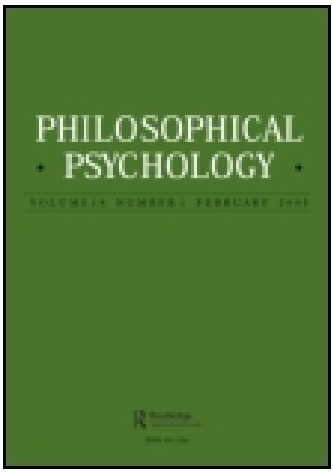

Philosophical Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/cphp20

Identity, control and responsibility: The case of Dissociative Identity Disorder

Jeanette Kennett & Steve Matthews
Published online: 24 Jun 2011.

To cite this article: Jeanette Kennett & Steve Matthews (2002) Identity, control and responsibility: The case of Dissociative Identity Disorder, Philosophical Psychology, 15:4, 509-526

To link to this article: http://dx.doi.org/10.1080/09515089.2002.10031978

PHILOSOPHICAL PSYCHOLOGY, VOL. 15, NO. 4, 2002

Identity, control and responsibility: the case of Dissociative Identity Disorder

JEANETTE KENNETT & STEVE MATTHEWS

ABSTRACT *Dissociative Identity Disorder (DID) (formerly known as Multiple Personality Disorder) is a condition in which a person appears to possess more than one personality, and sometimes very many. Some recent criminal cases involving defendants with DID have resulted in "not guilty" verdicts, though the defense is not always successful in this regard. Walter Sinnott-Armstrong and Stephen Behnke have argued that we should excuse DID sufferers from responsibility, only if at the time of the act the person was insane (typically delusional); otherwise the presumption should be that persons with DID are indeed responsible for their actions. We find their interpretation of DID and of the way in which the requirements for criminal insanity relate to this condition worrying and likely to result in injustice to DID sufferers. Our thesis is that persons with DID cannot be responsible for their actions if the usual features of the condition are present. A person with DID is a single person in the grip of a very serious mental disorder. By focusing on the features of DID which have, as we argue, the effect of deluding the patient, we try to show that such a person is unable to fulfill the ordinary conditions of responsible agency (namely, autonomy and self-control).*

> … for this curious child was very fond of pretending to be two people. "But it's no use now," thought poor Alice, "to pretend to be two people! Why, there is hardly enough of me left to make one respectable person!" (Lewis Carroll, 1865, *Alice's adventures in wonderland*, third last paragraph of Chapter 1; quoted in Hacking, 1995, p. 18)

1. Introduction

Dissociative Identity Disorder (DID), formerly known as Multiple Personality Disorder (MPD), raises a thorny problem about moral responsibility: if someone with DID acts in a morally or legally bad way, can we hold this individual responsible when it is claimed that the accused personality is not the personality who

Jeanette Kennett, Department of Philosophy, Monash University, Wellington Road, Victoria 3800, Australia, email: jeanette.kennett@arts.monash.edu.au; Steve Matthews, School of Humanities and Social Sciences, Centre for Applied Philosophy and Public Ethics, Charles Stuart University, Locked Bag 588, Wagga Wagga, NSW 2678, Australia, email: stmatthews@csu.edu.au

ISSN 0951-5089/print/ISSN 1465-394X/online/02/040509–18 © 2002 Taylor & Francis Ltd
DOI: 10.1080/0951508021000042??

acted [1]? Walter Sinnott-Armstrong and Stephen Behnke argue we can; we present reasons for denying that we can. The paper will proceed as follows. In the next two sections, we set up some positions and distinctions before presenting and criticizing Sinnott-Armstrong and Behnke's view. We then present the case for our thesis. We finish with some comments on the nature of the responsibilities we think DID sufferers do have once they have acknowledged the implications their condition has for their behavior.

2. How many persons?

Consider the case of Eve (O'Kelly & Mackless, 1956, p. 27).

> [A] demure, retiring individual, Eve White was quiet, industrious, and, as … [her therapists Thigpen & Cleckley] … put it "in some respects almost saintly." During the course of therapy … [Thigpen & Cleckley] … were led to suspect the existence of another [alter] personality … This new personality, Eve Black, had been co-existing with Eve White since childhood … Eve Black's behavior was the opposite of White. Black was shrewd, rowdy, and provocative; she enjoyed joking and pranks. Uninhibited and frank, Black lived for the moment. Furthermore, Black was aware of the existence of White while Eve White knew nothing of her other personality. Black delighted in placing Eve White in embarrassing positions: "When I go out and get drunk" … Eve Black once said … "*She* wakes up with the hangover. She wonders what in the hell's made her so sick."

After some initial skepticism, the condition that Eve and others suffered from gained an entry into the *Diagnostic and statistical manual of mental disorders*, or DSM, as Multiple Personality Disorder (MPD). The criteria for diagnosis of the condition were later refined and the name of the condition changed to Dissociative Identity Disorder [2]. This condition is real in the uninteresting sense that there are in fact individuals like Eve who exhibit its characteristics as set out by the DSM. In this paper, we assume that the condition is genuine in this sense, and that these characteristics constitute a distinct disorder. We acknowledge, however, another debate over whether we should individuate fully formed alter personalities in roughly the same way we individuate ordinary persons, or whether we should regard a person with the condition as a single person, and understand the alters as aspects of that person.

Consider the following two positions. Realists would argue that we should treat alter personalities as though each was a separate person, because each *is* a separate person. There may be independence of wills, separateness of character and disposition, and partitioning of experiential memory sets. This position entails that each alter is a separate agent, morally autonomous, a subject apt for praise and blame, a potentially independent social actor with moral and legal rights and responsibilities. We can regard realists as committed to what we will call the *Multiple Persons thesis*.

Those with a genuine commitment to this thesis will occupy the default position that body-sharing alters are not responsible for each others' actions.

