Is Blameworthiness Forever?

Andrew C. Khoury (Arizona State University) & Benjamin Matheson (Stockholm University)

Abstract: Many of those working on moral responsibility assume that "once blameworthy, always blameworthy." They believe that blameworthiness is like diamonds: it is forever. We argue that blameworthiness is not forever; rather, it can diminish through time. We begin by showing that the view that blameworthiness is forever is best understood as the claim that personal identity is sufficient for diachronic blameworthiness. We argue that this view should be rejected because it entails that blameworthiness for past action is completely divorced from the distinctive psychological features of the person at the later time. This is because on none of the leading accounts of personal identity does identity require the preservation of any distinctive psychological features, but merely requires some form of continuity. The claim that blameworthiness is forever should therefore be rejected. We then sketch an account of blameworthiness over time, and consider two objections.

Key words: blame, diachronic blameworthiness, moral responsibility, personal identity, time

Words: 9978

Many of those working on moral responsibility assume that "once blameworthy, always blameworthy." They believe that blameworthiness is like diamonds: it is forever.[1] Basic assumptions are often those most in need of scrutiny, and that is our task here. We will argue that blameworthiness is not forever; rather, it can diminish through time. We begin by considering why one might be initially attracted to the view that blameworthiness is forever. We then show that this view is best understood as the claim that personal identity is sufficient for diachronic blameworthiness: that one is now blameworthy for a past blameworthy act so long as one is personally identical with the author of the past act. We then consider the leading accounts of personal identity and note that on none of these accounts does identity require the preservation of any distinctive psychological features, but merely requires some form of continuity. This sets the stage for our objection to the claim that blameworthiness is forever: it entails that blameworthiness for past action is completely divorced from the distinctive psychological features of the person at the later time. But, we argue, this is implausible. The claim that blameworthiness is forever should therefore be rejected. We then sketch an account of blameworthiness over time, and consider two objections.

1. The View that Blameworthiness is Forever

Suppose that Shady commits some crime for which he is blameworthy in virtue of meeting the relevant control and epistemic conditions on blameworthiness. If Shady is blameworthy for committing the crime, then it's natural to think he'll still be blameworthy for it the following day. And the same is true for every subsequent day: Shady will be blameworthy for the crime tomorrow, the day after that, next week, next year, and so on. According to Neal Tognazzini:

---

[1] We owe this analogy to John Martin Fischer.

[An] important question that deserves attention, I think, is how to account for the characteristics that, once one temporal part comes to have them, "infect" future temporal parts of that person. Moral responsibility seems to be such a characteristic. When Shady commits his crime at t_1, not only is he responsible at t_1 for robbing the bank, he is also responsible at all subsequent times for robbing the bank at t_1 (2010: 160).

The view that blameworthiness, and moral responsibility more generally, is forever is implicit in the thought of many theorists. One of the contexts in which it sometimes rises to the level of an explicit commitment is within "ledger" theories of moral responsibility (e.g. Feinberg 1970; Zimmerman 1988; Haji 1998). According to such theories, moral responsibility is to be understood along the lines of a metaphorical "ledger of life." To be blameworthy is to have a debit on one's ledger, and to be praiseworthy is to have a credit on one's ledger. And it's typically thought that, as it were, the entries on one's ledger are made in permanent ink. Consider Joel Feinberg:

Moral responsibility, so conceived, is liability to charges and credits on some ideal record, liability to credit or blame (in the sense of "blame" that implies no action). Just as it is, as we say, *"forever to the credit"* of a hero or saint that he performed some noble act, so a man can *forever be "to blame"* for his faults (1970: 30; our emphasis).

In this way, it's natural to think that a person is forever blameworthy for her sins. Of course, what 'forever' entails depends on how long the person continues to exist. So there's a link between the claim that blameworthiness is forever and the notion of personal identity. The following sort of case might lend credence to the view that personal identity is the relation that governs blameworthiness over time.

Two twins, Barry and Larry, are arrested. The detectives have video evidence that a person who looks just like the twins robbed a liquor store, killing its owner in the process. As it happens, Barry is the culprit. Barry, however, manages to convince the detectives that he's innocent and that his brother, Larry, is the guilty party. Larry has no alibi and is eventually found guilty of Barry's crime. Not only is Larry punished for Barry's crimes, he is also subject to the blame that Barry deserves.

Reflection on this case and those like it suggest that personal identity is the relation that governs blameworthiness, and moral responsibility more generally, across time. First, the fact that Larry isn't blameworthy for Barry's action suggests that personal identity is a *necessary* condition for blameworthiness for an earlier act. That is, a person can be blameworthy for a past act only if that person is personally identical with the author of that earlier act. This explains why Larry isn't blameworthy for Barry's crimes. We set aside this claim in what follows.[2]

Second, the fact that Barry *is* blameworthy for his past crimes suggests that personal identity is a *sufficient* condition for blameworthiness for past action. Some ways of understanding this sufficiency claim are less plausible than others. According to David Shoemaker:

> No one could plausibly think identity could ever be a sufficient condition for MR [moral responsibility] anyway—there are surely epistemic and control conditions one must meet in addition to identity—so identity can serve as no more and no less than a necessary condition for MR" (2016: 308).

Putting the claim in terms of blameworthiness, it's wildly implausible to believe that the *mere* fact that personal identity holds between A and B is a sufficient condition for B's (or A's) blameworthiness. This view entails that any persisting person, as such, is necessarily

---

[2] Though see Shoemaker (2012) for an argument against this necessity claim. Our focus is on a particular sufficiency claim, to be specified shortly.

blameworthy. No one would be tempted by such a view because everyone accepts that there are

various synchronic conditions – conditions that must be satisfied at the time of action – on

blameworthiness (such as the control and epistemic conditions).

But this isn't to say that there's no interesting sense in which personal identity might be

thought to be sufficient for blameworthiness. Rather than thinking that personal identity is

sufficient for blameworthiness *simpliciter*, one might think that personal identity is sufficient for

diachronic blameworthiness. Diachronic blameworthiness contrasts with synchronic

blameworthiness. Synchronic blameworthiness is the extent to which an agent *at the time of the*

action is blameworthy for the action. Diachronic blameworthiness is the extent to which an agent

at some later time is blameworthy for an earlier action (see Khoury 2013; Matheson 2014).[3] The

claim that personal identity is sufficient for diachronic blameworthiness is the claim that *given*

some earlier act for which some agent is synchronically blameworthy in virtue of satisfying the

control and epistemic conditions at that earlier time, some agent at a later time is diachronically

blameworthy for that earlier act *if* the later agent is personally identical with the earlier agent.

