

JAEGWON KIM
ON MENTAL CAUSATION

Nebojsa Delic

A Thesis

Submitted to the Faculty of Graduate Studies

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF ARTS

Department of Philosophy

Faculty of Arts

University of Manitoba

Winnipeg, Manitoba

© November 2003

THE UNIVERSITY OF MANITOBA
FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION

JAEGWON KIM
ON MENTAL CAUSATION

BY

Nebojsa Delic

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of
Manitoba in partial fulfillment of the requirement of the degree
Of
MASTER OF ARTS**

Nebojsa Delic © 2004

**Permission has been granted to the Library of the University of Manitoba to lend or sell copies of
this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell
copies of the film, and to University Microfilms Inc. to publish an abstract of this thesis/practicum.**

**This reproduction or copy of this thesis has been made available by authority of the copyright
owner solely for the purpose of private study and research, and may only be reproduced and copied
as permitted by copyright laws or with express written authorization from the copyright owner.**

ABSTRACT

Events are causes for Kim. He believes that there are events as particulars as well as types. When a mental event is a cause of a physical movement, it presumably is so in virtue of belonging to a mental type. On the assumption that mental and physical event types are distinct in conjunction with the assumption that causation is closed under physical events, Kim argues that it follows that mental events cannot be causes. His remedy consists in the identification of mental event tokens and types with physical tokens and types. This identification gives mental events the status of causes, and it provides mental types with explanatory relevance.

For Davidson, on the other hand, causation is one thing and explanation another. Mental events may be reasons but they remain causes. That a mental event is a cause and that it falls under a causal law emerges when it is re-described as a physical event. And when an event is described as a reason its proper framework of explanation is the explanatory pattern of psychology. This means that neurological and psychological explanations remain distinct even though they are about one and the same event particular.

Kim is open to at least two forms of critique. His understanding of closure seems too strict, and it seems to beg the question against forms of non-reductivist physicalism. Closure can be understood in such a way that mental events can also be causes of physical events. The second form of critique has to do with Kim's conflating of relations that cannot be confused. The identity relation is symmetrical whereas both, the supervenience and the realization relation, are asymmetrical.

DEDICATION

TO LILIANA

ACKNOWLEDGMENTS

My thanks to my thesis committee members, Drs. Tim Schroeder, Jack Bailey and Jason Leboe, for their support and valuable input while this work was in progress.

My very special thanks to my thesis advisor, Dr. Tim Schroeder, whose knowledge, suggestions, and enthusiasm while reading hundreds of pages, carried me through a maze of philosophical reasoning to a happy ending in the form of this thesis.

Finally, my most heartfelt thanks and appreciation to my wonderful wife, who always supported me and believed in me.

TABLE OF CONTENTS:

ABSTRACT	i
DEDICATION	ii
ACKNOWLEDGMENTS	iii
TABLE OF CONTENTS	iv
INTRODUCTION	1
CHAPTER ONE	
Events and Causation	5
CHAPTER TWO	
Kim on Mental Causation	31
CHAPTER THREE	
Davidson on Mental Causation and Kim's Replies	52
CHAPTER FOUR	
Critiques of Kim's View	85
BIBLIOGRAPHY	118

INTRODUCTION

Ordinarily we tend to believe that the goings-on in our minds make us do things. Intuitively we want to say that our desire for a coke made us go to the fridge and fetch it. The desire is thought to be an attitude of our minds and not of our brains; it is thought to be mental and not physical. So when we talk about mental causation, what we have in mind is the fact that it is our desire for a coke that makes us go to the fridge and get it. We don't want to say that what made us go to the fridge was something that happened in our brains, although what happens in our brains may in some sense be relevant to what is going on in our minds.

So, generally speaking, the problem of mental causation refers to whether the diverse happenings (such as desires, beliefs, intentions, decisions, pains, wants, hopes, and so on) in our minds really cause us to act in various ways. It may also be the case that we were wrong all along in believing that mental happenings make us do things. If that should be the case, then what we are left with is goings-on in our brains that cause us to perform all our movements. This would be a very depressing conclusion since, if true, it would wipe away the ground for believing that we are free agents. On that account we would just be machines determined by forces inside our brains and factors within our environment.

The following work will investigate (in chapter 2) how our intuition that mental happenings are causes fairs in Jaegwon Kim's *Mind in a Physical World*. This monograph is interesting and important because it points to problems with regard to mental happenings as causes. Specifically he argues that there may be reason to believe that desires, beliefs, intentions, and the rest of our mental happenings, cannot cause us to

do anything. He believes that this appears to be a consequence of the fact that our world is a physical world. What this means is that all goings-on in our world are physical. And if all causes in our world are physical, then it seems that desires, hopes, thoughts, being mental, are no causes at all. If this were true all the causal work of our bodily movements would be initiated by whatever is going on in our brains.

This is a conclusion that Kim is drawn to. He provides arguments, which seem to show that beliefs, fears, knowledge, and the like cannot cause us to do anything, because these happenings are mental. And in our world, which is physical, it is physical goings-on in our brains that are causes of our bodily movements and not the mental happenings. Mental happenings are excluded as causes in a purely physical world. But this conclusion, if true, and some of us would agree, seems unacceptable because it would mean that we have never been free agents, and that we have always been puppets in the causal network of brains and environments.

But there is no need to despair, because Kim has a remedy. He argues that the mental happenings are identical with the physical goings-on in our brains. And since they are identical with these goings-on, they are causes. Say my desire for a coke is identical with the firing of a-fibers in my brain. This identity guarantees that my desire is a cause of my going to the fridge. Kim's good news is that mental causation is saved as physical causation. The bad news is that our minds have no causal powers independent from the causal powers of brains. In other words, our minds (meaning the mentioned mental happenings) cannot cause us to act in various ways just as minds but always as being identical with brains. And since mental and physical properties are identical, on Kim's view, we don't only act as physical beings.

Unfortunately, Kim's view seems subject to two decisive criticisms (see chapter 4). The first criticism attacks Kim's argument which attempts to prove that happenings such as desires, beliefs, hopes, and so on cannot be causes if they are understood as being purely mental entities. The proponents of this particular critique argue instead that mental happenings can also be causes. If Kim's argument turns out to be faulty, then he cannot use it to say that mental happenings cannot be causes of our behavior. And the second criticism attacks Kim's identifying the mind with the brain. Kim is confused if he simultaneously holds his belief in mind-brain identity as well as his beliefs that the mind depends on the brain.

Prior to elaborating on the two mentioned criticisms of Kim's views in chapter 4 I will be concerned with the explanation of Kim's understanding of mental causation in chapter 2 and an explication of Donald Davidson's views on the same topic in chapter 3. The problem of mental causation – as regarded by both philosophers (Kim in chapter 2; and Davidson in chapter 3) – depends on how both philosophers understand the problem of causation per se. And in chapter 1 the reader will find explanations and comparisons of both Donald Davidson's and Jaegwon Kim's views on causation.

The debate about causation (in chapter 1) includes matters regarding what things are causes and effects. Both philosophers agree that it is events (happenings or goings-on above) that are causes and effects. It is the event of the gun's firing that causes the event of Peter's death. But there are also differences between the two philosophers regarding the matter of causation, which will have repercussions later on in chapters 2, 3, and 4.

The reason why Donald Davidson's views matter and why they are contrasted with Jaegwon Kim's (especially in chapter 1 and chapter 3) has to do with the fact that

Davidson is the philosopher that Jaegwon Kim always responds to. It was Davidson who created the program of the philosophy of mind in the 1970's. He was considered a leading figure in the debate about mental causation. It was Davidson who argued that our reasons (beliefs, desires, intentions, hopes, etc.) can also be causes. He argues that my desire for a coke explains my behavior of going to the fridge. But my desire for a coke is also a cause of my going to the fridge when it is described as a firing of the a-fibers of my brain.

So this is all about believing and desiring and acting, and how the view Kim has developed about how the first two cause the third, while clever, just won't work.

CHAPTER ONE

Events are things like falls, crashes, births, deaths, wars, striking of matches, burning of fire, sunrises and sunsets, explosions, and so on. I will explain that the main reason why Davidson believes that events exist has to do with considerations in logic. Davidson thinks that events can be named: ‘the birth of Peter’, ‘the death of Paul’, ‘the explosion of the plane’, ‘the striking of the match’, etc. They stand in causal relations to one another: the striking of the match caused the lighting of the match. I will then elaborate on what it means when physical events in causal relations instantiate laws. This will be followed by Kim’s understanding of events, the causal relation of events, and the instantiation of laws. A crucial difference between the two philosophers’ views is that Davidson believes that there are events as particular entities whereas Kim believes that there are events as particular entities as well as that there are types or kinds of events. This difference will be of importance for Kim’s (chapter 2) and Davidson’s (chapter 3) philosophies of mind.

Sentences such as ‘Jimmy played the guitar’, ‘Peter ran a marathon’, ‘the plane crashed into the building’, ‘the star exploded’, ‘the apple fell’, seem to be about things as well as about happenings or changes. Such an intuition may be explained on a common sense level in the following vein. In the case of ‘Jimmy played the guitar’ it could be said that there was this person called Jimmy, and there was this thing we usually call a guitar, as well as that there was something such that it was a playing of the guitar by Jimmy. Similarly, there was a star and there was an exploding of it; or, the apple and the falling of it; or, the plane, the building, and the crashing of the plane into the building.

According to Donald Davidson, if we think clearly enough about these sorts of sentences, we can get a very clear understanding of happenings, and so of causes and effects.

Usually there are names or descriptions that refer to or are about things that they name or describe, such as ‘Jimmy’, ‘guitar’, ‘plane’, ‘building’, ‘star’, and so on. But no such names or descriptions seem available for events such as playing, running, and crashing, because verbs (played, ran, crashed) do not seem to name any entities. But Davidson believes the contrary, that such ‘singular terms’, as he calls them (1980, essay 8, p. 164), are at hand. ‘Jimmy’s playing the guitar’ or ‘the playing of the guitar by Jimmy’, as well as ‘the plane’s crashing into the building’ or ‘the crashing of the plane into the building’ are phrases that do explicitly name events.

But “...the existence of these singular terms is of uncertain relevance until we can firmly connect such singular terms with sentences like ‘Vesuvius erupted in 1906’” (1980, essay 8, p 164). This means that a connection has to be found between ‘the playing of the guitar by Jimmy’ and ‘Jimmy played the guitar’. On the one hand, there are these singular terms that seemingly refer to events, but on the other hand, no traces of such are visible in ‘Jimmy played the guitar’. If no connection can be found between phrases that seemingly refer to events and no-event statements like ‘Vesuvius erupted in 1906’ or ‘Jimmy played the guitar’, then events may not even exist.

The reason why it may be thought that there is no connection between no-event sentences (‘Jimmy played the guitar’) and event phrases (‘The playing of the guitar by Jimmy’) can be traced to the fact that in sentences of the first sort there are only two names, namely ‘Jimmy’ and ‘guitar’, and a verb (such as ‘played’) that relates those two names to one another. In this case the verb is not considered to name a particular entity in

the world. As Davidson says, "...we would normally suppose that"..." 'Jimmy played the guitar' ..."consisted in two names and a two-place predicate. I suggest, though, that we think of "...‘played’ ... “as a three-place predicate ...” (1980, essay 6, p. 118). The verb ‘played’ is the predicate. ‘Jimmy played the guitar’ is understood as this: ----played----. We see two places. ‘Jimmy’ occupies the first, and ‘the guitar’ occupies the second. But Davidson suggests that the predicate ‘played’ should relate three objects, or it should have three places: ‘Jimmy’, ‘the guitar’, and the event of ‘the playing’.

This suggestion means that the desired connection between no-event and event-sentences can be had if events as entities are assumed to exist. And in the case of our example there are then three things in the world named by the different parts of the sentence ‘Jimmy played the guitar’. On Davidson’s view, there are three entities here: Jimmy, the guitar, and his playing the guitar. The same idea is expressed in the following: “In this way we provide each verb of action or change with an event-place; we may say of such verbs that they take an event-object” (1980, essay 8, p. 167).

Davidson’s event-introduction is simultaneously followed by a different way of analyzing the logical form of sentences such as ‘Jimmy played the guitar’. A predicate that would be analyzed as having two places occupied by names such as ‘Jimmy’ and ‘guitar’ has now three places naming an event (playing) as well. The structure of the whole sentence does not depend solely on the names (Jimmy, guitar) referring to those entities that they denote, but also on an additional element not found in the sentence-composition, namely ‘the playing’.

The usual analysis of the logical form of the sentence ‘Jimmy played the guitar’ as a two-place predicate conceals, according to Davidson, the real logical form of that

sentence. For this sentence to be about an event the sentence should be read and rendered as ‘There is an event x such that x is a playing of the guitar by Jimmy’, or $(\exists x) (\text{Played}(x, \text{guitar}, \text{Jimmy}))$ (1980, essay 6, p.118). On this account it seems obvious that there is a connection between the sentence ‘Jimmy played the guitar’ (a seemingly no-event sentence) and the phrase ‘Jimmy’s playing the guitar’, because they are about events.

It is now time to focus our attention on what I believe to be the motive for why Davidson accepts events as entities in the world. Davidson does not seem to want to accept the idea of the existence of events without explaining the reasons why events may exist. The intuition that events may exist is not enough for him. That events exist Davidson wants to prove by pointing to a problem in logic and by trying to overcome it. I will call it the entailment problem. There are sentences in ordinary language that entail others that are parts of the former and larger sentence. And as Davidson says: “This requires, it would seem, that the patent syntactical fact that the entailed sentence is contained in the entailing sentence be reflected in the logical form we assign to each sentence” (1980, essay 8, p. 166).

Let us try to explain the problem using an example. Assume that the following sentence is true. ‘Bob walked through the streets of Berlin at 2 p.m.’ It is a sentence that seems to consist of the following components: Bob, walked, through the streets of Berlin, at 2 p.m. The word ‘Bob’ plays the grammatical role of the subject of the sentence; ‘walked’ indicates a verb that expresses what Bob did, an action of Bob’s; ‘through the streets of Berlin’ is the prepositional object indicating a place where Bob’s walk occurred; ‘at 2 p.m.’ has the grammatical role of being a temporal object.

If Bob walked through the streets of Berlin at 2 p.m. then it is also the case that Bob walked through the streets of Berlin, and it is also true that Bob walked. The larger and first sentence entails the second and the third sentence. The reason why the first sentence entails the second and the third is because of the grammatical elements all three sentences share in common. The second sentence has almost all grammatical components that the first has except the expression playing the grammatical role of a temporal object. The third and shortest sentence shares two grammatical components with both the first and the second sentences. This obviously has to do with the fact that the third sentence is entailed by both the first and the second one. The entailment relation of entailing and being entailed seems to depend on the fact of sharing grammatical components in common, at least in this case (1980, essay 8, p. 166).

So far I have only explained the “patent syntactical fact that the entailed sentence is contained in the entailing sentence...” (1980, essay 8, p. 166). But now we have to show how the syntactical fact of entailment is “...reflected in the logical form we assign to each sentence” (1980, essay 8, p. 166). In other words, we have to show what the logical forms of the three sentences above are and that the entailment relation holds among these sentences just because they are logically related.

It is at this point that the entailment problem emerges. The reason lies in the fact that “...the usual way of formalizing these sentences does not show any such feature” (1980, essay 8, p. 166). This means that logic has no way of showing that the three sentences share a common logical feature or structure. And this also means there is a mismatch between syntax on the one hand and logic on the other. According to syntax the

first and largest sentence entails both the second and the third due to the having of grammatical components in common.

The reason why logic cannot account for the entailment relation is due to how these three sentences are usually analyzed. According to Davidson, logic "...directs us to consider the first sentence as containing an irreducibly three-place predicate 'x walked through y at t' while the second contains the unrelated predicate 'x walked through y'" (1980, essay 8, p.166). We can formalize 'x walked through y at t' as: $Wxyt$. The predicate 'W' has three objects, namely x, y, and t. But 'x walked through y' we formalize as: Zxy . The predicate 'Z' has two objects instead of three, namely x and y. This means that ' $Wxyt$ ' and ' Zxy ' are logically distinct relations. Since they are distinct ' $Wxyt$ ' does not logically entail ' Zxy '.

Davidson believes that there is a way to dissolve the problem, that there is a way to provide for the logical structure of one sentence to entail the logical structure of the other. He proposes to "...legitimize our intuition that events are true particulars by recognizing explicit reference to them, or quantification over them, in much of our ordinary talk" (1980, essay 8, p. 166). He wants to "...provide each verb of action or change with an event-place; we may say of such verbs that they take an event-object" (1980, essay 8, p. 167). This means that verbs such as fall, crash, explode, play, capsize, run, and so on, can introduce events as entities existing out there in the world. On this view, there are not only things like Jimmy, planes, buildings, canoes, but also things like falls, crashes, explosions, etc. While assuming that events exist, Davidson simultaneously adjusts the logical analysis of statements that are about events in that events can have variables such as 'x' and 'y' ranging over them. Now it is possible to say 'There is an

event x...' or ' $\exists x (x\dots)$ '. Quantification over events, as it is called, allows us to express the fact that there are events.

The entailment problem that logic faced because of its ways of formalizing sentences such as 'Bob walked through the streets of Berlin at 2 p.m.' in terms of 'x strolled through y at t' is now overcome by Davidson's proposal. Our example can now be rendered as 'There is an event x such that Bob walked x, x took place in the streets of Berlin, and x was going on at 2 p.m.'. This also helps to express the fact that several qualities are predicated of such an event. We say (or predicate) of such an event that 'Bob walked it', that 'it took place in Berlin', and that 'it was going on at 2 p.m.'.

Davidson relies on the following simple logical entailment. 'Bob walked through the streets of Berlin at 2 p.m.' we can express as: $\exists x (Px \ \& \ Qx \ \& \ Rx)$. 'Bob walked through the streets of Berlin' can be rendered: $\exists x (Px \ \& \ Qx)$. And 'Bob walked' is expressible as: $\exists x (Px)$. We express the logical entailment relation as this: $\exists x (Px \ \& \ Qx \ \& \ Rx) \Rightarrow \exists x (Px \ \& \ Qx) \Rightarrow \exists x (Px)$. It seems that the repetition of the 'x' after the brackets (in the quantificational form of the statement) – after $\exists x (\dots)$ or after 'There is an event x' - allows for the required entailment to occur. The statement 'There is an event x such that Bob walked x, x took place in the streets of Berlin, and x was going on at 2 p.m.' clearly entails the two shorter statements 'There was an event x such that Bob walked x, and x took place in the streets of Berlin' as well as 'There was an event x such that Bob walked x'.

I believe that the main reason why Davidson wants to introduce events as entities in the world stems from the entailment problem shown to exist in the domain of the analysis of the logical form of our examples as well as the need to remedy it. In other

words, the restoration of the entailment of the logical form of one statement by the logical form of another can take place only on the assumption that events exist. So, the ontological assumption that there must be events is driven by logical considerations regarding the entailment of some statements sharing a common logical structure (parts).

Having solved the problem of the logical entailment of statements by introducing events, we can now use singular terms to refer to events in a sentence like ‘Bob’s falling off the ladder caused Bob’s breaking of his leg’. Here we have phrases such as ‘Bob’s falling off the ladder’ as well as ‘Bob’s breaking of his leg’, which refer to particular events, in terms of one being the cause of the other. Thus we are finally ready to start talking about causes. Some of Davidson’s examples are ‘The short circuit caused the fire’ and ‘The flood caused the famine’ (1980, essay 7, p. 155). ‘The short circuit’ describes an event that is the cause of the event described as ‘the fire’; ‘The flood’ describes an event that is the cause of the event described as ‘the famine’.

Events as causes and effects exist out there in the world, according to Davidson. They remain the same regardless of how they are described. Some ways of describing them seem covert and general (‘The fact that there was a short circuit caused it to be the case that there was a fire’), others are overt and they seem to refer to particular events (‘The short circuit caused the fire’), and some appear more detailed than others. To explain what I mean by the latter I will use Davidson’s example ‘The cause of this match’s lighting is that it was struck’ (1980, essay 7, p. 155). In other words, ‘The striking of the match caused the lighting of the match’. Two singular terms (‘the striking’ and ‘the lighting’) referring to events name or describe events that are related to one another as cause and effect. It is one thing to say that one event was the cause of another,

and yet it is another thing to say how events are ‘characterized’, according to Davidson (1980, essay 7, p. 155). It is possible to broaden the description of the cause by saying that in addition to the striking of the match the striking itself was against an appropriate surface. Also, the match was dry. It is highly unlikely that a wet match can be lit. There was enough oxygen; without oxygen the match cannot catch fire.

In the case of our example all these characterizations of the cause hold. This particular match was struck against an appropriate surface; it was dry; there was enough oxygen. Regardless of whether the characterization of the cause is broader or narrower it still remains true that this particular striking was the cause of this particular lighting. The characterization of the cause as ‘the striking’ is partial but the causing event itself is not partial, because it is true for this event that there was enough oxygen; that the match was dry; and so on. The striking of the match was the whole cause, whereas the description of the cause as ‘the striking’ was partial (1980, essay 7, pp. 155, 156). To make things easier to understand let me propose an example. I am sitting at my kitchen table. I would like to light a candle. There is a brand new matchbox on the table. I inspect the surface against which matches are usually struck. I decide that this surface is pretty rough. From experience I know that a rough surface is important for a good match to catch fire. I pull out a match and disappointingly realize that it is just a piece of wood without the red and round top covered by phosphorus. I grab another one and gently touch its top. The phosphorus crumbles. I draw a third match and repeat the previous procedure. This one is all right; the phosphorus does not crumble. I know that a very slight touching of the matchbox surface will not be helpful in the match’s catching fire. So I decide to apply

some force in pressing the match against the matchbox surface. I finally move the good match across the matchbox surface. The match catches fire.

