Personal Identity: 
Functionalism and the M-Relation

MATT NUGENT is a senior at the University of Pennsylvania where he is a philosophy major with a political science minor. His primary areas of philosophical interest are Hegelian and Marxist thought, but other interests include epistemology and philosophy of religion. When not at school, he resides in Maine and Boston.
My aim in this paper is to reconcile what I consider the two most plausible theories for personal identity, by showing that Sydney Shoemaker’s functionalist argument provides a relation among mental states that fits perfectly with John Perry’s $M$-relation. By showing that Shoemaker’s brain state transfer device would not be person-preserving, I show that it is also capable of placing restrictions that limit the functional states in Shoemaker’s theory to a single body, thus reconciling completely his theory with Perry’s.

The problem of personal identity arises from efforts to define the criteria necessary to establish a common identity between two spatial-temporally distant objects. Perry’s theory of personal identity as found in *Identity, Personal Identity, and the Self* is an adaptation of H.P. Grice’s view, which is itself a reformulation of and improvement upon the traditional Lockean memory theory. Grice’s project is to discover the relation between sets of simultaneously occurring experiences of a single person, or total temporal states (t.t.s.), and in doing so account for the maintenance of a person’s identity over time. Perry summarizes Grice’s theory as follows:

There is a sequence of t.t.s.’s (not necessarily in the order they occur in time and not excluding repetitions), the first of which is $A$ and the last of which is $B$, such that each t.t.s. in the sequence either (i) contains, or would contain given certain conditions, a memory of an experience contained in the next or (ii) contains an experience of which the next contains a memory, or would contain a memory given certain conditions.

(Perry 85)

Although Grice’s formulation does account for the transitivity of identity, a failure of Locke’s theory generally thought to be damning, there exist three well known charges of circularity, all potentially fatal to the former’s theory. The first charge suggests that Grice’s reliance on veridical memory presupposes personhood. In the event that a t.t.s. contains an apparent memory of a past experience belonging to a previous t.t.s., there is nothing to ensure the accuracy or legitimacy of the memory; it could, for instance, be the work of a skilled hypnotist. However, if one asserts that the memory is veridical, then one is assuming that the t.t.s. containing the memory and the t.t.s. containing the remembered experience belong to the same person. The other two charges of circularity revolve around the claim that “[each t.t.s. in the sequence either contains, or] would contain given certain conditions.”

First, suppose person $X$ witnessed a mugging in West Philadelphia at two P.M., which we will call t.t.s. $A$. Later that night, person $X$ fell into a dreamless sleep. It is safe to say that while $X$ is sleeping but not dreaming, $X$’s sleeping t.t.s. contains little in the way of experiences, have no memories of past experiences, and have no experiences to be remembered in future t.t.s.’s. However, that t.t.s. can still belong to a longer chain because, according to Grice, they would contain memories if the conditions were different. For instance, if $X$ was woken up and asked about the mugging, he would be able to describe it in vivid detail. But to make this claim is to say that given certain circumstances $X$ would have had different t.t.s.’s while asleep than he would have he been awake. In doing so, Grice is evoking the notion of personhood by ascribing a common identity to both the sleeping person (to whom a relatively empty t.t.s belongs) and the person who was rudely awakened (to whom the t.t.s. containing a memory of the mugging would belong). Secondly, one can contend that “[t].t.s. $A$ would contain, if the same person who has $A$ had been awakened and asked or if the same person who has $A$ had not just taken a powerful drug or … a memory of an experience contained in $B$” (Perry 92). One would once again be invoking the concept of personal identity in their analysis and would therefore be proposing a circular argument.

In order to remedy these charges of circularity, Perry attempts to better define what is meant by memory. He outlines three conditions: one to be met at the time of the remembering, one to be met at the time of the experience to be remembered, and one that denotes what the link between the two t.t.s.’s must be (Perry 92-3). The conditions are as follows:

1. If $A$ is identical with $B$, and $B$ is identical with $C$, then $A$ must be identical with $C$.
2. Here, ’$A$’ refers to the t.t.s. of a sleeping or unconscious person, and ’$B$’ refers to the t.t.s. of an awake and conscious person.
My aim in this paper is to reconcile what I consider the two most plausible theories for personal identity, by showing that Sydney Shoemaker’s functionalist argument provides a relation among mental states that fits perfectly with John Perry’s M-relation. By showing that Shoemaker’s brain state transfer device would not be person-preserving, I show that it is also capable of placing restrictions that limit the functional states in Shoemaker’s theory to a single body, thus reconciling completely his theory with Perry’s.