This view isn't a nonstarter in the way that the view that personal identity is sufficient for

blameworthiness *simpliciter* is, and we believe it's the most plausible substantive construal of the

view that "blameworthiness is forever." It can be put more carefully as:

> *Blameworthiness is Forever* (BIF): If a person A at t_1 is blameworthy to degree *d* for
>
> committing act X at t_1, then an individual B at t_2 is blameworthy to degree *d* for act
>
> X if A is personally identical with B.[4]

---

[3] Note that the same sense of blameworthiness is appealed to in both notions.

[4] We formulate our claims here, and largely in what follows, in a way that naturally accords with a four-dimensionalist ontology in which the subject of an attribution of blameworthiness is a person *at a time*. But nothing in our argument requires the adoption of four-dimensionalism. There are other ways of temporal indexing. For example, rather than attaching the temporal index to the subject one may hold that the temporal index attaches to the predicate, such that a given alleged n-place predicate is really an (n+1)-place predicate (e.g. 'X is round' is properly understood as 'X is round at t'). Alternatively, one may hold that instantiation itself is temporal, or appeal to a temporal sentential operator. Which account one prefers will depend on one's favored solution to the problem of temporary intrinsics (see Lewis 1986: 203-205). We remain neutral on this question here. If the reader is averse to four-dimensionalism, she should feel free to read the claims according to her preferred account of temporary

Suppose that a person commits some action at t_1 and is blameworthy to a particular degree for doing so. BIF says that an individual at some later time t_2 is blameworthy to that same degree for that earlier action if that later individual is personally identical with the earlier person. So long as personal identity holds diachronic blameworthiness freely transfers through time undiminished. It's in this sense, proponents of BIF hold, that blameworthiness is forever.[5]

There's another sense in which one might understand the claim that blameworthiness is forever – namely:

The Trivial Sense in which Blameworthiness is Forever (TRV): If a person A at t_1 is blameworthy to degree *d* for committing act X, then it is forever true that A at t_1 is blameworthy to degree *d* for X.

TRV is trivial because it follows trivially from the fixity of the past. The past is fixed, so what's true of the past today will also be true of it tomorrow, and so on. TRV doesn't express a substantive view about the way in which blameworthiness functions over time while BIF does. We will argue that BIF is false but we aren't committed to rejecting, nor do we have any quarrel with, the trivial claim expressed in TRV.

It's not always clear whether those who think that blameworthiness is forever have BIF or TRV in mind. We suspect that some may fail to distinguish the two and that this may account for some of the appeal of BIF. For example, consider Michael Zimmerman:

properties.
[5] We speculate that many theorists are disposed to accept BIF. Locke encourages such a reading when he says, "In this personal identity is the basis for all the right and justice of reward and punishment" (1694: 46). Other theorists who appear to explicitly endorse or commit themselves to BIF include Feinberg (1970), Glannon (1998, 2002), Tognazzini (2010), and Zimmerman (1988). Coleman and Sarch (2012) have come very close. Moreover, we think that the relative dearth of discussion on the relation between responsibility and time in the literature is plausibly explained by a relatively widespread disposition to accept something like BIF.

We may ask: if *S* wills *e* at *T* (that is, roughly, if *S* decides at *T* that *e* should occur) and is directly culpable for this, *at what time* is he culpable for this? In this context, what this question amounts to is this: at what time is it correct, or true to the facts, to blame *S* inwardly for willing *e* at *T*? The answer clearly is: at that time at which it is true that (to speak tenselessly) *S* wills *e* at *T*. Now, it is relatively uncontroversial that, if *S* wills *e* at *T*, then it is true at *T* and forever thereafter that *S* wills *e* at *T* (1988: 40).

As we read Zimmerman, he is here appealing to the relatively uncontroversial claim that the past is fixed. If some agent acts at a particular time, then it's true at all subsequent times that that act was committed by that agent at that earlier time. Zimmerman goes on:

Thus at least this much may be said to be true: if S wills *e* at *T* and is culpable for this, then he is culpable at *T* and *forever thereafter* for it (1988: 40; our emphasis).

It appears to us that Zimmerman has here inferred BIF from TRV. What follows uncontroversially from the fact that *S* wills *e* at *T* and the fact that the past is fixed is that it's true at *T* and forever thereafter that [*S* wills *e* at *T*]. And if *S* is culpable at *T* for willing *e* at *T*, then it's true at *T* and forever thereafter that [*S* is culpable *at T* for willing *e* at *T*].[6] This is simply TRV. But Zimmerman, it seems, is making a stronger claim. He's claiming that it's true that [*S* is culpable *at T and forever thereafter* for willing *e* at *T*]. As we interpret Zimmerman, 'he' in the consequent of the above quoted conditional refers to *S* as either wholly present or temporally extended subject, and not a time-slice of *S* such as *S* at *T*. If so, then Zimmerman's above claim involves a commitment to BIF.[7] This illustrates the ease with which one can slip between TRV

---

[6] We here follow Zimmerman in treating the predicate rather than the subject as temporal. See fn.4.
[7] Note that Zimmerman (1988: 69, fn.3; 167-168) isn't entirely insensitive to our concerns.

and BIF. Because TRV is uncontroversial, it's easy to think that BIF is as well. But BIF isn't equivalent to TRV and it doesn't follow from it. Consider the property of being a fast swimmer. If S is a fast swimmer at T, then it follows from the fixity of the past that it's true at T and forever thereafter that [S is a fast swimmer at T]. But it clearly doesn't follow that it's true that [S is a fast swimmer at T and forever thereafter]. The property of being a fast swimmer is a property that can diminish and is ultimately lost across time (at death, if not before). The property of being blameworthy, we argue, is as well.

The reason that there is a sense in which blameworthiness is ultimately lost across time is because attributions of blameworthiness have existential import. That is, a claim of the form 'S at T is blameworthy' is true only if S at T exists. This doesn't, however, conflict with the coherence of posthumous blame. It's simply to say, roughly, that the subject of a true attribution of blameworthiness must exist at the time *to which* blameworthiness is attributed. But the subject needn't exist at the time *at which* blameworthiness is attributed. For example, assuming that Hitler met the conditions of blameworthiness, it's true now and forever that Hitler in early 1945 is blameworthy, for he existed then and instantiated the properties that ground blameworthiness. The referent of the subject of the attribution is Hitler as he existed in early 1945. But Hitler in 2018 is no more blameworthy than the present king of France is bald. Thus, when T refers to a time at which S does not exist then the description 'S at T' fails to secure reference, and hence any statement of the form 'S at T is blameworthy' isn't true.