This striking of the match by me is the cause of the match's lighting. This particular striking can be characterized or described in a less partial or a broader way by mentioning all the things I said in my small story including the presence of oxygen. I can infer, even without the knowledge of the laws of the science of chemistry, that a match will light if it is a good match, if the matchbox surface is rough, if there is enough oxygen, and if the match is moved across the matchbox surface with some force. So if I know all these descriptions of the situation (that my match is good, that there is enough oxygen, and so on) before I actually strike the match against the matchbox surface I can conclude that the match is most likely going to catch fire. I can claim this on the basis of these descriptions that I am aware of and a rough causal story based on past experience and observation. And it is now that we can better understand Davidson's claim that "...we must distinguish firmly between causes and the features we hit on for describing them, and hence between the question whether a statement says truly that one event caused another and the further question whether the events are characterized in such a way that we can deduce, or otherwise infer, from laws or other causal lore, that the relation was causal" (1980, essay 7, p. 155).

Even without knowing the laws of the science of chemistry but provided that I know that there exists a cause event – such as the match's striking – and given the above very broad description of this cause event, I can infer that the lighting of the match (the effect event) could not have occurred without the match's striking. This means that the striking was enough or 'sufficient', as Davidson would say, for the match's lighting. This

also means that all the other descriptions such as ‘that there was oxygen’, ‘that the match was a good one’, ‘that the matchbox surface was rough’, even conjointly but without the description ‘the striking of the match’ are not enough or sufficient to help me infer that the match will light. This shows that for a cause to be a cause it must be sufficient for the effect it brings about (1980, essay 7, p.158).

The effect event can be described in a broader way as well, according to Davidson. Maybe instead of the short description ‘the match’s lighting’ I can broaden it by saying ‘that a flame is almost immediately present’, ‘that the flame emits heat and light’, or ‘that the air around the phosphorus head changes’, ‘that the phosphorus head is not cool but glowing’, ‘that a particular smell is present usually only experienced when phosphorus is burning’. All my descriptions may be very clumsy and non-scientific, but this should not prevent us from realizing that, when the effect event is given a broader description, the cause event description ‘the striking of the match’ is one of several components equally important for the effect of the lighting to occur. In Davidson’s words, the cause event is ‘necessary’ as well (1980, essay 7, p. 158). Viewed from a broad description of the effect event the presence of oxygen is as necessary as a good match, and a rough matchbox surface is as necessary or important as the striking of the match against the matchbox surface. So it turns out that, when the cause and effect events are described non-partially or broadly, the striking of the match is sufficient as well as necessary for the effect event to occur.

We can now formulate a causal law for the striking and the lighting of the match. The law is a linguistic entity. It consists of two statements tied by conjunction (&). The first statement expresses sufficiency the second one necessity of the cause event by

utilizing descriptions of cause and effect events. I think it would not be wrong to explain Davidson's point (1980, essay 7, p. 158) as this: Whenever there is a striking of a good match in enough oxygen, with some force, against a rough matchbox surface (and so on), at some time, then there is an event of a lighting of a match, which occurs later, and the first event is a cause of the second, and (this 'and' conjoins the 'sufficiency' statement with the 'necessity' statement, which follows now) whenever there is a lighting of a match the flame of which emits heat and light, the phosphorus head of which is not cool but glowing, the particular smell of which is present (and so on) occurring later, then there is a striking of a match that occurred earlier and the striking of a match was a cause of the lighting of the match (1980, essay 7, p. 158).

If we know the event described as 'the striking of the match at time t' to exist, and if we also know the causal law consisting of the sufficiency and necessity statements, then we can infer the singular causal statement 'the striking of the match at time t' caused 'the lighting of the match at time t' (later). 'The striking' and 'the lighting' of the match are physical events, or we describe them as physical events. And the former causes the latter. Davidson says that "...where there is causality, there must be a law: events related as cause and effect fall under strict deterministic laws" (1980, essay 11, p. 208). This means that when the physical events 'the striking' and 'the lighting' are related to one another as cause and effect events they instantiate the above causal law. In other words, they realize it. The expression of the singular causal statement implies that there is a causal law whether we know the sufficient and the necessary conditions or not. That one event is the cause of another is left unexplained. Davidson says, "...I have abjured the analysis of the causal relation" (1980, essay 7, p. 158).

It is now time to turn to Jaegwon Kim's analysis of events. The events of Bob's falling, the plane's exploding, and Jimmy's playing the guitar Kim analyzes as follows: $[(\text{Bob}, t), \text{falls}]$; $[(\text{plane}, t), \text{explodes}]$; $[(\text{Jimmy}, t), \text{plays}]$ (1973, p. 222). Whatever stands within these brackets – e.g. Bob, t, falls - are elements out of which this particular event of Bob's falling is constituted. All elements together are the event itself. The particular event of Bob's falling exists only if Bob falls at t. In other words, it exists only if the object Bob has the property of falling at time t, or, if the thing Bob exemplifies (as Kim says) the property of falling at time t. Kim says: "Events, therefore, turn out to be complexes of objects and properties, and also time points and segments, and they have something like a propositional structure; the event that consists in the exemplification of property P by an object x at time t bears a structural similarity to the sentence 'x has P at t'" (1973, p. 222).

Kim believes that this 'structural similarity' between the event $[(\text{Bob}, t), \text{falls}]$ and the sentence 'Bob has the property of falling at time t' is the reason why sentences such as 'x has P at t' are often used to refer to, describe, represent, or specify an event (1973, p. 222). But it is not only such sentences as 'Bob has the property of falling at t' or 'The plane has the property of exploding at t' that can be used as tools in describing or representing events. Kim also says that 'gerundial nominals' of sentences can also refer to events. Examples of these are 'the falling of Bob', 'this falling of Bob', 'Bob's falling at t', 'the exploding of the plane', or 'this exploding of the plane', 'the plane's exploding'.

Every particular event has a 'constitutive object(s)', a 'constitutive attribute (property)' and a 'constitutive time' (1973, p. 223). Taking one of my examples, Bob

would be the constitutive object of the event; falling would be the constitutive property; and t would be the constitutive time of the event. An event can also have several constitutive objects. If this is the case, the constitutive property is a relation of the constitutive objects at some particular time t. An example would be this: a book of mine is standing between my coffee cup and the loudspeaker of my computer. Let 'a' stand for the coffee cup, 'b' for my book, and 'c' for my loudspeaker. Kim represents this state as $[(a, b, c, t), (2) \text{ stands between } (1) \text{ and } (3)]$. This corresponds to 'b stands between a and c at t' (1973, p. 223). The standing of one thing between two other things may not be an event in the narrow sense, in terms of it being a change of some sort, but this does not bother Kim since events in his sense of object(s) exemplifying properties at times is as broad as to include changes, states, and conditions (1973, p. 222). So, the standing of my book between my coffee cup and my loudspeaker is a triadic event, just because it has three constitutive objects related in some way to one another at a particular time. And Bob's falling is a monadic event, just because the event consists of one constitutive object (Bob) exemplifying one property (falling) at some particular time.

Having elaborated Kim's view on how events are to be analyzed and what they are constituted of I now want to turn to event causation and why one event and not another is the cause of a particular effect event. Kim would not doubt that 'Bob's falling' is the cause event of 'his leg's breaking'. It is not 'Jimmy's playing', nor 'the dog's barking' that can be viewed as the cause of 'Bob's leg breaking', but rather 'his falling'. Since Peter, Paul, and Mary, or for that matter anybody else, could be substituted for Bob in the falling and leg breaking incident; and since the times of the fallings and breakings

could also be different, it makes sense to claim that there is some kind of connection between fallings and leg breakings.

For particular fallings and breakings to be fallings and breakings they need to be instances of types of fallings and breakings. Besides particular fallings and breakings Kim requires ‘generic events’ (1973, p. 226). Kim says: “...the requirement of constant conjunction for causal relations for individual events is best explained in terms of lawlike correlations between generic events” (1973, p. 226). Particular events as causes and effects are, as the adjective suggests, particular. There is nothing general about them. In opposition to that, laws or lawlike relations, regardless of however they are conceived, as well as the entities that stand to one another in lawlike relations are general and not particular entities. So in order to account for constantly conjoined particular events in causal relations lawlike correlations between types of events are needed.

Kim believes that these event types or generic events are the ‘constitutive properties’ of events (1973, p.226). He says: “It follows that each event falls under exactly one generic event, and that once a particular cause-effect pair is fixed, the generic event that must satisfy the constant conjunction requirement is uniquely fixed” (1973, p. 226). The ‘constant conjunction requirement’ has to do with causal laws. Causal laws are about the relationship of properties (generic events) and not event descriptions as Davidson holds. Kim wants to be able to say that whenever there is this unique kind of falling then this property will be constantly followed by this type of leg breaking. Constant conjunction is the fact that one and the same type property always follows another unique type of property.

Not only do events have constitutive properties the having (instantiation or exemplification) of which by an object at a particular time constitutes that particular event – such as ‘falls’ in $[(\text{Bob}, t), \text{falls}]$ – but they themselves also exemplify properties (1973, p. 226). Bob may fall at home, at school, at work, and so on. If Bob falls at home, then the event $[(\text{Bob}, t), \text{falls}]$ exemplifies the property of occurring at home. If he falls at school, then the same event exemplifies the property of occurring at school.

Kim is concerned about how particular events can satisfy the constant conjunction requirement. Generally speaking, he believes that particular events can satisfy this requirement only if we invoke type or generic events (constitutive properties of events). Particular events – such as this particular falling of Bob – are instances of a type event; they fall under that event; or, they are subsumed by that type. Once we have recognized, categorized, and characterized these type events, it is those types then that are constantly conjoined; it is those types that stand in lawlike correlations to one another. And the individual events of cause and effect satisfy the requirement of constant conjunction just because they are instances of types of events under which they fall.

The event of Bob’s falling at 3:00 p.m. causes the event of Bob’s leg breaking at 3:01 p.m., but Mary’s falling at 3:00 p.m. does not cause Mary’s leg breaking at 3:01 p.m. And, Paul’s falling at 4:00 p.m. may or may not cause Paul’s leg breaking at 4:01 p.m. Bob breaks his leg; Mary does not; and Paul may or may not break his leg. These examples only go to show that it is difficult to fix the constitutive properties or generic events for cause and effect events. What is it about Bob’s falling that makes it a leg breaking, and that is not present in Mary’s and maybe Paul’s case? A plausible answer could be that we have not correctly categorized and characterized (described) the generic

events that are lawfully correlated with one another. Kim may say that the mere falling does not cause the leg breaking. It is only a specific type of falling that causes the leg breaking. Generally speaking, there must be a ‘unique’ generic event (constitutive property) under which cause and effect event fall respectively (1973, p. 227). But again, it is difficult to say exactly which those generic events are.

There are (a) times at which Bob’s leg is not broken (times other than 3:01 p.m. or t’); there are people whose legs – maybe even after falls – are not broken; and there are other leg breakings occurring at 3:01 p.m. that are causally unrelated to Bob’s fall (1973, pp. 228-229). This kind of reasoning leads Kim to believe that for ‘Bob’s fall at 3:00 p.m.’ – or ‘x’s being F at t’ – to be causally related to ‘Bob’s leg breaking at 3:01 p.m.’ – or ‘y’s being G at t’ (in this case x is identical with y) – it must be so in virtue of some relation ‘R’ holding for x, t, y, t’ (1973, p. 228). In order to broaden the explanation of the proposal I will modify Kim’s example from the same page. John and Paul are convicted serial killers. They are standing at the execution wall awaiting their doom. Mary and Peter are executioners. They are standing 10 meters away from their targets. Mary is aiming at John; Peter is aiming at Paul. The rifles of both fire at exactly 3:00 p.m. Mary’s bullet kills John; Peter’s bullet kills Paul. For Mary’s shot to be causally related to John’s, and not Paul’s, death, “...this must be so in virtue of some relation R holding for x, t, y, t” (1973, p. 228), where, in my example ‘x’ stands for Mary, ‘t’ for 3:00 p.m., ‘y’ for John, and ‘t’ for 3:01 p.m.

Once Kim accepts that there must be some, as he says, ‘pairing relation R’ (1973, p. 229) he realizes that not only do the appropriate objects and times have to be related to one another so that, say, the event of Mary’s shooting at 3:00 p.m. is ‘the’ cause of the

event of John's dying at 3:01 p.m., as well as that John's dying at 3:01 p.m. is 'the' effect event of Mary's shooting at 3:00 p.m., but also that we have to make a decision as to what would help determine the choice of the appropriate pairing relation. Since it is a shooting (say, constitutive property F) by Mary at 3:00 p.m. that causes John's dying at 3:01 p.m. (say, constitutive property G of the effect event), it seems plausible to use constitutive properties of cause and effect events in determining the appropriate pairing relation for cause and effect event. It makes sense to say this, because it was definitely not Mary's singing (another possible constitutive property) that caused John's death.

Kim has three suggestions about what could determine the choice of the appropriate pairing relation, or how else the idea of a pairing relation could be theoretically presented so that only Mary's shooting is 'the' cause of 'the' effect of John's dying. I do not plan to elaborate on all three options, which Kim presents. I only intend to choose one of these options, which he favors. While doing that I will try to paraphrase in my own words what seems involved in his suggestion. All three options face some difficulties, which I will not elaborate on. My motive for talking about the choice of the appropriate pairing relation has to do with my conviction that Kim's analysis of events as an object(s) x having a unique property (or relation) at time t – [(x, t), P] – is stricter than Davidson's mainly because of the requirement that the constitutive property of an event must be a 'unique' property for that particular event. Also, Kim seems to be heading several steps towards an outright analysis of causation, even though he officially does not look at his presented work as an analysis of causation: "It is best, therefore, to look upon the tentative accounts of Humean causation in this section not as full-fledged analyses of causation, but rather as approximations to the broader notion of

subsumption of events under a law, an idea that forms the foundation of the Humean, or nomological, approach to causation" (1973, p. 235). Even if his suggestions turn out not to be full-fledged analyses of causation, he still seems to be saying more about causation than Davidson in his essay 7; remember, Davidson "...abjured the analysis of the causal relation" (1980, essay 7, p. 158).

Let us return to Mary's rifle shot causing John's dying. Kim's proposal goes something like this. The rifle shot of Mary alone is not the cause of John's dying. Mary could shoot as many times as she wants without killing John. The cause event is rather Mary's rifle shot of a rifle in such-and-such a spatiotemporal relation to John whose death it causes (1973, p. 233). This cause event is not a 'monadic event'; it is not $[(\text{rifle}, t), \text{fires}]$, it is rather, as Kim says on the same page, a 'compound event' of $[(\text{rifle}, t), \text{fires}]$ and $[(\text{rifle}, \text{John}, t), R]$. So now Kim can say that the compound event $[(\text{rifle}, 3:00 \text{ p.m.}), \text{fires}]$ and $[(\text{rifle}, \text{John}, 3:00 \text{ p.m.}), \text{spatiotemporal relation } R]$ is the cause of the event $[(\text{John}, 3:01 \text{ p.m.}), \text{dying}]$ provided (i) that the compound cause event and the effect event of John's dying exist; (ii) there is a law under which the events in question fall due to the fact that their constitutive properties and their constitutive relations are constantly conjoined generic (type) events. The law expressed in (ii) is general; it uses constants such as 'x' and 'y' for things, capital letters such as 'F' and 'R' for properties and relations, 't' for times, ' Δt ' as a time difference indicator where this time difference is identical with $t' - t$ (which in our case could probably be understood as 3:01 p.m. – 3:00 p.m.). According to Kim, the law in (ii) is this: $(x) (y) (t) \{[(x, t), F] \text{ exists} \& [(x, y, t), R] \text{ exists} \rightarrow (\text{then}) [(y, t + \Delta t), G] \text{ exists}\}$, where $\Delta t = t' - t$ (1973, p. 233).

A mark of this approach is that the constitutive object of the effect event (say, John, in ‘John dying at 3:01 p.m.’) is now also a constitutive object of the cause event, because the constitutive objects of the compound cause event are the ‘rifle’ and ‘John’, who is spatiotemporally related (R) to the ‘rifle’ (1973, p. 234). Kim believes that cause events will most likely be ‘relational generic events’ (1973, p. 234), since cause events will have at least two constitutive objects due to the fact that the constitutive object of the effect event is also one of the constitutive objects of the cause event.

Having talked about Davidson’s and Kim’s understanding of events including how events are to be described, it is now time to try to find some examples which would attest to how Davidson’s and Kim’s views are different. Let us use an example that appears in Kim’s mentioned essay as well as in Davidson’s essay 8. History books say that ‘Brutus stabbed Caesar’ or that ‘Brutus killed Caesar’. The question now becomes whether these two sentences are about one or two events. For Kim these sentences would be about two different events. Remember, “Events, therefore, turn out to be complexes of objects and properties....” and that “...they have something like a propositional structure” (1973, p. 222). Events are the havings of properties or relations by objects at times. Both events are relational or dyadic (1973, p. 223). The object Brutus is stabbing-related to the object Caesar at some time t; or, [(Brutus, Caesar, t), stabs]. And in the case of the other event it is the object Brutus that is killing-related to the object Caesar at some time t; or, [(Brutus, Caesar, t), kills]. The constitutive relations (properties) of both events are different. For events to be the same they must be instances of one unique property (relation). So, having different constitutive properties or generic events makes Brutus

stabbing Caesar at some time and Brutus killing Caesar at some time two different events.

Davison's view is different. First of all he allows only 'singular terms' ('the falling of Bob') to explicitly refer to events. No sentences such as 'Brutus stabbed Caesar' refer to or describe particular events. The previous sentence is only a general statement about events. It only says that there was 'an' event that was a stabbing of Caesar by Brutus. This general statement only means that there was 'at least' one but possibly even more than one of these stabbings. What is needed is, as Kim would say, a 'gerundial nominal' such as 'Brutus' stabbing of Caesar', or a gerund prefixed with the definite article ('the') such as 'the stabbing of Caesar by Brutus'. These describe events immediately.

The second point Davidson would make is that there are not two events here. 'The stabbing of Caesar by Brutus' and 'the killing of Caesar by Brutus' are one and the same event differently described. Davidson agrees, "not all stabbings are killings" (1980, essay 8, p. 171). He also agrees, on the same page, that "...the death (of Caesar) is not identical with the stabbing (killing)" of him. Caesar's death is really a different event from his being stabbed / killed, because the death occurred later. Regardless of the fact that not all stabbings are killings and the fact that Caesar's death is a separate event, it is still the case, says Davidson here, that this particular stabbing "...was in fact, though of course not necessarily, identical with Brutus' killing of Caesar" (1980, essay 8, p. 171). Caesar died in fact because Brutus stabbed / killed him.

Once Brutus stabbed Caesar he killed him, even though Caesar's death, as a separate event, occurred later. It is not as if Brutus has to perform an additional action

(actions are a subspecies of events for Davidson) of killing after he stabs Caesar. The killing is identical with the stabbing in this case. The event (action here) remains the same. If the event is described as ‘the stabbing of Caesar by Brutus’, then the event is described from the point of view of the cause of Caesar’s death. And if the event is characterized as ‘the killing of Caesar by Brutus’, then this is done so from the perspective of the effect of the stabbing, having the death of Caesar in mind. Davidson believes that those who think that ‘the stabbing’ and ‘the killing’ are two different events confuse “...a feature of the description of an event and a feature of the event itself”. He also says, on the same page: “The mistake consists in thinking that when the description of an event is made to include reference to a consequence, then the consequence itself is included in the described event” (1980, essay 3, “Agency”, p. 58). The event of ‘Caesar’s dying’ remains a different event from the event the effect of which it is. We still retain the freedom to describe the cause of his death in terms of its consequence or effect without changing the fact that this description is of the cause event.

Similar considerations hold for another of Davidson’s examples. I intend to kill a space traveler. I pour poison in the water tank of his space vessel at 3:00 p.m. on January 23rd. The traveler reaches Mars on January 28th. He then takes a drink and dies immediately (1980, essay 8, p. 177). Davidson says that ‘my pouring the poison’ is identical with ‘my killing the traveler’, even though the death of the traveler occurred five days after my pouring of poison. Assuming that the traveler dumps the content of the water tank upon arrival on Mars – he does not even have to know that the water was poisoned – and survives, we have no justification in saying that the ‘pouring of poison’ was ‘the killing of the traveler’. Let us assume that the traveler dies after drinking the

poisoned water on January 28th. We are less justified in calling the event a killing prior to the death of the traveler. But once he dies we are fully justified in calling the event a killing. The longer we wait for the consequence of death to occur the higher the chance of interference of other events, happening in between the poison pouring and the possible dying. This means that the description of the poison pouring as a killing depends, as Davidson says, on the ‘directness of the causal connection’ between cause and effect event (1980, essay 8, p. 177). The closer the effect event of dying is to the cause event of poison pouring the more justified we are in calling it a killing as well.

