Then, 'M = P' will rightly be said to be a contingent statement of identity. So be it; but, then, the statement is *not* identifying *pain* with anything, but merely attributing a contingent property to the object in question.⁶

So far the argument is stated for 'M' and 'P', where 'M' designates a mental *type* and 'P' designates a physical *type*, but the argument is true for anything designated by 'M' that is a mental *type or token* and anything designated by 'P' that stands for a physical *type or token*.⁷

We conclude that using Kripke's argument that every statement of identity of the kind that states that

⁵ As Fitch says: "On Kripke's view, true contingent identity statements...are not "contingent identities". What is contingent about these statements is not the claim of identity, but rather attributions of contingent properties to the object(s) involved in the identity claim." Fitch, G. W., *Saul Kripke*, (Montreal and Kingston: McGill-Queen's University Press, 2004) p. 90

⁶ See footnote 5.

⁷ Kripke, S. A., *Naming and Necessity*, (Cambridge: Massachusetts: Harvard University Press, 1980), pp144- 155.

there is an identity between a mental token or type and a physical token or type is false. This means that the mental is not reducible to the physical. So, it seems KMA establishes some form of dualism, KMA dualism. The next thing we want to show is that if KMA is correct, then it leads to the practical irrelevance of the mental with respect to the physical, and actions in particular. This is essential to the project of section 3, showing that KMA leads to absurdity. It is not true that every dualism leads to this kind of practical irrelevance⁸, so it is essential to demonstrate that KMA dualism does. Therefore, we need to show that KMA dualism does lead to practical irrelevance. In order to show that practical irrelevance is a consequence of KMA, we need a definition of practical relevance/irrelevance, and so this is what we turn to next in 2.1. Following with an argument that connects the KMA with practical irrelevance in 2.2.

2. 1 .

Here is a rough idea about what practical relevance amounts to: Something which is useful, can be said to be something which has some kind of practical relevance defined in relation to some physical practice or process. For the purposes of this paper here is an attempt to capture this intuition:

What is "practically relevant" is something that adds a difference to a physical practice or process such that that difference is something that is understood as either (a) causing a change in the practice or process in question, or (b) registering some other non-causal difference in the practice or process in question where that difference can be designated or described.

What is "practically irrelevant" is something that either adds no difference to a physical practice or processes, or a difference to a physical practice or process such that that difference is not something that is understood as either (a) causing a change in the practice or process in question, or (b) registering some other non-causal difference in the practice or process in question where that difference can be designated or described.

We can understand this better by turning to a now traditional form of thought experiment, a "twin Earth" thought experiment. A "twin Earth" is somewhat like a possible world in that it is taken to be as much like our Earth as possible with a slight difference that we are asked to imagine. However, it isn't just any other possible world, rather it is a possible world that is taken as actual for some twin-being.

Consider the designator "water." "Water" designates H₂O in all possible worlds for an Earthling. Whereas, "water" designates XYZ for a twin-Earthling. Imagine that there is a structural difference between the objects that the two terms refer to; there are no other palpable differences. There is, to be sure, some difference between the objects in question and a difference in the terms used by the terrestrial and his extra-terrestrial twin. There is a structural and semantic difference. The question is, from the practical standpoint, whether or not there is any practical relevance. Does the structural or semantic difference add a practical difference to the relevant practices and processes?

⁸ For example, "interactionism," by definition, would not.

Let us say that the physical practices and processes amount to the set of all things that we do with water (e.g., bathing, drinking, filling and firing water pistols, creating energy, etc.) Then, there is no such difference to the physical practices or process in question. (This is by definition, since I said there was no other palpable difference than the structural difference.) And thus it seems the differences in question are practically irrelevant.

You might suppose that the structural difference in question does make a difference to the *beliefs* in question. But still the particular difference in the beliefs in question is not sufficient to bring about a change in the relevant physical practices and processes such that the difference marks a causal or describable or designatable difference in such practices and processes. It would be irrational for someone who had been using water for all the purposes we usually use water for to give up using it for such purposes on the basis that its structure was found to be different to the one she thought it had, if the change is such that it *only* renders her beliefs false or alters the nature of the beliefs that she has without any other corresponding change in the environment in which the relevant physical practices and processes occur. There simply is insufficient reason for any difference to be added to the physical realm in terms of the set of physical practices and processes that we have stated are relevant.