Given this, we see no reasonable interpretation of the claim that blameworthiness is forever other than BIF or TRV. While TRV is trivial, BIF is distinct from and doesn't follow from it. This, of course, does nothing to show that BIF is false. In order to assess the plausibility of BIF we must first discuss the notion of personal identity to which it appeals.

2. Personal Identity, Continuity, and Distinctive Psychology

2.1 The Logic of Numerical Identity

Personal identity is a kind of numerical identity. Numerical identity is a relation that can only hold between a thing and itself. It stands in contrast to the notion of qualitative identity. Two things are qualitatively identical when they share the same qualities. Entities that are qualitatively identical may not be numerically identical, and entities that are numerically identical may not be qualitatively identical. Two cars on the dealer's lot may be qualitatively identical. They may be the same make, model, color, and so on. But they aren't numerically identical; there are two cars, not one. And a car that's white at one time may be numerically identical with a car that's black at a later time; a car isn't destroyed by a coat of paint.

The relation of numerical identity has two important logical features. First, numerical identity is a *non-scalar* relation; it's all-or-nothing in the sense that it doesn't admit of degrees. Any two objects are either the same entity or they aren't; there is no middle ground. Second, numerical identity is a *transitive* relation. If A is numerically identical to B, and B is numerically identical to C, then A is numerically identical to C.[8]

Since numerical identity has these logical features and personal identity is a form of numerical identity, personal identity must also have these features. We believe that this accounts for much of the appeal of BIF. It's easy to see why it would seem plausible that blameworthiness doesn't diminish across time if one thought that personal identity was sufficient for diachronic blameworthiness. Because personal identity is a transitive relation that either holds completely or not at all, it would be natural to think that diachronic blameworthiness is transitive and either holds completely or not at all. And this is just to say that blameworthiness is forever, so long as personal identity holds. We will present our case against this view in the next section, but first we must introduce the leading accounts of personal identity.

2.2 Approaches to Personal Identity

---

[8] Numerical identity is also symmetric and reflexive.

There are two dominant approaches to personal identity: one physical and one psychological. According to the physical approach, personal identity consists in some kind of physical continuity. One popular version of this approach identifies the relevant form of physical continuity as continuity of a biological organism's vital functions.[9] Consider:

> *The Biological Criterion* (BIO): Person A at t_1 is identical with some individual B at t_2 if and only if B is biologically continuous with A.

According to BIO, you are a biological organism and you continue to exist so long as that biological organism continues to exist (Olson 1997; DeGrazia 2005). On BIO a person could survive the removal of her cerebrum just as she could survive the removal of one of her kidneys. In either case, the person's vital functions, controlled by the brain stem, continue to operate and the biological animal that is the person continues to exist.

The psychological approach, on the other hand, holds that identity isn't a matter of any particular physical relation. Rather, it's a matter of the continuation of one's psychology. Consider:

> *The Psychological Criterion* (PSY): Person A at t_1 is identical with some individual B at t_2 if and only if B is uniquely psychologically continuous with A.[10]

According to PSY, you aren't a biological organism but a mind (see Parfit 1984: 207). If your cerebrum were removed and placed into an empty-headed body, you would survive as the

---

[9] What we say about the biological criterion will also apply, *mutatis mutandis*, to other substance-based accounts of personal identity, including those that are versions of the so-called 'simple' view of personal identity (e.g. Swinburne 1984) and the bodily continuity theory (e.g. Thomson 1997).

[10] The uniqueness requirement is necessary because psychological continuity can hold one-many, but personal identity can only hold one-one.

person composed of your cerebrum and the new body because your pre-op psychology is uniquely continuous with the psychology of this person.[11]

2.3 Continuity and Distinctive Psychology

On neither of these broad approaches to personal identity does identity entail the preservation of any distinctive psychological features. By distinctive psychological features we mean to refer to the psychological features that tend to vary across persons. Most people believe that two plus two equals four, but it's not the case that most people have a current intention to write on the duration of blameworthiness. The latter, but not the former, is the sort of thing we have in mind when we speak of distinctive psychological features (see Parfit 1984: 300-301, 515 fn.6, 517 fn.17).

It's obvious that identity doesn't entail the preservation of such distinctive features according to BIO. On BIO, personal identity consists in the persistence of the biological animal that you are, and the persistence of this animal doesn't require the persistence of any psychological features. If you were put into a persistent vegetative state, you would survive without any psychology whatsoever, according to BIO. You would survive because, though your psychology perished, your vital biological functions continued to operate. Thus, on BIO, personal identity doesn't require the preservation of any distinctive psychological features because it doesn't require the persistence of psychology at all. This is obvious and uncontroversial.

What is less obvious but which should be no more controversial is that this is also true on PSY, and we think that some of the appeal of BIF may be partly due to confusion here.

---

[11] One might be inclined to think that four-dimensionalism offers another alternative account of personal identity, but this isn't so. The four-dimensionalist must appeal to a criterion of personal identity like BIO or PSY in order to explain what unites the different temporal parts of a single person. Four-dimensionalism, as such, provides no such explanation. See Shoemaker (2012).

Personal identity, on PSY, doesn't require the persistence of any distinctive psychological features.

On PSY, as developed by Derek Parfit (1984: 204-207), personal identity consists in unique psychological continuity. Psychological continuity consists of overlapping chains of strong psychological connectedness. Psychological connectedness concerns the holding of some direct psychological connections across time. Direct connections between memory and experience are the most familiar. A memory is directly connected to an experience if the memory was caused, in the right way, by the experience and the memory is of that very experience.[12] Other examples include the relation between an intention and the act that carries it out, and the relation between a belief at one time and a belief at another time. These states are directly connected when the earlier one caused the later one, in the right way, and when they have the same content. Similar remarks apply to other psychological states such as desires. Psychological connectedness concerns the persistence of psychology across time.

Personal identity cannot simply consist in psychological connectedness. This is because connectedness doesn't fit the logic of numerical identity. First, connectedness is a *scalar* relation. This is because there can be more or less direct connections across time. Connectedness can hold to any degree depending on the number of direct connections (and, perhaps, on the strength of the individual connections).

The notion of personal identity, however, isn't scalar. This is why PSY must appeal to a threshold of connectedness that is *enough* for identity to hold over a given period of time. According to Parfit, while we can't say with any precision what the threshold *is*, we can say that the threshold is *surpassed* when the number of direct connections is "at least half the number that

---

[12] Locke (1694) is typically associated with the view that personal identity completely consists in the psychological connections between memory and experience, though it is not clear that this was actually his view (see Strawson 2011). Proponents of PSY cannot appeal to the direct connections between memory and experience insofar as the notion of memory presupposes personal identity, as Butler (1736) argued. Proponents of PSY must appeal to the broader notion of quasi-memory (Parfit 1984), which is like memory except that it does not presuppose identity. A similar point applies to other psychological states.

hold, over every day, in the lives of nearly every actual person" (1984: 206).[13] When there are

enough such connections, call this *strong psychological connectedness*. Unlike psychological

connectedness, strong psychological connectedness is a threshold concept and for that reason is

non-scalar.