Kim may answer that the poisoning and the killing, as well as the stabbing and the killing, remain distinct events having different constitutive properties (generic events) exemplified, just because it is one thing to explain why I poisoned the traveler and yet another to explain why I killed him (Davidson uses a different example backed by the same reasoning in 1980, essay 8, pp. 170-171). My reason for killing him could be the fact that the traveler double-crossed me in an important business deal. I want him dead, and I don’t care how I kill him. But my reasons for poisoning his water tank may be different. I want to avoid prosecution by the authorities. My main goal is to use a very inconspicuous method of killing him, because I believe that this is the cleanest manner of getting rid of him. Since the explanations are different, the events explained must be different.

Davidson agrees that my reasons for killing the traveler and poisoning him may in fact be different (1980, essay 8, p. 171). But he believes that the difference in explanation does not show that the events described are different. He argues that there is no additional event (action) that I have to perform after the water tank poisoning. I do not have to

perform a killing as well, because my poisoning the tank is my killing the traveler. Davidson says: "And explanation, like giving reasons, is geared to sentences or propositions rather than directly to what sentences are about ..." (1980, essay 8, p. 171). The reasons that I give for the killing and the poisoning – or, in other words, the explanations that I provide – are sentences, and, according to Davidson, different sentences can refer to the same event.

This kind of reasoning leads me to believe that Davidson does not believe in a tight connection between event causation and event explanation. Events are causes and effects of one another regardless of how we name or describe them or how we explain them and their causal relation. I can describe an event as 'Bob's falling off the ladder', while a physicist may give a different description of the same event. I can also explain the same event as being the effect of 'Bob's slipping', while, again, a physicist may explain the cause-effect relationship among those events in a totally different way. It also seems that both are good explanations depending on the context of the utterance of the sentences of explanation. Even if I knew the physicist's explanation for the causal relation between those events, it would not be good to use the physicist's event descriptions and causal explanation when I am with my friends who neither know the physicist's descriptions nor his explanations. So it seems that, on a Davidsonian account, there can be many event descriptions referring to one and the same event as well as many explanations of the causal relation of events, those, of course, depending on the purpose and context of communication.

I think that Kim differs from Davidson, in the sense that he believes that there is a strict connection between event causation and event explanation. I believe that Kim

wants to say that the event of ‘Bob’s falling’ – [(Bob, t’), falls] – as the effect event of ‘Bob’s slipping’ – [(Bob, t), slips] – can be explained by the proposition ‘Bob had the property of slipping at time t’. It seems to me that causation and explanation are connected (in Kim’s view) because events exhibit a ‘propositional structure’ and because an effect event is explained by its real cause. The having of this ‘unique’ slipping property at t is not only the real cause of “Bob’s fall” but it also is the real explanation of the effect event. So an explanation is causal just because there is an underlying cause event having a propositional structure of one thing x having or exemplifying a unique property. Real explanations may be propositions or sentences but they nevertheless mirror what properties or generic events are correlated. In my example it is the unique property / generic event that I have called ‘slipping’ and the unique property ‘falling’ that seem to be correlated. Both events exemplify generic events / properties respectively. And because both events, as instances, fall under their unique (types) generic events, they do satisfy the constant conjunction requirement. Just because these events are instances of generic events, they do fall under causal laws in virtue of the fact that generic events stand in a lawlike correlations to one another.

An important difference between Kim’s and Davidson’s views on events consists in the fact that Kim introduces types of events in addition to individual events. Davidson is committed openly to the existence of particular events, even though he does make the distinction between ‘features of the description of an event and a feature of the event itself’ (1980, essay 3, “Agency”, p.58). Regardless of whether the phrase ‘feature of an event’ may or may not point towards general entities such as types of events, it still remains the case that Davison does not undertake attempts to analyze the structure of

events in terms of them being instances of properties of objects had at particular times. Particular events remain the same even though they are describable in many ways.

Kim differs just because he demands a strict analysis of the structure of events in terms of objects having properties at times, even though it is hard to pin down ‘the’ properties (types or generic events) that are had by objects at times and whose instances are particular events. Since he believes that there must be unique properties instanced in every particular event and in its causal relation to other particular events, it seems that there must be the right description of an event and of its causal relation to another event, even though we do not know how to correctly categorize and describe these events and their causal relation. Davidson’s many descriptions of one event, and its many descriptions of its causal relation to another event, are unacceptable because only one unique property is exemplified in the cause and effect event respectively. An event that is described in so many ways, and provided that events are identical with the having of properties at times by objects, must be an instance of different properties at the same time. An event cannot be ‘the moving of the finger’, ‘the pressing of the trigger of the gun’, ‘a shooting’, and ‘a mercy killing’, at the same time. There must be one right description depicting the right property the instance of which is identical with the event, according to Kim.

That there are generic events on top of particular events, and that the former have a say in causation has consequences for Kim’s understanding of the mind as well as on mental causation (chapter 2). Davison’s understanding of events as unanalyzed particulars and his rejection of generic events determines his philosophy of mind and mental causation (chapter 3).

CHAPTER TWO

In the following essay I plan to show how the reasoning from the first chapter about Kim's understanding of events, causation, and explanation reflects on his views on mental causation. I will investigate what it means for Kim when we say that a mental event, such as my desire for a coke, causes me to go to the fridge. I will be concerned with matters of whether and how mental events can be causes in a world that we believe to be physical.

It is now time to talk about Kim's understanding of mental causation. It was said that events are causes and effects. When talking about mental causation we want to be able to say that mental events are causes and effects. I will mention examples to show what I mean. The mental event of my desire for a coke causes me to go to the fridge and get a can of it. My belief that the assassins are at my door causes me to jump out of my bedroom window in an attempt to save my life. I decide to write this essay, and this decision causes me to start to write it. Or, suppose that I am left alone in a jungle to survive without anybody's help for several days. Imagine the first scenario to be the case that I do not have any knowledge of how to survive in these circumstances. The likelihood that I will die from thirst, hunger, snakebites, or many other things is very high. But let us imagine the second scenario. I have knowledge how to survive in the jungle. This knowledge will most likely cause me to survive. And the knowledge may in fact be something other than what underlies it neurally.

All these examples seem to be instances of mental causation. I think that we can provide a rough Kim-style analysis of the structure of the mentioned event particulars in

causal relations. Let us assume that Kim's analysis of the structure of events is correct. Bob desires a coke; this desire causes him to go to the fridge: [(Bob, t), desires a coke] causes [(Bob, t'), going to the fridge]. Or, Bob believes the assassins are at the door; this belief causes Bob to jump out of his bedroom window: [(Bob, t), belief that the assassins are at the door] causes [(Bob, t'), jumping out of the window].

The question now becomes whether non-physical mental events can be causes distinct from neural events. We want Bob's desire for a coke to be the real cause of Bob's going to the fridge. Bob's belief that the assassins are at the door must be the real cause event for the effect event of Bob's jumping out of his bedroom window. This kind of reasoning should hold for any mental events. Kim says: "For the only way in which I believe that we can understand the idea of causal explanation presupposes the idea that the event invoked in a causal explanation is in reality a cause of the phenomenon to be explained" (1999, chapter 3, p. 64); "...causal explanation of an event that invokes another as its cause can be a correct explanation only if the putative cause really is a cause of the event to be explained" (1999, chapter 3, p. 75); "Realism about explanation should at least cover causal explanation" (1999, chapter 3, p. 76). This means that in order to explain the effect event of going to the fridge we have to mention its real cause, namely the mental event of my desire for a coke. The same holds for the other example. To explain the effect event of my jumping out of the window I have to invoke the real cause of this event, namely my belief that the assassins are at the door.

But now one may wonder why the cause event can explain the effect event. Why is it that going to the fridge and jumping out of the window are explainable by the desire for a coke and the belief that the assassins intend to kill me respectively? I think that the

answer has to do with the following. That particular event of my going to the fridge cannot be explained by my belief that the earth is round, because it was not caused by this particular belief. Neither can the effect event be explained by my intention to write this essay, nor any other mental event. Only my desire for a coke can in this case be the cause and the explanation of my going to the fridge. There is something about the event [(Bob, t), desire for a coke] that makes this and only this event the cause of my going to the fridge and an explanation of that effect event. I think that this something is the generic type [(x, t), desire for a coke], or in other words, it is the property of being a desire for a coke. The particular event of [(Bob, t), desire for a coke] – Bob's desire for a coke – is an instance of the property of being a desire for a coke. As an instance of that property (or as an event particular) it causes me to go to the fridge. But this event causes me to go to the fridge in virtue of it being an exemplification of the mentioned property.

For the sake of argument, let us assume that being a desire for a coke is the unique property (see my chapter 1), which is instantiated in the example, the instance of which is the real cause of my going to the fridge as well as the real explanation of that effect event. It is in virtue of this unique property that the event the desire for a coke is the cause of the going to the fridge. The latter event is also an instance of a ‘unique’ property, namely that of the property of being a going to the fridge. So it seems that cause and effect events are ‘the’ cause and ‘the’ effect of one another due to the properties instances of which they are.

But the main question is: Can mental events, such as the desire for a coke and the belief that the assassins are at the door, be causes and effects? Kim doubts whether mental events, if they are distinct from physical events, can be causes and effects; in

other words, whether the mental property of being a desire for a coke – viewed as a property distinct from some neural property - really has a role to play in causation. In Kim's words: "Suppose then that mental event m, occurring at time t, causes physical event p, and let us suppose that this causal relation holds in virtue of the fact that m is an event of mental kind M and p an event of physical kind P" (1999, chapter 2, p. 37). If the event of the desire for a coke is an instance of the non-physical mental kind (or property) of being a desire for a coke, and if this event were to causally interact with a physical event, say the movement of my body to the fridge, it would be a case of "...a clear violation of the causal closure of the physical domain" (1999, chapter 2, p. 37).

The 'causal closure of the physical domain' can be explained as this. Every physical event has other physical events in its ancestry and posterity. This means that non-physical events cannot be causes and effects of physical events. Or, every physical event has only physical causes and only physical effects. The principle of the causal closure of the physical domain threatens to exclude non-physical mental events as causes. This means that my desires, beliefs, intentions, decisions, and so on, cannot cause me to do anything. This is not a promising conclusion. Intuitively, some of us believe that our minds have causal relevance, that the mental events of our minds cause us to bring about changes in the physical world. Neither Kim nor Davidson (chapter 3) seem to be comfortable with the looming consequences of the principle of the causal closure of the physical domain.

Kim has a remedy for the unwanted conclusion. The following discussions will pertain to Kim-style ways of dealing with the challenge of the exclusion of the human mind when it comes to matters of causation and explanation. I think that our next task

must be finding ways of bringing the mind closer to the brain so that we can provide the mind with the causal and explanatory relevance, which we intuitively believe it has. We have to find a relation that would link the mind and the brain in such a way that mental causation is not only possible but also that mental causation does not violate the principle of the causal closure of the physical domain.

There is a relation that seems to bring the mind closer to the brain. This relation is known as the supervenience relation, or the relation of the dependence of mental on physical properties. As Kim says: "Supervenience is standardly taken as a relation between two sets of properties, the supervenient properties and their base properties" (1999, chapter 1, p. 9). The word 'supervenience' usually means dependence. When one property supervenes on another it depends on it, or it is determined by it. Since our focus is on the relationship between the mind and the brain, I will an example of this dependency relation: the mental property of being a belief that the assassins intend to kill me supervenes / depends on (or is determined by) a neural property, which I will call 'being a c-fiber firing'. The mental property of being a belief that the assassins intend to kill me is the supervenient property. The neural property of being a c-fiber firing is the subvenient or base property.

The thesis of mind-body property supervenience is this: "Mental properties supervene on physical properties, in that necessarily, for any mental property M, if anything has M at time t, there exists a physical base (or subvenient) property P such that it has P at t, and necessarily anything that has P at a time has M at that time" (1999, chapter 1, p. 9). Suppose I have or instantiate the mental property of being a belief that the assassins intend to kill me ('M' above). Whenever anybody has or instantiates that

kmental property at some time then that person has or instantiates the neural property of being a c-fiber firing ('P' above) at the same time. And whenever anybody has the neural property of being a c-fiber firing at some time then that person has the mental property of being a belief that the assassins intend to kill me at the same time.

Any changes in our beliefs, desires, intentions, decisions, and so on, reflect changes in our brains. There cannot be any changes of the mind independently from changes in the brain. The supervenience relation between the mental and the physical is "...an asymmetric dependence of the mental on the physical..." (1999, chapter 1, p. 6). This means that the mental property of being the belief that the assassins intend to kill me depends on the neural property of being a c-fiber firing, but not vice versa. It is not the case that the mental property determines the neural property.

This approach provides a link between the mind and the brain, so that there are now conditions for mental-to-physical-event causation and causal explanation. The assumption is that the mind does not exist independently from the brain, and that neural changes underlie mental changes. If we suppose that mental properties supervene on physical properties will this allow mental events, as belonging to their respective mental kinds, to be causes of neural and other physical events? Somebody who believes in the causal relevance of the mind would want to say that desires, beliefs and other propositional attitudes are responsible for the movements of bodies that instantiate these mental properties.

But it seems that our worries about mental causation cannot be avoided via the introduction of the supervenience relation. I have the mental property of desiring a coke at some time. If the mental property supervenes on the neural property, then the former is

presumably a property different in kind from the latter one, because the supervenience relation is a relation of dependence of one property on another. If this is so, then it is also the case that the instances (or events) of these respective properties are distinct as well. But on Kim's view, the event, which is the instance of the mental property, cannot cause the physical event of the movement of my body's leg. The desire for a coke is an instance of a mental kind. That event is not a physical event; it is not an event of a physical kind. As an event of a mental kind the desire for a coke cannot be the cause of the physical movement of my leg. This causal relation is excluded by the principle of the causal closure of the physical domain. If the desire for a coke were to miraculously cause the physical movement of my leg then this would constitute a breach of the causal closure of the physical domain. This reasoning holds if events are instances of mental kinds and if these mental kinds are distinct from neural kinds. From this it follows that the introduction of the supervenience of the mental on the physical has brought more problems in the form of the breach of the closure principle.

It becomes clear that the supervenience property-relation of the mental on the physical, as stated in Kim's quote, cannot provide mental events with grounds for them to be causes of physical events. It seems that the supervenience relation only amounts to the following claims: (a) mental properties depend in some sense on physical properties; (b) mental properties are distinct from neural properties; (c) mental and neural properties covary; (d) mental properties do not have causal powers, since the principle of the causal closure of the physical domain precludes mental events from being causes. The consequence of (d) is that mental properties end up being shadows or epiphenomena without any causal work to do.

Kim, too, is dissatisfied with the supervenience thesis. He states that this thesis is not an explanation of the mind-body relation at all. He says that the supervenience thesis “...merely states a pattern of property covariation between the mental and the physical and points to the existence of a dependency relation between the two. Yet supervenience is silent on the nature of the dependence relation that might explain why the mental supervenes on the physical” (1999, chapter 1, p. 14). Why is it that whenever the neural property of c-fiber firing is instantiated the mental property of being a belief that the assassins are at the door is exemplified as well? It does not suffice to just say that the mental supervenes on the physical ‘in some sense’. We need to specify in exactly what sense mental properties depend on physical properties.

If my desire for a coke and my belief that the assassins intend to kill me are to occur at all, or if the respective mental properties, instances of which these two mental events are, are to be realized at all, they “...must be physically realized” (1999, chapter 1, p. 19). Neural properties must somehow take over the causal powers of mental properties for the latter to have causal and explanatory relevance. If the mental property is to do any causal work at all, if it is to determine the mental virtue in which mental events can be causes, then the mental property itself (say, being a belief that the assassins intend to kill me) must be realized by a neural property. The causal work of the mental property must be done by a neural property. This reasoning seems to be driven by the requirements implicit in the principle of the causal closure of the physical domain.

Let us combine some of our findings so far. The mental (attitude) property of being a belief that the assassins intend to kill me is supervenient on the neural property of being a c-fiber firing in that the latter realizes the former. Since the mental property

supervenes on the neural property, it is ‘second order’ in relation to the neural property, because the mental property depends on the neural property in that the latter realizes it (1999, chapter 1, 19-27). And, the mental property is explained as a relation. It is in-between properties of being a seeing of the assassins, and being an escape behavior (rather than some other behavior). As a property, my belief follows after the property of seeing the assassins, but it precedes the property of being an escape behavior. So, being a belief that the assassins intend to kill me is (a) a ‘second order’ (b) ‘relational’ or ‘functional’ property.

Kim says: “Functionalism takes mental properties and kinds as functional properties, properties specified in terms of their roles as causal intermediaries between sensory inputs and behavioral outputs, and the physicalist form of functionalism takes physical properties as the only potential occupants, or “realizers”, of these causal roles. To use a stock example, for an organism to be in pain is for it to be in some internal state that is typically caused by tissue damage and that typically causes groans, winces, and other characteristic pain behavior. In this sense being in pain is said to be a second-order property: for a system x to have this property is for x to have some first-order property P that satisfies a certain condition D , where in the present case D specifies that P has pain’s typical causes and typical effects” (1999, chapter 1, p. 19).

Let us see how this reflects on my example. My belief is a ‘causal intermediary’ between the type (or property) of seeing the assassins (input) and the type of behavior (escaping) (output). For the sake of argument we may say that whenever there is a seeing of the assassins, then the instance of that property causes the instance of my belief property, which in turn causes the instantiation of the behavior property (here escaping).

Kim's 'physicalist form of functionalism' says that the belief property supervenes on the neural property of c-fiber firing by way of the latter being the realizer of the former. All the causal work is done by the c-fiber firing. To be a real causal intermediary an event must be an instance of a physical property in order to satisfy the principle of the causal closure of the physical domain. So, strictly speaking, being a belief that the assassins intend to kill me, understood functionally, is not the causal intermediary, it is rather the neural property of being a c-fiber firing that is the causal intermediary between seeing the assassins and escaping. The second-order functional property of being a belief that the assassins intend to kill me only specifies the conditions under which any realizer property plays the role of the mental property of being a belief that the assassins intend to kill me. These conditions can be expressed as: for a neural realizer to be 'the' realizer of the mental property of being the belief that the assassins intend to kill me the neural realizer must have this belief's typical causes and typical effects.

For the sake of the argument, let me just say that being a seeing of the assassins (a physical property) is a typical cause of my belief, and being an escaping (also a physical property) is a typical effect of that belief. In my case, it is the property of the c-fiber firing whose instance plays this causal role, according to the property realization relation (the fact that the neural property realizes the mental property). The defining characteristic of that belief, or, what makes that belief that belief, is the satisfaction of the causal role of having this belief's typical causes and typical effects. The causal specification in terms of the having of the same input and the same output is what essentially determines the content of that belief. The open space in between (same input) ... (same output) can be filled by almost any property provided that this property is a physical property and that it

has causal potentials to do the job of being a realizer of that belief. Kim says: "Whether or not a given property qualifies as an occupant of a specified role – that is, whether or not it is a realizer of a functional property – depends essentially on its causal / nomological relations to other properties, not on its intrinsic character. Intrinsic characters do matter of course, but only because of their capacity to get causally hooked up with other properties" (1999, chapter 1, p. 21).

A stone in our world cannot have the belief that the assassins intend to kill it even when faced with the assassins. The stone has no appropriate sense organs to see the assassins; it also lacks motor systems that would enable it to escape. And it has nothing like a neural network, some mechanism of which could turn out to have some appropriate realizer property. The stone is not made out of material that is appropriate for the having of beliefs. So what kind of material an entity is made of is relevant when it comes to whether it can have that belief or other mental attitudes. But it is possible that not only brains and bodies made up of flesh and bones can have that belief. Maybe there are life forms on some distant planet, which are not made out of flesh and bones and whose brains do not consist of grey matter, that can have the belief that the assassins intend to kill me.

This is an empirical question, which we can, in principle, settle by observation and experiment. The mentioned life form should have equivalents of information gathering sense organ type of mechanisms, as well as information processing devices like our brains, and motor systems enabling that life form to exhibit escape behavior. If the life form has these mechanisms and devices, then it can have the belief that the assassins

intend to kill it, regardless of the fact that its brain-like structures are not made out of grey matter and its body not made out of flesh and bones.

Slight variations within the neural wiring of specimens of one and the same species can still allow for having same belief. There is a mechanism in my brain that has the property of being a c-fiber firing. This property is for me the realizer of the belief that the assassins intend to kill me. This property is appropriately related to the properties of seeing the assassins and escaping from them. My brain is causally wired this way. But Paul has a mechanism in his brain that has the property of being an x-fiber firing. Paul's realizer property of x-fiber firing is appropriately related to seeing the assassins and escaping from them. Paul can have the belief that the assassins intend to kill him. As long as Paul has a mechanism that has some property, which has the causal powers or potentials to get hooked up with properties of seeing the assassins and escaping from them Paul can have the mentioned belief.