The problem of personal identity arises from efforts to define the criteria necessary to establish a common identity between two spatial-temporally distant objects. Perry’s theory of personal identity as found in *Identity, Personal Identity, and the Self* is an adaptation of H.P. Grice’s view, which is itself a reformulation of and improvement upon the traditional Lockean memory theory. Grice’s project is to discover the relation between sets of simultaneously occurring experiences of a single person, or total temporal states (t.t.s.), and in doing so account for the maintenance of a person’s identity over time. Perry summarizes Grice’s theory as follows:

There is a sequence of t.t.s.’s (not necessarily in the order they occur in time and not excluding repetitions), the first of which is A and the last of which is B, such that each t.t.s. in the sequence either (i) contains, or would contain given certain conditions, a memory of an experience contained in the next or (ii) contains an experience of which the next contains a memory, or would contain a memory given certain conditions.

(Perry 85)

Although Grice’s formulation does account for the transitivity of identity, a failure of Locke’s theory generally thought to be damning, there exist three well known charges of circularity, all potentially fatal to the former’s theory. The first charge suggests that Grice’s reliance on veridical memory presupposes personhood. In the event that a t.t.s. contains an apparent memory of a past experience belonging to a previous t.t.s, there is nothing to ensure the accuracy or legitimacy of the memory; it could, for instance, be the work of a skilled hypnotist. However, if one asserts that the memory is veridical, then one is assuming that the t.t.s. containing the memory and the t.t.s. containing the remembered experience belong to the same person. The other two charges of circularity revolve around the claim that “[each t.t.s. in the sequence either contains, or] would contain given certain conditions.”

First, suppose person X witnessed a mugging in West Philadelphia at two P.M., which we will call t.t.s. A. Later that night, person X fell into a dreamless sleep. It is safe to say that while X is sleeping but not dreaming, X’s sleeping t.t.s. contains little in the way of experiences, have no memories of past experiences, and have no experiences to be remembered in future t.t.s.’s. However, that t.t.s. can still belong to a longer chain because, according to Grice, they would contain memories if the conditions were different. For instance, if X was woken up and asked about the mugging, he would be able to describe it in vivid detail. But to make this claim is to say that given certain circumstances X would have had different t.t.s.’s while asleep than he would had he been awake. In doing so, Grice is evoking the notion of personhood by ascribing a common identity to both the sleeping person (to whom a relatively empty t.t.s belongs) and the person who was rudely awakened (to whom the t.t.s. containing a memory of the mugging would belong). Secondly, one can contend that “[t].t.s. A would contain, if the same person who has A had been awakened and asked or if the same person who has A had not just taken a powerful drug or … a memory of an experience contained in B” (Perry 92). One would once again be invoking the concept of personal identity in their analysis and would therefore be proposing a circular argument.

In order to remedy these charges of circularity, Perry attempts to better define what is meant by memory. He outlines three conditions: one to be met at the time of the remembering, one to be met at the time of the experience to be remembered, and one that denotes what the link between the two t.t.s.’s must be (Perry 92-3). The conditions are as follows:

1: If A is identical with B, and B is identical with C, then A must be identical with C.

2. Here, ‘A’ refers to the t.t.s. of a sleeping or unconscious person, and ‘B’ refers to the t.t.s. of an awake and conscious person.
A’ remembers e if and only if
(1) A represents the past occurrence of an event of type E;
(2) B witnessed e; and
(3) B’s witnessing of e is M-related to A’ s representation of the past occurrence of an event of type E.
(4) e is of type E
(5) A believes (1)-(4)
Perry then defines the M-relation as that of recollection. This is defined:

A recollects e if and only if:
(1) A represents the past occurrence of an event of type E;
(2) B witnessed e, and e is of type E; and
(3) B and A are the same live human body.