Although we agree with their position on responsibility, we reject the metaphysics of the Multiple Persons thesis. We favor what we will call the *Single Person thesis*: someone with DID is an individual human person whose psychiatric symptoms, we would argue, are akin to a species of global self-delusion. So-called alter personalities are not to be regarded as metaphysically separate entities from the person, but rather count as altered states of that person [3].

More details of this idea will emerge as we proceed, but for now we can put the gist of our argument on moral responsibility and DID in relation to this view. We think there are strong reasons for not regarding the DID sufferer as responsible for the bad actions performed when in, as we shall term it, an alter state. Those who would hold the sufferer responsible for those actions, while holding the Single Person thesis, have not paid due attention to the nature and condition of the (single) person they would hold responsible. Their attitudes have been, we suggest, covertly influenced by the Multiple Persons thesis, which has held sway over popular imagination and, as in the case of Eve, is implicitly invoked in almost every description of persons with the disorder. If, instead, we *begin with* the view that there is just one person in these cases and proceed to ask questions about the nature of the disorder and the person who suffers it, we come up with answers that challenge the notion that such a person could fulfill the ordinary requirements for responsible agency.

Our argument against holding someone responsible for the bad actions committed when in an alter state is simply that the patient does not possess the relevant capacities of judgment and control with respect to that state. In particular, moral responsibility over time quite generally depends on the having of psychological connections that facilitate forms of self-control. If these connections are absent, through no fault of the individual concerned, then we are not justified in holding that individual morally responsible. A so-called alter who commits a morally reprehensible act is the same person as a different alter who shares the same body. But we regard alters as persons in altered states, and no more; at the time of acting, so we will claim, the person with DID is deluded about who s/he is, and so it is, in the morally relevant respects, merely *as if* there was another person in control at this time. On the Multiple Persons view, by contrast, it is claimed there really is another person in control at that time so there is someone who may be a fit subject of blame. In virtue of the delusional character of the disorder, we can say that the lack of psychological connections between different stages of a person prevents that person from effectively exercising self-control, and so the individual should not be held morally responsible for the act [4].

3. Hosts and alters

Our support for the Single Person thesis places us at odds with the loaded language of "alter personality" and "host personality" found in much of the realist literature on MPD/DID. We believe the use of this language is partially responsible for the

perpetuation of the Multiple Persons thesis we think our opponents have tended to smuggle into their analyses. We think it important, therefore, to specify a way of understanding these terms compatible with our more deflationary single-person metaphysics.

A so-called alter personality is a state of the one person in which only the person's concept of *self* has been replaced, distorted, and diminished. In support of this claim we note that alter personalities do not, for example, exhibit symptoms of global amnesia, a condition that can disorient a person, leave them paralyzed in the present, reduce them to near total dysfunction [5]. On the contrary, when in an alter state, patients continue to survive in their environment. Their beliefs about all manner of worldly things are preserved and continue to be utilized. They may well continue to know where the train station is, may continue to know how to use the washing machine, may continue to know who the Prime Minister is, may continue to write and add up, and so on. In short, it seems that the continuity-of-self problem the DID patient has is in a sense monothematic. Only those bits of information directly or indirectly associated with experiential memory are lost; the rest is preserved. This fact seems far more compatible with the Single Person thesis than its alternatives. It suggests that, in addition to the preservation of bodily continuity, there is a significant degree of psychological continuity across altered states [6].

The interpretation of the term "host personality" is problematic. Is the host the personality mostly in control of the body? Is it the one with maximum control over the switching process? Is it the personality who bears the same name as the name on the birth certificate, and who identifies with that birth person? We circumvent these problems by stipulation. What we will mean by "host personality" is what might best be described as the *historical person*. The identity of the historical person is determined bodily, relationally and conventionally. There will be a fact of the matter concerning the bodily identity of a patient with DID—one simply traces a line of physical continuity back to the person's birth. There will be a fact of the matter concerning the birth name of a patient with DID—one simply checks the public record. And there will be many other facts, independent of the subjective point of view of the patient, which determine who the person is. Most salient among these will be the views of those around the patient—family, friends, doctors, and perhaps other professionals. These many facts we regard as sufficient for identification and individuation of the historical person, and it is the historical person we will regard as the referent of the term "host personality."

4. Sinnott-Armstrong and Stephen Behnke

In 1993, a woman—Denny-Shaffer—removed a baby, some blood, and a placenta from a hospital. She contacted her boyfriend and family in order to convince them that she had given birth to the baby herself, and over the next few weeks looked after it. Denny-Shaffer was caught and charged with kidnapping. In her defense it was argued that an alter personality had committed the crime and that the host personality—Bridget—was not aware of what had happened. The defendant was found not guilty. The court found that:

> [A]t the time of the abduction, her dominant or host personality was not in control so as to cause commission of the offense, [she] was not aware that an alter personality or personalities were the cognizant parties controlling the physical actions; that as a result of the defendant's severe mental disease or defect, the host or dominant personality was unable to appreciate the nature and quality of wrongfulness of the conduct which the alter or alters controlled. [7]

Sinnott-Armstrong and Behnke argue that the court's finding was in error. In presenting a test for deciding criminal responsibility in cases of DID they argue, in line with the Single Person thesis, that the individual Denny-Shaffer-cum-Bridget was criminally responsible for kidnapping the baby. The test they propose takes its cue from a court ruling in another case, the Grimsley case, where Grimsley's alter, Jennifer, drove under the influence of alcohol. The court held Grimsley responsible, saying:

> There was only one person driving the car and only one person accused of drunk driving. It is immaterial whether she was in one state of consciousness or another, so long as in the personality then controlling her behavior, she was conscious and her actions were a product of her own volition. (cited in Sinnott-Armstrong & Behnke, 2000, p. 302)

Armstrong and Behnke agreed with the ruling in this case as well as the reasons cited. What conditions for responsibility do they require then? Returning to the first case, suppose we are now attempting to judge whether to hold the host personality Bridget morally responsible for removing the baby from the hospital. Sinnott-Armstrong and Behnke think Bridget should be held morally responsible if (1) Bridget and the alter Denny-Shaffer are the same person, and (2) Denny-Shaffer was not criminally insane at the time of acting. Although we agree with Sinnott-Armstrong and Behnke that there are not two (or more) separate persons in these cases, we dispute the second claim [8]. We think that, given there is a single person throughout the series of events involved in the case, the delusive character of the disorder provides an excusing condition for moral responsibility. The individual is indeed criminally insane at the time of acting.

