

AUTONOMY AND THE EMOTIONS

CHRISTINE TAPPOLET

Université de Montréal

ABSTRACT

Can actions caused by emotions be free and autonomous? The rationalist conception of autonomy denies this. Only actions done in the light of reflective choices can be autonomous and hence free. I argue that the rationalist conception does not make room for akratic actions, that is, free and intentional actions performed against the agent's best judgement. I then develop an account inspired by Harry Frankfurt and David Shoemaker, according to which an action is autonomous when it is determined by the agent's most central cares, where cares are defined in terms of emotional dispositions.

Keywords: freedom, autonomy, emotions, care, weakness of will

The emotions we feel determine a great many of our actions. The fear I experience when being threatened by a bear can cause me to run away; the love I feel for a friend can make me give her a present; compassion can prompt me to help a blind person to cross the road. In certain cases, emotions interfere with free agency. For example, an agoraphobic whose fear makes it impossible for him to leave his home can hardly be said to be a free agent. A traditional question is whether all the actions caused by the experience of an emotion are like this or whether they can be, and even sometimes are, free.

This question is related to issues about moral responsibility and blame. It is natural to think that if an action lacks freedom, then we cannot be held morally responsible for it. If the agoraphobic's failure to go to an appointment is not free, it is not clear that he can be considered morally responsible for not having done so. The fault seems to lie not with him but with his condition. Also, it is far from clear that we should blame the agoraphobic for missing his appointment.

Again, it rather seems that it is his condition that is to blame. Quite generally, it is often considered that a crime committed under the influence of strong emotions or passion – what is called “crime passionnel” in French – is less blameworthy, morally and legally, than cold-blooded crime.

To answer the question whether or not actions caused by emotions can be free it is necessary to have a better idea of what free action requires. As Gary Watson (1987) underlines, freedom issues have traditionally been divided into two according to the scholastic distinction between *liberty of indifference* and *liberty of spontaneity*, a distinction which plays an important role in David Hume’s thinking (T, 2.3.2). These two questions are a) that of the availability of alternative possibilities and b) that of autonomy (or self-determination). Any conception of free action has to account for the fact that an action appears to be free if and only if it is true a) that the agent could have done otherwise, and b) that it is really the agent himself who is the source of the action.

It could be thought that a quick argument gets one to the conclusion that actions caused by emotions lack freedom. Agents who act under the spell of emotions cannot do otherwise, so how could they be free? This argument, however, is far from convincing. It is plausible that most emotions only have an impact on the kind of options that are considered by an agent. In particular, negative emotions such as anger, fear or jealousy appear to reduce the number of options that are considered, while positive emotions such as joy seem to have the opposite effect. As such, this is something that does not constitute a threat to free agency. Such emotions simply direct the agent’s attention toward a set of options and make them salient, without fully determining the exact course of action, so that choice remains necessary. However, suppose that our emotions fully determine what actions are performed. This does not yet mean that the actions in question lack freedom. As a compatibilist would underline, the absence of alternative options does not entail that the action is unfree. According to this admittedly modest conception of freedom, it is sufficient for free agency that the agent would have acted differently had his motivation been different. The compatibilist can thus say that although the action of the agent is fully determined by an emotion, so that he could not have acted differently given this emotion, it is nonetheless true that had the emotion been different, or had the agent not had this emotion, he would have acted differently. Since I want to concentrate on the question whether emotions threaten free agency, and not on the question whether free agency as such is possible, I shall assume that a compatibilist account of this kind is true.

What about the autonomy issue? Quite generally, one can contrast two radically different conceptions of autonomy, which come with radically different verdicts as to whether emotions and free agency are compatible or not. According to the view that is most common in philosophy, and that one could call the *rationalist conception*, emotions as such do not allow for autonomous actions. Emotions are claimed to be passive states, which are external to our essence as rational agents.

In a particularly telling passage, Laura Waddell Ekstrom wonders why one should think that our passions, which she considers on a par with whims and impulses, do not allow us to act autonomously. She worries that the rationalist conception might be too intellectualist. Here is her answer: “In my view, we do not act autonomously in acting on passions, whims, and impulses because these overtake us; we are generally passive with respect to them, they do not engage our understanding or capacity for reflective evaluation. We make our lives more our own by examining such impulses and by acting in accordance with our evaluations” (2005, p. 160).

The intuition at the heart of the rationalist conception is that only actions done in the light of a choice that follows the verdict of deliberation could be autonomous and hence free. It is thanks to reflection, a process that results in value judgements capable of justifying our actions, that we are able to determine our actions ourselves. That is why we could summarize the rationalist conception by the slogan: the agent is constituted by his reason.