(Perry 95-7)

Unaided recollection, i.e. recollection requiring only that t.t.s. A exists without the added impetus of some additional information, constitutes the M-relation that connects various t.t.s.’s across time. In Perry’s formulation of the memory theory of personal identity, there are no claims of common personal identity between A and B, and the aforementioned claims of circularity are circumvented. The original Lockean memory theory is embedded in Perry’s theory through his reliance on memory, but whereas Locke intended to produce a theory not at all reliant of any form of bodily identity, Perry does require a common body to be present across time.

Sydney Shoemaker in Personal Identity proposes a functionalist theory for personal identity that relies on the psychological connectedness of various person-stages. Two person-states (or t.t.s.’s) are said to be psychologically connected if one contains a mental state that ‘stands in the appropriate relation of causal dependence to a state contained in a previously occurring mental state. Two person-states can then be said to belong to the same person if and only if:

(I) they are connected by a series of stages such that each member of the series is directly connected, psychologically, to the immediately preceding member, and (II) no such series of states which connects them ‘branches’ at any point, i.e., contains a member which is directly connected, psychologically, to two different stages occurring at the same time.

(Shoemaker 90)

A mental state within these person-stages, according to the functionalist view, plays a functional role in that, in response to certain stimuli, it causes certain behaviors and other mental states, which can be beliefs, character traits, desires, habits, talents, memories, attitudes, and preferences. These other mental states are definable in terms of their place within a complex network of causally related states. For example, suppose a person hears specific notes of a song over the radio and, from that information, forms a belief that a song by his favorite band is being played. Furthermore, this belief causes him to be happy and dance. Within this example, certain auditory stimuli resulted in the formulation of two mental states; one being the belief that his favorite song was playing, the other being this person’s happiness. The belief and happiness then caused the person to dance, i.e. the mental states created in response to the auditory stimuli were the middle steps in a causal chain that resulted in the behavioral output of dancing. Part of the functionalist view’s appeal is in its ability to conform to both materialism and dualism, and, as will be shown later, it is compatible with Perry’s view of personal identity, despite the fact that it does not explicitly state a physiological requirement for identity.

In order for a group of mental states to exhibit causal connectedness resulting in further mental states and behavioral outputs, it is necessary for the states to be co-personal; they are, therefore, illustrative of the synchronic unity between the elements of an individual’s person-stage. The functional role of mental states also explains the diachronic unity of mental states and person-stages. For instance, whereas the auditory stimuli in the aforementioned example resulted in a belief, happiness, and dancing as immediate responses, a week later the person may be left with a memory of his happiness and dancing and a belief that seven days ago he heard his favorite song, all of which are mental states defined by their causal relations to previously occurring mental states. These causal relations explain the appropriate causal dependence various mental states must have across time in order to be part of two psychologically connected person-stages or for a series of person-stages to be psychologically continuous. Assuming there is no branching of person-stages, a person is that which is composed of the psychologically continuous person-stages.

Perry’s and Shoemaker’s views are reconcilable as functionalism would be an adequate formulation of Perry's M-relation, if we mandate that functional states be confined to a single physical embodiment. As Perry calls for B’s witnessing of e to be M-related to A’s representation of an event of type E, the stimuli presented while witnessing e might produce various mental states that causally lead to B’s future representation of an event of type E. For example, a person A may possess a memory of skiing at Killington three
A' remembers e if and only if
(1) A represents the past occurrence of an event of type E;
(2) B witnessed e; and
(3) B's witnessing of e is M-related to A's representation of the past occurrence of an event of type E.
(4) e is of type E
(5) A believes (1)-(4)

Perry then defines the M-relation as that of recollection. This is defined:

A recollects e if and only if:
(1) A represents the past occurrence of an event of type E;
(2) B witnessed e, and e is of type E; and
(3) B and A are the same live human body.

(Perry 95-7)

Unaided recollection, i.e. recollection requiring only that t.t.s. A exists without the added impetus of some additional information, constitutes the M-relation that connects various t.t.s.'s across time. In Perry's formulation of the memory theory of personal identity, there are no claims of common personal identity between A and B, and the aforementioned claims of circularity are circumvented. The original Lockean memory theory is embedded in Perry's theory through his reliance on memory, but whereas Locke intended to produce a theory not at all reliant of any form of bodily identity, Perry does require a common body to be present across time.