In what sense is the individual to be regarded this way? Since we regard the offending alter and Bridget as the same person, and since we regard Bridget as the one and only person in the whole affair, we think that the term "Denny-Shaffer" is simply shorthand for a state of Bridget in which she "lacks substantial capacity either to appreciate the criminality (or wrongfulness) of [her] conduct or to conform [her] conduct to the requirements of the law" (extract from the legal standard of sanity in the US Model Penal Code). The word "she" here applies to Bridget, who indeed is unaware of the wrongfulness of the actions in question when she is in her Denny-Shaffer manifestation. Given this understanding of the case it is misleading of Sinnott-Armstrong and Behnke to regard Denny-Shaffer as the person we should be testing for criminal insanity. We think that to do so is to covertly appeal to aspects of the Multiple Persons thesis. The right, and only, person to test is Bridget—a point

Sinnott-Armstrong and Behnke overtly accept—and when Bridget is in her alter state as Denny-Shaffer she does indeed count as criminally insane.

So to sum up, we think that Sinnott-Armstrong and Behnke are right to hold the Single Person thesis, but once you are so committed we think it important to consider fully the logical consequences for people who exhibit the phenomena of DID. In particular, you are not entitled to help yourself to those aspects of the Multiple Persons thesis that support a positive conclusion on the question of moral responsibility. According to a supporter of the Multiple Persons thesis, Denny-Shaffer, *qua* so-called alter, is an autonomous individual person and given the facts of the case (that she had access to the semantic information that her action was against the law and she was fully cognizant of what she was doing) is the morally responsible one. Again, one needs to look closely at what the word "she" refers to. It seems to us that it refers to the patient with DID, Bridget, and Bridget in her Denny-Shaffer state quite literally does not know what she is doing.

In the light of our support of the Single Person thesis we now turn to an analysis of DID in terms of delusion.

5. DID and delusion

MPD/DID was not recognized as a genuine condition until quite recently. In the 1950s, Thigpen and Cleckley faced a barrage of denial and criticism over their account of Eve. Arguably, it took the popularization of the condition—following the book and film *Sybil* in 1976—to engender more serious interest [9]. Interestingly, the book was written only after Sybil's therapist, Dr. Cornelius Wilbur, felt sufficiently snubbed by the psychiatric community to "go public."

To be taken seriously MPD needed, *inter alia*, an entry into DSM, published periodically by the American Psychiatric Association. It was not until 1980 that Multiple Personality Disorder was able to occupy its own taxonomic position in the DSM. The diagnostic criteria for it were given as follows:

(1) The existence within the individual of two or more distinct personalities, each of which is dominant at a particular time.
(2) The personality that is dominant at any particular time determines the individual's behavior.
(3) Each individual personality is complex and integrated with its own unique behavior pattern.

DSM criteria do change. They changed in DSM-III-R (1987), and again in DSM-IV (1994). The 1994 criteria were:

(1) The presence of two or more distinct identities or personalities or personality states (each with its own relatively enduring pattern of perceiving, relating to and thinking about the environment and self).
(2) At least two of these identities or personality states recurrently take control of the person's behavior.

(3) Inability to recall important personal information that is too extensive to be explained by ordinary forgetfulness.
(4) The disturbance is not due to the direct physiological effects of a substance (e.g. blackouts or chaotic behavior during Alcohol Intoxication) or a general medical condition (e.g. complex partial seizures). *Note*: In children the symptoms are not attributable to imaginary playmates or other fantasy play.

The important change to the 1994 criteria relevant to our purposes is in relation to (1). The earlier criteria are plainly metaphysically loaded insofar as they presuppose the existence of discrete personalities within a single human being. It is an easy jump from the idea of a personality to all of the moral and legal baggage of an autonomous person, and these criteria certainly invite such a risky inference. The 1994 criteria specify the option of the far less extravagant choice of a personality state. Indeed, in some circles following DSM-IV, the idea of multiple persons is no longer *de rigueur*, and the new official title of "Dissociative Identity Disorder" helps permit the less extravagant reading. The chair of the committee responsible for DSM-IV, David Spiegel, remarked that, "There is a widespread misunderstanding of the essential psychopathology in this dissociative disorder, which is failure of integration of various aspects of identity, memory, and consciousness. The problem is not having more than one personality; it is having less than one personality" [10].

In line with this, Stephen Braude (1995) argues that there is evidence to suggest that so-called alter personalities are mere person-fragments, and not in the sense of having fully rounded personalities that last for short periods, but in the sense that the character of the alters is unidimensional, or at least extremely under-developed. According to Braude (1996, p. 51):

> [T]he traits and abilities manifested by or latent in the pre-dissociative personality begin to get distributed throughout the members of the personality system. Moreover, as alters proliferate, they apparently become increasingly specialized, and one is less likely to find any personality having the complexity or range of functions presumably possessed by the subject prior to the onset of splitting.

This phenomenon Braude refers to as "attribute-distribution" and "attribute-depletion." It makes the Multiple Persons thesis look implausible indeed. If we regard an individual body-person as supplying only a finite number of psychological attributes for the formation of a single person, then attribute-depletion is precisely what you would predict in cases where dissociation occurs. Our moral theorizing should thus track this more plausible interpretation.