Although the notion of strong connectedness is non-scalar, it can't be what identity

consists in because it's not a transitive relation. If A is strongly connected to B and B is strongly

connected to C, A may not be strongly connected to C.[14] In order to attain transitivity, a criterion

of identity can't simply appeal to strong connectedness; it must appeal to psychological *continuity*

because, unlike strong connectedness, this relation is transitive (Parfit 1984: 206-207). Continuity

is made up of overlapping chains of strong connectedness; continuity is the ancestral relation of

connectedness. If there are overlapping strong connections between A and B, and there are

overlapping strong connections between B and C, then there are overlapping strong connections

between A and C.

So while connectedness helps to explain identity on this view, it isn't an adequate

criterion of identity because it's scalar and non-transitive. In order to arrive at a psychological

criterion that has the appropriate logic, proponents of the psychological approach must make

connectedness fit the logic of identity. Appeal to the notion of strong connectedness yields a

non-scalar relation. And appeal to continuity, composed of overlapping chains of strong

connectedness, yields a transitive relation.

Importantly, while the appeal to psychological continuity yields a relation with the

appropriate logic, it does so only at the cost of entailing that this psychological relation doesn't

necessarily involve the preservation of any distinctive psychological features (or, equivalently, the

presence of any distinctive direct connections). You are, presumably, psychologically continuous

with your two-year-old self. Your three-year-old self remembered many of the experiences of

---

[13] Parfit is commonly misunderstood on this point.
[14] This is a familiar point owing to Reid's (1785) Brave Officer case.

your two-year-old self, your four-year-old self remembered many of the experiences of your three-year-old self, and so on. The same holds true for other psychological connections. So there are overlapping chains of strong connectedness between you now and your two-year-old self. But, of course, you have few if any memories of your two-year-old self. The same holds true for your other psychological states. So you aren't now strongly connected to your two-year-old self, though you are continuous with your two-year-old self. Psychological continuity, it's crucial to see, doesn't entail the persistence of any distinctive psychological features *whatsoever*.

3. Blameworthiness is Not Forever

Suppose that Brian at t_1 commits an act X for which he's blameworthy. Just prior to t_2 Brian is involved in a car accident. While the accident results in severe brain damage, he remains, at t_2, both biologically and psychologically continuous with himself at t_1. Proponents of BIF are committed to holding that Brian at t_2 is blameworthy for X. On both BIO and PSY, personal identity holds between Brian at t_1 and Brian at t_2. This fact, coupled with BIF, entails that Brian at t_2 is blameworthy for X. And this would be so *even if* the brain damage that he suffered was severe enough to permanently undermine his status as a moral agent, because on neither BIO nor PSY does the persistence of a person entail the persistence of the capacities that ground moral agency. Thus BIF, coupled with BIO or PSY, entails that a non-moral agent may be blameworthy. This strikes us as confused. As Michael McKenna notes:

> A morally responsible agent…is accountable for her conduct. She is one who can be held morally responsible for what she *has done*…Most simply, one is a morally responsible agent just in case she is, *in the here and now sense*, capable of moral action for which she could be morally responsible. This is a near tautology" (2012: 11-12; our emphasis).

14

Thus, a person at t can be appropriately held responsible for a past act only if she is a moral agent at t. If Brian at t_2 has the psychological capacities of a two-year-old, then he isn't, and can't, be blameworthy. Blameworthiness, whether synchronic or diachronic, requires that the subject be a moral agent, and the notion of a persisting person that BIO and PSY aim to capture simply isn't the same as the notion of a persisting moral agent.

Even if this objection could be met in some way, there is a further problem for BIF *granting* the assumption that the individual to whom diachronic blameworthiness is attributed is a moral agent. Consider Leon.[15] His life has been dedicated to the search for the fabled "elixir of life" – a drink rumored to provide immortality. On his travels, he discovers a mysterious island. The community on this island is just like any other human community, but with one difference: it's not uncommon for people on this island to live beyond the age of 500. Some have theorized that this is because of the water they drink there, which comes from an ancient well. Whatever the truth is, everyone in the community drinks this water and they all typically live beyond the age of 500. The people in this community exhibit psychological change across time just as we do. But because they live so long, these changes become much more pronounced: the psychology of a person in this community at one time may not resemble their earlier psychology whatsoever.

Leon drinks from this well. He goes on to live an extraordinarily long life, and over the course of his life undergoes the same sort of psychological change across time that the island's inhabitants do. So, Leon at age 500 is biologically and psychologically continuous with but psychologically *disconnected* to Leon at age 250; Leon at age 250 is biologically and psychologically continuous with but psychologically disconnected to Leon at age 50. Suppose that Leon at age 30 (Leon-30) – the age he drinks from the well – has the psychological profile of a ruthless conquistador. Imagine that Leon-30 commits a reprehensible act M – namely the massacre of the island's native inhabitants. And imagine that Leon-30 satisfies the conditions on synchronic

---

[15] This is an adaption of a case from Matheson (2014: 330) and is inspired by Parfit (1984: 302-303).

blameworthiness for doing so. So Leon-30 is synchronically blameworthy for M. Now fast forward 500 years. Leon is now 530 years old. He has gone through many characters during his life. Suppose that Leon-530 is a perfect time-slice psychological twin of your favorite moral saint. Leon-530 shares *no* distinctive psychological features with Leon-30 but remains psychologically and biologically continuous with him. Now, is Leon-530 blameworthy for M?

We contend that Leon-530 is *not* blameworthy for M. Why? Because Leon-530 shares none of Leon-30's *distinctive* psychological features, and these distinctive psychological features were essential to him committing M. While these two stages of Leon might be connected by the relations of psychological and biological continuity, they aren't connected by any direct distinctive psychological connections themselves. So the psychological states (e.g. beliefs, desires, or intentions) from which M stemmed aren't shared by Leon-530. Further, having the full psychological profile of your favorite moral saint, Leon-530 has no memory of the action M; there's simply no lasting trace of the "springs of action" that gave rise to M. Given this, we find it intuitively compelling to think that Leon-530 isn't blameworthy for M.