2. 2.

Now, we are going to argue that if Kripke's argument is true i.e., that every statement of identity of the kind that states that there is an identity between a mental token or type and a physical token or type is false, then the mental has no practical relevance. Specifically we will talk about *pain* and related physical practices and processes (assuming that *pain* is mind representative in the sense of mind given by Chalmers' "type-II" nomenclature). We are going to take actions as the relevant physical practices and processes in question. (We assume that any other physical practice or process may be substituted for actions.) Given this, the argument follows like this: All actions are physical. Actions have causes. Physical things only have physical causes i.e., the physical world is causally closed. All actions, since they are physical things, have physical causes. Pain is not physical by KMA. Thus, if I have a pain, then it follows that I have something that is neither:

- (a) A cause of an action; nor
- (b) A component of an action.

But using Kripke's argument we can strengthen the claim in the following way. We should conclude on the basis of Kripke's argument that pain *necessarily* has no physical reality. And this means that it cannot be identified in a way that would allow us to say that at least it has some physical *correlate*.

Recall, first, that Kripke's argument establishes that every statement of identity of the type that states that there is an identity between a mental token or type and a physical token or type is false. Now let's turn

our attention to correlation. As a general rule of thumb let us say that X is correlated with Y only if X differs with Y over a number of cases.⁹ Applying this to our situation, we say that M (a mental phenomenon of the type-II kind) is correlated (*not identified*) with P* (a physical phenomenon) only if M differs with P* over a number of cases (say, experiencing subjects). But if this is so, then no M is correlated with P* because it is false that that M differs with P* over any number of cases. And this is precisely because there is no valid identity statement of the 'M = P' type: When we want to correlate some mental phenomena, M, with some physical phenomena, P*, we actually correlate some physical phenomena, P, with the physical phenomena in question, P*. So, for example, when I want to correlate *feelings of pain*, M, with *stimulation of your nerve endings*, P*, what gets correlated is your *pain behaviour, reports, physiological changes* etc., P, with P*. But there is no valid identity statement of the 'M = P' type, where 'M' designates M and 'P' designates P, therefore, it is false to infer that M is being correlated with P* (we cannot substitute one for the other). So, ultimately, what we are stuck with is a correlation between physical phenomena.¹⁰

We think then, for the reasons given, it follows that we stand no chance of *correlating* the mental with the physical. Thus, if pain is mental, we claim that, given KMA, pain is neither:

- (a) Correlated with something that is the cause of an action; nor
- (b) Correlated with something that is a component of action.

Thus pain neither has a presence in the physical realm nor a legitimate correlation with a physical phenomenon. Previously we said: What is "practically irrelevant" is something that either adds no difference to a physical practice or processes, or a difference to a physical practice or process such that that difference is not something that is understood as either (a) causing a change in the practice or process in question, or (b) registering some other non-casual difference in the practice or process in question where that difference can be designated or described. If pain has no physical presence and no correlation with anything that is physical, then pain is not and has no connection with anything that is the cause or component of the physical practices and processes in question. Therefore, it adds no difference and it is practically irrelevant. On the assumption that what goes for pain, an archetype of type-II phenomena, goes for the mental as a whole, we extend the conclusion to the mental.

We have argued that if KMA is true, then there is no valid identity statement of the kind that ties mental types or tokens to physical types or tokens. This seems to constitute a dualism. Not every dualism

⁹ For example, where the X axis represents growth in GDP, the Y axis birth rate, and the cases, countries (England, Japan, Australia etc.)

¹⁰ David Chalmers suggests that we can "assume" that, say, reports of consciousness really do report consciousness, though he understands that "It's not guaranteed to be true" and goes on to say that it "introduces an element of danger into the process [of correlation]." He justifies the assumption and the element of danger that it introduces on the basis that "it seems to be an element that people can live with." (Chalmers, D.J., *The Problem of Consciousness*, at <http://consc.net/papers/montreal.html> extracted 12/04/2008 p. 6/9) This is a compromise and nothing more, *given* the problem that we have stated.

leads to practical irrelevance. But, if we have argued correctly, KMA dualism is a dualism that renders the mental practically irrelevant to physical practices and processes. We now consider the charge of absurdity.