The comments of Spiegel and Braude prompt us to think about DID in quite different ways to that presupposed by the realist literature on the subject. How should we understand what happens to a person with the condition when, as it is said, an alter personality takes over in the host? It may be fruitful to analyze the condition as involving episodes of delusion about the self. We will not argue for this in detail here but we will point to some commonalities between DID and other psychiatric disorders involving delusions surrounding identity. A defensible account

of moral responsibility across time needs to treat like cases alike. We think that accounts which would treat DID sufferers as *prima facie* responsible for the actions of an alter personality have, among other things, failed to notice the similarities between this condition and other serious psychiatric conditions which *prima facie* exempt sufferers from responsibility or mitigate their responsibility, for actions performed while affected. It is worth taking time then to compare DID with disorders involving delusion, since it is widely accepted—and accepted by Sinnott-Armstrong and Behnke—that agents cannot exercise appropriate levels of autonomy and control over actions undertaken while in a delusional state.

If the sufferer is just one person as we maintain, then it seems clear that at various times, when in an alter state, she has false or incomplete beliefs about the person she is; she has a radically changed or restricted outlook on her projects, commitments and responsibilities, her likes, her dislikes, her relations to other people and so forth. She is Eve Black, not Eve White, she is a different age, maybe only a child, she may be a different sex, and she has a different history. Such beliefs will be resistant to challenge and revision despite, to quote DSM-IV on *delusion*: "What almost everyone else believes and despite what constitutes incontrovertible and obvious proof or evidence to the contrary" (1994, p. 765). By anyone's standards, these periodic changes in self-ascription go well beyond, and cannot be compared to, everyday changes of states of consciousness or mood in normal people. The right comparison is with the delusional disorders.

DID is not the only disorder that affects one's beliefs about the nature and identity of self and others. A number of monothematic delusions, as well as some of the delusions found in schizophrenia, do seem to center on matters of identity and personality. The sufferers have false beliefs about their own or others' identities, and often an altered set of attitudes, dispositions, commitments, and so forth in relation to this change. In the Capgras delusion, for example, the agent believes that a close relative, commonly a spouse or parent, has been replaced by a duplicate. They admit that the person targeted resembles the loved one in every way but claim that it is not in fact *their* loved one. In the Fregoli delusion, by contrast, sufferers believe that familiar people in disguise are following them. Patients suffering from such paranoid delusions typically have an integrated set of unsupported beliefs and attitudes about the person they believe is following or persecuting them which may present a kind of mirror image to the DID sufferers' integrated delusive beliefs about the person they are, their identity, projects, traits, motivations, etc., when they are in an alter state.

Schizophrenic patients when delusional may believe, say, that they are secret agents or they may be confused about their identity in relation to others or about the identity of the thoughts in their head. In the case of the delusion of thought insertion they are aware that thoughts are occurring in their mind but believe that these are not *their* thoughts; they are the projected thoughts of another person. These thoughts lack "my-ness" just as the thoughts of an alter, even when known to the host through the phenomenon of co-consciousness or "looking on," lack my-ness for the host [11]. Questions surrounding identity do not have a secure or enduring answer for many schizophrenics. It is notable, however, that in all these cases, as in

cases of DID, the bulk of the patients' everyday working knowledge of the world and their non-experiential memory seems unaffected by the presence of the delusion [12].

People with bipolar disorder seem to us to present a particularly useful comparison case to DID. They experience wide mood-induced variation in their self-conception that can tip over into delusions, the contents and manifestation of which have much in common with described alter states.

DSM-III (p. 259) has this to say on alters: "the individual personalities are nearly always quite discrepant and frequently seem to be opposites," and gives as instances "a quiet retiring spinster" and "flamboyant promiscuous bar habitué" (Hacking, 1991, p. 848). Ian Hacking (1991, p. 857) notes that "alters are typically stock characters with bizarre but completely unimaginative character traits, each one a stereotype or one might say TV type who readily contrasts with all the other characters." This is of course exactly what the attribute-depletion and distribution model put forward by Braude would predict. A similar narrow concentration of personality attributes is also seen in the extremes of manic depression. It was not uncommon in the past for sufferers of MPD/DID to be diagnosed as manic-depressive—a situation that perhaps continues [13]. The case of Bernice R, discussed by Ian Hacking, is noteworthy here. The case arose in an era where the emergence of multiple alters was not encouraged by therapy. Bernice had just one regular alter, or in the terminology we adopt, alter state. The description of her when in that state by the superintendent of the state hospital to which she was committed in 1926 is worth quoting.

> [T]he "Polly" personality had been in the ascendancy almost continually; … her conduct had been such that she was regarded by the committing physicians, … as one suffering from an attack of Acute Mania; … she was brought to the State Hospital "bound hand and foot to a stretcher, and was held down by a canvas restraining sheet; and … she was boisterous and noisy with much loud childish talk and swearing." (Hacking, 1991, p. 865)

Such boisterousness and lack of inhibition are common features of accounts of alter personalities and are often seen in bipolar sufferers at the onset of the manic state. The feelings of omnipotence that commonly accompany the hypomanic state can lead the bipolar sufferer to believe that they are especially talented, wealthy, important, immune to danger and to ordinary requirements and so forth: that is, they attribute to themselves qualities and characteristics which are significantly different from those they normally acknowledge as belonging to them. They too may change from being teetotalers to heavy drinkers, or from being mild mannered and accommodating to being irritable and demanding perfectionists. They may in this state present as more arrogant, outgoing, charming, promiscuous or shallow. In addition, they may suffer (fairly stock standard) delusions relating to identity [14].