The rationalist conception of autonomy corresponds to what Stefaan Cuypers (2000) calls the ‘Platonic’ intuition, which he relates to the tradition of the Enlightenment and according to which autonomy depends on lucidity and reflectiveness. As Cuypers underlines it, there is another conception of autonomy that corresponds to the intuition, which he characterises as ‘Nietzschnean’ and which goes back to the Romantic tradition. According to this second conception, reflection is a threat to authenticity and autonomy. Given the traditional opposition between reason and emotions, it is natural to maintain that, for the Romantic conception, emotions are central to autonomy. An attractive idea is that the emotions we experience are intimately related to who we are, for they reveal what we care for. According to this suggestion, the different emotional dispositions we have would determine who we are. Thus, the emotions would constitute the very condition of autonomy.

I shall start with a general objection against the rationalist conception of autonomy, namely, its incompatibility with weakness of the will. I shall then discuss the view defended by David Shoemaker, Stefaan Cuypers and Harry Frankfurt, according to which our cares are essential to our autonomy. I argue for a duly pruned-down version of this view, which gives up the claim that we have to endorse or recognize our cares, as well as the claim that such cares produce volitional necessity.

1. An objection against the rationalist conception of autonomy

According to what is often considered to be the paradigmatic rationalist conception of autonomy, namely Gary Watson’s account (1975), autonomous actions are actions that are justified by value judgements. As Watson writes: “Only evaluations can give one *reasons* to oppose first-order desires, and when and only when agents’ behaviour

expresses their evaluations are they sources and ‘authors’ of (because they ‘authorized’) their behaviour.” (1987, p. 149) Using Susan Wolf’s (1993) notion of the real self, that is, the part of the self that has to determine an action for it to be autonomous, here is how one could summarize Watson’s account:

(1) A motivation is part of the real self of an agent iff it corresponds to the agent’s evaluation.

As Watson himself has underlined (1987), there are a number of problems with this suggestion. The problem I would like to underline, however, is that this account produces counter-intuitive results with respect to *akratic* actions; in other words, actions which manifest weakness of the will. Such actions are performed against the better judgment of the agent, such as when I light a pipe and start smoking it while judging that, all things considered, it would be better not to smoke this pipe, since I am convinced it would be better to stop smoking altogether. It is generally acknowledged that an agent who acts against his judgement can be free. Indeed, akratic action is often defined as a free and intentional action, which is performed in spite of the judgement that, all things considered, another action would be better.¹ However, according to Watson’s account, such an action cannot be free: it clearly lacks autonomy. The reason for this lack of autonomy is simply that the motivation that produces the action is not part of the real self of the agent since it opposes the agent’s judgement.²

The question whether akratic actions can be free or not is controversial. On both sides of the debate, however, it is agreed that there are at least apparent cases of free and intentional actions that are performed in spite of the judgement that all things considered another course of action would have been better. In the absence of strong reasons to deny that they are possible, it would seem that the possibility of such actions has to be acknowledged.

Here is what could be considered such a reason. It could be argued that there is a good reason to think that akratic actions lack autonomy, and hence that they would also lack freedom. Is it not obvious, after all, that akratic agents lack self-control? And is not self-control central to autonomy? This is the intuition that is behind Alfred Mele’s claim that autonomy is essentially the absence of weakness of will (Mele 1995).³ However, contrary to what is widely accepted, it is important to distinguish self-control and autonomy. Self-control, or at least self-control as usually understood, is control exercised

¹ See Davidson 1963 and Mele 1987.

² Since Watson also considers that it is not possible to freely act against one’s better judgement (1977), this objection should not worry him. In a nutshell, the argument he gives is that akratic actions cannot be distinguished from compulsive actions, for one cannot choose to act against one’s better judgement. So, it has to be concluded that the agent was unable to resist the rebellious desire. The problem is that it is far from clear that choice has to follow the agent’s better judgement. Note that Watson nonetheless seems to consider akrasia to be a problem for his account of autonomy (1987).

³ Given that Mele (1987) also claims that akratic actions can be free, it seems he has to deny that free agency presupposes autonomy. But see Mele 2002 for the claim that akratic actions can be both free and autonomous.

by the agent's reason over his desires and other first-order appetites. Self-control is hence the opposite of weakness of will. An agent exercising self-control acts according to his value judgements. By contrast, autonomy is the determination of the action by what really constitutes the agent, i.e. by his real self. Whether it is the opposite of weakness of will depends on what the real self of the agent is considered to be.