From this we see that the same mental attitude can be realized by entities of different makes and properties of various kinds as long as those properties are appropriately causally hooked up with seeing and escaping. This means that mental attitudes are multiply realizable. But mental attitude properties depend primarily not on the make of the bodies nor the type of the physical properties that instantiate them, but on causal relations holding for the properties involved. The instances (physical events) of causal properties of seeing, neural realizers (x, y, z, etc.), and escaping must always follow one another. In other words, the causal properties must be constantly conjoined. Expressed yet differently, the causal laws of the world in question must be such that the

physical events of seeing the assassins always cause some realizers to be instantiated, which in turn cause the same behavior (here escaping).

We will have the same belief (I, Paul, and the alien) as long as the causal laws of our world are held constant. But if the laws were different, if the causal properties were differently arranged, that could affect our realizers' status as realizers of the belief that the assassins intend to kill me (1999, chapter 1, p. 23). Suppose that I have an exact duplicate in a world with different laws. In that world the following causal properties are constantly conjoined: seeing a red tomato, c-fiber firing, and escaping. The c-fiber firing does not seem to be the realizer of the belief that the assassins intend to kill me. It is the realizer of something else. Kim says: "... in worlds in which different laws hold at the level of M's base domain, thereby generating different causal structures in those worlds, P may fail to satisfy the functional specification definitive of M" (1999, chapter 1, p. 23). Because of the rearrangement of the causal properties, the c-fiber firing ('P' in the quote) of my exact duplicate does not seem to satisfy the causal specification for being the realizer of the belief that the assassins intend to kill me. The causal properties are not arranged the way they are in our world: seeing the assassins, c-fiber firing, and escaping.

The realization relation as an explanation of the mind-body property supervenience relation has led us to the conclusion that it is causal laws primarily that determine mental attitudes. Same laws for same entities mean same mental attitudes. But our worries expressed several pages ago are still not settled by this ending. Recall that we were concerned with providing mentality with causal relevance within the framework of the causal closure of the physical domain, which precludes mental events belonging to different kinds from breaching the closure. After introducing neural properties such as the

c-fiber firing as the realizer of my belief, it seems that the mental property of being a belief that the assassins intend to kill me has no role in causation and explanation. But this is a conclusion is something some of us wanted to avoid in the first place.

Kim believes that mental attitude properties are neither causally nor explanatorily excluded. His answer lies in the identification of mental attitude properties with their realizers: “The answer is that by definition, having M is having a property with causal specification D, and in systems like s, P is the property (or one of the properties) meeting specification D. For systems like s, then, having M consists in having P. It isn’t that when certain systems instantiate P, mental property M magically emerges or supervenes (in the dictionary sense of “supervene”). It is rather that having M for these systems, simply is having P” (1999, chapter 1, p.24)). To explain this, we could let ‘s’ stand for organisms (systems) like me, ‘P’ for the neural property of being a c-fiber firing, ‘M’ for the mental attitude property of being a belief that the assassins intend to kill me, and ‘specification D’ for the having of this belief’s same causes and same effects (‘same’ refers to properties or types of events constantly conjoined). Anybody just like me, living in a world just like ours, having the mentioned belief, has a physical realizer, the having of which is the having of the belief.

My belief being identical with its realizer strongly depends or supervenes on the causal laws (causal property arrangement) that prevail in the world I inhabit. This expresses Kim’s idea of the strong supervenience of functional mental properties on prevailing causal laws. Provided that causal properties of seeing the assassins, c-fiber firing of the realizer, and the behavior of escaping, are constantly conjoined, any organism like me will instantiate that particular belief whenever it instantiates the

property of the c-fiber firing. The causal powers of my belief (or any mental attitude property for that matter), because, if identical to c-fiber firing, it isn't second order, are not excluded. The identity of this mental property with its realizer precludes the exclusion of its causal powers. The identification allows the mental property to 'inherit' the causal powers of its realizer (1999, chapter 4, p. 110).

So it seems that the mind has a role to play when it comes to causation. Desires, intentions, decisions, thoughts, and so on, are causes not as free floating and distinct mental events unanchored in the physical domain but by being identical with physical realizer events. But what about explanation? Are mental attitude properties explanatorily excluded? No, the property identification and the 'inheritance principle' allow mental attitude properties to be explanatorily relevant: "Given that each instance of M has exactly the causal powers of its realizer on that occasion ("the causal inheritance principle"), all the causal / explanatory work done by an instance of M that occurs in virtue of the instantiation of realizer P1 is done by P1..." (1999, chapter 4, p. 110).

It does not matter that the same belief that the assassins intend to kill me (M) is identical with different realizer properties (say P1, P2; c-fiber firing, x-fiber firing, etc.). The mental property M inherits the explanatory relevance of its multiple realizers. It is in virtue of the property of the c-fiber firing that my belief can be explained. This means that to explain my belief I have to resort to the property of c-fiber firing. The same belief held by the alien is explained by resorting to his y-fiber firing. The explanatory relevance question of mental attitude properties is settled via something like this. Why do I instantiate the above belief at 3 p.m.? Because I instantiate c-fiber firing (P1) at that time, and c-fiber firing is a realizer of that belief in organisms like me. To have that belief is by

definition to have some property that satisfies the causal specification of having this belief's same causes and same effects. And in things like me it is the c-fiber firing that meets this specification. Why do things like me instantiate the belief that the assassins intend to kill me and not the desire for a coke whenever they instantiate the c-fiber firing? The answer is: For things like me it is the c-fiber firing that is the realizer of the belief but not of the desire for a coke (1999, chapter 4, pp. 111-112).

The identification of the mental attitude property with its specific realizer is not undesirable. It allows for mental attitude properties to have causal powers as well as explanatory strength. This is a way in which the mental can be saved as part of the physical domain. But the price to be paid here is that these properties have the same causal powers and the same explanatory strength as their physical realizers. Their causal powers and explanatory strength is ‘nothing over and above’ the causal and explanatory work their realizers perform (1999, chapter 1, p. 24). In other words, this means that Kim’s method of treating mental attitude properties as second-order relational (functional) properties defined in terms of the relation of first-order physical properties (seeing the assassins, c-fiber firing, and escaping) and their identification with their realizers, as a result of the need to explain the property supervenience relation in terms of the realization relation, leads to the reduction of mental attitude properties to their realizers, which seems to take away from mental properties the status of being independent properties different in kind from neural properties. This is something a nonreductivist physicalist would want to avoid.

Kim calls the method of looking at some property as a function or relation holding for first-order properties ‘functionalization’. He believes that functionalization is a

“...necessary condition for reduction” (1999, chapter 4, p. 99). In order to reduce a property (such as my belief, desire, and so on) we have to show that it is not an intrinsic property to a thing but that it depends on a relation to other properties outside of the organism that instantiates it. My desire for a coke is nothing I was born with; it is not part of me the way my genes are, for instance. The genes determine my intrinsic nature regardless of whether I will ever be related to coke cans or not.

Kim suggests that we should re-think whether mental properties really are intrinsic or not. It turns out that mental attitude properties are not properties intrinsic to the organisms that instantiate them. And his method to deal with them is to functionalize them and identify them with their realizers, which leads to their reduction. Kim says: “...to reduce a property M to a domain of base properties, we must first “prime” M for reduction by construing, or reconstruing, it relationally or extrinsically. This turns M into a relational / extrinsic property. For functional reduction we construe M as a second-order property defined by its causal role – that is, by a causal specification H describing its (typical) causes and effects. So M is now the property of having a property with such-and-such causal potentials, and it turns out that property P is exactly the property that fits the causal specification. And this grounds the identification of M with P. M is the property of having some property that meets specification H, and P is the property that meets H. So M is the property of having P. But in general the property of having property Q = property Q. It follows that M is P” (1999, chapter 4, pp. 98-99).

Kim believes that his ‘functionalization’ / ‘reduction’ holds for beliefs, desires, and intentions. But other aspects of human mentality resist reduction, in the manner of functionalization, to their physical subvenience (underlying) bases (of properties). Kim

believes that the mental properties that cannot be reduced are so-called ‘qualia’. In courses of philosophy of mind as well as in various textbooks ‘qualia’ are explained as ‘subjective feels’, or ‘the what it is like’ of particular feels. Say you eat a mango. Its taste is a subjective quality to you. You can distinguish that taste from the way a banana tastes. Smells, too, are to be counted among qualia. Take ammonia as an example. The subjective feel of its smell makes us cringe. We experience these qualities without being able to verbally mirror their feel.

Other such subjective qualities include the feel of sexual gratification, of emptying ones bladder and bowel. Yet others are various pains. A hit in the face feels different from a burn on the hand. One can attempt to ‘functionalize’ both as being identical with their realizers, which are intermediaries between tissue damage (input) and wincing, crying (output), and other pain behavior. But the specific subjective quality of those feels remains unreduced. It seems – at least according to Kim – that qualia are properties ‘intrinsic’ to the organism that feels it. Kim says: “...it seems to me that the felt, phenomenal qualities of experiences, or qualia, are intrinsic properties if anything is” (1999, chapter 4, p. 102).

If we cannot reduce qualia via functionalization to identities with first-order realizers, as was the case with beliefs, desires, and intentions, then we have no way of explaining the correlation of neural properties and qualia, which are supervenient on the neural properties. If Kim is correct, qualia are mental properties correlated with neural properties. But the idea of correlation does not explain much. It only states that whenever Bob has or instantiates neural property C he has the pain-quale (mental property) of being hit in the face. We can now ask: why is it that whenever Bob has neural property C he

also has this subjective pain experience? In the absence of maneuvers of reductive identification of neural C and the pain-quale, the only thing that we can say – according to Kim – is that the mental quale and the neural C property are distinct. Kim says: “Indeed the possibility of functionalization is a necessary condition of reduction. As I have already said, if both Mi and Pi are distinct intrinsic properties in their own right, replacing \leftrightarrow with = in the correlation $Mi \leftrightarrow Pi$ is entirely out of the question, and the correlation must be regarded as a brute fact that is not further explainable” (1999, chapter 4, p. 99).

The pain-quale is a mental property distinct from the neural C property. The repercussions of this are the following. Bob – a physical organism – simultaneously instantiates a physical (neural) and a mental property (quale). If it is true that our world is made up out of physical things and their physical properties, and if it is true that causes and effects are physical events, then the presence of a purely mental property in such a context seems out of place. Kim believes that properties have causal powers, meaning that their instances, namely physical events, are causes and effects. If the pain-quale is a distinct mental property, then it must have its own causal powers, then its instances must be causes and effects. The problem that now arises has to do with the causal relations among purely mental (quale) and purely neural events. If mental quale-events are causes, if it is the instance of the quale-property that makes Bob scream (physical behavior) when hit in the face, then this constitutes a breach of the fact that physical events have other physical events as their causes and effects, in other words, the fact that all causal relations in our world are explicable in terms of physical events exclusively. In addition to the mentioned breach of the causal closure of the physical domain by mental events (here

qualia), there is also the problem of ‘causal over-determination’ (1999, chapter 2, pp. 38–47). If it is true in our world that physical events are the causes of other physical events, then it is true that the instance of neural C causes Bob to scream when hit by thugs. But if the mental pain-quale-property has causal powers in its own right, if its instance (mental event) is a cause, then it too is a cause of Bob’s scream when hit.

An additional problem has to do with making scientific or theoretical sense of mental causation of physical events. We do not know what it means for a mental event to cause a physical event, if the mental event is non-physical and distinct from the neural event. We do not know how to investigate the causal relations among mental and physical events. In other words, these causal relations are unintelligible to us. Whereas causal relations among physical events do, in principle, not constitute difficulties for scientific or theoretical investigations, experiments, and observations.

If qualia cannot be reduced – and this is Kim’s conclusion in his quoted monograph – the problems that they create are manifold. There is (a) the inexplicable property correlation among qualia and neural properties, which must be taken as ‘a brute fact’ of nature. If qualia have their own causal powers, then this leads to (b) the breach of the ‘causal closure of the physical domain’ holding for physical events. It also invites (c) the problem of ‘over-determination’ in conjunction with the aforementioned ‘breach’. These problems are yet enlarged by (d) the fact that mental-to-physical causal relations are unintelligible to us, provided that the mental event is of a non-physical kind distinct from an event of a neural kind.

From chapter 1 it emerged that properties have a role in causation for Kim. In chapter 2 I showed how Kim's understanding of events as structured entities has led him to strongly tie causation and explanation. Under the pressure of the principle of the causal closure of the physical domain he reductively identified mental attitude properties with neural realizers in order to save mental causation but only as physical causation. Qualia failed to be reductively identified with their realizers.

The following chapter (3) will be concerned with Davidson's understanding of mental causation and explanation. It will be shown that Davison differs from Kim on those questions as a consequence of his understanding of events as explicated in chapter 1.

CHAPTER THREE

In this chapter I plan to explore Davidson's theory of mind. After explaining Davidson's views about the mind I intend to contrast Kim's and Davidson's accounts. This will be conducted in the form of objections and replies. I will conclude with my own reasoning why Davidson's approach should be favored over Kim's.

Davidson's philosophy of mind is driven by the motive that there is no conflict between necessity or determinism in nature and human freedom (1980, essay 11, pp. 207-208). His is an attempt at reconciling apparent contradictions between the rule of deterministic causal laws of nature subsuming physical event descriptions and the domain of human mentality explicable not in terms of physical causal laws but in its own terms, in terms of reasons rather than causes. Our hope, though, is that our minds are relevant in some sense in a purely physical world.

There are three principles which Davidson accepts. First, mental and physical events interact (1980, essay 11, p. 208). An example could be the following. It is minus forty degrees, and Bob lets his dog go outside. He hears the news about how cold it is. This causes him to become concerned about the well being of his pet. Driven by his belief that the dog is cold as well as his desire to save his dog's life, Bob opens the deck-door and lets the dog in. The sound coming from the radio causes Bob to open the door. This is a case of physical-to-mental event causation.

The reverse pattern of causation is also possible, according to Davidson. Say Mary wants to meet Paul. She knows where she can meet him at the university. She

decides to meet him there. Wanting, knowing, and deciding are mental attitudes. In Mary's case these instances are instrumental in making Mary go to the university.

So Davidson accepts that there is causation between mental and physical events, or that mental and physical events interact. This is his principle (1). But he also believes that only physical events are causes and effects, similar to Kim's causal closure principle. These physical events fall under strict causal laws of nature. And this is his principle (2). He calls it the 'Principle of the Nomological Character of Causality': "...where there is causality there must be a law: events related as cause and effect fall under strict deterministic laws" (1980, essay 11, p. 208). The expression 'Nomological Character' refers to something having no exceptions. Events in our world are physical. Every physical event has other physical events that preceded it and physical events that will follow it. Strict causal laws subsume events described as physical events and not events described as mental.

But Davidson's principle (3) says: "...there are no strict deterministic laws on the basis of which mental events can be predicted and explained" (1980, essay 11, p. 208). When it comes to the mental domain and its event descriptions in terms of belief, memory, perception, thought, desire, decision, intention, judgment, knowledge, and so on, it must be said that events described as mental do not fall under, or are not subsumable by, or do not instantiate strict causal laws. One and the same event can be given a neural or a mental description. These descriptions are different in kind, one being mental and the other neural. And since the descriptions of the same event are distinct in kind there is also a difference in terms of whether we explain the same event neurophysiologically or psychologically. When we explain the same event

neurophysiologically we invoke neural (physical) descriptions of it. Described as the neural event of the c-fiber firing it is explainable as the effect of the b-fiber firing of the brain. This means that the event described as the c-fiber firing is explained causally when it is described as that neural event. But when the same event is given a mental description no causal explanation is forthcoming of that event. The event that was described as the c-fiber firing is now described as my belief that the assassins intend to kill me. When the event is so described it needs another mental description to be explained. The event described as the belief that the assassins intend to kill me cannot be explained by the neural description of the b-fiber firing. It can be explained by, say, my seeing the assassins at the door. The mental event description of my seeing the assassins at the door is a reason for me to have the belief that the assassins intend to kill me. When the same event is described as the b-fiber firing it is a cause. But when it is described as my seeing of the assassins at the door it is a reason. Because the event descriptions are different in kind the explanations of one and the same event differently described are distinct as well.

The reasons, which I invoke to explain my belief could be the following. I have this particular belief because I see or perceive through my living room window that Angelo and Jovani are approaching my house. I know that these men are the cold-blooded killers of our local mafia unit. I remember that they killed Toni last week. I judge what the most likely method of killing me will be. I can do this because I know their trademark ways. The former uses knives, the latter baseball bats. All these reasons, plus my knowledge that I am a mafia member and that I have double-crossed the boss, need to be invoked to explain my particular belief. All these mental events taken together can

make sense only against the yet broader background of many more of my thoughts, beliefs, and actually the whole of my mentality or psychology.

Davidson says: “Beliefs and desires issue in behavior only as modified and mediated by further beliefs and desires, attitudes and attendings, without limit. Clearly this holism of the mental realm is a clue both to the autonomy and to the anomalous character of the mental” (1980, essay 11, p. 217). The autonomy of the mental pertains to the explanatory independence of mental explanations from neural explanations. And the anomalous character of the mental has to do with the absence of strict causal laws with regard to mental event descriptions.

Explanations of physical goings-on are different from and incommensurable with explanations of mental goings-on. It is as if they constitute distinct explanatory schemes or ‘patterns’ (Davidson’s word in 1980, essays 11, 12, and 13). Since mental and physical explanations are distinct, it is not the case that, when a mental event is related to a physical event, they fall under a causal law. Davidson says that there are “...no more than rough correlations between the psychological and physical phenomena” (1980, essay 12, p. 231). This means that when an event is described as ‘the belief that the assassins intend to kill me’ it is not best explained as an event described neurally, but its explanation is to be found by invoking other mental event descriptions. This also means “...there are no precise psychophysical laws” (1980, essay 12, p. 231). Or, there are no regularities describable in which physical descriptions correlate with mental descriptions. When we talk about the relationship between the mental and the physical, we do not talk in terms of laws of cause and effect. But when we re-interpret the mental event ‘the belief that the assassins intend to kill me’ as the physical event of ‘the c-fiber firing’, then we can say

that the latter event can be a lawful cause or an effect of another event physically described.

Causation is a relation that obtains between events no matter how they are described. But when the events in question are described in the physical vocabulary then they fall under strict causal laws of nature. When we talk about the relationship between ‘mental characteristics’ (1980, essay 11, p. 214) of, for instance, ‘being a belief’ (or ‘being a desire’, ‘being an intention’, and so on) and ‘physical characteristics’ of ‘being a c-fiber firing’, we can only say that mental characteristics depend on physical characteristics. Davidson says: “Although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect” (1980, essay 11, p. 214).

Davidson believes that the supervenience relation of mental characteristics on physical is an explanation of the relationship between mental and physical characteristics. This is a point of disagreement with Kim’s views. But before I engage in a Kim-style attack on Davidson’s understanding of mentality, I still need to show whether and how Davidson can reconcile the claims (1) that mental events interact with physical events, (2) that only events describable physically instantiate strict and deterministic causal laws, and (3) that there are no strict laws on the basis of which we can predict and explain mental events.

How can mental and physical events interact at all - which is the content of principle (1) - if principle (2) also holds against the backdrop of principle (3)? Recall that strict causal laws are instantiated only by events physically described. In order for mental and physical events to interact (1) there would have to be a strict law subsuming mental and physical events in causal relations, according to principle (2).

Davidson sees no problem here: "...causal interaction deals with events in extension and is therefore blind to the mental-physical dichotomy" (1980, essay 11, p. 215). This means that events interact causally with one another regardless of how we describe them and regardless of how we explain them. The 'mental-physical dichotomy' pertains only to how we describe and explain events. Causation and explanation are two different things. Let us add another quote: "Anomalous monism ...allows the possibility that not all events are mental, while insisting that all events are physical" (1980, essay 11, p. 214). If all events in the world are physical then mental events must be physical too. And this is exactly what Davidson concludes later on and what principle (2) demands: "The demonstration of identity follows easily. Suppose 'm', a mental event, caused 'p', a physical event; then, under some description 'm' and 'p' instantiate a strict law. This law can only be physical...But if 'm' falls under a physical law, it has a physical description; which is to say it is a physical event" (1980, essay 11, p. 224). A mental event can be causally related to a physical event (and vice versa) when the mental event is correctly describable as a physical event. And now both events physically described do fall under causal laws.

This means that my belief that the assassins intend to kill me does cause my feet to move. But the physical movement is not causally explained by invoking the belief-

description of the same event. For the physical movement of my body to be causally explained the (physical) neural description of the same event has to be mentioned. The physical movement is causally explained by the c-fiber firing of the brain. So now we can say that whenever the c-fiber firing event is triggered, the movement of legs follows as a matter of law, or without exception.

Davidson's identity theory of mental and physical events leaves causation to physical event descriptions thereby satisfying the assumption that the physical goings-on in our world are strict and deterministic, exceptionless instances which are subsumed by laws of nature. One part of the nonreductivist physicalist's intuition is thereby satisfied. Now another of his intuitions requires accommodation. And this is the belief that we are free agents. We want to be able to say that and explain why we as minds do not fall prey to the necessity of causal laws of nature, in other words, that freedom of decision and action is a part of our existence.