3.

We established above that KMA dualism renders the mental practically irrelevant with respect to the physical. Now, this may seem like an unwanted result for it leads to a real concern for philosophers past, for example, William James¹¹, and philosophers contemporary, for example, Jerry Fodor. For Jerry Fodor it threatens the very essence of our understanding about who we are and our place in the world as a whole. For example, he says:

...if it isn't literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching, and my believing is causally responsible for my saying...if none of that is literally true, then practically everything I believe about anything is false and it's the end of the world.¹²

We have established the practical irrelevance of the mental with respect to the physical for KMA dualism. Practical irrelevance, as defined above, means a lack of casual efficacy, so it isn't "responsible" for the things that Jerry Fodor requires it to be "responsible" for. So, if we accept KMA dualism and the practical irrelevance of the mental with respect to the physical that comes with it, and remember by the mental we are referring to things like itching (which Fodor explicitly mentions), then we are threatening in part what philosophers like Jerry Fodor consider to be essential to their world view. This represents a real common sense concern on the part of many philosophers today. So, one likely argument against KMA dualism is the argument from absurdity, which we are now in a position to state given the results of section 2

- (a) Let's assume that KMA dualism is true, which means the mental is not physical and the mental is practically irrelevant with respect to the physical, which means it does not cause anything physical.
- (b) We assume that the mental is whatever causes us to act in certain ways, in order to meet the demands of practical relevance.
- (c) Further, we assume that physical things, like acts, only have physical causes i.e., the principle of physical causal closure.
- (d) Therefore, the mental is physical (by (b) and (c)).
- (e) Therefore, the mental is physical and the mental is not physical (by (a) and (d)).

¹¹ James, W., "Are we Automata" in *Mind* 4, pp.1 -22

¹² Fodor, J. A., "Making Mind Matter More" quoted in Kim, J., "Mental Causation" in *Op Cit.* Chalmers (eds.) (2002) p.171

We have derived an absurdity from our set of assumptions, and, therefore, we are within our rights to blame one of the assumptions for the ensuing absurdity, and so the possibility of rejecting KMA dualism as the offending assumption is a live possibility. Affirming, for example, (b) by appealing to common sense, and (c) by appealing to science, would be sufficient reason to get rid of (a). This represents our first argument against KMA. Well, then, what exactly is wrong with KMA? I think we should blame the premise that Kripke starts with, which says that we can imagine or conceive that the mental and the physical are separable.

However at best this amounts to nothing more than an indirect attack on KMA and an even less direct attack on the premise we have blamed. Furthermore, it should be clear that this objection is no objection to the epiphenomenalist (dualist or physicalist), for the epiphenomenalist is content to deny the causal efficacy of the mind altogether. For these reasons we detail another objection to the premise in question (now that we have come to suspect its legitimacy): We will try to show that the mental and physical are, in fact, *inseparable*. This amounts to a frontal attack on the premise we are trying to cast doubt upon.

4.

Let us imagine four worlds in each of which different states of being are admitted:

In W1: Both physical phenomena and type-II phenomena are present.

In W2: Physical phenomena are present, but type-II phenomena are absent.

In W3: Physical phenomena are absent, but type-II phenomena are present.

In W4: Both physical phenomena and type-II phenomena are absent.

The same set of type-II phenomena and the same set of physical phenomena are either present or absent from the worlds in question. Further, W1 can be thought of as the actual world. The actual world is a repository of the phenomena in question and whatever set of relations exist between them, if any.

Our hypothesis is that the self or person exists only when type-II phenomenon and physical phenomenon coexist. If this is correct, we will have shown that when we consider the mind in relation to the self, it is not in fact imaginable that the mental and the physical can be separated. This will mean that Kripke's starting premise can be rejected. Our argument is stated in such a manner as to appeal to the dualist and physicalist alike and we try to highlight this where possible.