To assimilate self-control to autonomy is to assume that value judgements of an agent have to rule over his other mental states. This rationalist assumption should not be simply taken as a fact. There are good reasons to question it. To see this, it is useful to consider the well-known example of Mark Twain's character, Huckleberry Finn.⁴ At one point during Huck's journey down the Mississippi River, he has the opportunity to give his friend Jim, a runaway slave, back to his owner. Although he thinks that this is what he should do, his conscience prevents him doing so. This example has been used to illustrate the claim that it can be more rational to act against one's better judgement than to follow this judgement⁵. According to the usual reconstruction of the case, Huck judges that all things considered he should give Jim back to his owners, but he nonetheless decides not to do it. Although Huck's judgement is the outcome of his deliberation, it is easy to imagine that his decision is more rational than the one to give Jim back to his owners. One just has to suppose that there are different considerations that Huck would have taken into account if he had thought harder about the case – after all, Jim was his friend, Jim trusted him, and Jim desired to be free. Thus, the judgement that Huck arrived at was hasty. Given this, it would not be particularly rational to act upon it. On the contrary, it would have been more rational not to follow such a judgement. Once it is accepted that Huck's action is more rational than the one that would have followed his better judgement, even a rationalist could accept that the action is also autonomous. As we shall soon see, however, other cases show that autonomy and rationality can part company.

This intuition is confirmed by the fact that, as Arpaly and Schroeder (1999) note, it is natural not only to praise Huck's action, but Huck himself. Huck might not be as admirable as someone who overtly condemns slavery, but it seems clear that Huck has a better character than someone who would not have suffered from weakness of will and who would have turned Jim in. As Arpaly and Schroeder write: "One does not only think helping Jim was a good thing for Huckleberry to do, but regards Huckleberry's action as saying something important about Huckleberry's self. The reader of Twain's novel tends to see Huckleberry's action not as an accidental good deed done by a bad boy, but as indicative of the fact that Huckleberry is, in an important sense, a good boy, a boy with his heart in the right place" (1999, p. 163). However, on Watson's account, one would have to conclude that Huck was not really himself when he failed to turn Jim in. Given his better judgement, one would have to conclude that Huck's real self was in

⁴ See Bennett 1974; McIntyre 1993; and especially Arpaly and Schroeder 1999.

⁵ See McIntyre 1993.

favour of slavery. His action would be praiseworthy, but he himself would not. According to this account, Huck's action is not performed by his real self. So, it is hard to see how Huck could be praiseworthy. At best, one can welcome his lack of control, given the doubtful ideal to which he sticks.

Now, as Arpaly and Schroeder underline, this seems to be a general problem for hierarchical theories of the real self, whether they invoke higher-order desires (Frankfurt 1971, Stump 1988), value judgements (Watson 1975), preferences concerning one's first-order desires (Lehrer 1990), the desire that one's desire be rational (Velleman 1989) or else policies concerning which desires have to count as reasons (Bratman 1999). The problem is that all these theories subscribe to rationalism. They divide the self into (contemporary equivalents) of reason and appetites and they identify the self with reason. According to such theories, it is always reason (or its contemporary substitute, such as higher-order desires, value judgements, etc.) which wins the day when there is conflict between reason and appetite. However, as Arpaly and Schroeder write, "it is clear that any Real Self theory that identifies the agent's Real Self with her Reason rather than her Appetite, regardless of technical details, is not going to do justice to cases such as Huckleberry Finn's." (1999, p. 170) For what can be considered to be Huck's rational part tells him to turn Jim in, while it is his appetite that prevents him doing so. In so far as Huck's real self is considered to be (a contemporary equivalent of) his reason, Huck cannot be praiseworthy for his action. This seems plainly false.

Now, we have seen that Huck's action is more rational than the action that he would have performed had he followed his better judgement. Therefore, it might be thought that what is important for autonomous action is that the action reflects the reasons the agent has, whether he acts on his better judgement or not. The intuition behind this thought is that by reflecting the reason an agent has, the action would not merely be the effect of some non-rational cause; something which can be thought incompatible with free agency. As cases of irrational akrasia show, however, this will not do. If one allows for free action, performed against a better judgement that reflects the agent's reasons, it has to be agreed that one can act autonomously while acting irrationally. One can be oneself while being irrational.

If one thinks of the fact that Huck is motivated by the emotions he feels, such as sympathy and love, it is tempting to think that autonomy can side with emotions. Let us thus turn to a radically different account of the real self, according to which emotions are at the heart of autonomy.