We require for this principle (3), the no-law or anomaly principle about the mental. Mental event descriptions may not instantiate strict causal laws, but this does not exclude psychological explanation. In order to explain psychological happenings (say my belief that the assassins intend to kill me) we do not offer physical descriptions (say the c-fiber firing). We need to stay within the domain of the psychological in our explanations of mental events. The belief that the assassins are at my door can be explained by appeal to other mental events such as my seeing them through the living room window, my recognizing or remembering that they are killers employed by the local mafia, my knowledge that I am a mafia member and that I have double-crossed the boss, and others. The belief that the assassins are at my door distinguishes itself as this particular belief in

relation to the other mentioned mental event descriptions. I think it would not be wrong to say that my particular belief is a relational entity, whose content depends on the network or context of the mental event descriptions that I claim to be the explanations or reasons for the occurrence of the belief that the assassins are at the door.

The context in which this particular belief is embedded seems to be even broader. The mental events I invoke to explain the belief that the assassins are at my door belong to the whole of my psychology, to all other mental attitudes characteristic of me. The content of any of those ascriptions of mental attitudes of mine depends on relations to other mental events. The whole system of my mental attitudes (about myself and the world I live in) constitutes the broad explanatory background of my belief that the assassins are at the door. The mental descriptions would not make sense in isolation from one another. For me to believe that the assassins are at the door I need to have other beliefs about assassins. I need to have beliefs about what they do and how they differ from other members of the mafia and ordinary people. I need to know that this profession is not an ordinary one; that it also is against the law. This now means that I have to believe that some actions are legal while others illegal, and so on.

The story could be continued like this for any mental event. Davidson says; “There is no assigning beliefs to a person one by one ...” and “...we make sense of particular beliefs, as they cohere with other beliefs, with preferences, with intentions, hopes, fears, expectations, and the rest” (1980, essay 11, p. 221). I think it would not be wrong to explain the idea of ‘coherence’ as the interdependence of mental events with others within the system of the mentality of a person viewed as one whole.

So it seems that mental descriptions are made for one another as members of the holistic pattern of a person's mentality. It is persons that have beliefs, desires, and all other mental attitudes. It is not as if some neural event of my brain has the desire to save its life by jumping out of the bedroom window after seeing the assassins at the door. It is I (this particular person), who behaves in the described way. My out-of-the-window-jumping behavior can be explained by appeal to the mental events of believing that the assassins are at the door and desiring to save my life. These explanations of my behavior figure against the background of the tacitly invoked pattern of mentality of the person who does the explaining of my behavior. It could be I who explain this behavior or somebody else. Regardless of who does the explaining, it still remains the case that the mentality (as a pattern) of the person who does the explaining is tacitly invoked.

Why is it that all those who explain my behavior come up with similar if not identical explanations, provided that they know a lot about me? It is so because they believe that I am similar to them when it comes to mentality (1980, essay 12, p. 239). They believe that my behavior is consistent, because their behavior is largely consistent. My escape behavior is consistent in the light of reasons, namely in the light of my belief that the assassins are at the door and my desire to save my life. Whenever I have reason to believe that somebody is threatening my life I will exhibit escape behavior. And this is what many people who explain my behavior hold too.

Suppose that instead of jumping out of the window I open the door and sing a song when the assassins are at my door. The people seeking to explain my behavior would be puzzled. Knowing everything about my situation, and knowing that I have good reasons to run (they believe that the two men are assassins and they believe that I should

desire to try and save my life.) they cannot explain my behavior. My behavior makes no sense in the light of the mentioned reasons. It is not consistent with these reasons. And my singing behavior stops making any sense at all if we cannot explain it by appealing to my beliefs and desires.

But consistency does not seem to be only a matter of whether behavior is consistent with reasons. It holds for mental events as well. Some belief can be consistent with other beliefs that I hold. For instance, my belief that the war against Iraq is an unnecessary endangerment of peace and stability in the world is consistent with my belief that acceptance of the diversity of belief-systems makes excellent foreign policy and diplomacy, as well as my belief that this acceptance provides a good chance for the survival of all parties involved.

Besides coherence and consistency there is yet a third element characteristic of the mental domain. This is rationality. My singing in the light of danger can probably not be classified as rational behavior. This is, I think, the reason why a breakdown of the explanation of my behavior occurs. Those that are supposed to explain my behavior fail to do so. To them my behavior makes no sense; it is irrational. Similar considerations hold for my belief that Bush is irrational in the light of the reasons I mentioned above. Bush could reply that his belief that the war is necessary is rational in the light of his reasons, namely his belief that the free world must be protected, that the values of the free world have to be realized all over the world because only those values are correct.

If we look upon my and Bush's stories as theories then "...between these theories there may be no objective grounds for choice" (1980, essay 11, p. 222). Interpretation of behavior and attributions of mental attitudes (beliefs, desires, intentions, etc.) can be

diverse. In other words, sometimes one person will be interpretable as having belief system A or system B, and that there will be no objective grounds for saying one interpretation is right and the other wrong. This goes to show that the predictions and explanations of mental events are not as precise or accurate as is the case with physical events, which are the subject of the physical sciences.

So there are three elements characteristic of the mental domain. One was coherence: "There is no assigning beliefs to a person one by one on the basis of his verbal behavior, his choices, or other local signs no matter how plain and evident, for we make sense of particular beliefs only as they cohere with other beliefs, with preferences, with intentions, hopes, fears, expectations, and the rest" (1980, essay 11, p. 221). The other two were consistency and rationality: "My point is that if we are intelligibly to attribute attitudes and beliefs, or usefully to describe motions as behavior, then we are committed to finding, in the pattern of behavior, belief, and desire, a large degree of rationality and consistency" (1980, essay 12, p. 237).

Coherence, consistency, and rationality as characteristics of the mental domain "...have no echo in the physical theory..." (1980, essay 12, p. 231). Explanations of physical events have different features: "It is a feature of physical reality that physical change can be explained by laws that connect it with other changes and conditions physically described" (1980, essay 11, p. 222). I think that an example may be something like this. There is a physical event, which I will call 'the movement of the leg'. This is an event I want to explain. I find out that before there was the movement of the leg there was only one change relevant to bring about the leg movement. It was the event of the c-fiber firing, which occurred in the brain of the specimen I investigated. I investigate many

such leg movements individually. I conclude that each individual leg movement is a consequence of an event physically described as ‘the c-fiber firing’. I have also conducted these experiments under identical circumstances, say leg movements of human bodies standing upright and influenced by earth’s gravity rather than leg movements of dogs influenced by earth’s gravity or leg movements of human bodies affected by neural disease x. Now I can formulate a law-like statement. Whenever there is an event physically described as ‘the c-fiber firing’ occurring under same conditions – they are of human rather than dog brains unaffected by disease x – another event physically described as ‘the leg movement’ follows as a matter of necessity. There are no exceptions for the causal relations of the two events physically described while same conditions hold.

The leg movement is explained by describing the causing event as the c-fiber firing of the brain and not by the mental description the belief that the assassins intend to kill me. The effect event of the leg movement can in principle be predicted and explained in advance with certainty provided that we know that the c-fiber firing exists under the conditions specified in the above law-like statement. But what we can do with events physically described we cannot do with events mentally described. We cannot explain and predict behavior of persons or mental events with the kind of certainty characteristic of physical events: “...nor can we expect ever to be able to explain and predict human behavior with the kind of precision that is possible in principle for physical phenomena” (1980, essay 12, p. 230).

Somebody knowing a lot about my situation and psychology (the equivalent of the specified necessary conditions in the case of the c-fiber firing) including that I have

the belief that assassins are at my door (equivalent to the sufficient condition for me jumping out of the window) cannot with certainty predict and explain what I am going to do. It is likely that I may try to escape, but it is also possible that I may act differently under mentally identical circumstances. This unpredictability means that whoever tries to explain my behavior will not know in advance whether the necessary conditions for my behavior will hold (those conditions being the knowledge of my situation and psychology, e.g., that I am a mafia member, that I double-crossed the boss, that these two men are killers, etc.). Say Peter wants to explain and predict my behavior. He spots the assassins first. He knows that I am a mafia member, that I have double-crossed the boss, and so on; in other words, he knows the conditions that seem to be necessary for my escape-behavior. On the basis of that, Peter speculates: Once Bob sees the assassins, he will believe that they are at his door and that they intend to kill him, and he will desire to escape. These mental event descriptions are Peter's predictions and probable explanations of what I will do. Peter's predictions will be correct since in my example I do jump out of the window. But my behavior could have easily been different once I spot the assassins. Peter's prediction regarding my behavior would then have been incorrect. Peter cannot with certainty say how I will behave, if he only knows psychologically-described facts about me.

Say we think of my belief that the assassins are at the door and my desire to save my life as the sufficient condition (or that which is enough for my behavior; recall my example from my chapter 1, the striking of the match against the matchbox surface was the sufficient condition for the lighting of the match) of my out-of-the-window-jumping-behavior. Now add to this the fact that necessary conditions for the same behavior must

obtain too. I said that, among other things, they could be ‘knowledge that I am a mafia member and that I have double-crossed the boss’, ‘memory of what happened to past traitors’, and so on. Davidson holds that one cannot fix these necessary conditions for my behavior in advance: we cannot “...determine in advance whether or not the conditions of application are satisfied” (1980, essay 12, p. 233). Peter, who tries to explain and predict my behavior, may know a lot of my reasons, but he does not know, before I behave in some way or other, whether these (necessary conditions for some behavior) reasons obtain for me at the moment I see the assassins at the door. And since he is not sure whether the necessary conditions for the out-of-the-window-jumping behavior are satisfied, he cannot predict my behavior in advance with certainty. It is only in hindsight that he can say: Before Bob saw the assassins I predicted that he will jump out of the window once he saw them. I was wrong. Once he saw the assassins Bob did believe that they intend to kill him and he had the desire to escape (sufficient condition for the out-of-the-window-jumping behavior), but he froze out of fear instead. So it must be that his psychological attitude towards authority was one of the conditions among others necessary for his freezing behavior. Even though Peter knew beforehand that Bob’s psychological attitude towards authority is one of the necessary conditions for Bob’s behavior Peter could not have with certainty predicted that the description of the psychological attitude towards authority would have more weight over other necessary conditions in this instance of Bob’s behavior.

The fact that we cannot predict mental events and behavior with certainty amounts to Davidson’s principle (3). There are no strict laws that would subsume mental events in their relations to one another. But there are strict laws subsuming physical

events in their causal relations (2). Explanations of events physically described are qualitatively different from explanations of events mentally described.

The next question becomes: Can we at all relate these disparate explanatory schemes or patterns to one another? If so, what is the character of this relationship? Davidson thinks that we can relate physical and mental / behavioral event descriptions to one another. It is possible to say that c-fiber firing events are often followed by out-of-the-window-jumping behavior. The relation expressed by the words followed by is not that of causation. The c-fiber firing can cause the jumping-out-of-the-window behavior. But the description of the c-fiber firing is not related to the description of the behavior as a matter of natural law. We cannot say that whenever an event described as the c-fiber firing occurs it must be followed by an event described as the jumping-out-of-the-window behavior. And since events described in mental and behavioral terms do not instantiate strict laws, because laws are linguistic entities, there cannot be strict psychophysical laws correlating neural and mental / behavioral event descriptions either.

It may be often true that c-fiber firing is followed by jumping out of the window, but this does not mean that the events described in terms of qualitatively different vocabularies are law-like in the strict sense of causal laws. We, says Davidson, "...have no reason to believe it more than roughly true" (1980, essay 11, p. 216), and there "...cannot be tight connections between the realms if each is to retain allegiance to its proper source of evidence" (1980, essay 11, p. 222). The phrase 'allegiance to its proper source of evidence' means that in order to explain behavior or mental attitudes such as desire, beliefs, intentions, and so on, we must use mental event descriptions if we intend to properly explain them. And if we want to explain the c-fiber firing in the brain we need

to stay within the realm of physical event descriptions, which invoke the cause of the c-fiber firing. When we use event descriptions of the realm of behaviors, beliefs, desires, and so on, we interpret people as (roughly) rational agents. We can also interpret people as mechanisms described in physical terms (e.g. biological, chemical). But there is no guarantee that a system described in physical terms will also be describable in terms of the rational pattern of explanation; or, there is no guarantee that a system described in physical terms will map on to rational patterns.

The reason why we can use ‘roughly true’ generalizations or statements involving ‘the c-fiber firing’ (neural description) and the event described as ‘the jumping out of the window’ (a description of an action, which requires mental event descriptions to be properly explained) can be found in the presence of a connection that is less tight than causal laws would require. Davidson believes that mental and physical characteristics are related to one another at least via the relation of supervenience: “Although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect” (1980, essay 11, p. 214) and “Although, as I am urging, psychological characteristics cannot be reduced to the others, nevertheless they may be (and I think are) strongly dependent on them. Indeed, there is a sense in which the physical characteristics of an event (or object or state) determine the psychological characteristics...” (1980, essay 13, p. 253).

Now we know why jumping out of the window often follows c-fiber firing. The former is an action, which supervenes or is determined by a physical event. The underlying physical event (say, the leg movement) is the effect event of the c-fiber firing of the brain. The c-fiber firing re-described as the belief that the assassins intend to kill me remains the cause of the leg movement. But because the belief that the assassins intend to kill me supervenes on the c-fiber firing, is it possible to express statements such as: the leg movement followed the belief that the assassins intend to kill me.

So to be precise: "...it is descriptions of individual psychological events, not sorts of events, that are supevenient on physical descriptions. If a certain psychological concept applies to one event and not to another, there must be a difference describable in physical terms. But it does not follow that there is a single physically describable difference that distinguishes any two events that differ in a given psychological respect" (1980, essay 13, p. 253). I think that what he is saying here is that the description 'the belief that the assassins intend to kill me' supervenes on or is determined by, in some sense, the description 'the c-fiber firing'. Suppose the description 'being a belief that the assassins intend to kill me' applies to a particular event a. But this description does not apply to event b. The reason why is that there is some difference between events a and b describable in physical terms. But b may differ physically from a not in one physical feature alone; there may be several physically describable features of event b that distinguish it physically from event a.

Let me explain this with an example. Neural event a occurs in Bob's brain. For our purposes it suffices to say that the a-event is a neural or physical description of the occurring event. The mental description 'the belief that the assassins intend to kill me'

supervenes on the neural a-event description. This is the case for Bob. But for Peter things may be different. Peter has the same belief as Bob, but his belief that the assassins intend to kill me supervenes on the neural b-event description instead. This expresses the idea that the same mental event description can supervene on different neural event descriptions.

Besides beliefs Bob also has desires. There surely is a difference describable in neural terms between Bob's belief that the assassins intend to kill me and his desire to escape. His belief depends on the neural a-event description and his desire supervenes on the neural b-event description. Although Peter has the same belief and the same desire the neural event descriptions on which the former two supervene may differ: neural b-event description for the belief, and neural c-event description for the desire.

Supervenience of descriptions seems to allow us to formulate 'roughly true' statements or generalizations involving the c-fiber firing (a physical event) and the jumping out of the window (an action), even though there are no causal laws connecting mental and physical event descriptions. But Kim claims that if "...there are no laws connecting mental kinds or properties with physical ones", then "...there is no connection between our mental nature and our physical nature" (1999, chapter 1, pp. 4-5). Kim says that Davidson's is "...a negative thesis: it tells us how the mental is not related to the physical, and says nothing about how the two are related" (1999, chapter 1, p. 4).

Kim wants a 'positive' story about 'why' the mental is related to the physical the way it is. Davidson's supervenience thesis (1980, essay 11, p. 214) does not provide an answer to this question. Kim believes that "...supervenience concerns mental and physical properties and kinds (or "characteristics" or "respects", in Davidson's words),

not mental and physical tokens or Davidsonian events as structure-less particulars" (1999, chapter 1, p. 6).

The notion of supevenience is attractive in that it says that the mental depends on the physical but not vice versa, and that it seems "...consistent with the irreducibility of the mental to the physical" (1999, chapter 1, p. 6). Kim believes that supervenience "...can be viewed as defining minimal physicalism" (1999, chapter 1, p. 15), and nothing more. The thesis seems to amount to: "...a pattern of property co-variation between the mental and the physical and points to the existence of a dependency relation between the two. Yet supervenience is silent on the nature of the dependence relation that might explain why the mental supervenes on the physical" (1999, chapter 1, p. 14).

Kim thinks that the 'why' of supervenience is explained via 'physical realizationism'. Being a belief that the assassins intend to kill me supervenes on being a c-fiber firing because the property of being a c-fiber firing realizes the property of being a belief that the assassins intend to kill me. We can say that in individuals who belong to the same class (say humans, or anybody like Peter, or even anybody like Peter at 3 p.m. Sunday June 15th) it is the case that the physical property of being a c-fiber firing realizes the mental property of being a belief that the assassins intend to kill me. Provided that we are interested in organisms that have a relevantly similar physical constitution at a particular time and provided that we hold the laws of nature constant, from this it follows that the second order functional property of being a belief that the assassins intend to kill me can exist only if the physical property of being a c-fiber firing is instantiated. In other words: "From the definition of realization, it follows that P is sufficient for M (that is, if a system of kind S instantiates P at t, it instantiates M at t) – in fact, given the nomological

constancy just noted of the realization relation, it follows that P is nomologically sufficient for M" (1999, chapter 1, p. 23). This also means that having the belief that the assassins intend to kill me is having a physical property that satisfies the causal specification of having this belief's typical causes and typical effects. The neural realizer property is a causal intermediary in-between the physical properties of seeing assassins and escaping from them. The having of the belief 'simply is' (or is the same as, or, is identical with) the c-fibers firing, and nothing more (1999, chapter 1, p. 24). (For the readers who are now led to the belief that Kim is not a functionalist but a type-identity theorist see the upcoming chapter 4 on critiques of Kim's view.)

From this it follows that mental properties are nothing over and above physical properties and that they do not have causal powers in their own right. When applied to Davidson's views the reasoning that preceded implies that mental properties (or so interpreted: e.g., the belief that the assassins intend to kill me) are shadow-like entities: "But this apparently threatens the causal relevance of mentality: the fact that 'm' is a mental event – that it is the kind of mental event it is – appears to have no role in determining what causal relations it enters into", and "This seems to consign mental properties to the status of epiphenomena" (1999, chapter 2, p. 34). This means that beliefs, desires, intentions, hopes, and so on may exist but only as shadows, which are not causes. Our minds as minds cannot do anything in this world of ours, which is physical.

Kim's move to explain supervenience in terms of physical properties realizing mental ones cannot be accepted by Davidson because (a) the mental scheme of interpretation demands different concepts; it does not need concepts of neurophysiology. Belonging to a different domain, beliefs, desires, and so on, cannot be reduced to 'the c-

'fiber firing' or other neurophysiological concepts. The other point, (b), is that Kim is mistaken about what entities supervene on one another. It is not mental properties or 'sorts of events' that supervene on one another, but it is mental event descriptions that supervene on physical event descriptions (1980, essay 13, p. 253). There is only one event, which is interpretable either as the belief that the assassins intend to kill me or the c-fiber firing, depending upon the fact from within which scheme of interpretation that particular event is viewed.

Kim believes that anything other than the identity of mental and physical properties warrants a consequent (c) causal and (d) explanatory exclusion of mental properties. If there could not be a mental property without a neural property, or, if the causal capacities of mental properties depend on those of the neural ones, and, provided that mental and neural properties are distinct in kind in conjunction with the assumption that the world in which we live in is causally closed under physical events, then all causal work is done by events which are instantiations of neural rather than mental properties. The neural property of the c-fiber firing realizes the mental property of the belief that the assassins intend to kill me. Kim stresses that events are causes and effects in virtue of being instantiations of some properties. The in virtue of what (property) a cause is a cause (similarly for effect events) is very important for Kim's understanding of the causal relation of events. To explain what the cause of an effect was one has to appeal to the real virtue (property) which determines the fact that the event in question really was 'the' cause of some effect. Davidson disagrees with this view. Only dated unrepeatable events are causal entities. Events are causes and effects regardless of how they are interpreted, or, regardless of under what virtue (property) they are viewed. But for Kim, events

consist of three elements, which constitute an organic unit, namely the event itself. The structure of an event, Kim thinks, consists of an object (Bob) having a unique property (falling) at a time (t). This is a metaphysical commitment Davidson cannot share. The event, the falling of Bob at t, is only one way of interpreting it. There may be different ways of describing the same event without distorting the event itself, or without changing anything about the event. The event remains the same regardless of whether it is said of the event that it is ‘being a falling’, ‘being a dropping’, or ‘being a slipping’.

Since properties do not play a role in causation according to Davidson (c), it may be argued that Kim’s accusation that mental properties such as being a belief that the assassins intend to kill me are realized by physical ones misses the point. The assumption that mental properties are causally excluded by property-identity achieved through the realization relation is flawed since properties do not participate in causation anyway. Since causation is a matter of events and not of properties, properties cannot be causally excluded. In other words, there is nothing to be excluded there, Davidson may argue.