> The patient asserts that he is descended from a noble family. That he is a gentleman; he calls himself a genius, the Emperor William, the Emperor of Russia, Christ, he can drive out the devil. A patient suddenly cried out on

> the street that he was the Lord God … Female patients … are leading singers, leading violinists, Queen of Bavaria, [or the] Maid of Orleans … (Mondimore, 1999, p. 12)

The changes between the normal and manic states in bipolar disorder do not seem too far removed from the typical changes between the DID sufferers' normal and alter states; Eve Black, for example, may just as well manifest as the Queen of Bavaria from the point of view of judging Eve White's culpability for her actions. Nevertheless, despite the radical and cyclical alteration of mood, personality traits and self-ascription experienced by bipolar sufferers, all their actions are ascribed to just one individual. Now of course DID and bipolar disorder have their own distinctive features and causes, but we think a comparison between them is fruitful. Indeed, none of the disorders we have mentioned in which personality changes and periodic delusions surrounding identity are apparent generate an assumption even akin to multiplicity; they do, however, almost always encourage us to suspend our attitudes of praise and blame. An interpretation of DID which does not treat alters as separate and distinct personalities, but as delusion-like states of a single person with a serious psychiatric condition, does not lend itself to the hard line conclusion on moral responsibility espoused by Sinnott-Armstrong and Behnke.

6. DID, delusion, and responsibility

The empirical features of both the DID and non-DID conditions relevant to judgments of moral responsibility are as follows:

(1) The individual with the disorder cannot control the slide into the altered state.

Francis Mondimore tells us that persons with bipolar disorder do not have control of their mood states. "People with bipolar disorder cannot 'snap out of it' … Telling a manic person to 'slow down and get a hold of yourself' is simply wishful thinking; that person is like a tractor trailer careening down a mountain highway with no brakes" (1999, pp. 240–241). Braude says of individuals with DID that while some may learn to control switching in therapeutic settings, "In forensic or everyday settings … switching may remain generally spontaneous and uncontrollable" (1996, p. 44).

(2) The altered state may be one in which the sufferers' capacity to appreciate moral reasons will be diminished due, in part, to attribute-depletion and distribution.

As Braude notes, in virtue of such depletion, the person with DID is constitutionally incapable of drawing on all the reasons for action s/he would have if s/he were fully together and so in a state to draw on all the available (moral) reasons for action. "… [T]he multiple as a whole cannot judge actions in a suitably integrated and comprehensive way" (1996, p. 51). Our speculation is that the attributional biases in reasoning thought by some writers to be a key feature of delusion might arise in

these other cases too because of this kind of depletion of the normal range of characteristics and general disintegration of self-awareness (for example, paranoid personality types show a bias towards external attributions of negative events, depressives make internal attributions of negative events) [15].

(3) Even where the person at some times is capable of accessing the moral reasons and weighing them appropriately, they are unlikely to be able to control the actions they perform while in an alter or delusional state. There may be too few of the usual psychological connections between different stages and states of the person to permit the normal operation of autonomy and of self-control over time.

With respect to (2), we think that the abnormal states which mark the various disorders are comparable and that they are clearly states which impair the agent's moral capacities. The Sinnott-Armstrong and Behnke proposal entails that the alter states be regarded *prima facie* as states in which the agent's moral capacities are not impaired. But this is either manifestly implausible, since alter states are grossly abnormal and depleted states of a person, or it implicitly buys into the Multiple Persons thesis by regarding alters as having access to the full range of attributes required to exercise moral agency. And this is just false in light of the material we have presented.

We now turn our attention to (3). If we suppose that the *host* personality does have an adequate grasp of moral considerations, to what extent can s/he be held responsible for actions performed in an alter state? Our main argument with respect to the responsibility the DID sufferer bears for actions performed while in an alter state will focus on establishing the effects of the condition on the agent's opportunities to exercise and maintain autonomy with respect to self-control over time.

7. Autonomy and future selves

We see autonomy as centrally involving the notion of self-rule [16]. Autonomous agents are agents who possess the capacity to shape, plan and direct their own lives. Through their free, considered choices, they make their lives their own. The capacity for autonomy thus makes responsible agency possible. Decisions that commit one's future self, that one expects one's self to take notice of at the appropriate time, are especially connected with the agent's autonomy. Such decisions include the kinds of commitments we make every day and they may involve either ongoing projects and relationships of varying significance, or one-off events. I arrange to meet you for a drink after work on Friday, confident that on Friday I will remember and (other things being equal) feel bound by that arrangement. Or I enroll in and pay for a course, thereby committing my future self to attend classes and hand in assignments.

Both social and moral life would be impossible if we could not quite generally, in these ways, commit our (future) selves. An individual agent who could not make such commitments for the future, who could never be confident of remembering or feeling bound by both their important and more everyday decisions and commitments, would not be an autonomous agent. David Velleman (1997, p. 45) argues

that, as autonomous agents, we must "have the power of making future directed decisions that are effective, so that we can determine today what gets done by us tomorrow." However, these future directed decisions must not get us to act simply by causing future movements of our bodies, otherwise our later selves will lack autonomy. We exercise autonomy by making choices and decisions that our future selves will buy into but maybe wouldn't otherwise have made [17].

What our autonomous commitments normally do, then, is provide our future selves with reasons of a powerful kind. After all, they are *our* commitments, decisions and so forth. This does not mean that we can never change our mind. We may come to think, upon reflection, that we made a foolish decision, or circumstances may change requiring a revision of our commitments. My mother becomes ill, so I have to break my arrangement with you. I discover that I was not cut out for engineering and realize that I only enrolled in it to please my father, so I change to philosophy. But autonomous agents need reasons to reject their prior commitments.

On any conception of moral responsibility that relies on the possession of the capacity for autonomy DID sufferers cannot count as responsible. They suffer a loss of autonomy while in an alter state, since, as a result of falling into this state, they are either incapable of remembering their prior rational decisions and commitments, or they are incapable of buying into or being appropriately affected by the reason-giving force of them, since they cannot view them as their *own* decisions. Equally seriously, since we regard this single person, whose moral standing is at stake, as being the so-called host personality, we argue that s/he suffers a loss of autonomy *prior to* the onset of the alter state for s/he cannot control either the switch to that state or what happens while s/he is in that state.

8. Self-control and moral responsibility

In this section, we argue that in ordinary situations where a loss of autonomy is feared or predicted, strategies of diachronic self-control, that is, self-control exercised across time, can be autonomy-preserving. Because these strategies are widely available, agents can be held responsible for actions performed while in a state where their autonomy is diminished, e.g. for actions performed when drunk. However, we conclude that the nature of DID makes it practically impossible to deploy these techniques.