2. The care account of autonomy

The emotions we feel seem to have a particularly close connection to who we are. This is the idea that is at the core of David Shoemaker's account of autonomy. According to Shoemaker, our emotions constitute our real selves: "(...) the emotions we have make

us the agents we are” (2003, p. 94). The notion that is central to Shoemaker’s account is that of care. This notion is taken from Harry Frankfurt’s later work (1982, 1994). According to Frankfurt, care (or love) has less to do with feeling and thought than with the presence of a certain motivational structure, which shapes an agent’s preferences and guides and limits his actions (1982, 1994)⁶. By contrast, Shoemaker considers that cares are mainly emotional. Cares involve *conative* elements, for at least most emotions involve motivations. Cares also often, but not necessarily, involve evaluative elements. What is important to underline is that cares are not a kind of emotion, for as Shoemaker puts it: “Talk of caring is simply a way of referring to the range of emotional reactions one is expected to have with respect to the fortunes of the cared-for object” (2003, p. 94). To care for someone, or for something, consists in being disposed to undergo a certain number of emotions, depending on the good or bad fortune of whom or what one cares for. We feel sadness when things go bad, joy when things go well, hope that things will go well, fear that things will go bad, and so forth. By identifying what an agent cares for, it is possible to explain his emotional reactions.

According to Shoemaker (and in this he follows Frankfurt (1994)) our real self is constituted by our cares. However, not any care will do. They have to be irresistible. Let us look for a moment at Frankfurt’s own account. According to Frankfurt “(o)ur essential natures as individuals are constituted (...) by what we cannot help caring about. The necessities of love, and their relative order or intensity, define our volitional boundaries. They mark our volitional limits, and thus they delineate our shapes as persons”⁷ (1994, p. 138). These irresistible cares define what Frankfurt calls “volitional necessities”, that is, motivations that force a person to act in a certain way. Paradoxically, these motivations do not constitute obstacles to our freedom. Given that they are supposed to define the real self, irresistible cares make autonomous action possible. Here is how Frankfurt formulates this claim: “A person acts autonomously only when his volitions derive from the essential character of his will. (...) the unconditional commands of love are not, as Kant suggests, adventitious elements of a person’s will. They are essentially integral to it, for what a person loves is a defining element of his volitional nature. When he acts out of love, accordingly, his volitions do derive from the essential character of his will” (1994, p. 132).

Cares (or love, to take Frankfurt’s other term) differ from compulsions in that the latter are irresistible, but foreign, forces. The reason for this is that the agent does not identify himself with these forces (1994, p. 136). According to Frankfurt, it is thus necessary to endorse one’s cares, that is, to wholeheartedly identify with them, for autonomy to be possible. It follows from Frankfurt’s account that the agent who acts on a volitional necessity fully stands behind his action: he is wholehearted, to use Frankfurt’s expression.

⁶ See also Cuypers 2000 for the claim that caring is central to autonomy.

⁷ As Velleman 2002 argues, it is far from obvious that we have such volitional essences.

Wholeheartedness is thus necessary for free and autonomous action.⁸ The example Frankfurt uses to illustrate this claim is that of Martin Luther and his claim “Here I stand. I can do no other” (1982, p. 86. See also Dennett 1984, p. 133). Luther’s care for the idea of the Reformation made it impossible for him to make any other choice.⁹

What is important to note is that the identification with the volitional necessities which derive from our cares is not optional. Thus, an agent is unable not to identify with what is required by his cares. As Velleman notes, Frankfurt requires “not only that the agent endorses the motive constraining his will but that he be unable to help endorsing it. In such a case, the agent has a second-order inability: the inability to will any change in his inability to will” (2001, p. 334). In fact, Frankfurt claims that what happens if the agent modifies his cares is that he becomes a different person (Frankfurt 1994).

Just as Frankfurt does, Shoemaker claims that the cares that define the agent produce volitional necessities. And he also subscribes to the thesis that we are free and autonomous when we act on the basis of such volitional necessities: “(...) if free agency is a matter of my doing what is ultimately dependent on my cares, then free agency is volitional necessity. I act freely in performing X insofar as I cannot bring myself to do anything other than X” (2003, p. 106). The claim is not only that autonomy is possible when we are volitionally necessitated, but the stronger claim that autonomy *requires* volitional necessity.

By contrast with Frankfurt, Shoemaker does not require that the agent identifies with or endorses his cares. However, Shoemaker claims that the agent needs to be reflectively aware of his cares. An agent has to recognise having the cares in question. Shoemaker writes: “I am identified with those elements flowing from my cares, that is, what is authoritative in self-determination are those elements dependent on the emotional investments *I recognize myself to have* (2003, p. 112, my italics). Shoemaker praises the freedom of persons who are entirely conscious of their cares and who do not have to make choices, since their cares determine what they will do.