Davidson would disagree with Kim’s accusation that mental properties are (d) explanatorily excluded. Using Davidson’s language, we can say that the mental description (Kim would say property) is not excluded by the neural description of the same event. The event described as ‘the out-of-the-window-jumping behavior’ can be explained by invoking the reasons for it. And these are other psychological descriptions. But when the same event is to be causally explained then we have to invoke neural descriptions of it. Kim’s assumption that the belief that the assassins intend to kill me is explanatorily excluded from the out-of-the-window-jumping behavior (if Davidson is right about mental event types not being physical types) is again connected with his belief

that properties have a role to play in causation as well as that a real explanation is usually causal. Davidson would argue that, firstly, as already said, properties play no role in causation, and there is, therefore, nothing to be excluded there. And secondly, not all explanations are causal. The explanation of the effect event of ‘the leg movement of some body’ may be causal since it can invoke, say, the event described physically as ‘the c-fiber firing’. But the same event interpreted as ‘the c-fiber firing’ can be re-interpreted as ‘the belief that the assassins intend to kill me’. When this is done another kind of explanation is required, namely the one that invokes other mental attitudes, behavior, and the concept of a person. This kind of explanation is not causal per se but rather rational. This amounts to saying that, on Davidson’s view, we cannot conflate rational and nomological / causal explanation.

Kim would disagree. He makes no distinction between rational and causal explanation. To explain an effect means to invoke the real cause event of it. And the real cause event is the one the virtue of which was causally relevant to bring about the effect event. Kim wants to be able to say that it is in virtue of the property of being a belief that the assassins intend to kill me that the effect occurred. It is because of this and only this property that the cause event was the real cause of this and only this effect event. Any other properties such as my belief that it is sunny today should not influence the causal process of the two mentioned events.

And since Kim does not make a distinction between causal and rational explanation (but Davidson does make this distinction), he must reduce the mental property to its neural subvenience base in order to accommodate physicalism. Kim states that physical events have only physical causes and only physical effects. Mental and

physical properties as well as their instances (events) are now identical. If mental and physical events are identical, no breach of the closure on physical events can occur. The reductive identification of mental and physical properties is bad, one can argue, because mental properties end up having no distinct causal powers. But Kim sees a positive side in the reductive identification. Mental properties may not have independent causal powers, but they have causal powers nevertheless. Since mental properties are identical to neural properties, they have or inherit the same causal powers the neural properties have.

Kim's reductive identification seems to have yet another positive element. Since mental properties are not distinct from neural properties, it cannot be said that they are epiphenomena. Mental properties as epiphenomena can be understood as being distinct in kind from neural properties in conjunction with the fact that instances of physical properties are real causes and effects. This means that mental properties are no causal powers whatsoever. But on Kim's account they do have causal powers.

Kim can argue against Davidson in the following way. Davidson's mental properties / descriptions are distinct in kind from neural properties. All the causal work is done by physical events, which are instances of physical properties. The mental properties end up being epiphenomena having no causal powers whatsoever. Since only physical events can be causes and effects events, mental events are causally and explanatorily excluded. Kim can argue that this is what happens on the Davidsonian account. Kim could say that on Davidson's view mental and neural properties are just distinct in kind and that mental properties supervene on neural ones. And since only physical events can be causes and effects there cannot be any interaction between mental and physical events because that would amount to a breach of the causal closure of the physical domain.

Kim proposes a remedy. The aim would be to show that mental properties are not epiphenomenal, or that they do have causal and explanatory powers in a physical world. Kim believes that this can be achieved by physical realization. Mental properties (or at least intentional ones) are second-order functional / relational properties defined over their first-order realizers. The neural property of the c-fiber firing is the physical realizer of Bob's belief that the assassins intend to kill him. Holding the laws of our world constant it can now be said that whenever there is the same input (seeing of assassins) the neural c-fiber firing follows. And the instance of the latter causes some output. The laws of our world guarantee the constant conjunction of same types (properties) of input, intermediary neural properties (c-fiber firing), and same kinds of output.

But to save the causal powers of mental properties something else is needed in addition to physical realization, according to Kim. If mental and neural properties remain distinct in kind, if supervenience of mental on neural properties holds, if only physical events can be causes and effects, and if neural properties realize mental ones, it follows that mental properties have no causal powers whatsoever. Kim believes that the problem can be solved by identifying my belief that the assassins intend to kill me with the c-fiber firing of my brain. The mental and neural properties are not distinct in kind anymore. This means that the mental property inherits the causal powers of the neural property with which it is identified. Now the mental property has causal powers, and there is no threat of epiphenomenalism anymore. The mental property ends up being explanatorily relevant because it inherits the causal powers of the neural property. My jumping out of the window can be explained by invoking my belief that the assassins intend to kill me since this mental property is identical with the c-fiber firing of my brain. The mental event

falling under the type of being a belief that the assassins intend to kill me becomes a real cause of my jumping out of the window. And since that event is an instance of that type it can be said that the virtue of that type is relevant when it comes to explaining the type of effect (jumping out of the window) that issues. On Kim's account, the mind becomes a physical entity in a physical world.

Davidson could reply that the mental (my belief that the assassins intend to kill me) and the neural description (c-fiber firing) – Kim would say properties - cannot be identified because they belong to disparate explanatory schemes. An event described as the c-fiber firing requires a causal explanation. But when the same event is described as the belief that the assassins intend to kill me we need to invoke reasons. We then use descriptions of behavior and other mental event descriptions as well as the concept of a person. The mental event description may be dependent in some sense on the neural event description, but it remains distinct in kind. Even though the descriptions (Kim would say properties) are distinct (or not identical) from one another, the event they are descriptions of may still remain the same. So if Davidson talks about identity it definitely does not refer to properties. But Davidson does accept event-identity or token-identity. There is only one event here, which we have described causally as the neural c-fiber firing and rationally as the belief that the assassins intend to kill me.

From this reasoning it is obvious that Davidson cannot accept Kim's view that the neural property or description is identical with the mental property or description and that the neural property realizes the mental property. On Davidson's view there is a distinction between causation and interpretation or explanation. Causation is a matter of events regardless of how we describe them. And interpretation or explanation has to do

with how we describe events initially. Causation is something that happens out there in the world, and explanation is a matter of language. Davidson can argue that causation and explanation should not be confused or conflated. He would accuse Kim of doing exactly that.

Davidson would argue that the event described as ‘the c-fiber firing’ is also a mental event since it is interpretable as ‘the belief that the assassins intend to kill me’. And although the descriptions of the processes of the mind are not identical with the descriptions of the processes of the brain, Davidson would say that the mind still has its place in the physical world since some physical events can also be interpreted as mental. And if mental events are unconstrained by causal laws, then there is space for human freedom. Our decisions, deliberations, thoughts, beliefs, desires, and so on, are free from the necessity of causal laws, which does not tolerate exceptions. As persons we are free agents, but as physical bodies we cannot escape natural necessity.

Chapter 4 will consider two important arguments against Kim. But first I would like to make some points of my own comparing Kim to Davidson.

My main concern has to do with Kim’s identification of mental and neural properties. I think that attempts at identifying the mind with the brain should be rejected. The reason for the rejection of reductive moves could be something like this. When somebody asks why Bob jumped out of his bedroom window he does not want to hear a story about c-fiber firings. What he wants to know are the reasons, as Davidson would say, that were the driving force for Bob’s behavior. These reasons are always about things, people, behaviors, and other reasons. It seems that reasons have a different meaning for a person than whatever the meanings of their physical and neural substrates

may be, regardless of how the relation of supervenience and realization may be explained.

I can imagine two scenarios. Bob is prevented from seeing the assassins. But instead he is visually presented with a physical and neural description of the situation underlying Bob's seeing the assassins. The description involves an elaborate scientific story about perception, external factors that really exist (the assassins are really at the door), and a complicated neurophysiological story about Bob's brain. But all this, it seems, is not what Bob really needs. Bob must see the assassins in order that this information may mean something of importance to Bob as a person.

Once Bob sees the assassins, which is the second scenario, there is a host of contentful happenings in his mind. Beliefs, desires, fears, hopes, and other attitudes may cross his mind. Bob fears that he could lose his life. He hopes that he can avoid being killed. He believes that he must escape in order to save his life. He knows that if he does not escape he has no chance of surviving. All these contentful mental events have importance as they are had or instantiated by Bob the person. These mental events having mental content can mean what they mean only in relation to the person that has them and the broader environmental context (whether real or imaginary) the person finds himself to be situated in.

It appears that the relevance of mental events (having the content they have in relation to the person who has them) is of a different kind from whatever the relevance of physical and neurophysiological descriptions or Kim-style properties may be. The c-fiber firing has significance only when embedded in a physical and neurophysiological vocabulary but not in the vocabulary of contentful attitudes of the mind. Davidson's

approach preserves this distinction into different kinds of concepts needed in psychological and neurophysiological explanations. But Kim's reductive identification of Bob's belief and Bob's c-fiber firing fuses concepts, which should remain distinct, because their meanings differ in kind.

If Bob's belief (reason) and his c-fiber firing (cause) remain explanations distinct in kind, then perhaps it is possible to say that we are free agents. And on Davidson's account it seems that this is the case. By conflating mental and neural properties or descriptions Kim erases the independent meaning mental contents have. I believe that Kim's identification of M and P may actually eliminate mental contents altogether. What remains after the elimination are physical and neural processes. To me it seems that on such an account there is no space for freedom and agency as it is conceived by common sense. If I am right that on Kim's account mental contents do not count (unless they are identical with neural events), then it is the case that in whichever ways our bodies move, the movement is a result of causal processes among environmental and neural events.

I think that the reason why Kim ends up in this dead end for mental contents has to do with his acceptance of the closure principle in conjunction with his realism about causal explanation, or in other words, his conflation of causation and explanation, as well as his analysis of the structure of events as objects exemplifying properties at times. It is no wonder that on these three assumptions mental properties end up having no distinctive causal powers. The need to explain effects such as the jumping out of the window via its mental cause (which is an instance of a mental property) within the context of the causal closure of the physical domain leads to the fact that neural events exempt mental events as causes and to the loss of causal powers of mental properties. And Kim's remedy

(reductive identification conjoint with the causal inheritance principle) for the preserving of causal powers of mental properties only makes things worse by identifying properties or descriptions, which should remain distinct. On top of that, it seems to me that, even on Kim's understanding of what a property is, mental properties do not really exist at all. The reason for this is that for a property to be a property it must have causal powers in its own right. In other words, the instances of a property must be capable of causing effects other than instances of other properties. But since mental content properties are identical with neural properties, their instances have the same effects that the instances of neural properties have. And it should be clear to Kim that, on his own understanding of what a property is, his reductive identification eliminates mental properties altogether. Instead, in order to be consistent with his understanding of properties, Kim might have argued that mental properties are distinct from neural properties. The instances of mental properties cause effects other than the effects caused by instances of neural properties. But he ends up denying this because it would mean that mental events are independent causes, which would breach the principle of causal closure. My claim is that Kim has to be inconsistent about his own understanding of properties in order to be consistent with the requirements of the principle of the causal closure of the physical domain.

Davidson's approach is better because its implied assumptions – seemingly incompatible (see Davidson's three principles above) - are such that, when consequently thought through, they accommodate our common sense notion of mentality. On his approach we do act in accordance with reasons. Davidson manages to satisfy common sense assumptions about the relation between reasons and actions, because of his substantial disagreement with Kim with regard to two assumptions. Both, Kim and

Davidson, seem to believe in the closure principle, namely, that physical events (or so described, Davidson would say) are causes and effects of other physical events. But Davidson rejects Kim's causal explanatory realism. He instead makes a distinction between causation and explanation. (Recall that causation is a matter of events in extension regardless of how they are described. And different descriptions are used in explanations of different kinds.) This gives Davidson the advantage over Kim in that it allows the domain of the mental to be distinct from the domain of the physical, at least in some respect. Reasons remain explanations of actions or behavior. We can ask our usual questions and in return we receive answers we expect. Why did Bob jump out of his bedroom window? The answer is because Bob believed that the assassins intend to kill him and because he desired to save his life.

This ‘because’ is causal for Davidson (1980, essay 1, “Actions, Reasons, and Causes”). This is what common sense requires. Bob’s belief and desire made him do what he did. Davidson’s restriction on reasons being causes has to do with the notion of re-description. The reason is a cause, however it is described. But it is only evident that it is a lawful cause when physically described. I can see that there may be objections with regard to this claim. Some might argue that what we really want to say is that reasons as reasons (or described as reasons) must be the causes of actions and not neural events. Davidson’s answer could be that this may be so on the assumption of Kim’s causal explanatory realism, but not on the assumption of the split between causation and explanation. Davidson could also say that if reasons are causes described as reasons (that actions are explained by their real causes, Kim’s causal explanatory realism) then the closure principle would be violated, or, physicalism would be false. But if physicalism

holds, then reasons are either epiphenomena or Kim is right and reasons are identical with neural properties / descriptions.

It seems obvious that the dispute between approaches like Kim's and Davidson's cannot be settled, because of the different starting assumptions of both views, namely those of the split between causation and explanation (Davidson) and the fusion of causation and explanation (Kim). In order to settle the dispute we would need an account of causation, explanation, and their relation, which we would know to be true. But as far as I understand there is no such theory.

The split of causation and explanation helps Davidson preserve the distinction between the mental and the neural domain. But there is yet another tool, which is instrumental in this endeavor. And this is his liberalism when it comes to event descriptions. As already said, an event can be described as a belief, as something neural, chemical, or even physical (elementary physics) without changing anything about the event described. Kim cannot say this. He must say the opposite. If an event is differently described, or, if an event is an instance of a different type (property), then it is a different event altogether. Kim is forced to look for 'the' properties or descriptions instanced in the cause and effect event respectively. Even on his own admission, it is very difficult to find and know these unique properties that are instanced in causal relations of their events. It seems to me that Kim's hunt for 'the' causal properties may be an adequate tool within the context of his causal explanatory realism, because he needs to find the properties the causal powers of which are responsible for the bringing about of particular effects. But in the end what his project seems to lead to is the disappearance of mental content.

When the approaches of both philosophers are compared, it seems that Davidson's is superior to Kim's. The reason why I think so has to do with the fact that Kim's view seems to have more shortcomings than Davidson's overall. I have discussed some of my own thoughts on these shortcomings above, but more will follow in chapter 4. And if nothing more, Davidson's approach preserves the distinction between the mental and the physical.

In this chapter I explained Davidson's account of the human mind. I have contrasted Davidson's and Kim's views. I have also passed my judgment for why I think that Davidson's approach is superior to Kim's. My remaining task is to investigate whether and how the views of Kim, convincingly argued for in "The Mind in a Physical World", can be exposed to a more convincingly damaging critique. Chapter 4 will focus on explanations of some possible shortcomings of Kim's understanding of the mind.

CHAPTER FOUR

I plan to talk about two ways of responding to Kim's views expressed in his "Mind in a Physical World". The first approach I intend to explain exploits flaws in Kim's supervenience argument, which Kim designed to show that if mental event causation were independent from physical event causation or if it were an addition to causation understood as being closed under only physical events, it would make no sense in a purely physical world. The result will be that there is a place for mental event causation even though the world is physical. This view holds that mental properties are distinct from neural properties, that they depend on neural properties, and that the instances (events) of these mental properties are causes nevertheless, even though the properties from which they issue are dependent on neural properties. The second approach argues that Kim is not really a functionalist about human mentality. Kim holds that attitudes of the mind are defined as the having of same causes and same effects in conjunction with the fact that mental properties supervene on and are realized by neural properties. It will become obvious that Kim cannot coherently hold these assumptions while simultaneously identifying mental properties with their neural realizers.

Crisp and Warfield perceive four steps in Kim's supervenience argument, which is meant to show that mental causation ends up being unintelligible in our physical world, if mental properties are distinct from physical properties. Their intention is to disrupt two main premises in Kim's argument showing them to be incorrect and thereby leaving open the possibility of property dualism or non-reductivist materialism with regard to mental

causation. The authors shorten Kim's ten steps of the supervenience argument (Kim, 1999, pp. 38-47) into four essential steps:

P1: Either Supervenience holds or it fails.

P2: If it fails to hold, then, if Closure and Property Dualism, mental causation is "unintelligible".

P3: If it holds, then, if Closure and Property Dualism, mental causation is "unintelligible".

C1: Therefore, if Closure and Property Dualism, then mental causation is "unintelligible" (2001, pp. 305-306).

Crisp and Warfield do not find any flaws in premise P1. This premise says that the mental property of being a belief that the assassins intend to kill me either depends on the property of being a c-fiber firing or not. If the mental property supervenes on the neural property, then it is tied in some sense to the goings-on in the brain. But if there is no supervenience of the belief on the c-fiber firing, then the realm of the mental is independent of the goings-on in the brain. The content of the Closure principle is that physical events have only physical causes and only physical effects. And Property Dualism is the belief that mental properties are distinct from neural / physical properties, for instance, that the property of being a c-fiber firings is distinct from the property of being a belief that the assassins intend to kill me. The former is a property of the brain of that particular organism, and the latter is a property of the mind of the same being. The mental property is different in kind from the neural property. It is usually believed that the mental property cannot be reduced to the property of the c-fiber firing.

Let us first see in more detail what premise P2 says before we investigate how Crisp and Warfield oppose it. This premise consists of three metaphysical assumptions.

(1) It presupposes that my belief (or for that matter any other mental phenomenon) does not depend on the c-fiber firing; (2) that physical events have only physical causes and only physical effects; (3) that mental properties (such as being a belief that the assassins intend to kill me) are distinct from neural properties. Under these three metaphysical assumptions mental causation makes no sense, it is “unintelligible”, Kim argues. If these distinct properties are related in some other way than supervenience suggests, then we cannot have a clue as to what type of relationship that might be. We would not know how to investigate the causal relations between mental and mental as well as mental and physical events.

But the most important drawback of this situation would be that mental causation represents a violation of the principle of the causal closure of the physical domain. This is (2) the metaphysical thesis that causation is a matter of physical events. So, on the assumptions in P2 (- Supervenience + Closure + Property Dualism) it follows for Kim that if mental to physical causation exists, then this implies that the causal closure of physical events is “breached” (Kim, 1999, p. 40). In other words, if mental properties are distinct in kind from neural properties, and if mental properties do not supervene on neural properties, then, if mental to physical event causation exists, the principle that physical events are causes and effects is falsified. It then is not true that causation is closed under physical events.

Crisp and Warfield disagree that the principle of the causal closure of the physical domain is breached and falsified on the assumptions that mental and physical properties

are distinct in kind and that mental properties do not supervene on neural properties. They imagine a world in which the mental and the physical domain are independent and in which they run in parallel (Crisp and Warfield, 2001, p. 307). My belief that the assassins intend to kill me is always accompanied by some neural property. The relation of accompaniment is not that of supervenience. The mental property does not depend on the neural property. Since it is only accompaniment of mental and neural properties that is the case in this scenario, it is consistent to say that my c-fiber firing can occur without my belief that the assassins intend to kill me. Since supervenience is absent we cannot say that necessarily if the c-fiber firing occurs then the belief occurs as well. Even in the absence of my belief that the assassins intend to kill me is it possible to say that there may be a physical law that subsumes neural and other physical events in causal relations. This causal law may state that there is some physical event (say the c-fiber firing) that precedes the physical event of my jumping out of the window.

The reason why Crisp and Warfield believe that there is no breach and falsification of the causal closure of the physical domain is the following. Under the above assumptions of the absence of supervenience and the holding of closure and property dualism, the instance of my belief that the assassins intend to kill me can be accompanied by some neural event. And there may even be a causal law that says that whenever there exists some neural event (it may even be the c-fiber firing event), then the physical event of my jumping out of the window follows as an effect. My belief that the assassins intend to kill me can be assumed to be accompanied by an instance of some neural property. Crisp and Warfield believe that accompaniment is not the same as supervenience. And whenever that belief of mine, as an instance of a mental property, is

the cause of the physical event of my jumping out of the window, there exists some neural event, which is causally sufficient for the physical event of my jumping out of the window to occur. Since there is some neural event that is causally sufficient for the physical event of my jumping out of the window, no breach or falsification of the causal closure of the physical domain is committed when my belief is also a cause of my jumping out of the window. We seem to have two causes of the physical event of my jumping out of the window. One is my belief that the assassins intend to kill me, and the other is the neural event of my c-fiber firing. But that does not matter to Crisp and Warfield. As long as there exists the event of the c-fiber firing subsumed by a causal law as the cause of the physical event of my jumping out of the window the physical domain of causes is not breached even when my belief is also said to be a cause of my jumping out of the widow.

The hypothetical case of a world of accompaniment of properties as well as mental to physical event causation on top of the existence of events sufficient to bring about desired effect events seems odd at best. Once there is a physical event (e.g. the c-fiber firing) sufficient for a physical effect event (e.g. my jumping out of the window), the mental event that allegedly also causes my jumping out of the window is irrelevant for the bringing about of the event of my jumping out of the window. But Crisp and Warfield do not seem to be bothered. All they are concerned with is to prove Kim wrong. They want to say that Kim is mistaken that closure is violated under (P2) the assumptions that supervenience fails, closure and property dualism hold (Crisp and Warfield, 2001, p. 307). They argue that Kim is wrong in believing that mental to physical event causation violates the closure principle on the causal relations among physical events. In the

hypothetical world mental events (instances of my belief that the assassins intend to kill me) can be causes of physical events (instances of my jumping out of the window) without violating the causal closure of physical events provided that there are physical events sufficient (my c-fiber firing) to bring about the physical effect events that the mental events are alleged to cause as well. Since there exist physical cause events sufficient for the physical effect events, the mental event causing the same physical effect event does not breach the causal closure principle.