Now a word about the worlds (this should help to stave off misunderstandings): W1 is neutral between a dualist interpretation (one that does not identify type-II and physical phenomena) and a physicalist interpretation (one that does identify type-II and physical phenomena). W2 and W3 are worlds which will be outright rejected by the physicalist, W3 is a world that will be rejected by some dualists, I think. W4 seems to be an impossible world. We use these world states in order to state our hypothesis: that it is impossible to conceive of world states in which both type-II phenomena and physical phenomena do not

co-exist, given we take the self or person as primitive. Therefore, it should be clear that rejection of the possibility of a world state will not harm our argument, for *a fortiori* there will not be a self or a person in such a world.

From the outset we will assume the primitiveness of the self or the person to the philosophy of mind, as P. F. Strawson asked us to do. Once we accept his request we find that that the mental and physical are in fact inseparable. We attempt to examine the worlds above in to order to make this out. I think it seems obvious that a self or person can exist in W1. W1 shares the same type-II phenomena and the same physical phenomenon as this world, so it seems obvious that a self or a person can exist. It seems even more obvious that that the self or person cannot exist in W4; that seems to be an out and out case of death; the state of being imagined in fact seems impossible, so W4 may be said to be an impossible world. Now, I want to argue, because it may not be obvious to many readers of this essay, that in W2 and W3 that the self or person cannot exist, I want to say that in such a state of being the self or the person is as dead as the self or the person in W4, and I want to emphasize this is acceptable to both dualist and physicalist alike.

Let us consider world W2, we offer two arguments for thinking that the self is lost when type-II phenomena are not present. First, type-II phenomena have been defined as subjective phenomena. To give us a sense of what this means we can list the phenomena that Chalmers does when he introduces us to his term, they include: "visual experience, auditory experience, emotional experience, imagistic experience, and so on."¹³ Therefore, when type-II phenomena are lost, subjective phenomena (experiences and feelings) go with them. Then, if the self exists only if it has subjective phenomena (experiences and feelings), as one of a conjunction of logical necessities, then the self does not exist when its type-II phenomena fails to be present. So the question now becomes is subjective phenomena really necessary to selfhood? This question is surely to be answered in the affirmative because it is part of the very meaning of the self that she have a subjective nature: subjectivity, experiences and feelings, and, therefore, what we are referring to as type-II phenomena.

Let me note that *both* dualists and physicalists can accept that subjectivity is part of the meaning of the self. What they disagree about is whether subjectivity in the sense defined can be reduced to the physical. The point can be put like this: for the physicalist world W2 = world W1 (if W2 is the same as W1 in all physical respects), but for the dualist world W2 ≠ world W1 (even if W2 is the same as W1 in all physical respects). It is not that they must disagree that subjectivity (experience and feeling) is essential to the self, rather they disagree on whether W2 is different from W1. *Given our definition of W2* (which says that type-II phenomena is absent from W2, and remember we are assuming that the physical phenomena in any of the worlds we are considering, if present, is the same as the physical phenomena in W1, and thus the actual world), the physicalist denies W2 is a possibility, the dualist accepts W2 is a possibility. Either way they can agree that subjectivity (experience and feeling) is essential to the self, and that in such a world the self

¹³ Chalmers, D.J. (1998), *The Problem of Consciousness* at <http://consc.net/papers/montreal.html> extracted 2008/04/12

is not realized—for the physicalist because W2 is impossible, for the dualist because the world lacks type-II phenomena.¹⁴

Second, even if one could formulate a sense of subjectivity in the world we are considering, it would be no ordinary sense or concept of subjectivity. That is, it would refer to nothing that we would understand as subjective precisely because there is no subjectivity as we understand it in the world being described. The residents of such a world, therefore, would not be able to understand what we meant when we referred to pain, itching, feelings of nostalgia, being in a funk, etc. as we understand such references. They, then, would not be able to understand the self in the way that we do. And if they did have a term “self”, the meaning of that term would be different from the meaning of our term. Thus, we can say, that by the lights of our ordinary sense of self, there is no self in W2.

Again, we can put the case in terms of how the physicalist and dualist will understand the argument. For the physicalist recall subjectivity is identical to certain physical realities. Therefore, if subjectivity doesn't exist the physical realities cannot exist, and W2 is impossible. Therefore, a sense of subjectivity contrived for a world in which neither kind of phenomena are present, a world like W4, is either absurd or radically different to the sense of subjectivity the physicalist gives us. Even if we imagined that given subjectivity is missing in W2 that the relevant physics is missing in W2, but not that all physics is missing in W2 such that a sense of subjectivity can be contrived for some set of physical realities in W2, we would be talking about two different senses of subjectivity: the one the physicalist maintains (subjectivity (= experience and feeling) = the relevant physics) and the contrived sense (subjectivity (≠ experience and feeling) = some other physics).