A point that needs to be underlined is that we usually have different, sometimes competing cares. Shoemaker thus claims that freedom and autonomy depend on acting on one’s strongest care: “(i)f we attach free agency to willing action, then, and willing action consists in the action I genuinely want to do, and what I genuinely want to do depends on what I care most about in any particular situation, then free agency is grounded in

⁸ As Velleman (2006) argues, there are reasons to think that the ideal of wholeheartedness is doubtful, since it often involves the unhealthy repression of motivations that should be considered part of ourselves.

⁹ Frankfurt seems to allow that a wholehearted agent can have conflicting motivations, but what happens in such cases is that the desires that opposes what the agent cares for gets segregated: “Wholeheartedness does not require that a person be altogether untroubled by inner opposition to his will. It just requires that, with respect to any such conflict, he himself be fully resolved. This means that he must be resolutely on the side of one of the forces struggling within him and not on the side of any other.” The question is whether one can still speak of conflict when the agent has resolutely taken side. (1999, p. 100)

care. To the extent that what I do does not ultimately depend on my strongest care(s) at the time of action, I am unfree” (2003, pp. 103-104).

Putting these different claims together, we obtain the following definition:

(2) A motivation is part of the real self of an agent at the time of action iff it results from a care, which the agent either endorses or recognises himself to have, which produces volitional necessity, and which is the strongest one at the time of action.

I think there is a grain of truth in this account. However, as it stands, it clearly will not do. A first problem is that it would make Huck’s case less than free. Let us suppose that Huck helps Jim because he cares for him and that this is his strongest care at the time of action. It is clearly wrong to claim that Huck endorses his care: given his better judgement, he would rather get rid of any attachment to Jim. In the same way, Huck is clearly not wholehearted about his action and his motivation. Moreover, it is not clear Huck has to recognize himself as having this disposition towards Jim for him to act autonomously. In fact, one can well suppose that Huck is moved by a strong desire to help Jim without realizing that he cares for Jim. It thus seems that one can attribute an action to an agent while the agent is not aware of the care that moved him to action. This is confirmed by the fact that it is well-known that a person can love someone without realising this. We frequently fail to realise how much we care for someone before separation teaches us better; but this does not mean that before becoming aware of the fact that her feelings are engaged, the agent’s actions are not the deeds of her real self. She might not realise this, but that is neither here nor there.

Another problem comes from the requirement that the cares in questions come with volitional necessities. If the endorsement of a care is not necessary, then being unable to help endorsing it cannot be necessary either. Huck neither endorses his care for Jim, nor is he unable to help endorsing it. In the same way, it just seems false to say that Huck is unwilling to will any change in his inability to will. Given his better judgement, he surely would want to change his motivation. Moreover, it is not clear that we should say, as Shoemaker suggests, that a free agent’s motivations have to be irresistible, in the sense that he cannot bring himself to do otherwise. It is clear that Huck’s motivations to help Jim are stronger than his motivation to refrain from helping. However, it certainly makes sense to say that he could have resisted his motivation had he taken the right steps at the right time.

Thus, it seems the care account has to be stripped down to the following claim:

(3) A motivation is part of the real self of an agent at the time of action iff it results from the care that is the strongest one at the time of action.

Here are a few examples that can be considered paradigmatic cases of autonomous action according to this account: deciding to tell the truth in spite of a promise one made because what one cares about most is truthfulness; spending time with a sick friend because one is strongly attached to this friend; moving to a different city in order to

live with the person one loves. To many, the suggestion that these are paradigm cases of autonomous actions will seem deeply wrong. Acting out of attachments or cares, they will claim, are just cases of determined, non-autonomous actions. To what extent should we believe this?

Before considering the main objections to this account, let me start by underlying one of its virtues. It will be clear, given the above discussion, that one advantage of the care account is that it entails that akratic actions are autonomous. As we have seen, it is plausible to say that Huck's action is motivated by his strongest care. Had he not cared that much for Jim, he would not have hesitated to give Jim back to his owners. More generally, it is plausible to claim that each time we freely act against our better judgement, we are motivated by what we care for most. Consider the akratic smoker. In this case too, it seems true that in spite of his better judgement what the smoker cares most for is the pleasure involved in smoking his pipe.