If one were to think that Leon-530 is blameworthy for M in virtue of the fact that he's personally identical with Leon-30, then one is committed to thinking that diachronic blameworthiness is a purely extrinsic property. It's a purely extrinsic property, on this view, in the sense that being diachronically blameworthy at t for a past act entails nothing whatsoever about one's distinctive psychological features at t. But it's hard to see how such an extrinsic property could be of intrinsic practical relevance. This, we take it, is the basic thought motivating Parfit's famous claim that "identity is not what matters". It seems that whether blame is appropriately directed at a person at a time t must depend in *some* way on what that person is like at t. It's not clear how, on such a view, our blaming practices could fulfil their characteristic functional role, for blameworthiness seems to involve a particular kind of criticizability – that is, a blameworthy individual is criticizable in light of a particular kind of flaw. A moral saint,

however, isn't so criticizable because she, by definition, lacks any such flaws. We find it hard to believe that blameworthiness could be extrinsic in the way entailed by BIF.

One might resist by raising a number of related questions at this point. First, wouldn't it be appropriate for a victim to blame Leon-530 for M? Second, doesn't Leon-530 have a reason to apologize for M? Third, wouldn't it be appropriate for Leon-530 to feel guilt for M? And fourth, doesn't Leon-530 have a duty to compensate victims of M?[16] It may be thought that intuition favors affirmative answers to these questions, and that this entails that Leon-530 remains blameworthy for M.

We agree that it may be appropriate for a victim to blame Leon-530, but we think the sense of appropriateness is merely epistemic. If your friend has overwhelming evidence that you, contrary to fact, are guilty of spreading a vicious rumor about her, then her blame of you is epistemically appropriate; it accords with her evidence. But it doesn't accord with the facts; you aren't actually blameworthy for spreading the rumor. Similarly, given that Leon-530 is personally identical with Leon-30 this may be evidence that Leon-530 is diachronically blameworthy for M.[17] And so the victim's blame of Leon-530 for M may be epistemically appropriate. But this doesn't entail that Leon-530 is blameworthy.

Further, the reasonable appearance of blameworthiness, we think, can provide the seemingly blameworthy agent with a reason – and perhaps even a duty – to respond in particular ways.[18] In the case above, you may have an obligation to explain yourself to your friend in a way that acknowledges that her blame of you is epistemically reasonable, to acknowledge the harm, and to express regret at its occurrence. Similar remarks may apply to Leon-530 if he is aware that others reasonably see him as blameworthy. Whether or not we wish to describe such responses

---

[16] Thanks to an anonymous referee for raising these issues.
[17] This would depend on how the details of the case are filled in. To foreshadow some points to come, in the actual world personal identity is some evidence of the persistence of distinctive psychological features and so can be evidence of diachronic blameworthiness.
[18] For development of this idea in the collective context, see Radzik (2001).

as genuine apology is beyond the scope of this paper, but it's clear that the existence of reasons to respond in such ways doesn't entail blameworthiness.

Related considerations can also help to explain why it *might* seem appropriate for Leon-530 to feel guilt. As Daniel Jacobson argues in the context of Bernard William's famous example of the lorry driver who non-negligently runs over a child, guilt can be both admirable and irrational: "…There are grounds for self-scrutiny whenever one is the causal agent of a bad outcome. An admirable person in a real-life situation would be prone to search for grounds of self-reproach: he cannot help himself to the stipulation that he is blameless" (2013: 114). There may be similar grounds for self-scrutiny when one is personally identical with a past blameworthy agent. Acknowledging Jacobson's point that the epistemic bar for such self-scrutiny may be irrationally low for admirable agents, we emphasize that there must be *some* ground for such self-scrutiny to be admirable rather than pathological. Perhaps it would suffice if Leon-530 knew, or at least believed, that he was psychologically continuous with Leon-30. He may have grounds, in effect, to survey his psychological profile for any direct connections with Leon-30, however few and however weak. And he may feel admirable but irrational guilt in the belief that there must be some such connection. But if the case is truly imagined as we have described it, then there will be no such grounds whatsoever because we stipulate that Leon-530 is a total (and not merely evaluational) time-slice psychological twin of your preferred moral saint. And, we assume, your preferred moral saint doesn't believe that she is psychologically continuous with Leon-30.

Finally, we agree Leon-530 may have a duty to compensate victims of M. However, this doesn't entail that he is diachronically blameworthy for M. It's plausible that compensatory duties may arise simply because one has benefitted from past wrongdoing. And so Leon-530 may have a duty to compensate the victims of M insofar as he benefitted in some way from M, just as any beneficiary of slavery may have a duty to compensate victims or their descendants. Thus, the

claim that Leon-530 is not blameworthy for M is consistent with the possibility that he has a compensatory duty.

The case of Leon is, we admit, quite outlandish. The reason that we are appealing to such a case is because we want to explore the plausibility of BIF. BIF says that blameworthiness freely transfers through time so long as personal identity holds. But on both BIO and PSY personal identity can hold in the absence of the holding of any direct distinctive psychological connections. Thus, we consider the limiting case in which though personal identity holds, there are no such direct connections whatsoever. This is why we focus on a case of, *ex hypothesi*, complete psychological change, unrealistic as it is.[19] Such cases illustrate just how implausible the implications of BIF are. According to BIF a person at a later time may be blameworthy for a past action *irrespective* of what she is like at the later time. Thus, BIF implies that a person who is a time-slice psychological twin of your preferred moral saint, your dear grandmother, or even *you yourself as you are now* may be diachronically blameworthy for the worst action you can dream up. This is difficult to accept.
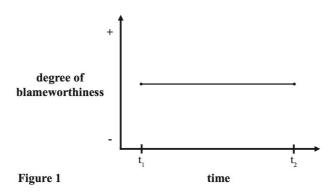

4. Diachronic Blameworthiness

In the previous section we argued that Leon-530 is not blameworthy for the action of Leon-30, because Leon-530 does not at all resemble Leon-30 psychologically – they share no distinctive psychological features.