Sydney Shoemaker in *Personal Identity* proposes a functionalist theory for personal identity that relies on the psychological connectedness of various person-stages. Two person-states (or t.t.s.'s) are said to be psychologically connected if one contains a mental state that 'stands in the appropriate relation of causal dependence to a state contained in a previously occurring mental state. Two person-states can then be said to belong to the same person if and only if:

(I) they are connected by a series of stages such that each member of the series is directly connected, psychologically, to the immediately preceding member, and (II) no such series of stages which connects them 'branches' at any point, i.e., contains a member which is directly connected, psychologically, to two different stages occurring at the same time.

(Shoemaker 90)

A mental state within these person-stages, according to the functionalist view, plays a functional role in that, in response to certain stimuli, it causes certain behaviors and other mental states, which can be beliefs, character traits, desires, habits, talents, memories, attitudes, and preferences. These other mental states are definable in terms of their place within a complex network of causally related states. For example, suppose a person hears specific notes of a song over the radio and, from that information, forms a belief that a song by his favorite band is being played. Furthermore, this belief causes him to be happy and dance. Within this example, certain auditory stimuli resulted in the formulation of two mental states; one being the belief that his favorite song was playing, the other being this person's happiness. The belief and happiness then caused the person to dance, i.e. the mental states created in response to the auditory stimuli were the middle steps in a causal chain that resulted in the behavioral output of dancing. Part of the functionalist view's appeal is in its ability to conform to both materialism and dualism, and, as will be shown later, it is compatible with Perry's view of personal identity, despite the fact that it does not explicitly state a physiological requirement for identity.

In order for a group of mental states to exhibit causal connectedness resulting in further mental states and behavioral outputs, it is necessary for the states to be co-personal; they are, therefore, illustrative of the synchronic unity between the elements of an individual's person-stage. The functional role of mental states also explains the diachronic unity of mental states and person-stages. For instance, whereas the auditory stimuli in the aforementioned example resulted in a belief, happiness, and dancing as immediate responses, a week later the person may be left with a memory of his happiness and dancing and a belief that seven days ago he heard his favorite song, all of which are mental states defined by their causal relations to previously occurring mental states. These causal relations explain the appropriate causal dependence various mental states must have across time in order to be part of two psychologically connected person-stages or for a series of person stages to be psychologically continuous. Assuming there is no branching of person-stages, a person is that which is composed of the psychologically continuous person-stages.

Perry's and Shoemaker's views are reconcilable as functionalism would be an adequate formulation of Perry's M-relation, if we mandate that functional states be confined to a single physical embodiment. As Perry calls for B's witnessing of e to be M-related to A's representation of an event of type E, the stimuli presented while witnessing e might produce various mental states that causally lead to B's future representation of an event of type E. For example, a person A may possess a memory of skiing at Killington three

---

3: In Perry's example, 'A' is now referring to a live human body and not a t.t.s.
years ago. This memory would be a representation of an event of a certain type. Person B may have skied at Killington three years ago, and would have thereby witnessed an event e. If A’s current representation was M-related to B’s having witnessed an event, and A and B possess the same live human body, then it could be said that A and B are the same person. Likewise, if A’s representation of an event of type E was appropriately caused by B’s witnessing of e, then A’s memory could be a mental state existing in a complex network of causally related states that began with B’s witnessing of e. The only difference between theories is that while Perry suggests A and B must have the same live human body in order for personal identity to be actualized, Shoemaker believes that only by mandating that A and B share a common body, or brain, can personal identity be realized in humans. He contends that functionalism provides a sufficient theory of personal identity abstract enough to account for personal identity in some species who are able to have psychologically connected person-stages, but not possess a brain. For Shoemaker, it is the “functional account of mental states, and the associated psychological continuity account of personal identity” that portray the ‘real essence’ of personal identity (Shoemaker 128).