Let us turn to the objections which can be mounted against this account. Maybe the most obvious problem is that there seem to be cases of autonomous actions that do not depend on any cares. If so, it would be just wrong to claim that our real selves are constituted by our cares. As Shoemaker (2003) notes, there are many actions that have apparently nothing to do with our emotions and our cares. Consider a standard morning: you get out of your bed, you stretch, you take your shower, you drink your tea, and so on. All these actions seem perfectly intentional and can be explained in terms of desires and beliefs, but they do not seem to relate to emotions or to cares. Shoemaker suggests that such actions are just like those of the wantons described by Frankfurt. A wanton is characterised by the fact that he has no desires about what ought to motivate him: "Not only does he pursue whatever course of action he is most strongly inclined to pursue, but he does not care which of his inclinations is the strongest" (1971, p. 17). So, according to Shoemaker, in doing such things as getting out of bed, we act as unreflective beings who do not care what our wills are. Such actions would thus lack autonomy. There is, hence, no problem in admitting that they are not grounded in cares. Is this true?

What is controversial about this reply is that Shoemaker assumes that only what could be called 'reflective' actions can be autonomous. This rationalist assumption is difficult to square with an account that puts care at the heart of autonomy. As we have seen, cases of akratic action show that actions that are not the outcome of deliberation or reflection, but are actually contrary to what we think is the better course of action, can be autonomous. Moreover, Shoemaker's reply entails that it is not the agent himself – his real self – who gets out of the bed in the morning and who stretches. Thus, such an action could not count as free. This is difficult to believe. It is true that to get out of the bed in the morning or drinking one's tea, usually does not require much thought. This, however, does not mean that the action in question is not free and hence autonomous.

I do not think that this makes for a deep problem for the care account, however. There is in fact a rather obvious suggestion that allows handling such cases. The suggestion is simply that contrary to first impressions, such actions bear a relation to our cares. The agent who more or less automatically gets up in the morning can have adopted this habit because he cares for a certain way of life. This seems confirmed by the fact that if something threatened this way of life – a sudden paralysis that would prevent the agent from moving from his bed, for instance – it is clear that this would result in emotions of frustration and even despair, thus revealing that there was a care involved. The same seems true of most habitual actions. These actions are related to habits the agent has formed because of things he consciously or unconsciously cares for.

However, is it not possible that an agent acts autonomously while motivated by reason alone? Obviously, this would have to be denied by the advocate of the care account. He has to argue that when we are motivated by the conclusion of deliberation, our motivation derives at least indirectly from some care. Shoemaker (2003) argues that the moral agent who acts out of respect for the moral law in fact cares for this law. This is certainly what seems to follow from the fact that a moral agent feels guilty when he fails to do what he thinks he ought to do. In the same way, one could point out that someone who acts against his better judgement on a prudential matter usually has negative feelings, such as guilt or shame. He thus seems to care for the prudential principles he embraces. In fact, it is plausible to argue that, at bottom, what is crucial to prudential deliberation is that the agent cares for himself.

If this is on the right lines, cares are a necessary condition for autonomy. Are they also a sufficient condition? Compulsive actions, such as that of the unwilling addict, take us to another kind of problem, that is, cases of non-autonomous actions that appear to derive from our strongest cares. The question is how an account which does not require the endorsement of care deals with compulsion. One possibility is to claim that the motivations of compulsive agents are independent from his cares. This is what Shoemaker suggests: “The desires stemming from what (the unwilling addict) reflectively cares most about here and now are impotent, and his will is thus impeded from the inside. The same goes for unwilling kleptomaniacs and others who are moved by some internal compulsion; motivational forces utterly independent from things they care about” (2003, p. 102).

The problem is that it seems difficult to deny that even the unwilling addict cares for his high. It is difficult to see why this care should not be considered to be the strongest one, given its impact on action. This is even more obvious when one thinks of the fact that the addict will experience a certain number of strong emotions, depending on whether he obtains the drug he craves for or not. In fact, Shoemaker acknowledges that the addict cares to have his drug, but he claims that the reason he does is that this would allow him to get rid of his desire. What the unwilling addict in fact cares for is to get rid of his desires. The care of the addict would motivate the agent to take his drug only in so far

as this is instrumental to getting rid of the desire. The unwilling addict would thus have an instrumental desire to take the drug, based on his care to get rid of the drug-desire; but he would not care to have the drug as such. Thus, one could say that the desire to have the drug does not belong to the real self of the agent.

This story is difficult to believe. The suggestion that the unwilling addict does not care for the immediate and pleasurable affects of his “high”, but only tries to satisfy a desire he would prefer not to have, seems doubtful. It would seem more plausible to acknowledge that different cares are competing. On the one hand, there is the care for the drug, which is in fact a care for the pleasure of the high, and on the other hand there is the care for a life free of drugs. The first care wins the day. Does this mean that the unwilling addict’s action is autonomous after all?

