

ScholarWorks@GSU

Human Flourishing and Autonomy as Passive

Authors	Taylor, Gerald
Citation	Taylor, Gerald. "Human Flourishing and Autonomy as Passive." 2015. Thesis, Georgia State University. https://doi.org/10.57709/7138199
DOI	https://doi.org/10.57709/7138199
Download date	2026-05-10 06:05:52
Link to Item	https://hdl.handle.net/20.500.14694/11973

HUMAN FLOURISHING AND AUTONOMY AS PASSIVE

by

GERALD DAVID TAYLOR

Under the Direction of Dr. Eddy Nahmias

ABSTRACT

Most prominent accounts of autonomy are *active accounts*, which means they hold that an agent can be autonomous with respect to a given action only if that agent has appropriately sanctioned that action. Active accounts, however, are vulnerable to the regress problem, since it seems that the required sanctioning actions are themselves just actions that must be sanctioned. *Passive accounts* hope to avoid the regress problem by eschewing the notion that autonomous action requires agential sanction, but face in its place what I call the *incompleteness problem for passive accounts*. Here, I evaluate one passive account, recently defended by Sarah Buss, and argue that it can avoid the incompleteness problem only if it is supplemented by a satisfactory account of the development of autonomy. I then suggest that one development account, offered by Ishtiyaque Haji and Stefaan Cuypers, is an especially good supplement. My conclusion is that Buss's account, supplemented in the way I suggest, can avoid the incompleteness problem.

INDEX WORDS: Autonomy, Active, Passive, Regress, Incompleteness, Flourishing, Manipulation, Development, Buss, Haji, Cuypers

HUMAN FLOURISHING AND AUTONOMY AS PASSIVE

by

GERALD DAVID TAYLOR

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2015

Copyright by
Gerald David Taylor
2015

HUMAN FLOURISHING AND AUTONOMY AS PASSIVE

by

GERALD DAVID TAYLOR

Committee Chair: Dr. Eddy Nahmias

Committee: Dr. Andrew I. Cohen

Dr. Nicole Vincent

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2015

ACKNOWLEDGEMENTS

First, I'd like to thank my thesis advisor, Eddy Nahmias, for being an exceptional mentor to me during my time here at Georgia State. I also would like to thank the other members of my thesis committee, Andrew I. Cohen and Nicole Vincent, for providing very helpful feedback and encouragement as I completed this project.

Second, I'd like to thank my fellow grad students, especially Nate Houck and Jay Spitzley, for being my intellectual sparring partners. Our conversations over the past couple of years have helped me to refine my ideas (and to discover new ones), as well as to become a better philosopher overall.

Finally, I'd like to thank the faculty at the Ohio State University, especially Sigrún Svavarsdóttir, who helped to nurture and encourage my interest in philosophy when I was just getting started. I can say without reservation that I would not have even made it to Georgia State, much less to the completion of this thesis, without their aid.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	v
1 Introduction: Autonomy, Active and Passive.....	1
2 Motivating the Passive Account.....	3
3 Buss on Manipulation.....	7
3.1 The Human Flourishing Condition	7
3.2 Buss’s Passive Treatment of Manipulation.....	10
4 The Development of Autonomy	18
4.1 A ‘Two-Stage’ Account of the Development of Autonomy	19
4.2 Bringing Together Buss and Haji & Cuypers	23
5 Conclusion: Toward a <i>Flourishing Account</i> of Autonomy.....	30
REFERENCES.....	32

1 Introduction: Autonomy, Active and Passive

Suppose that Smith, faced with a choice between action *A* and action *B*, chooses *A*. To say that Smith did so autonomously presupposes at least two things about him. First, that Smith was *in control* of himself when he chose *A*; Smith had a certain set of motivational pro-attitudes (MPAs) – beliefs, desires, principles, etc. – and acted in accordance with those MPAs. Second, that the set of MPAs that guided Smith's choice to perform action *A* were *his* to begin with; that is, his MPAs were not the result of, for instance, hypnotic suggestion or some other form of psychological manipulation. Accounts of personal autonomy, then, have at least two goals: to explain when an agent is *in control* of his actions, and also to explain when an *agent himself* is in control of his actions. The current inquiry will focus primarily on the latter goal.

What, then, makes it the case that the MPAs that guided Smith's decision were *his* MPAs? Most accounts of autonomy hold that the relevant set of MPAs would be Smith's only if he directly sanctioned or endorsed it, perhaps by reflecting on it and deciding whether or not to accept it. Call accounts that require agential sanction *active accounts*. Active accounts are problematic, however, because they seem to make autonomy impossible for us to attain. After all, if we suppose that an action is autonomous only if it were, say, reflectively endorsed, then it would also be true that the endorsement, which is itself an action, must be sanctioned. And there seems to be no way to terminate the impending regress.