Let's accept that sufferers cannot, in their everyday life, control or prevent the switch into the alter state. Is it really, true however, that they cannot now control what they will do in the alter state? Once sufferers understand their condition and become aware that they will at times, maybe even at fairly predictable times, fall into such a state, can they not exercise some control over what they do in that state? If they can, then they may be responsible for what they do in an alter state. That state may provide, as Sinnott-Armstrong and Behnke seem to suggest, no more of an excusing condition than many other circumstances in which I predictably ignore my past decisions.

Sometimes we have good reason to fear or predict that at time t in the future a decision that we make now, or a policy we have adopted, may not be adhered to

because of the presence of temptation or of conditions that will adversely affect our judgment. Suppose I have adopted the sensible policy of not driving after drinking at the local hotel with my friends on Friday evenings. I decide that I should let my teetotal friend drive me home. But past experience warns me that once drunk my confidence in my driving ability soars and I am likely to take the wheel. I am always full of regrets the next day and resolve never to do such a thing again—until the next time … Plainly, these are circumstances where my rational autonomous decisions have an insufficient hold on me. Nevertheless, this kind of failure will not excuse me from responsibility for drink-driving.

What can we do when we perceive this kind of future threat to our autonomy? What I need here are some explicit strategies of self-control to restrain myself from actions that I do not generally endorse and invariably regret. I need to take action *now* to ensure that I do not perform the disvalued action *then*. I need to institute specific strategies of self-control.

Diachronic self-control takes two basic forms. First, the agent may act to change his motivations, his internal circumstances we might say, so that when the time comes he no longer wants most to perform the disvalued action, whether it be having another drink and then driving or eating an extra helping of dessert. He can do this directly, say by attending Alcoholics Anonymous or Weight Watchers meetings, or indirectly, by making a public bet or commitment that would raise the cost of such actions or by promising himself a reward for sticking to his resolution. The indirect strategies work by providing him with an additional motivation for refraining from the disvalued action. Internal strategies of self-control may thus be seen as augmenting the normal operation of autonomy across time. If these strategies seem unpromising, the agent can restructure his future external circumstances to prevent the desire from arising or so as to make it more difficult to act on. So he might leave the car at home on Fridays or arrange to go to the movies and out for coffee rather than to the local hotel. Both kinds of strategies are commonly engaged in by agents seeking to preserve their future directed autonomy and they are often effective.

The issue here is whether these strategies are available to the person with DID. We think that, for the most part, they are not. Consider first the internal strategies outlined. When I exercise self-control by acting to change my motivations or by arranging incentives for my future self, it is crucial to the success of such exercises, first, that the strategies I have set in place are remembered by my future self, and second, that I then take myself to be the same person that I take myself to be now. These conditions are not met in DID. I might not remember my dutiful attendance at AA while in an alter state; if I do know of it I may regard it as having nothing to do with me [18]. And the promise of a new book or a trip to the opera are unlikely to cut much ice in my hard drinking barfly alter state. Even if I am aware that in an alter state I am prone to drink-driving, here and now I may not be able to promise myself the very rewards that might work then, e.g. a bottle of scotch and a night on the town if I refrain from drink-driving, since these too are things I don't want myself to do.

Now some of the mechanisms of diachronic self-control might seem, in contrast

to the ordinary influence our autonomous decisions have on us, to aim at *binding* one's future self, at causing by, as it were, remote control, the actions we aim to perform then, rather than providing reasons to and for one's future self. Leaving my car at home to prevent myself from drink-driving is a case in point. Perhaps, then, a model of self-control along these lines is appropriate for the kinds of psychological separation we see in cases of DID. Perhaps sufferers can exercise self-control across time even if they know that the later, alter self they are attempting to control would not respect their policies and decisions. And if sufferers can exercise such control but do not, it may be that they are properly held responsible.

Two issues arise in this connection. First, if the sufferer were capable of some of the techniques of remote self-control, would s/he then count as acting autonomously, and so responsibly? Second, is s/he really capable of exercising the techniques?

On the first point, we note that in the ordinary case techniques of remote self-control can fail to be autonomy-enhancing since they may bind our future selves against our (then) better judgment [19]. However, the circumstances of DID are plainly extraordinary, and the binding of someone in a dangerous or disturbed alter state might not be regarded as posing a threat to the sufferers' autonomy, if when in that state, s/he is not capable of accessing and appreciating the available moral reasons. So perhaps the exercise of remote self-control in the case of DID would be an exception to the norm and involve the enhancement of autonomy. Nevertheless, this point loses much of its force once it is recognized that DID sufferers are largely incapable of even this form of control.

The reasons for this incapacity should be evident. Resolutions to, for example, leave my car keys or my wallet at home will have little impact on whether I drive drunk or not unless the switch into an alter state takes place *after* I have set out to meet my friends. The time lock on the liquor cabinet strategy can work but this is a form of self-control that is limited in availability and easily circumvented. Perhaps the time lock on the liquor cabinet works for me because I am usually averse to paying the higher prices that the late opening licensee down the road charges. But in an alter state I may not be at all averse to his prices and so the strategy will fail.

In general, it is necessary to (though not sufficient for) the success of the strategies of diachronic self-control that we have outlined that an individual thinks of himself or herself as being the same person, with all that that implies, at the different times. This condition for success is not met in cases of DID.

9. The responsibilities of sufferers with DID

Though we deny that sufferers can exercise autonomy and control over the actions they perform in the alter state, and so deny that they should bear criminal responsibility for those actions, we do think that they have some responsibilities with regard to their condition, once they know about it, and that they acquire some moral responsibility for the actions performed while affected if they have not discharged these responsibilities, at least on the assumption that the host personality is not so depleted as to be unable to appreciate moral reasons.