This immediately suggests an account of diachronic blameworthiness. Diachronic blameworthiness concerns the persistence of distinctive psychological features. The persistence of distinctive psychological features is a matter of whether and the extent to which those distinctive features are psychologically connected to the psychology of a later individual. On this view, diachronic blameworthiness is a function of the extent to which the later person is

---

[19] A recent study (Harris, Brett, Johnson, and Deary 2016) suggests that many of us, however, do exhibit significant personality change over the course of our lives. For some discussion of diachronic blameworthiness in the context of more ordinary cases see Khoury (2013).

distinctively psychologically connected, in the relevant way, to the author of the act. If a person A at t_1 is blameworthy to degree d for act X which occurs at t_1, then an individual B at t_2 is blameworthy for X to degree d if B is maximally psychologically connected to A. This is represented in Figure 1, which maps an agent's blameworthiness for an action across time:
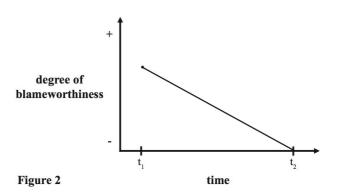

Figure 1

At t_1 the agent is blameworthy to a moderate degree for committing the act. At all points in time between t_1 and t_2 the agent is still blameworthy to that exact same degree for the act. BIF holds that blameworthiness across time is always like this. We've argued that BIF is false, but we readily admit that it's possible for a person's diachronic blameworthiness to be exactly equal to her synchronic blameworthiness for the same act. This will occur when there is maximal relevant psychological connectedness, which is to say that the agent exhibits no relevant psychological change across time.[20]

But it's also possible that, in a given case, there are no distinctive direct psychological connections between two points in time. When there's no distinctive psychological connectedness, there's no diachronic blameworthiness. This is the case with Leon-530. He, by
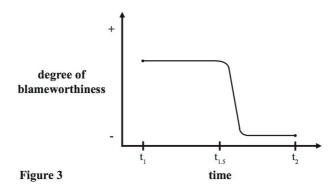
---

[20] We understand the notion of maximal (relevant) psychological connectedness as follows: A is maximally (relevantly) psychologically connected to B if and only if all of A's (relevant) psychological states are directly connected to B's psychology and B has no (relevant) psychological states that aren't directly connected to A's psychology.

stipulation, shares no distinctive psychological connections with Leon-30, and for that reason isn't diachronically blameworthy for the actions committed by Leon-30 even though he's personally identical with Leon-30. Furthermore, if there's diminished relevant psychological connectedness, then we think there's diminished blameworthiness (*pace* Glannon 1998; 2002). Consider Figure 2:

Figure 2

At t_1 the agent is blameworthy to a high degree for committing the act. At t_2, the agent isn't blameworthy for the earlier act. On the view we suggest here, this is because there's no relevant psychological connectedness between t_1 and t_2. This is the case with Leon-530. At the times between t_1 and t_2, there's diminished blameworthiness for the act, because there's diminished psychological connectedness of the relevant sort.

The view we suggest is straightforward. Diachronic blameworthiness is a function of distinctive psychological connectedness of the relevant sort. Our point is primarily structural. We take no stand on the precise nature of the function and we don't claim that diminishing blameworthiness must occur at a constant rate. Consider Figure 3:

Figure 3

In this case, the agent's blameworthiness for the action X which occurred at t_1, remained constant between t_1 and $t_{1.5}$. But some significant event occurred just after $t_{1.5}$ which results in a great loss of distinctive psychological connections. Perhaps, for example, the agent experienced a profound moral epiphany such that distinctive psychological connections hold to only a small degree, after $t_{1.5}$. We can imagine a version of Leon's case like this. Rather than imagining that the relevant psychological changes are gradual and ultimately total, as expressed in Figure 2, it may be that they occurred over a relatively short duration but weren't quite total as expressed in Figure 3.[21]

While we've focused on a hypothetical case in which we stipulate that the relevant psychological change is total, in the actual world it will often be difficult to judge the degree of psychological change in a given case (even in the first-personal context); we're epistemically limited beings, after all. This goes some way, we think, to explain the ambivalence we may feel towards the apparently reformed inmate, the old man living a quiet life in Canada who is discovered to have been a Nazi guard half a century earlier, or the friend guilty of past betrayals who claims to have turned over a new leaf. Such ambivalence can be explained, at least in part, by uncertainty about the relevant psychological change. Still, it's worth emphasizing, we are

---

[21] While we won't dwell on it here, we note the additional conceptual space for an agent's blameworthiness to increase over time. See Khoury (2013: 742).

sometimes confident that, for example, no relevant change has taken place in the criminal despite an apparent religious conversion or that a friend has indeed truly mended her ways.

One might worry that this view sets the bar too low by making it too easy for wrongdoers to escape blame. For example, it might be thought that one could become less blameworthy for an atrocity simply by doing a bit of charity work. But note, first, it's not these outward redemptive acts, as such, that matter to an individual's diminishing diachronic blameworthiness. Rather, it's the change in her underlying psychology. Redemptive acts can serve as evidence of such a change, but what matters is a relevant change in the psychology itself. Because these psychological states (e.g. beliefs, desires, and concerns) aren't directly subject to one's will, one cannot lessen one's diachronic blameworthiness simply through a few minor acts. What's required, instead, is deep psychological change. Thus, we don't think that this account is inappropriately lenient.[22] Moreover, we reemphasize that our claims have been structural in that we don't take any stand about the precise nature of the function relating degree of connectedness to degree of diachronic blameworthiness (e.g. it may not be one of direct proportion).

This is admittedly only a brief sketch of the nature of diachronic blameworthiness, but we think this view is clearly preferable to BIF. It implies that if one doesn't *at all* resemble, psychologically, the author of a past blameworthy act, then one isn't blameworthy for that earlier act even if one is personally identical with the author of the act. In addition to the precise nature of the function, further questions present themselves. Which psychological connections are the relevant ones? Is diachronic blameworthiness a function of the total psychological connections that hold across time, or does some subset of those connections get special weight?[23] And is

---

[22] Alternatively, one might argue that this view is inappropriately harsh insofar as it holds that maximal relevant psychological connectedness entails that blameworthiness isn't diminished. We thank an anonymous referee for raising this issue. We lack the space here to explain why we don't find the conception of blameworthiness implicit in this objection plausible.

[23] One of us has suggested that the connections that matter concern those between the psychological elements implicated in the execution of the earlier action and the later person's psychology. See Khoury (2013).

there an additional requirement that the earlier action can be narratively integrated (in some sense) with the later agent?[24] We won't try to settle these questions here. Instead, we offer the account of diachronic blameworthiness sketched here as a model for more fully developed accounts. Such accounts will have to take into consideration not only the possibility of diminished blameworthiness, but should also address the upshots for related phenomena, such as apology, guilt, and forgiveness.[25]

5. Objections

We've argued that blameworthiness is *not* forever, in the sense expressed by BIF, because there are cases in which the following is true: (a) The relation of personal identity holds between the earlier and later persons and (b) the later person is not at all blameworthy for the blameworthy act of the earlier person. The case of Leon, we argued, is one such case. Here we discuss two further objections. The first objection comes from metaphysics and challenges the truth of (a). The second comes from ethics and challenges the truth of (b).