Shoemaker proposes, in an attempt to better define his view of personal identity, the thought experiment of the brain-state-transfer device. The device exists in a world in which the inhabitants, which we will assume are exactly like us, must have their brain-states transferred to a different body that is exactly like their original every few years, (due to radiation that slowly damages their bodies). Their old bodies are then immediately destroyed upon completion of the transfer. They use the word ‘person’ in exactly the same manner as do both the pre and post-body-transfer possess the same legal rights and are considered by society to be identical. Shoemaker argues that this device, which we will assume works correctly every time, is person-preserving. He claims that “the physical realization of a mental state requires the existence of a physical ‘mechanism’ whereby it stands, or is capable of standing, in the functionally appropriate causal relations to other mental states of the same person” (Shoemaker 110). Initially, this statement seems to condemn the use of the BST device as a person-preserving procedure, since the brain states are being duplicated in a different body and must exist, at some point, outside of both bodies. Therefore, it seems that there must be a time when the mental states of the pre-body-transfer patient are not realized in any body at all.

Shoemaker proposes that the BST device can itself constitute an adequate physical mechanism that allows for the realization of mental states, and therefore the continuation of appropriate causal relationships among mental states that constitute a person. I, however, take issue with this claim. First, it is necessary to consider what is meant when it is said that a mental state is ‘realized.’ If a person is sitting at his desk and typing a paper, his concentration is most likely, or at least ideally, focused on the paper; memories of the 1996 Olympic Games, for example are most likely not at the forefront of his consciousness. Assuming the person possesses memories of the games but is not currently representing them in any way, I believe the mental states that constitute his memories are not being realized. My claim, which if correct will show Shoemaker’s BST device is not person-preserving, is that the realization of mental states is achieved when a mental state is recognized as present by its possessor. For instance, the mental state of pleasure is not realized in a person immediately after that person breaks his arm because that person would not be feeling pleasure. Although he may know what it is like to feel pleasure, no causal process is realizing pleasure. Just as a physical mechanism is necessary to store mental states, i.e. latent memories or character traits not currently being exhibited, one is also necessary to realize them; however, this latter claim demands certain capabilities of the mechanism. For mental states to be realized within a mechanism, the mechanism must be able to recognize or reflect upon the mental states being recognized. In other words, the mechanism must possess self-awareness or consciousness.

The BST device, it is safe to assume, is not self-aware and possesses no knowledge of or unaided access to the mental states contained within it. Therefore, it would not be possible for the mechanism to recollect, unaided or not, a particular event or realize a particular mental state. Because there would be no continuity of consciousness, the mental states held within the BST device and the mental states posited in the cloned body would be copies of those found in the original body. These copies would not be properly causally connected to the stimuli beginning the causal chain of mental states or the causal chain of mental states itself because the copied mental states would, in essence, be reconstructions of mental states qualitatively identical to the originals, but not part of the same potentially continuous consciousness. Therefore, the BST device does not act as an adequate mechanism to physically realize mental states. If it is not possible to transfer mental states from one mechanism to another, then we must put constraints on Shoemaker’s theory similar to those of Perry’s; namely, we must restrict mental states to a single body, so long as we seek the preservation of the subject under consideration.

Both Perry’s and Shoemaker’s theories of personal identity possess great merit as non-dualistic theories of personal identity in their ability to avoid circularity and their conformance to naturalistic concepts of personal-
years ago. This memory would be a representation of an event of a certain type. Person B may have skied at Killington three years ago, and would have thereby witnessed an event e. If A's current representation was M-related to B's having witnessed an event, and A and B possess the same live human body, then it could be said that A and B are the same person. Likewise, if A's representation of an event of type E was appropriately caused by B's witnessing of e, then A's memory could be a mental state existing in a complex network of causally related states that began with B's witnessing of e. The only difference between theories is that while Perry suggests A and B must have the same live human body in order for personal identity to be actualized, Shoemaker believes that only by mandating that A and B share a common body, or brain, can personal identity be realized in humans. He contends that functionalism provides a sufficient theory of personal identity abstract enough to account for personal identity in some species who are able to have psychologically connected person-stages, but not possess a brain. For Shoemaker, it is the *functional account of mental states, and the associated psychological continuity account of personal identity* that portray the 'real essence' of personal identity (Shoemaker 128).