Sufferers need to acknowledge that there is a sense in which the actions they perform while in the alter state are their own actions. It is the same sense of ownership we might feel upon being informed of the actions apparently done while we were sleepwalking [20]. Even if DID sufferers are not morally responsible for the bad actions they perform in an alter state it is fitting that they should feel some measure of concern and responsibility for those actions, in much the same way as we expect the truck driver to feel particularly bad about the death of a child who ran into his path even though he could not have helped hitting the child. This is what Williams (1981) calls "agent-regret." Williams notes that the sentiment of agent-regret is not restricted to voluntary actions. Though we feel sorry for the truck driver, "That sentiment … presupposes, that there is something special about his relation to this happening, something which cannot merely be eliminated by the consideration that it was not his fault" (Williams, 1981, pp. 27–28). The point is that neither the truck driver nor the DID sufferer is in the same position, morally speaking, as a bystander: we will find it disturbing if either moves too quickly to distance or excuse themselves.

What follows from this observation? Insofar as sufferers have some understanding of their condition and of the undermining effects the condition has on their autonomy, they have a responsibility to seek to restore their autonomy by seeking treatment for their condition. Insofar as patients acquire, through treatment, a degree of co-consciousness, or by any other means, some knowledge of the features and characteristics they display in the alter state or of the intentions and projects they have while in that state, they have a responsibility to try to prevent any projected bad acts by reporting the intention, warning potential victims, or to the best of their ability, making sure that they cannot perform them—say, by checking in to hospital. Of course, their attempts may be ineffective. But the attempts must be made or they will bear some responsibility for those of their bad acts that they might have prevented. This is implied by our view that there is just one person whose acts (and character) are to be evaluated. Whether or not their attempts are effective, they may take on the burden of explanation and perhaps of some kind of apology or recompense to those adversely affected. In Williams' words, "An agent conscious that he was unintentionally responsible … might still feel that he should do something, not necessarily because he could actually compensate … but because … his actions might have some reparative significance other than compensation" (1981, p. 29). How we judge the character of sufferers may ultimately depend on their response to knowledge of those bad actions that they were unable to prevent.

10. Conclusion

The causes of DID do not create extra persons, but rather have the effect of fragmenting an existing one. In this respect, it is like other serious mental illnesses. We hope elsewhere to explain and defend this idea in more detail. In this paper, we have largely presupposed it, and once that stance is taken, the moral implications turn out to be far more complex and interesting than the implications of the

Multiple Persons thesis. No special moral theoretical difficulty is posed by the Multiple Persons thesis for questions of moral responsibility, just practical difficulties along similar lines to punishing one of the persons and not the other in a criminal case involving Siamese twins. The peculiarity of the situation is that the metaphysically interesting case of multiple persons is the morally simple one, and vice versa. But then this is hardly surprising once we recognize the way DID undermines the autonomy and control of persons over the different aspects of their lives. Such gross interference is morally disabling, and so we think DID hardly leaves even a single "respectable" or responsible person, let alone more than one, as Carroll's Alice aptly observed.

Acknowledgements

Versions of this paper have been read at the *Time and Ethics* conference, Otago University, February 2001, and at the Centre for Applied Philosophy and Public Ethics, Australian National University, Canberra, April 2001. We thank audiences at these talks for useful input, as well as the action editor of *Philosophical Psychology* and a referee for their helpful comments.

Notes

[1] We will in general adopt the acronym "DID" to refer to the condition, unless it is historically apt to adopt its former name.

[2] We support the change to DID, and in a later section will rehearse some of the arguments which motivate it. To have MPD presupposes in the minds of many that a single body may contain literally multiple persons. However, we think that a cautious approach is needed, especially in a legal context, to such a radical ontological shift. One of the most sustained and careful accounts of the skeptical approach to this issue can be found in Spanos (1996).

[3] There are three possible answers to the question "How many persons?"—none; one; or, two or more. So, for completeness, we really should mention a third position. According to an *eliminativist* position, alter personalities are not persons because there are no persons *simpliciter*, just bundles of experiences; it is for this reason that individuals with DID cannot be held responsible for the crimes of the alters that share their bodies. The eliminativist position is about as extreme as one can be on questions of personhood and moral responsibility over time. If it is true, then it makes redundant any account we might put forward concerning the latter, but for different fundamentally skeptical reasons. Hume (1978/1739, p. 252) defended this idea; also, one interpretation of Derek Parfit's (1984) work is that there are no persons ontologically speaking, even if a necessary presupposition of conventional life is that there are.

[4] The conditions for moral responsibility we put forward are quite general, and are designed to clarify some of the problems and complications involving DID cases now emerging, particularly in the legal sphere. Of course, we do not suggest that the considerations we adduce necessarily apply to all defendants *claiming* DID whose cases require careful individual scrutiny.

[5] See Eisen (1989) for description of such a case.

[6] Braude (1995, p. 164) argues similarly that since there is significant overlap of basic personal skills we are entitled to infer the existence of a single underlying entity. Hacking (1995, p. 229) has responded by pointing out that such an inference may be premature. Perhaps for each basic skill there is a "subsystem" or "module" preserved through the transition from personality to personality. The different personalities are like changing "heads of state" that work collaboratively with these preserved subsystems. Perhaps this is so, but our model is less profligate, and we think more

plausible. The maintenance of abilities and retention of basic knowledge requires the preservation of information over time. We think it makes more sense to think of this information as continuing to operate, or to be available, in the same person, in contrast to the idea of its being transferred and newly presented to a different personality, a decidedly more cumbersome process than the ordinary one we assume to be true.

[7] *U.S. v. Denny-Shaffer*, 2 F.3d 999 (10th Cir. 1993) at 1016. NB. Quoted in Sinnott-Armstrong and Behnke (2000, p. 302).

[8] We agree, then, on the Single Person thesis. However, we strongly disagree with their method of argumentation for reaching this thesis. One of us elsewhere has put the case against this method.