This is an unusual notion of causation, which allows a mental event parallel with a neural event to be a cause of a physical event just because it happens to be accompanied by a neural event. This notion of causation should not worry Kim because –as already said above – the neural event is sufficient for causing my jumping out of the window, but the mental event is not sufficient for it, or necessary. The mental event has no causal relevance whatsoever. If the mental event did not accompany accidentally the neural event of the c-fiber firing it could not be thought of as being an additional cause of my jumping out of the window. But it is questionable whether the mental event of my belief that the assassins intend to kill me can be a cause at all under the assumed scenario. Ex hypothesi the mental and neural realm are independent from one another. This means that the mental event can, in principle, occur without the occurrence of the neural event and vice versa. Assume that the neural event of the c-fiber firing had occurred without the co-occurrence of my belief that the assassins intend to kill me. The c-fiber firing would have caused my jumping out of the window. Since the mental event would not have co-occurred with the c-fiber firing it would not have been a cause of my jumping out of the window. So it seems that mental events are no causes at all.

Probably aware of the above reasoning the authors say the following: “For we have supposed (and this supposition is consistent with denying supervenience) that every instance of M occurs together with an instance of some physical property P. And we see no reason for thinking that it couldn’t also be the case that there is a causal law for each P to the effect that the instantiation of P is causally sufficient for the instantiation of P*. Thus we see no reason for thinking that the conjunction of the denial of Supervenience, Property Dualism, and the existence of mental to physical causation implies a violation of Closure”(Crisp and Warfield, 2001, p. 307). To apply this reasoning to our example, every instance of my belief that the assassins intend to kill me is accompanied by a physical property (it may even be the c-fiber firing). And there is a physical laws, which says that whenever the c-fiber firing occurs then the jumping out of the window occurs as an effect event.

It may be true that the authors’ thinking is consistent with the absence of supervenience, but it is also true that they are saying that “...every instance of M occurs together with an instance of some physical property P”. To me this seems as imposing regularity of co-occurrence of M (my belief that the assassins intend to kill me) and P (my c-fiber firing). Whatever this regularity of the co-occurrence of mental and physical events may be one could argue that it needs to be explained or explicated. Of course, the authors can always claim that these psychophysical regularities are brute and that as such they are built into the fabric of the world we live in. But still, mere regularities don’t support counterfactual claims, but causal relations do.

But Crisp and Warfield do not seem to accept my proposal about brute psychophysical regularities. Such a proposal would ultimately falsify physicalism. They

are rather interested in allowing mental to physical event causation without the breach of the closure principle. So what seems to have happened is that Crisp and Warfield may have misunderstood Kim's formulation of the closure principle. The authors are aware of this possibility (Crisp and Warfield, 2001, p. 307). They believe that there is evidence to the effect that Kim's interpretation of closure is different. Kim says: "One way of stating the principle of physical causal closure is this: If you pick any physical event and trace out its causal ancestry or posterity, that will never take you outside the physical domain. That is, no causal chain will ever cross the boundary between the physical and the nonphysical" (Kim, 1999, p. 40).

The statement "...no causal chain will ever cross the boundary between the physical and the nonphysical..." seems to be much stronger than the version of closure that Crisp and Warfield suggested. They said that when my belief is a cause of my jumping out of the window then there exists some neural event (the c-fiber firing), which is nomologically sufficient for my jumping. But Kim says that there is no causal chain from my nonphysical belief to my physical jumping. He says that the boundary between the nonphysical and the physical cannot be crossed in terms of event causation. On this interpretation of closure no non-physical event can ever be a cause, or, a mental event must be physical to be a cause.

Crisp and Warfield name Kim's interpretation of the closure principle "exclusion" (Crisp and Warfield, 2001, p. 307). Exclusion is such a strong requirement for event causation that under no circumstances can mental events have any role to play with regard to causation. This is the result of the assumption that only physical events can be causes. Mental events are excluded from the start, unless they are also physical.

On the interpretation of causal closure as exclusion mental causation (as non-physical) cannot exist regardless of whether we talk about Kim's premise P2 or P3. Recall that P2 assumed that mental properties do not supervene on neural properties, that mental properties are distinct in kind from neural properties, and that causation is closed under physical event causation. Closure interpreted as exclusion says that there cannot exist a causal chain from the non-physical mental to the physical. The consequence of this is that only my c-fiber firing can be the cause of my jumping out of the window. My belief has no say from the beginning. Recall that P3 of Kim's supervenience argument assumed that mental properties supervene on neural properties, that mental properties are distinct from neural properties, and that closure is a matter of physical events only. Closure interpreted as exclusion leads without further arguments to the same conclusion as in premise P2. My belief may now supervene on my c-fiber firing but it still remains the case that there cannot exist a causal chain from the nonphysical to the physical. Crisp and Warfield conclude that, because Kim interprets closure as exclusion, Kim's argument for premise P2 of his supervenience argument fails. The failure is due to the "overdetermination" (Crisp and Warfield, 2001, p. 308) implicit in the interpretation of closure as exclusion, which states that there cannot be non-physical mental causation.

This is how Crisp and Warfield deal with premise P2. Now we turn to P3. This premise consists of two parts. In part (1) Kim intends to show that mental to mental causation implies mental to physical causation. In part (2) he offers to explain that mental to physical causation is unintelligible on three metaphysical assumptions: + Supervenience + Property Dualism + Closure. Let us first see what Kim's understanding of part (1) of P3 consists of as well as Crisp and Warfield's replies.

Let us keep in mind that the assumptions of P3 are: + Supervenience + Property Dualism + Closure. For the sake of convenience I will write ‘SDC’ from now on. This abbreviation will refer to these three assumptions. The first step (1) of premise P3 is to show that mental to mental causation implies mental to physical causation. My belief that the assassins are at my door (mental event) causes another mental event, namely my belief that the assassins intend to kill me. This would be a case of mental to mental causation. We could abbreviate that as M causes M*, as Kim does (Kim 1999, p. 42). But premise P3 assumes that supervenience holds. So, since supervenience holds, it must be the case that M* (being a belief that the assassins intend to kill me) has a neural supervenience base. I follow Kim and say that this base property is P* (being a c-fiber firing). There is a question to be asked now. How did the mental property M* come to be instantiated on this occasion?

There seem to be two explanations. (i) M (being a belief that the assassins are at my door) occurred first and then it caused M*. (ii) P* (being a c-fiber firing) was instantiated on this occasion. P* is the supervenience base of M*. Whenever P* is instantiated M* occurs as well (Kim, 1999, p. 42). Kim perceives an explanatory tension here. He views M and P* as competing explanations of the occurrence of M*. He cannot accept two explanations of one and the same effect event (or property instance) M* that appeal to two different properties. In my chapter 1 I have talked about how Kim thinks events are causes due to them being instances of unique properties. The properties that are causal may be simple or complex, but it still remains the case that there can be only one property in virtue of which an event is ‘the’ cause of another event. An event cannot

have more than one “complete and independent explanation”. This is “the principle of causal / explanatory exclusion” (Kim, 1993, p. 239).

If both M and P* are responsible for the occurrence of the mental property M*, then the mentioned principle is violated. His intent is to reconcile the claim (i) that mental property M is a cause of M* with the claim (ii) that P* is the supervenience base of M*. He achieves this reconciliation by saying that the mental property M (being a belief that the assassins are at the door) causes the neural property P* (being a c-fiber firing) to be instantiated. M causes M* indirectly by causing its supervenience base P*. And the instantiation of the subvenient (P*) and supervenient (M*) is simultaneous, or they are co-instantiated. Whenever P* is instantiated M* is too.

Crisp and Warfield fail to see any explanatory tension with regard to M and P* as explanations of M*. They have three replies. (a) Recall that Kim’s principle of the causal / explanatory exclusion requires that an effect event can have only one independent and complete explanation. But Crisp and Warfield doubt whether either of the two properties, M or P*, are really independent and complete explanations of M*. If they are correct, if M and P* are in some sense incomplete and dependent on something else, then they are not competing explanations of M*. This means that in this case Kim’s principle of the causal / explanatory exclusion may not be violated.

Let us first investigate why M (being a belief that the assassins are at my door) is not a complete and independent explanation of M* (being a belief that the assassins intend to kill me). This has to do with our assumption of supervenience in premise P3 of Kim’s supervenience argument. My belief M is not independent because it is a supervenient property depending itself on its base neural property P (say the b-fiber firing of my

brain). M occurs only if P does. And it seems "...natural to suppose that P is also a cause of P*" (Crisp and Warfield, 2001, p. 310). And since the mental property M depends on P it cannot be an independent explanation of an instance of P*.

Now let us see why P* (being a c-fiber firing) may not be an independent explanation for the occurrence of an instance of M* (my belief that the assassins intend to kill me). For P* to occur P must occur first. P* depends on the occurrence of P. P* does not seem to be an independent explanation of M* for it depends on P. P seems to be a cause of P*. If P is a cause of P* and if P* therefore depends on P then it may be that only P can be regarded as an independent causal explanation of the occurrence of M*, because only the instance of P seems to be the real cause of an instance of M* (Recall that an explanation is causal for Kim only if the real cause of an effect is invoked. When we say that the property P is the real explanation of the property M* we imply that the instance of P –an event of course – is the real cause of an instance of M*.) So it seems that neither M nor P* qualify as "independent" explanations of M*'s occurrence.

Another reason (b) why Crisp and Warfield believe that there is no explanatory tension between M (being a belief that the assassins are at my door) and P* (being a c-fiber firing) as competing explanations for the occurrence of M* (being a belief that the assassins intend to kill me) has to do with the fact that M and P* are explanations of "different kinds": "The instantiation of M provides a diachronic, causal explanation for M*'s occurrence. But the instantiation of P* yields a synchronic, non-causal explanation for M*'s occurrence" (Crisp and Warfield, 2001, p. 310). This means that one and the same thing, namely M*, can be given different explanations depending on what questions are initially stated. We could ask why the instance of M* occurred. Such a question

would require a causal explanation. We would have to say that M*'s instance occurred because M's instance occurred first. In other words, the M-event is the cause of the M*-event. But we could also ask how the instance of M* occurred. This question would require a different kind of explanation than the previous one. We answer by saying that M*'s instance occurred at the same time as the instance of M*'s supervenience base occurred, namely P*. Or, in other words, M*'s instance occurred because P*'s instance occurred. The 'because' in the previous sentence is not causal but it is rather an expression of a non-causal, synchronic relationship of the simultaneity of the instantiations of the properties of P* and M*. It may be that we are explaining the occurrence of one and the same thing, namely M*, but we are not explaining one and the same relation. One of these relations is causal and the other one is non-causal. Invoking M when explaining M* means giving a causal explanation, and invoking P* constitutes giving an explanation in terms of supervenience.

To clarify our reasoning let us start with P*. The relationship between the properties M* and P* is that of supervenience. M* supervenes or it is dependent on P*. When P* occurs M* occurs automatically. These two properties are co-instantiated. They occur simultaneously. This also means that P* is not causally related to M*. But when we look at the relationship between the properties M and M* we have to say something else. M occurs before M* does. M and M* are not instantiated at the same time as was the case with P* and M*. This means that M is causally related to M*.

P* explains the occurrence of M* in terms of the fact that P* is the supervenience base of M*. M explains the occurrence of M* in terms of the fact that M is the cause of M*. Kim himself makes this distinction: "...the relation between base properties and

supervenient properties is not happily construed as causal. For one thing, the instantiations of the related properties are wholly simultaneous, whereas causes are standardly thought to precede their effects..." (Kim, 1999, p. 44). But he fails to see that there is no tension between the explanations of M^* by M and P^* . He does not seem to want to acknowledge that M and P^* are explanations different in kind, because of the fact that they explain different relations.

We are still concerned with Kim's step (1) in his premise 3 of his supervenience argument. Step (1) intends to show that mental to mental causation implies mental to physical causation. Crisp and Warfield have shown that there is no tension between M and P^* as explanations of M^* . But it seems that this is only a minor loss for Kim. If it turns out that Kim has another way of showing that mental to mental causation implies mental to physical causation, then he has won the war against Crisp and Warfield even though he may have lost the battle when it comes to the question of the alleged tension between M and P^* as explanations of M^* .

Indeed, there is a way that may show that mental (M) to mental (M^*) causation implies mental (M) to physical (P^*) causation. On the assumption of supervenience M^* depends on P^* (Recall that supervenience is a relation between properties: "Note, in particular, the fact that supervenience concerns mental and physical properties or kinds (or "characteristics" or "respects", in Davidson's words), and not mental and physical tokens or Davidsonian events as structureless particulars"; Kim, 1999, chapter 1, p. 6). An instance of M^* cannot occur without the occurrence of an instance of P^* . In order for an instance of M^* to occur an instance of P^* must occur. So to make an instance of M^* occur we must cause an instance of P^* . Kim says: "To cause a supervenient property to

be instantiated, you must cause its base property (or one of its base properties) to be instantiated" (Kim, 1999, p. 42). That is Kim's independent justification for his step (1) in premise P3, that mental to mental causation implies mental to physical causation.

The strategy of Crisp and Warfield is to suggest two different readings (Crisp and Warfield, 2001, p. 311) of Kim's principle that to cause a supervenient property to be instantiated one must cause its base property to be instantiated. Crisp and Warfield's third (c) attack against Kim's step (1) of his premise P3 has to do with these two formulations or readings of Kim's principle. I will use my examples in order to explicate these formulations.

Recall that Kim's reconciliation of the explanatory tension between M and P* as competing explanations of M* consisted of saying that M (being a belief that the assassins are at my door) causes M* (being a belief that the assassins intend to kill me) by causing P* (being my c-fiber firing). So the first, (1.), formulation of Kim's principle goes approximately like this. In order for my belief that the assassins are at my door (M) to cause a supervenient property, namely my belief that the assassins intend to kill me (M*), the belief M itself must cause one of M*'s base properties (namely P*) to be instantiated. Crisp and Warfield believe that this formulation is Kim's preferred interpretation of the principle.

The second formulation of the principle is slightly different. In order for my belief that the assassins are at my door, M, to cause a supervenient property M* (my belief that the assassins intend to kill me) to be instantiated, the instantiation of one of M*'s base properties (namely P* - the c-fiber firing) must be caused. It seems that this formulation does not involve M as an immediate cause of P*, or it does not state that M

itself must cause P^* . It is rather said that the ‘instantiation of one of M^* ’s base properties (P^*) must be caused’. This formulation does not say that M must cause P^* ; it says that P^* ’s instantiation must be caused. Something does cause P^* to be instantiated but it is not M directly.

Crisp and Warfield believe that on the second formulation it can be claimed that not one property (namely M) is responsible for both M^* and P^* . I think that the reason for this claim has to do with what I have said in my previous paragraph. And this is the fact that M causes the supervenient property M^* in that P^* ’s instantiation ‘is caused’. This formulation does not say that M causes P^* ; P^* ‘is caused’ but most likely not by M . Crisp and Warfield believe that this scenario is “...clearly consistent with there being mental to mental causation without mental to physical causation” (Crisp and Warfield, 2001, p. 311). This means that my belief that the assassins are at my door (M) can be a cause of my belief that the assassins intend to kill me (M^*) without M being a cause of the instance of the c-fiber firing (P^*). The supervenient property M causes another supervenient property, namely M^* , to be instantiated. And the subvenient neural property P causes another subvenient property, namely P^* , to be instantiated.

Crisp and Warfield suggest that M (being my belief that the assassins are at my door) could be viewed as a relational or functional property occurring when I see assassins (Crisp and Warfield, 2001, p. 311). M then causes M^* , namely my belief that the assassins intend to kill me. But M is a second-order relational property. It is realized by a neural property, namely the b-fiber firing. The subvenient property of being a b-fiber firing (P) causes the c-fiber firing (P^*). I am inclined to say that there are two causal

processes running in parallel. One is, call it, the subvenient causation, from P to P*; the other is, call it, supervenient causation, from M to M*.

Crisp and Warfield's third (c) attack on step (1) of P3 of Kim's supervenience argument has to do with the previous discussion. The two authors believe that Kim's claim "To cause a supervenient property to be instantiated, you must cause its base property (or one of its base properties) to be instantiated" (Kim, 1999, p.42) does not "specifically motivate" either one of the two formulations of the principle that to cause a supervenient property one must cause its subvenient property to be instantiated (Crisp and Warfield, 2001, p. 311). Since two formulations of the principle are possible, and neither one of the two formulations is specifically motivated, Crisp and Warfield conclude that Kim has not given a "compelling independent justification he promised for thinking that M causes M* by causing P*". The first step of Kim's argument for P3 is "... a shaky step" (Crisp and Warfield, 2001, p. 311).

Now we will turn our attention to Kim's second step (2) in premise P3 of his supervenience argument. In this step Kim intends to prove that mental to physical causation is unintelligible given SDC (Supervenience, Property Dualism, Closure). For the sake of argument Crisp and Warfield grant Kim the first formulation of the principle that to cause a supervenient property one must cause its supervenience base to be instantiated (Crisp and Warfield, 2001, p. 312), even though they have shown that Kim's first step (1) is shaky. So let us assume that the mental M (being a belief that the assassins are at my door) causes P* (being a c-fiber firing). Given that supervenience holds (SDC of P3), it seems to follow that the mental property M also has a supervenience base, namely P. Kim argues that P (being a b-fiber firing) is sufficient for M (being a belief that

the assassins are at my door), and M for P^* (being a c-fiber firing). From this it follows then that P is sufficient for P^* . But if both, M and P, are sufficient for P^* , then P^* has too many causes (Kim, 1999, p.43). When one event like P^* has too many causes, it is, as Kim says, overdetermined. This is the first (a) problem, which Kim detects in his second (2) step (mental to physical causation is unintelligible given SDC) of his premise P3.

But it is not only in this instance that causal overdetermination occurs. Kim believes that every case of mental causation (be it mental to mental or mental to physical causation) is a case of overdetermination if mental and physical properties are distinct. Say that M causes M^* (a case of mental to mental causation). Recall that step (1) of P3 claimed that to cause a supervenient property such as M^* to be instantiated one must cause its supervenience base P^* to be instantiated. This means that M must cause P^* in order to cause M^* . But M itself has a supervenience base P, which is also causally sufficient for P^* . P^* , as the supervenience base of M^* , is overdetermined by M and P. These consequences follow as well in the case of mental to physical causation. This is when M is said to cause P^* . M has a supervenience base P. Both, M and P, are sufficient for P^* , and the latter is overdetemined.

But besides the problem of overdetermination (a) there is yet another one (b), namely that the causal closure of the physical domain is breached. Kim argues that in a world, which is just like ours in all respects except that P does not occur the mental property M ends up being the only cause of the neural property P^* (Kim, 1999, p. 45). We assumed that M and P are both causes of P^* in our world. We then assumed that P (the supervenience base of M) does not occur in the imagined (but almost identical with ours) counterfactual world. It then follows that the instance of the mental property M

breaches the principle that physical events have only other physical events in their ancestry and posterity therewith falsifying the closure principle.

Kim concludes that (a) because of the problem of causal overdetermination and (b) the problem of the breach of closure, mental to physical causation is unintelligible if it invokes distinct properties (Kim, 1999, pp. 44-46). Kim believes that the only plausible explanation is this. M is dependent on P. M* is dependent on P*. Only an instance of the physical property P can be the real cause of an instance of the physical property P*. The relations between the instances of M (being a belief that the assassins are at my door) and M* (being a belief that the assassins intend to kill me) and M and P* (being a c-fiber firing) are not really causal. They are only by-products of the real causal relation holding between P and P*.

Crisp and Warfield have two points in reply to Kim's steps (1) and (2) of his P3 premise. (i) They believe that those who adhere to property dualism (that mental properties are distinct in kind from neural properties) as well as the closure principle (that physical events are caused by physical events) must be committed to accepting causal overdetermination, as was the case when P* was caused by both M and P (Crisp and Warfield, 2001, p. 313). I think that the reason why they must bite the bullet and accept causal overdetermination of P* by M and P has to do with how they understand the principle of the causal closure of the physical domain. They do not believe that a physical event such as P* has *only* physical events as causes. This was what Kim believed. Their formulation of the closure principle is weaker than exclusion (Kim's version of the principle). It is weak enough to allow mental events (such as an instance of M) to be causes as well. Their formulation of closure is this: "Every caused physical event has a

physical cause" (Crisp and Warfield, 2001, p. 305). It seems that this formulation does not explicitly state that every caused physical event must have only physical causes. So M too can be a cause of P*. It only seems to say that as long as there is some physical event (presumably P), P* will be caused to be instantiated by it. And since supervenience of M on P holds, M too is a cause of P*. So mental to physical causation seems to be saved on this version of closure (Crisp and Warfield, 2001, pp. 313-314).

Crisp and Warfield believe that provided that supervenience of M on P is the case closure is not violated as long as there exists a physical event such as P (when M is also instantiated), which causes P*. Thus the event of mental kind M is also a cause of P*. But if we assume that supervenience fails conjoined with the absence of some physical property when the instance of M occurs, or, if we imagine a world in which mental and physical properties such as M and P are distinct in kind and in which physical events cause physical events but in which P fails to occur at the time when M also occurs, then any case of mental to physical causation (e.g. M to P*) constitutes an instance of the violation of the closure principle. It seems that only the assumption of supervenience can save mental to physical causation. Crisp and Warfield agree with Kim that in worlds in which supervenience of the mental on the physical fails (-S+D+C), mental to physical event causation in the absence of an instance of a physical property violates the closure principle therewith falsifying the claim that every caused physical event has a physical cause (Crisp and Warfield, 2001, p. 314).