For the dualist the argument should be even easier to accept. To put it in Nagelian terms, subjectivity amounts to what it feels like to be you. What it feels like to be you is constituted by your type-II phenomena, by your experiences and feelings or by the way we experience and feel. A world minus type-II phenomena is a world in which there is nothing it feels like to be you, so the sense of subjectivity defined in or for that world must be radically different to the sense that Nagel tells us it has for us.¹⁵

For reasons such as these we conclude that W2 is a world in which the state of being considered surely results in the loss of the self or person for both physicalist and dualist. We now turn to the state of being which we envisaged in W3.

There are three arguments for why the self would be essentially diminished in a world like W3. The first argument simply denies that a world like W3 could exist. If we understand experience as dependent on a physical basis, as most philosophers of mind do, then the state of being in W3 could not possibly arise. Rather, it is the same as the state of being (or non-being) in W4. And, thus, the person certainly cannot occur. But should we understand experience as dependent on a physical basis? Well, we might not, but

¹⁴ You might have noted the possibility or not of W2 revolves around the possibility or not of Zombies. We are saying, either way, if they are possible (as the dualist maintains) or impossible (as the physicalist maintains) they do not have selfhood.

¹⁵ See Nagel, T., “What Is It Like To Be A Bat”, *Philosophical Review* LXXXIII, 4, (October 1974), pp. 435-50

most philosophers of mind do: the epiphenomenalist (dualist or physicalist), the reductive-physicalist and the physicalist must!

The second argument turns to criteria of personal identity. We note two of four from Searle:

1. "Spatio-temporal Continuity of Body", by which he means that a person at 90 is continuous in space and time with the infant she was some 90 years previous.
2. "Relative Temporal Continuity of Structure", by which he means that a person who has wholly changed bodily form is still the person she was before the total change.¹⁶

These first two criteria are spatial-temporal criteria involving the continuity of the body (condition 1) and the continuity of structure (condition 2). But we have said that there is no physical reality in W3, and, therefore, it follows that there is no spatial-temporal reality and no continuity of body and no continuity of structure. Further we can note that *objects* are what are identified as the same persons. It is plausible to consider objects in physical terms. Therefore, again, since there is not physical reality, there are no persons. For these reasons, the relevant criteria for the self cannot be satisfied and, therefore, there is no self in W3.

The last argument is set at the conceptual level. When we imagine that only type-II phenomena are present and physical phenomena are lacking, anything that can be given a physical definition, also, falls out of the picture. Type-II phenomena as we have understood the term, and in line with how the term appears in Chalmers, does not include functionally definable phenomena. It is generally thought that beliefs, desires, and conceptual phenomena in general can be understood as functionally definable phenomena, and it is generally considered that functionally definable phenomena are to be understood, at least in part, in physical terms.¹⁷ Both physicalist and dualist can accept this. But, if this is so, a problem arises. Given that only type-II phenomena exist and there are no physical phenomena, then the intentional and attitudinal states referred to above are missing. This is to say that there is no belief or no sense of the type that identifies or misidentifies a pain; no belief or sense that relates to the *same* or *different* type-II phenomena, and no belief or sense that unifies or discriminates internal "my" or external "your" type-II phenomena from each other etc. There is no sense of a subjective nature, no sense of unity "I" and no sense of difference "you". Further neither the residents nor we can simply impose a sense of self on the residents of this world from the outside because there simply is no outside on which to base that sense of self! It is clear that any sense of self that can be predicated of anything in this world will be totally alien to our way of thinking; in fact, it seems to be impossible.