5.1 The Objection from Metaphysics

The first objection maintains that, contrary to what we've argued, personal identity actually fails to hold in the cases we discuss. For example, one might argue that on the true account of personal identity Leon-530 isn't identical with Leon-30 and so the fact that Leon-530 isn't blameworthy for the act of Leon-30 fails to show that BIF is false. Note first that this objection involves a highly controversial metaphysical commitment – namely, the rejection of the two leading accounts of personal identity, BIO and PSY, both of which rule that Leon-530 is personally identical to Leon-30. Next, notice that if one holds that Leon-530 isn't personally identical to Leon-30, then either personal identity is severed at some sharp determinate point in

---

[24] For a suggestion along these lines, see Matheson (2014).
[25] For some initial investigations see Khoury (2013: 742-745). For an account of forgiveness that draws on our view, see Milam (manuscript).

time or it gradually reduces over some duration. Both of these options are especially problematic when conjoined with BIF.

Recall that Leon-30 has the psychological profile of a ruthless conquistador, whereas Leon-530 has the psychological profile of your favorite moral saint. Imagine that Leon's psychology changes at a constant rate such that the difference between any two neighboring person-stages is small and the same as the difference between any others. This may, for example, involve the severing of a single psychological connection. The difference between any two consecutive person-stages, then, is minuscule. Yet, if personal identity is broken at some moment in time between Leon-30 and Leon-530, then one of these minuscule differences must make all the difference. That is, it must make the difference between personal identity holding and failing to hold. This view, in conjunction with BIF, implies that there are two neighboring person-stages that differ psychologically only to an extremely small degree but that the first is *fully* blameworthy for M while the other is *not at all* blameworthy for M. This is strange: how could such a small difference make such a large difference? Furthermore, it's not clear how any point could be singled out as *the* point in which personal identity is severed in a non-arbitrary way. If personal identity is arbitrary with respect to what matters for blameworthiness across time, then personal identity isn't the relation that governs blameworthiness across time.

Alternatively, one might be tempted to reject the view that there is some sharp determinate point that marks the boundary of personal identity, and that instead personal identity gradually diminishes across time. The problem with this view is that it runs amok the logic of numerical identity which is standardly conceived as an equivalence relation: a binary relation that is symmetric, reflexive, and transitive. Such a view of personal identity, then, would involve either rejecting the standard conception of numerical identity or rejecting the notion that personal identity is itself a matter of numerical identity. Many have thought that the former view,

that numerical identity is sometimes vague, is incoherent (see Evans 1978).[26] And it's not clear that development of the latter option would yield an account that is a genuine alternative to accounts of personal identity involving this standard conception, as opposed to simply an account of something else entirely that presupposes such an account (Schechtman 1996 may be such an example). In any case, our point is that diachronic blameworthiness just doesn't follow the logic of the standard conception of numerical identity.

5.2 The Objection from Ethics

One might instead wish to defend BIF by appealing to considerations about the *ethics of blame* (on this topic see Smith 2007; Coates and Tognazzini 2011; and Zimmerman 1988). There is a difference between the conditions on blameworthiness and the conditions on all-things-considered appropriate *blaming*. And it might be claimed that it would simply be all-things-considered inappropriate to actually blame Leon-530 even though he remains blameworthy. Once this distinction is brought into the open, the objection goes, the problems for BIF evaporate.[27]

In response, first note that blameworthiness involves a sense of appropriateness (or fittingness, warrant or merit) that we call "worthiness" – that is, being blameworthy makes one a worthy target of blame, and so provides a *pro tanto* reason to blame. But a *pro tanto* reason to blame a person doesn't necessarily entail that person may be blamed all-things-considered. Hence the distinction between worthiness and all-things-considered appropriateness, where the latter means that there is sufficient justification to blame. Factors that might affect the latter include epistemic considerations (e.g. whether there is evidence of blameworthiness), standing

---

[26] Note that this is distinct from Sider's (2001) view that personal identity is indeterminate. Sider's claim is that that it is indeterminate whether BIO or PSY is the true criterion because, in effect, both are tied for first place as candidates for what we mean when we speak of personal identity. So, indeterminacy is generated because personal identity is an imprecise designator.

[27] This objection is implied by Coleman and Sarch (2012). It has also been suggested to us by Patrick Todd and Daniel Miller.

considerations (e.g. whether potential blamers satisfy non-hypocrisy or non-complicity conditions), and desirability considerations (e.g. blame would be undesirable if would lead to a disaster).[28] Being a worthy target of blame thus refers to a property of an individual, one which holds independently of these further considerations relevant to all-things-considered appropriate blaming.

We claim that Leon-530 isn't blameworthy because he isn't a worthy target of blame in the sense gestured to above. The objector may or may not grant that we have fixed upon the correct sense of appropriateness implicated in the notion of blameworthiness. Suppose she resists. She may agree that Leon-530 isn't a worthy target of blame as we're using the term 'worthy,' but insist that this isn't the sense of appropriateness implicated in the notion of blameworthiness. This dispute quickly threatens to be merely verbal. Perhaps the objector is simply using 'is not an all-things-considered appropriate target of blame (for a past action)' as we are using 'is not (diachronically) blameworthy', and she is using 'is (diachronically) blameworthy' as we would use the phrase 'is personally identical with the author of the (synchronically) blameworthy act.' While the objection, so understood, may be merely verbal, we think the objector's use of the term 'blameworthy' risks stretching it beyond recognition. This is because it isn't clear that the objector's use of the term 'blameworthy' entails that blame of the subject is worthy (i.e. appropriate) *in any sense at all*. If so, she should choose a less misleading term to refer to the notion she has in mind.

Suppose, instead, the objector grants that we have fixed upon the correct sense of appropriateness, worthiness, implicated in the notion of blameworthiness. She claims that while it would be all-things-considered inappropriate to blame Leon-530, he is in fact a worthy target of blame. Here the dispute isn't verbal but substantive. Clearly, the objector will need a plausible explanation as to why blaming the blameworthy in this case isn't all-things-considered

---

appropriate. Rather than speculating on what such an explanation might look like, we focus on the objector's positive claim that Leon-530 is blameworthy. Notice that the objector's task will be easier the more minimal the conception of blame to which she appeals. That is, if blame is, for example, understood as an overt sanction (e.g. punishment) then it will be relatively more difficult to sustain the claim that Leon-530 is a worthy target of blame than if blame is construed more minimally such as, say, simply a belief or emotion. So, if we can provide an adequate response to the objection by only appealing to a relatively minimal account of blame, then our response will be all the stronger.

We contend that this can be done by appealing to the view that blame has intentional content. On this view, blame involves a constituent thought or representation and blame is worthy just when that thought is true or that representation is accurate. For example, on some cognitivist accounts (e.g. the ledger theory), to blame is simply to make a particular kind of appraisal. Blame is worthy just when that appraisal is accurate. On some accounts in the Strawsonian tradition, blame is identified with a reactive emotion such as resentment. Emotions, in general, represent their objects to be a certain way and the object is a worthy target of the emotion just when that representation is accurate. Resentment, qua emotion, then, is worthy just when its constituent representation is accurate (see Graham 2014; Rosen 2015).