In order to claim that the unwilling addict is not moved by his real self, one has to claim that the care for a life without drugs is the strongest care and thus defines the real self, though the motivation that comes with the care for the drug is stronger. How could it be that the strongest care is not the one that comes with the strongest motivation? What one would intuitively want to say is that what is necessary for a care to constitute the real self is that it be central to the personality of the agent, whatever its causal efficacy with respect to a particular action or type of action. What could this mean?

The suggestion I would like to make, and which is inspired by Nomy Arpaly and Tim Schroeder (1999), is that the centrality of a care depends on the integration of this care with the other mental states of the agent. This is what Arpaly and Schroeder, who explain the notion of integration in terms of the concept of depth, say about of beliefs and desires: “A belief or desire is deep insofar as it is a powerful force in determining the actor’s behavior, deeply held, deep-rooted. Deep beliefs tend to resist revision (...) and deep desires tend to be satisfied with preference over shallower desires (...)” (1999, p. 173). The integration of a care manifests itself by its effect on motivation, but also by the number of causal links between the care and other mental states of the agent. The more a care is central, the more it will determine what the agent desires and what he thinks and feels. So, one particular care could have more causal links to what an agent thinks and what he feels than another care, but be less powerful than this second care with respect to a particular motivation. In other words, the advantage of introducing the distinction between the *centrality* and the *motivational strength* of cares is that one can claim that a motivation, which depends on a care that is on the periphery of the agent’s psychology can be motivationally stronger, on a given occasion, compared to the care which constitutes the agent’s real self. As a result, the care that determines if an action is autonomous or not is not necessarily the one which comes with the strongest motivation. It is the one that is strongest in the sense of being most central to the agent’s psychology.

Finally, let us consider the so-called manipulation problem. This is a problem that all autonomy accounts have to deal with. It is clear that our emotional dispositions are

shaped by our genes as well as by the socio-cultural background. However, cares can also be the result of manipulative processes, such as when one is the victim of hypnosis or brainwashing. Shoemaker's reply to the manipulation objection consists in the claim that our cares are ours in so far as they result from well-functioning cognitive mechanisms. These mechanisms include critical reflection and self-evaluation, as well as mechanisms which allow our cares and our motivations to be sensitive to our thoughts. That these mechanisms have been tampered with would explain why the victim of hypnosis or of brain-washing would not be autonomous when he acts on his cares.

One point that has to be underlined is that it would be wrong to claim that our cares need to be the result of reflection. We acquire our cares long before we are able to consider our own reactions in a critical way. Consider our attachment to the place in which we have grown up, for instance. At least our initial cares depend in part on our environment. They are not a matter of reflective choice. Moreover, as Shoemaker himself underlines, reflection is always done on the basis of a care. It is because we care to be a certain kind of person that we engage in a process of self-critical reflection. It is because we care for the outcome of our actions that we deliberate about what to do.

However, it seems to make sense to claim that for our cares to constitute our real autonomous selves, they have to have developed naturally, on the basis of a natural unfolding of our emotional and cognitive capacities. It has to be underlined that the concept of naturalness is used here in a normative way. One might object that a natural unfolding of our emotional and cognitive capacities thus seems to be nothing else than whatever process leads to autonomy. Now, it has to be agreed that there might be no general way to characterise cases in which there is manipulation in order to distinguish them from naturally grown cares. What is clear, however, is that certain ways to acquire cares are manipulative: if the cares an agent develops are the result of neurosurgery, hypnosis or brainwashing they undoubtedly fail to constitute the agent's real self.

The upshot is that in order to take into account this last point, as well as the distinction between centrality and motivational strength, we have to modify the care account in the following way:

(4) A motivation is part of the real self of the agent iff it results from the most central care of the agent, and this care has developed naturally.

Given this, an action would be autonomous if it is determined by the strongest, and thus most central care of the agent, provided this care has naturally developed.