Shoemaker proposes, in an attempt to better define his view of personal identity, the thought experiment of the brain-state-transfer device. The device exists in a world in which the inhabitants, which we will assume are exactly like us, must have their brain-states transferred to a different body that is exactly like their original every few years, (due to radiation that slowly damages their bodies). Their old bodies are then immediately destroyed upon completion of the transfer. They use the word 'person' in exactly the same manner we do and both the pre and post-body-transfer possess the same legal rights and are considered by society to be identical. Shoemaker argues that this device, which we will assume works correctly every time, is person-preserving. He claims that "the physical realization of a mental state requires the existence of a physical `mechanism' whereby it stands, or is capable of standing, in the functionally appropriate causal relations to other mental states within the same person" (Shoemaker 110). Initially, this statement seems to condemn the use of the BST device as a person-preserving procedure, since the brain states are being duplicated in a different body and must exist, at some point, outside of both bodies. Therefore, it seems that there must be a time when the mental states of the pre-body-transfer patient are not realized in any body at all.

Shoemaker proposes that the BST device can itself constitute an adequate physical mechanism that allows for the realization of mental states, and therefore the continuation of appropriate causal relationships among mental states that constitute a person. If, however, take issue with this claim. First, it is nec-

ecessary to consider what is meant when it is said that a mental state is 'realized.' If a person is sitting at his desk and typing a paper, his concentration is most likely, or at least ideally, focused on the paper; memories of the 1996 Olympic Games, for example are most likely not at the forefront of his consciousness. Assuming the person possesses memories of the games but is not currently representing them in any way, I believe the mental states that constitute his memories are not being realized. My claim, which if correct will show Shoemaker's BST device is not person-preserving, is that the realization of mental states is achieved when a mental state is recognized as present by its possessor. For instance, the mental state of pleasure is not realized in a person immediately after that person breaks his arm because that person would not be feeling pleasure. Although he may know what it is like to feel pleasure, no causal process is realizing pleasure. Just as a physical mechanism is necessary to store mental states, i.e. latent memories or character traits not currently being exhibited, one is also necessary to realize them; however, this latter claim demands certain capabilities of the mechanism. For mental states to be realized within a mechanism, the mechanism must be able to recognize or reflect upon the mental states being recognized. In other words, the mechanism must possess self-awareness or consciousness.

The BST device, it is safe to assume, is not self-aware and possesses no knowledge of or unaided access to the mental states contained within it. Therefore, it would not be possible for the mechanism to recollect, unaided or not, a particular event or realize a particular mental state. Because there would be no continuity of consciousness, the mental states held within the BST device and the mental states posited in the cloned body would be copies of those found in the original body. These copies would not be properly causally connected to the stimuli beginning the causal chain of mental states or the causal chain of mental states itself because the copied mental states would, in essence, be reconstructions of mental states qualitatively identical to the originals, but not part of the same potentially continuous consciousness. Therefore, the BST device does not act as an adequate mechanism to physically realize mental states. If it is not possible to transfer mental states from one mechanism to another, then we must put constraints on Shoemaker's theory similar to those of Perry's; namely, we must restrict mental states to a single body, so long as we seek the preservation of the subject under consideration.

Both Perry's and Shoemaker's theories of personal identity possess great merit as non-dualistic theories of personal identity in their ability to avoid circularity and their conformance to naturalistic concepts of person-
al identity. However, Shoemaker’s view is incomplete, as is demonstrated by its failure to achieve what Shoemaker claims it is capable of, i.e. not rely on any one physical mechanism for the realization of mental states. In order to reconcile Shoemaker’s functionalist view with its actual capabilities, it is necessary to constrict the theory slightly in the addition of a further demand that functional states be realized in a common physical mechanism. However, it is important to note that Shoemaker’s functionalist view relies on one physical mechanism for the realization of mental states only if the definition of ‘realization’ proposed in this paper is accurate. Although Shoemaker claims that the brain and the central nervous system are the mechanism in which personal identity is recognized in humans, he believes functionalism is abstract enough to account for identity in some alien species not possessing either mechanism. Though this may be true, it in no way allows for the transfer of mental states from one body to another, or the existence of mental states outside of their original mechanism.

Received January 2007
Revised April 2007

REFERENCES


ROBERT KUBALA is a sophomore in the Presidential Scholars Program at Boston College. Besides Plato, he is also interested in contemporary philosophy of mind. In his free time, he plays piano in several chamber ensembles and performs on the organ at Catholic liturgies. He would like to thank Gary Gurtler, S.J. and especially Robert Wood for their comments on earlier versions of this paper.