Briefly, we note that the idealist may object to all three of the arguments presented above. However, an idealism that would answer the three objections would need to persuade us that the mental could exist

¹⁶ Searle, J. *Mind*, (Oxford and New York: Oxford University Press, 2004)

¹⁷ Block, N. *Troubles with Functionalism* extracted from <http://w3.uniroma1.it/cordeschi/Articoli/block.htm> for example, p. 1 and p.4

without a physical base; would need to define space, time, body, structure and objects in non-physical terms; and would need to tell us how type-II phenomena produces and maintains attitudinal and intentional states and the conceptual base in general. That is, the idealist needs to tell us how all the *physical* requirements are to be met in a world in which none of the physics exists and she needs to do this in a convincing way and she can only appeal to type-II phenomenon. I do not say that this is impossible; however, I do not think there is any idealism of this kind held by any of the philosophers of mind writing today. Furthermore, the arguments presented here are aimed at the dualist and the physicalist, and I do not think they would accept an idealist line of attack on the arguments that have preceded us if the idealism in question is restricted in the way stated. So, for the general community of philosophers of mind our arguments will not be denied for idealistic reasons.

For the reasons stated we conclude that W3 is a world in which the state of being considered results in the loss of the self or person. We summarize our results below.

The state of being in W2 is bad because we lose type-II phenomena. The state of being in W3 is bad because we lose physical phenomena. If we lose either type-II phenomena or physical phenomena the self is lost or essentially diminished. But, this shows that the conjunction of type-II phenomena and physical phenomena is necessary to the self, and that the denial of one is sufficient to negate the self. We conclude that the only world in which the self or the person exists is the world in which both type-II and physical phenomena co-exist. We posit the following condition to reflect this result: *The self or the person exists in a world, W, only if there is the conjunction of type-II phenomena and physical phenomenon in W.*

And, now, it is clear that if we understand the notion of the self or the person as primitive to our theory of mind, KMA dualism does not get going because we cannot imagine or conceive worlds in which the mental and the physical are separable, to deny either would be to deny the self. And, so, Kripke's thought that we can conceive of some possible world in which the presence of the physical is unaccompanied by the presence of the mental is a non-starter. *Of course, that does not mean that the mental is identical to the physical.* (So the dualist, too, can accept our arguments against KMA). And, also, *it does not mean that they are not identical.* (So the physicalist, too, can accept the arguments against KMA.) Nothing has been said about the specificity of the fashion in which the mental appears with the physical. Further argument is required to establish just what kind of relation (an identity relation or some other) exists between the two.

To sum up, in this paper we considered Kripke's Modal Argument, which we have shown leads to a form of dualism, KMA dualism. We argued that KMA dualism, in turn, led to the practical irrelevance of the mental with respect to the physical. We argued for this in order to develop an objection to KMA based on the result cited. That objection argued from the absurdity of the result to the rejection of KMA dualism. We blamed Kripke's beginning premise, which said that it was imaginable or conceivable that the mental and the physical were separable. We recognised that this argument was an indirect attack on KMA and may not appeal to the epiphenomenalist. So we also raised an objection that attacked the premise more directly. By taking the person as primitive, we argued that the mental and physical were not separable. Without this

assumption KMA could not get started and so we took our objection as blocking KMA dualism from the start. We concluded that this did not mean the mental was identical to the physical, nor that this was not the case. That question cannot be answered in this paper.

(ルーク・マリク 哲学哲学史・博士後期課程)

Two Ways To Reject Kripke's Modal Argument: Absurdity, and The Primitiveness of the Person Luke Malik (Osaka University)

There are many arguments against identifying the mental and the physical. One of these arguments comes from the work of Saul Kripke. In *Naming and Necessity*, his *tour de force*, we find Kripke's Modal Argument against the identity theory of mind, the theory that advances the notion that the mental is identical to the physical. Kripke starts from the premise that we can imagine or conceive possible worlds in which, say, the so-called physical correlate of pain could exist without the sensation of pain. From this premise he argues to the conclusion that mental tokens or types are not identical to physical tokens or types. And, thereby, defeats the notion of identity that the identity theorist advocates. However, in this paper, we show there is reason to doubt the aforementioned premise: First, Kripke's argument ends up leading to a form of dualism that denies the practical relevancy of the mind with respect to the physical, which may lead to absurdity, the blame being put on the premise. Second, if we take the self or the person as primitive, then it is not possible to imagine worlds in which the mental and the physical are separable, so Kripke's opening gambit is just false. So, we offer an indirect and direct attack on Kripke's contribution to the philosophy of mind.

【キーワード】

Kripke, Mind, Dualism, Absurdity, Persons