Conclusion

We have seen that rationalist accounts of autonomy fail to account for free actions that are contrary to the agent's better judgement. By contrast, care accounts appear to have no problem with cases of weakness of will. As we have seen, there are reasons to be

doubtful of some claims made by Frankfurt and Shoemaker. However, it nonetheless seems plausible that the real self of an agent is constituted by his most central cares, in so far as these cares have developed naturally; that is, in so far as they are not the result of manipulations. Given that cares are nothing but a set of emotional dispositions, it follows that there is no general opposition between autonomy and emotions. On the contrary, if an emotion is involved in a network that constitutes an agent's most central cares, the action caused by this emotion is autonomous. Apart from the emotions that belong to cares that are not central, such as in the case of compulsion, the only emotions that would make for an exception would be those that do not depend on cares. However, it is not clear that emotions can be unrelated to cares.*

** I am grateful to Christine Clavien, Christopher Cowley, Josep Corbi, Paul Bernier, Stéphane Lemaire, Colin McLeod, Anne Meylan, Martin Montminy, Kevin Mulligan, Ruwen Ogien, Claude Panaccio, Fabienne Pironet, and Sophie Rietti for comments on various versions of this paper.*

REFERENCES

- Arpaly N. & T. Schroeder (1999), "Praise, Blame and the Whole Self", *Philosophical Studies*, 93, pp. 161-188
- Bennett J. (1974), "The Conscience of Huckleberry Finn", *Philosophy*, 49, pp. 123-134
- Bratman M. (1999), "Identification, Decision, and Treating as a Reason", in Bratman 1999, pp. 185-206
- Bratman M. (1999), *Faces of Intention*, Cambridge: Cambridge University Press
- Cuyppers S. E. (2000), "Autonomy beyond Voluntarism: In Defense of Hierarchy", *Canadian Journal of Philosophy*, 30, pp. 225-256
- Ekstrom L. W. (2005), "Autonomy and Personal Integration", in Taylor 2005, pp. 138-157
- Fischer J. & M. Ravizza eds. (1993), *Perspective on Moral Responsibility*, Ithaca: Cornell University Press.
- Frankfurt (1999), *Necessity, Volition, and Love*, Cambridge: Cambridge University Press
- Frankfurt H. (1971), "Freedom of the Will and the Concept of a Person", *Journal of Philosophy*, 68, pp. 5-20
- Frankfurt H. (1982), "The Importance of What We Care About", in Frankfurt 1988, pp. 80-94
- Frankfurt H. (1988), *The Importance of What We Care About*, Cambridge: Cambridge University Press
- Frankfurt H. (1994), "Autonomy, Necessity, and Love", in Fulda & Horstmann 1993, pp. 433-447 ; also in Frankfurt 1999, pp. 129- 141
- Frankfurt H. (1999), "The Faintest Passion", in Frankfurt 1999, pp. 95-107
- Fulda H.F. & R.P. Horstmann eds. (1993), *Vernunftbegriffe in der Moderne: Stuttgarter Hegel-Kongress 1993*, Stuttgart: Klett-Cotta
- Hume D. (1939), *A Treatise of Human Nature*, Oxford: Oxford University Press
- Taylor J.S. ed. (2005), *Personal Autonomy*, Cambridge: Cambridge University Press

- Lehrer K. (1990), *Metamind*, New York: Oxford University Press
- McIntyre A. (1993), "Is Akritic Action Always Irrational", in Flanagan & Rorty 1993, pp. 379-400
- Mele A.R. (1987), *Irrationality. An Essay on Akrasia, Self-Deception and Self-Control*, Oxford: Oxford University Press
- Mele A.R. (1995), *Autonomous Agents: From Self Control to Autonomy*, Oxford: Oxford University Press
- Mele A.R. (2002), "Autonomy and Akrasia", *Philosophical Explorations*, 5, pp. 207-216
- Flanagan O. & A. Rorty eds, (1993), *Identity, Character and Morality*, Cambridge, Mass.: MIT Press
- Shoemaker D. (2003), "Caring, Identification, and Agency", *Ethics*, 114, pp. 88-118
- Stump E. (1988), "Sanctification, Hardening of the Heart, and Frankfurt's concept of Free Will", *Journal of Philosophy*, 85, pp. 395-420
- Velleman D. (2006), *Self to Self*, Cambridge: Cambridge University Press
- Velleman D. (1989), "Epistemic Freedom", *Pacific Philosophical Quarterly*, 70, pp. 73-97
- Velleman D. (2006), "Identification and Identity", in Velleman 2006, pp. 330-360
- Watson G. (1975), "Free Agency", *The Journal of Philosophy*, 72, pp. 205-220
- Watson G. (1977), "Skepticism about Weakness of Will", *Philosophical Review*, 86, pp. 316-339
- Watson G. (1987), "Free Action and Free Will", *Mind*, 96, pp. 145-172
- Wolf S. (1993), "The Real Self View", in Fischer & Ravozza 1993, pp. 151-170

Received: November 3, 2006
Accepted: December 17, 2006

Université de Montréal
C.P. 6128, succursale Centre-ville
Montréal QC H3C 3J7
christine.tappolet@umontreal.ca