[9] *Sybil*, directed by Daniel Petrie, produced by Lee Rich (Lorimar Productions), and based on the book by Flora Rheta Schreiber (1973).

[10] Hacking (1995, p. 18).

[11] See Campbell (1982) and Stephens and Graham (2000) for discussion of thought insertion in schizophrenia.

[12] An excellent source for information on the Capgras delusion, Fregoli delusion, schizophrenia, and other disorders in relation to delusion can be found in a special issue of *Mind and Language*, *15*, 2000. There has been a renewed attempt to develop a coherent theory of delusion (see Davies & Coltheart, 2000). Earlier attempts include Bentall *et al.* (1991), Garety and Freeman (1999), and Frith (1992).

[13] DSM-IV notes that there is controversy over differential diagnoses between DID and "a variety of other mental disorders including schizophrenia and other psychotic disorders, bipolar disorder with rapid cycling, anxiety disorders, somatization disorders, and the personality disorders" (p. 487). This controversy is related to skepticism about whether the condition as described is a distinct disorder at all and so whether all alleged DID sufferers are either frauds or suffering from some other psychiatric condition. We think DID is a distinct disorder despite having features in common with other psychiatric disorders but we cannot address this skeptical issue here.

[14] The cultural imprimatur on the content of delusions and the explanation patients might give of their experiences is readily apparent in other delusions that center on issues of identity and personality, including bipolar disorder, paranoid schizophrenia and the Capgras delusion.

[15] For example, see Coltheart and Davies (2000).

[16] See Dworkin (1988, p. 6) for a discussion of the ways in which the term has been applied and Kennett (2001, pp. 129–133) for a discussion of the limits of this account of autonomy for understanding moral agency.

[17] See Velleman (1997, pp. 45–48).

[18] Of course, in the alter state I may still have access to the semantic information that drinking and driving is against the law. This would be enough to convince Sinnott-Armstrong and Behnke that I should be held responsible for the actions performed in that state. Our argument here is that the alter state itself prevents the agent, that is, the host, from exercising effective self-control over herself while in that state.

[19] See Velleman (1997) and Kennett (2001, pp. 148–149).

[20] Inability to recall important information that goes beyond ordinary forgetfulness forms part of the criteria for diagnosis of DID so the comparison with the sleepwalker is apt. But even supposing that there were no amnesic barriers between the different states, our claim that the responsibility of the sufferer is mitigated still stands so long as she (the host) is unable, for the reasons we have given, to control what happens in the alter state.

References

AMERICAN PSYCHIATRIC ASSOCIATION. (1987). *Diagnostic and statistical manual of mental disorders* (DSM-III-R). Washington, DC: American Psychiatric Association.

AMERICAN PSYCHIATRIC ASSOCIATION. (1994). *Diagnostic and statistical manual of mental disorders* (DSM-IV). Washington, DC: American Psychiatric Association.

BENTALL, R.P., KANEY, S. & DEWEY, M.E. (1991). Persecutory delusions: an attribution theory analysis. *British Journal of Clinical Psychology*, *30*, 13–23.

BRAUDE, S. (1995). *First person plural: multiple personality and the philosophy of mind*. London: Routledge.

BRAUDE, S. (1996). Multiple personality and moral responsibility. *Philosophy, Psychiatry, and Psychology*, *3*, 37–54.

CAMPBELL, J. (1982). Schizophrenia, the space of reasons, and thinking as a motor process. *Monist*, *82*, 609–625.

DAVIES, M. & COLTHEART, M. (2000). Introduction. In M. DAVIES & M. COLTHEART (Eds) *Pathologies of belief*. Oxford: Blackwell.

DWORKIN, G. (1988). *The theory and practice of autonomy*. Cambridge: Cambridge University Press.

EISEN, M.R. (1989). Return of the repressed: hypnoanalysis of a case of total amnesia. *International Journal of Clinical and Experimental Hypnosis*, *37*, 107–119.

FRITH, C.D. (1992). *The cognitive neuropsychology of schizophrenia*. Hillsdale, NJ: Erlbaum.

GARETY, P.A. & FREEMAN, D. (1999). Cognitive approaches to delusions: a critical review of theories and evidence. *British Journal of Clinical Psychology*, *38*, 113–154.

HACKING, I. (1991). Two souls in one body. *Critical Inquiry*, *17*, 838–867.

HACKING, I. (1995). *Rewriting the soul: multiple personality and the sciences of memory*. Princeton, NJ: Princeton University Press.

HUME, D. (1978/1739). *A treatise of human nature*. Oxford: Clarendon Press.

KENNETT, J. (2001). *Agency and responsibility: a commonsense moral psychology*. Oxford: Clarendon Press.

MONDIMORE, F.M. (1999). *Bipolar disorder: a guide for patients and families*. Baltimore, MD: John Hopkins University Press.

O'KELLY, L.I. & MACKLESS, F.A. (1956). *An introduction to psychopathology*. London: Random House.

PARFIT, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.

SCHREIBER, F.R. (1973). *Sybil*. Chicago: Regenery.

SINNOTT-ARMSTRONG, W. & BEHNKE, S. (2000). Responsibility in cases of multiple personality disorder. In J.E. TOMBERLIN (Ed.) *Action and Freedom, Philosophical Perspectives 14*. Malden, MA: Blackwell.

SPANOS, N.P. (1996). *Multiple identities and false memories: a sociocognitive perspective*. Washington, DC: American Psychological Association.

STEPHENS, G.L. & GRAHAM, G. (2000). *When self-consciousness breaks: alien voices and inserted thoughts*. Cambridge, MA: MIT Press.

VELLEMAN, J.D. (1997). Deciding how to decide. In G. CULLITY & B. GAUT (Eds) *Ethics and practical reason* (pp. 29–52). Oxford: Clarendon Press.

WILLIAMS, B. (1981). *Moral luck: philosophical papers 1973–1980*. Cambridge: Cambridge University Press.