On either such account, we claim that the intentional content of blame involves the attribution of a flaw to the subject. That is, blaming involves thinking or representing the subject to have a flaw. And a necessary condition of the subject being a worthy target of blame is the truth of this thought or accuracy of this representation. On this conception, the proponent of the objection from ethics – who claims that Leon-530 is in fact blameworthy – is committed to thinking that Leon-530 is flawed. But, by hypothesis, he is not.

6. Conclusion

We've argued that blameworthiness can vary over time – that is, blameworthiness is *not* forever. As noted, we don't hold that what an agent does in the future can erase the past. If an individual is blameworthy at a given time, it remains forever true that the agent was blameworthy *at that time*. But diachronic blameworthiness can come apart from synchronic blameworthiness. While we've focused on blameworthiness in this paper, we also believe that praiseworthiness is not forever. Indeed, we think this applies to moral responsibility in general, given that moral responsibility is normally taken as the extent to which one is praiseworthy or blameworthy.

Proponents of BIF hold that once an individual is blameworthy, she is forever blameworthy. They hold that we are eternally tarnished by sins of the past and that blameworthiness "infects" all of our future selves. Proponents of BIF must accept that no matter how much an individual changes, she is always blameworthy. We reject this. If an individual changes in the relevant way, she can become less blameworthy for a particular action. If an individual changes sufficiently, then she can stop being blameworthy altogether for a particular action. We think this is a very optimistic conclusion. Blameworthiness, unlike diamonds, is not forever.

References

Butler, Joseph (1736) 'Of Personal Identity'. In *The Analogy of Religion*, reprinted in Perry 1975, pp. 99-105.

Coates, D. Justin and Neal A. Tognazzini. (2012) 'The Nature and Ethics of Blame'. *Philosophy Compass*, 7, 197-207.

Coleman, Jules and Alexander Sarch (2012) 'Blameworthiness and Time'. *Legal Theory*, 18, 101-37.

DeGrazia, David (2005) *Human Identity and Bioethics*. New York: Cambridge University Press.

Evans, Gareth (1978) 'Can There Be Vague Objects?'. *Analysis*, 38, 208.

Feinberg, Joel (1970) *Doing and Deserving: Essays in the Theory of Responsibility*. Princeton: Princeton University Press.

Glannon, Walter (1998) 'Moral Responsibility and Personal Identity'. *American Philosophical Quarterly*, 35, 231-49.

Glannon, Walter (2002) *The Mental Basis of Responsibility*. Aldershot, UK: Ashgate.

Graham, Peter A. (2014) 'A Sketch of a Theory of Moral Blameworthiness'. *Philosophy and Phenomenological Research*, 88, 388-409.

Haji, Ishtiyaque (1998) *Moral Appraisability*. Oxford: Oxford University Press.

Harris, Matthew, Caroline Brett, Wendy Johnson, and Ian Deary (2016) 'Personality Stability From Age 14 to Age 77 Years'. *Psychology and Aging*, 31, 862-74.

Jacobsen, Daniel (2013) 'Regret, Agency, and Error'. In David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility: Volume 1* (Oxford: Oxford University Press), pp. 95-125.

Lewis, David (1976) 'Survival and Identity'. In Amélie Oksenberg Rorty (ed.), *The Identities of Persons* (Berkeley and Los Angeles: University of California Press), pp. 17-40.

Lewis, David (1986) *On the Plurality of Worlds*. Oxford: Blackwell.

Locke, John (1694) 'Of Identity and Diversity'. In *An Essay Concerning Human Understanding*, reprinted in Perry 1975, pp. 33-52.

Khoury, Andrew C. (2013) 'Synchronic and Diachronic Responsibility'. *Philosophical Studies* 165, 735-52.

Matheson, Benjamin (2014) 'Compatibilism and Personal Identity'. *Philosophical Studies*, 170, 317-34

Milam, Per (manuscript) 'Forgiveness as Change'.

McKenna, Michael (2012) *Conversation and Responsibility*. New York: Oxford University Press.

Olson, Eric (1997) *The Human Animal: Personal Identity Without Psychology*. New York: Oxford University Press.

Parfit, Derek (1984) *Reasons and Persons*. Oxford: Oxford University Press.

Perry, John ed. (1975) *Personal Identity*. Berkeley: University of California Press.

Radzik, Linda (2001) 'Collective Responsibility and Duties to Respond'. *Social Theory and Practice*, 27, 455-71.

Reid, Thomas (1785) 'Of Memory'. In *Essays on the Intellectual Powers of Man*, reprinted in Perry 1975 as 'Of Identity', pp. 107–12, and 'Of Mr. Locke's Account of Our Personal Identity', pp. 113–18.

Rosen, Gideon (2015) 'The Alethic Conception of Moral Responsibility'. In Randolph Clarke, Michael McKenna, and Angela Smith (eds.), *The Nature of Moral Responsibility* (Oxford. Oxford University Press), pp. 65-87.

Schechtman, Marya (1996) *The Constitution of Selves*. Ithaca: Cornell University Press.

Shoemaker, David (2012) 'Responsibility Without Identity'. *Harvard Review of Philosophy*, 18, 109-32.

Shoemaker, David (2016) 'The Stony Metaphysical Heart of Animalism'. In Stephan Blatti and Paul Snowdon (eds.), *Animalism: New Essays on Persons, Animals, and Identity* (Oxford: Oxford University Press), pp. 303-28.

Shoemaker, Sydney and Richard Swinburne (1984) *Personal Identity*. Oxford: Blackwell.

Sider, Theodore. (2001) 'Criteria of Personal Identity and the Limits of Conceptual Analysis'. *Philosophical Perspectives*, 15, 189-209.

Smith, Angela M. (2007) 'On Being Responsible and Holding Responsible'. *Journal of Ethics*, 11, 465-84.

Strawson, Galen (2011) *Locke on Personal Identity*. Princeton, NJ: Princeton University Press.

Thomson, Judith Jarvis (1997) 'People and Their Bodies'. In Jonathan Dancy (ed.), *Reading Parfit* (Oxford: Blackwell), pp. 202-29

Tognazzini, Neal A. (2010) 'Persistence and Responsibility'. In Joseph Campbell, Michael O'Rourke, and Harry Silverstein (eds.), *Time and Identity* (Cambridge, MA: MIT Press), pp. 149-63.

Zimmerman, Michael J. (1988) *An Essay on Moral Responsibility*. Totowa, NJ: Roman and Littlefield.