Kim, Crisp and Warfield believe that our actual world is a world in which the supervenience relation between mental and physical properties holds. From this it seems to follow that it is not possible that there could be a world just like ours but in which M

occurs and causes P^* without P also occurring and causing P^* . If the mentioned possible world is just like ours, then the supervenience of M on P is a matter of law. It cannot happen that M occurs without P , and that M alone causes P^* . It is not nomologically possible that the supervenient property M (being the belief that the assassins are at my door) occurs without the neural property P (being a b-fiber firing). The reason why this is not possible has to do with the fact that if the imagined world is exactly like ours this world must have the same laws of nature. But we have supposed that M occurs while P does not. This supposition violates our laws of nature, which mention that M depends on P and in which both cause P^* . From the point of view of causal laws or laws of nature the imagined world cannot exist.

Even if all mental properties except M were to depend on their neural supervenience bases all the time, and even if M usually supervenes on its P but just not that one time when P fails while M still occurs and causes P^* it would still remain the case that the counterfactual world is “nomologically impossible” (Crisp and Warfield, 2001, p. 314). It is not possible for causal laws to tolerate such exceptions.

Crisp and Warfield conclude that Kim does not offer a convincing argument (in his step (2) of P3) that overdetermination of P^* by M and P fails and that overdetermination is not a worthwhile option for non-reductivist accounts of mentality. They have substantially damaged Kim’s premises P2 and P3 by showing, for instance, that closure interpreted as exclusion seems a non-starter since it implies that an instance of a mental property (whether supervenience holds or fails) can never be a cause. In other words, closure interpreted as exclusion begs the question against the property dualist. In addition, Kim’s principle that to cause a supervenient property you must cause its base to

be instantiated does not seem to motivate either the stronger reading (that M directly causes the neural base property P*) or the weaker one (that the instantiation of P* must be caused, where it is not explicitly stated that it is M that causes P* directly). Crisp and Warfield do not deal with Kim's conclusion of his supervenience argument (Therefore, if Closure and Property Dualism, then mental causation is unintelligible.) because they believe that by damaging premises P2 and P3 they have created enough conceptual slack for non-reductivist accounts of mental causation (Crisp and Wardfield, 2001, p. 315).

The reason why I have chosen to talk about Crisp and Warfield's attacks on two premises (P2 and P3) of Kim's supervenience argument has to do with the fact that Kim's supervenience argument seems to be Kim's most important weapon against non-reductivist physicalists like Davidson. And if it can be shown – as Crisp and Warfield seem to have done – that Kim's weapon is faulty, or, that, figuratively speaking, it cannot discharge any bullets, it then deserves not to be called a properly functioning tool in the attack against nonreductivist accounts of the mind. From this it should then follow that Kim's attacks against various non-reductivist approaches in his book have no bite or substance at all, because, if anything, he has begged the question against nonreductivists and property dualists. It may be that Kim's premise P2 (- S + D + C) can survive Crisp and Warfield's attack against it. But since Kim's premise P3 of his argument begs the question against non-reductivist accounts such as Davidson's, who holds that + S +D + C is the case in our world, it can be claimed that Davidson's view is not shattered.

But it seems that it is not only Kim's supervenience argument that is problematic. Another form of attack points to flaws in Kim's functionalism. The reason why I intend to investigate these flaws has to do with my belief that if, as we have already done, we

disarm Kim's main weapon and if we conjointly show that there is something wrong with his functionalism, then it seems that we have no good reasons to accept Kim's view as superior to Davidson's in the mind-body debate. If the criticisms elaborated so far and those to follow have any strength at all, then, falling short of saying "*Mind in a Physical World*" ought to be committed to the flames, Kim at least has to re-think his position and present a new view, which will lack the flaws and maybe even inconsistencies of the mentioned monograph. So in tune with the mentioned motives and reasons I now turn to an explanation of an attack on Kim's functional identity theory. My intent is to show that this theory is incoherent at best; that Kim's view on the mind-body relationship is not really a functionalist theory, but that it rather is a form of type-physicalism. Let us start with the accusation of incoherence.

When we say that A is identical to B we realize that we can switch the order of the two terms A and B while their relation remains the same. We can, instead, say that B is identical to A. This is so because the identity relation is symmetrical. Kim has claimed that mental properties such as being a type of pain are identical with being a type of a first-order neural realizer (say being an a-fiber firing). But he also said that mental properties supervene on neural properties in that these mental properties are realized by their first-order neural properties.

Prior to launching an attack about the claim that his functionalism is incoherent, let us first recall how Kim argues for property identity, property supervenience, and property realization. It is good to give a recap of these ideas in order for us to better understand the following attack. Say I feel a pain in my foot. This pain occurs at this moment; it is a particular event, entity, or token. But this particular pain (event as an

object, entity, or token) is an instance of a type of pain. This type (or property) of pain is said to be identical to a neural property (say being an a-fiber firing). Let us have M stand for the type of pain (a phenomenal mental property), and let us have P stand for a neural property (being an a-fiber firing). Kim says: "...we construe M as a second-order property defined by its causal role – that is, by a causal specification H describing its (typical) causes and effects. So M is now the property of having a property with such-and-such causal potentials, and it turns out that property P is exactly the property that fits the causal specification. And this grounds the identification of M with P. M is the property of having some property that meets specification H, and P is the property that meets H. So M is the property of having P. But in general the property of having property Q = property Q. It follows then that M is P" (Kim, 1999, chapter 4, pp. 98-99).

Let us now recall what Kim understands under supervenience: "Mental properties supervene on physical properties, in that necessarily, for any mental property M, if anything has M at time t, there exists a physical base (or subvenient) property P such that it has P at t, and necessarily anything that has P at a time has M at that time" (Kim, 1999, chapter 1, p. 9). So when we are talking about the type of my foot pain we have to say something like this. If I experience pain in my foot it must be that I instantiate a neural property such that anyone identical to me living in a world that has causal laws identical to our world, he must experience pain or he must instantiate the mental property of being that type of foot pain.

Now we are reminded of Kim's understanding of property identity and property supervenience. Let us recall what the realization relation amounts to. We can again use the foot pain example. The type of pain I experience is best explained as a causal role.

We can say that whenever my foot is hit by a brick I immediately scream, groan, make grimaces, and maybe even swear. The locution ‘whenever’ points to the fact that types or properties (and not events or tokens) are related to one another. The type of stimuli (namely the impact of the brick with my foot) is followed by a type of foot pain, which in turn is followed by a type of behavior (groaning, crying, etc.). But to be more precise, the foot pain is a mental type. It is a property that is second order in relation to the type of stimuli (brick impact), the neural property instantiated at the same time, and the type of behavior. The latter three properties are physical. The foot pain as a mental type is also a functional property. This means that it is defined as a causal role over first order realizer properties (brick impact, a-fiber firing of the brain, groaning and crying). For me to have or instantiate the foot pain property is for me to have a neural property (say a-fiber firing) that meets the causal specification above. For somebody just like me living in a world whose laws are exactly like those of our world, having the same type of foot pain consists in having a-fibers firing. In Kim’s words: “...having M is having a property with causal specification D, and in systems like s, P is the property (or one of the properties) meeting specification D. For systems like s, then, having M consists in having P” (Kim, 1999, chapter 1, p. 24). The neural property of the a-fibers firing happens to be the realizer property of being a foot pain. For me to have the foot pain is for me to instantiate the property of being an a-fiber firing. Or, to have the foot pain is for me to have (or instantiate) the property of being an a-fiber firing.

Without the realizer property I could not have the pain provided that we hold the laws of our world constant. Or, in Kim’s words: “...P is sufficient for M...in fact, given the nomological constancy just noted of the realization relation, it follows that P is

nomologically sufficient for M" (Kim, 1999, chapter 1, p. 23). And realization understood in Kim's way actually explains the notion of the supervenience of mental on neural properties. An instance of M cannot occur without an instance of P. But what consequently happens is that realization understood as property identity ($M = P$) leads to the fact that P (it could be P_1, P_2, \dots) is sufficient for M provided that the laws are held constant. The constancy of laws allows that any P can be sufficient for the mental property M provided that each P has the causal powers to be appropriately hooked up with same inputs and same outputs. From this it follows that mental properties do not have causal powers in their own right.

Now we are prepared to critique Kim's functional identity theory as being incoherent, that it is really not a functionalist view at all but rather a form of type-physicalism. As already said Kim is committed to the following relations: identity, supervenience, and realization. As said above, identity is symmetrical. But supervenience is not. We can say that the mental property M supervenes on the neural property P, but we cannot say that the neural property P supervenes on the mental property M. It is believed (Kim included) that mental properties depend on neural properties and not the other way around. The dependence or supervenience relation, contrary to the identity relation, is asymmetrical. We cannot switch the terms M and P without falsifying the relation.

From identity and supervenience it follows that the mental property of being a belief M is identical with P (the neural property of being a c-fiber firing) and that M depends on P. Since M is identical to P and since M depends on P, M also depends or supervenes on itself: "...if M is identical to P, it cannot asymmetrically depend on P (or

itself” (Marras, 2000, p. 140). In other words, the simultaneous holding or accepting of property identity and property supervenience is incoherent. One cannot hold one and the other simultaneously. One can hold either one or the other but not both when it comes to the relationship between the mental and the physical.

Now let us have a look at the realization relation. We can say that the neural property P realizes the mental property M, but not the other way around. M does not realize P. Marras says: “Realization is not identity, if only because realization is an asymmetrical and, possibly, one-to-many relation, which identity obviously is not” (Marras, 2000, pp. 156-157). This means that functionalists usually believe that, in principle, it is possible to say that one and the same (type) mental property M is realized by different types of physical realizer properties. Hugo the Martian, Data the android, and Peter the human can all have the same belief M even though Peter instantiates the neural property of being a c-fiber firing, Hugo instantiates the neural property of being a z-fiber firing, and Data instantiates the network property of being an o-chip firing.

The hope of functionalists is to find an explanation for why and how different first-order properties (being a c-fiber firing, being a z-fiber firing, being an o-chip firing) realize the same mental property M. But one thing is sure, the realization relation cannot be explained, as Kim does, by identifying the property of being a belief M with all the different realizers. As already said, the reason is that the realization relation is asymmetrical and not symmetrical.

At the outset of his book Kim seems to be expressing his awareness about the fact that the realization relation is asymmetric. He says: “The key between mental properties (kinds, states, etc.) and physical properties was “realization” (or sometimes

“implementation”, “execution”, etc.): mental properties are “realized” or “implemented” by (or in) physical properties, though neither identical nor reducible to them” (Kim, 1999, chapter 1, p. 7). Kim does not say that the mental property realizes the neural property but rather the other way around. Although the quote is an explanation of how non-reductivist physicalists understand the realization relation it seems nevertheless the case that Kim implicitly understands that the realization relation is not symmetric but asymmetric. Or in other words, if a relation is truly asymmetric then it should be so for Kim or for a non-reductivist physicalist. Another thing we get from the quote is the fact that the realization relation does not imply the relation of identity. It is also said that the mental property is not reducible to the neural property. In short, neural properties may be said to realize (the realization relation needs explanation, but, of course, not in terms of identity.) mental properties, but mental properties are distinct from neural properties. And it is my belief that this formulation accepted by non-reductivist physicalists is correct since they do not conflate two incompatible relations, namely the symmetric identity relation and the asymmetric realization relation.

Let us look at a quote, which, I believe, is the point at which Kim commits the mistake of confusing different kinds of relations with one another: “This means that physical realizationism would give us an explanation of the supervenience thesis: the mental supervenes on the physical because mental properties are second-order functional properties with physical realizers (and no nonphysical realizers). And we have an explanation of mental-physical correlations. Why is it that whenever P is realized in a system s, it instantiates mental property M? The answer is that by definition, having M is having a property with causal specification D, and in systems like s, P is the property (or

one of the properties) meeting specification D. For systems like s, then, having M consists in having P. It isn't that when certain systems instantiate P, mental property M magically emerges or supervenes (in the dictionary sense of "supervene"). It is rather that having M for these systems, simply is having P. We might even say, using a familiar if shopworn reductive idiom, that having M, for these systems, is "nothing over and above" having P. Note that all of these explanations are nomological explanations – they depend on the fact that a certain set of laws prevail in our world, for these laws ultimately determine what physical properties are realizers of a given mental property" (Kim, 1999, chapter 1, p. 24).

Let us look at Kim's motives and reasoning starting from the top of the quotation. He seems driven by the need to explain the supervenience relation. He acknowledges that the supervenience relation is asymmetric: "Moreover the supervenience claim is physicalistically appealing: an asymmetric dependence of the mental on the physical is clearly implied..." (Kim, 1999, chapter 1, p. 6). Then he goes to say that the supervenience relation is explained by the realization relation. Mental properties are realized by physical properties: "...whenever P is realized in a system s, it instantiates mental property M...". This particular part of the above quotation (Kim, 1999, chapter 1, p. 24) as well as the previous quotation about how non-reductivists understand the realization relation (Kim, 1999, chapter 1, p. 7) taken conjointly lead us to believe that Kim understands the asymmetric character of the realization relation.

But when we continue to read the quotation from p. 24 it seems obvious that Kim conflates the asymmetric realization relation with the symmetric identity relation. He claims: "...having M is having a property with causal specification D, and in systems like

s, P is the property (or one of the properties) meeting specification D. For systems like s, then, having M consists in having P..." and "...having M for these systems, simply is having P" (Kim, 1999, chapter 1. p. 24). Obviously driven by the need to give an explanation of the realization relation Kim commits the mistake of fusing incompatible relations into one. Or, yet better, he thinks that he can explain the realization relation with the identity relation. But the realization relation is asymmetric; it cannot be explained by a symmetric relation. Kim's confusion amounts to the denial of both the realization relation as well as the supervenience relation holding between the mental and the neural. Since having M is having P, or since $M = P$, it follows that no asymmetric relations hold between M and P. If $M = P$, then it cannot be said that M depends on P. M does not depend on P because M is P. And M is not realized by P because M is P. The identity relation out-powers the supervenience as well as the realization relation. By conflating asymmetric with symmetric relations Kim actually contradicts everything he says or said about both the supervenience as well as the realization relation.

So it seems that Kim is mistaken and confused when he tries to explain the realization relation via the identity relation. He says that the physical properties (say) P, Q, and Z realize the mental property M just because M is identical with each of its realizers. This means, in effect, that if M is realized by P in that M is identical with P, then, having in mind the fact that the identity relation is symmetrical, P is identical with Q and Z as well. But the properties P, Q, and Z are not identical. One (P) is a property had by a human brain (Peter), another is (Q) a property instantiated by something (Martian's 'brain') made up of green as opposed to grey matter, and yet another one (Z) is had by a system of micro-chips (Data's 'brain'). A consequence of the fact that Kim

subscribes to the realization as well as the identity relation (see the discussion on the conflation of incompatible relations, previous paragraph) leads Kim to say that the realizer properties are identical with one another when they are not identical with one another. Identifying the mental property with its realizers leads Kim to this: $M = P = Q = Z$. But P is a property that differs from Q and Z ($P \neq Q \neq Z$).

If Peter, Hugo, and Data are alleged to have the same type of belief (say) that the assassins intend to kill me, but if their respective realizers differ, and if the mentioned type of belief is identical with its many types of realizers, then from this it follows that this is not the same type of belief anymore. M in Peter, M in Hugo, and M in Data have nothing in common, because the realizer properties are distinct from one another and because the mental property M is identified with each realizer property (Marras, 2000, pp. 155-156). The identification of M with P , Q , and Z violates a basic functionalist assumption. The assumption is that M (being a belief that the assassins intend to kill me) must remain the same mental property for all systems (Peter, Hugo, and Data) regardless of what realizes that mental property. On this account M is a separate property, supervenient on and realized by its many realizer properties. But Kim's identification leads to the fact that (a) M is not the same property. M is always something else because it is always identical with diverse realizers. And (b) Kim's identification move takes away the functionalist requirement that M be a separate but dependent property.

The mental property is type-identified with its many realizers. And this is no functionalism anymore. Marras says: "Kim's 'functional reductionism' strikes me as incongruous with the very spirit of functionalism" (Marras, 2000, p.154). The identification of M with its many realizers "...strikes me as type-physicalism with a

vengeance" (Marras, 2000, p. 155). But the biggest downside of the identification is that M seems lost as an independent property having its own causal powers. This means ...that we have lost M as a single, unified property..." (Marras, 2000, p.154). Marras perceives this to be an "eliminativist threat" (same page) in erasing M as a property, and therewith wiping away mental causation as non-existent. These worries are only increased by Kim's own words that M is "... a property we will have to learn to live without" (Kim, 1999, p. 106).

It would be an understatement that Kim is ambivalent on the subjects of the relation of mental and neural properties and mental event causation. Independently of what Marras says the incoherence in Kim's view is obvious. A look at some statements and ideas of Kim brings this to light. He claims that we have to learn to live without M (which seems to say that M should be eliminated as a property) (1). But he also intends to (2) save (The stress here is on save.) mental causation. Plus he mentions several times in *Mind in a Physical World* that (3) something is a property just in case it has its own independent causal powers. Irrespective of the incoherence about the simultaneous subscribing to identity, supervenience, and realization, it must be said that (1), (2), and (3) do not fit together.

By identifying M with P, Q, Z, and so on, there is nothing to be saved at all, because M does not exist as a separate property had by different physical systems. Being identical with its many realizers M does not have (3) its own causal powers. The claim that M 'inherits' the causal powers of each of its realizers only amounts to saying that there is no independent M property having its own causal powers.

From all the above it seems to me that Kim does not save mental causation. He eliminates it. The reason why I think that he eliminates mental causation altogether has to do with the fact that mental properties by being identified with their alleged realizers are no independent properties having causal powers in their own right. And maybe Marras is right that Kim's view is not a functionalist one but rather a version of type-physicalism. In order for Kim to be a functionalist he would have to drop the identity thesis about M and P ($M = P$). A functionalist must claim that $M \neq P$; that M supervenes on or is determined by P; and that various physical properties (such as P, Q, Z, and so on) realize the same (type) M (say, the belief that the assassins intend to kill me). In addition to that, stories about supervenience and realization would have to be fleshed out without, of course, an appeal to identity of mental and physical properties.

It has been shown that major aspects of Kim's theory exhibit weaknesses, flaws, and, sometimes, outright confusions. Mental causation does not really seem to be threatened by his supervenience argument. Mental events can be causes provided that there are neural events that underlie them, or provided that the mental properties they are instances of supervene on neural properties. Mental events may depend on neural events but they remain causes nevertheless. It also seems that one cannot coherently subscribe to mutually excluding relations at the same time. A true functionalist can claim that the mental supervenes on the neural, and that it is realized by the neural, but he cannot subscribe to the identity thesis about the mental and the neural at the same time.

BIBLIOGRAPHY

- Crisp, T.M., and T.A. Warfield. 2001. Kim's master argument. *Nous*, 35 (2):304-316.
- Davidson, Donald. 1980. Essay 1. Actions, Reasons, and Causes (1963). In *Essays on Actions and Events*, 3-19. New York: Oxford University Press.
- _____. Essay 3. Agency (1971). In *Essays on Actions and Events*, 43-61. New York: Oxford University Press.
- _____. Essay 6. The Logical form of Action Sentences (1967) Criticism, Comment, and Defence. In *Essays on Actions and Events*, 105-147. New York: Oxford University Press.
- _____. Essay 7. Causal Relations (1967). In *Essays on Actions and Events*, 149-161. New York: Oxford University Press.
- _____. Essay 8. The Individuation of Events (1969). In *Essays on Actions and Events*, 163-180. New York: Oxford University Press.
- _____. Essay 11. Mental Events (1970) Appendix: Emeroses by Other Names (1966). In *Essays on Actions and Events*, 207-227. New York: Oxford University Press.

- _____. Essay 12. Psychology as Philosophy (1974) Comments and Replies. In *Essays on Actions and Events*, 229-244. New York: Oxford University Press.
- _____. Essay 13. The Material Mind (1973). In *Essays on Actions and Events*, 245-259. New York: Oxford University Press.
- Kim, Jaegwon. 1999. The Mind-Body Problem: Where We Now Are. In *Mind in a Physical World*, 1-28. 2nd ed. Cambridge, Massachusetts: The MIT Press).
- _____. The Many Problems of Mental Causation. In *Mind in a Physical World*, 29-56. 2nd ed. Cambridge, Massachusetts: The MIT Press).
- _____. Mental Causation: The Backlash and Free Lunches. In *Mind in a Physical World*, 57-88. 2nd ed. Cambridge, Massachusetts: The MIT Press).
- _____. Reduction and Reductionism: A New Look. In *Mind in a Physical World*, 89-120. 2nd ed. Cambridge, Massachusetts: The MIT Press).
- Kim, Jaegwon. 1973. Causation, Nomic Subsumption, and the Concept of Event. *The Journal of Philosophy*, LXX (8): 217-236.
- Marras, A. 2000. Critical Notice. *Canadian Journal of Philosophy*, 30 (1): 137-160.