This regress problem has motivated thinkers like Sarah Buss (2012) to reject this requirement of active accounts – that autonomous actions require agential sanction – and accept instead a *passive account* of autonomy, according to which it isn't necessary for Smith to *do* anything in order for his decision to perform *A* to be considered *his* decision. This isn't to say that agential sanction is not an important part of what it means to be an autonomous agent. It is

merely to say that insofar as Smith *himself* plays the decisive role in determining his actions, he does so in a way that cannot be adequately described by recourse to agential sanction. What this entails, though, is that the onus is on the proponents of passive accounts to explain how it is possible for agents to act autonomously without having sanctioned their actions or the MPAs that guide them.

In this essay, I will offer a defense of passive accounts of autonomy. In particular, I will argue that Buss's passive account, based on what she calls the *human flourishing condition*, succeeds in providing the relevant explanation, but only if it is supplemented with an explanation of how autonomy develops in the first place.

The essay will proceed as follows. In Section II, I will review the motivations for passive accounts in more detail, and make the nature of the problem that they face more explicit. In Section III, I will present Buss's account, and suggest a way in which her account could be supplemented, namely, by assimilating a satisfactory account of the development of autonomy. In Section IV, I discuss one development account, put forward by Ishtiyaque Haji and Stefaan Cuypers (2007), and argue that it helps Buss's account avoid the problem that I outline in Section II. Finally, in Section V, I conclude the essay.

2 Motivating the Passive Account

To a first approximation, autonomy is the ability of agents to govern themselves according to some set of MPAs that are their own. Most attempts to analyze autonomy in this sense fall into one of two broad categories: structural or historical. Structural accounts of autonomy hold that an agent is autonomous with respect to an action just in case the set of MPAs that led to the performance of that action is appropriately arranged. One classic example of the structural account comes from Harry Frankfurt (1971), who argues that autonomy requires acting on a first-order desire (*e.g.*, a desire that *X*) that is both endorsed by a second-order desire (*e.g.*, a desire to desire that *X*) and actualized by a second-order volition (*e.g.*, a desire to be moved by the relevant second-order desire regarding *X*).¹

Historical accounts, on the other hand, hold that an agent is autonomous with respect to an action just in case that action was endorsed by the right kind of historical process. Mikhail Valdman (2011), for example, argues that an agent is autonomous with respect to an action just in case that agent has *engaged* with the MPAs that motivate it, where *engagement* entails “a process of deliberation whose purpose is to determine whether a desire is worth maintaining and worth acting on” (423).

It is important to note that ‘historical’ in this context does not refer to the simple passage of time. Rather, it connotes a special kind of causal participation by an agent in the production of her actions. A historical process, then, will be one by which an agent establishes the appropriate kind of causal relationship between herself and the production of her actions, regardless of how much time actually passes as the relevant causal relationship is established. Usually, this process

¹ Of course, Frankfurt wrote this essay as a contribution to the free will debate, rather than to the debate about personal autonomy. However, Frankfurt’s account is often presented within the personal autonomy debate as a paradigm example of the structural account. My inclusion of Frankfurt’s account here reflects that convention.

includes not the mere having of a desire, but a decision to have a desire that is arrived at by deliberative and self-evaluative processes.

Structural and historical accounts have obvious differences, but are united in their commitment to the idea that autonomous actions require agential sanction of some sort, either in the form of taking the appropriate attitudes or performing the appropriate actions. We can represent this idea as a commitment to the following principle:

Activity Principle (AP): an agent is autonomous with respect to an action only if the agent appropriately sanctions that action.

As we've seen, structural accounts cash out 'appropriately' in terms of suitably arranged sets of MPAs, while historical accounts do so in terms of the completion of historical processes. However, since both kinds of account accept AP in some sense or other, I will henceforth and freely refer to them as *active accounts*.

One problem with active accounts is that they seem to make autonomy impossible to attain. To see this, consider again Frankfurt's structural account, according to which an action is performed autonomously only if it was, among other things, actualized by a second-order volition. The required second-order volition is itself a desire (about another desire, of course, but a desire all the same), and it is unclear why it shouldn't have to be endorsed in just the same way that first-order desires do. We could call upon third-order desires or volitions, but the same sort of questions would arise. Why should we stop there? Why should it be the case that third-order MPAs can confer autonomy while second-order MPAs cannot? Without a principled answer to questions like these, there seems to be no good reason for terminating the regress at any particular point. And, so the objection goes, the fact that there is no reason to terminate the regress at any particular point is conclusive reason to reject the account that set the regress in motion.

A similar problem afflicts historical accounts of autonomy. Consider again Valdman's account, which involves a process of deliberation about one's MPAs, the purpose of which is to inform a decision about whether those MPAs are worth maintaining or acting on. The required deliberation, as well as the decision informed by that deliberation, are themselves just actions that can only be performed autonomously if they were preceded by the appropriate process of deliberative endorsement (and the same goes for this pre-endorsement endorsement). And again, there seems to be no principled reason to terminate the regress at any particular point in the agent's personal history and, as such, no reason to think that any particular instance of endorsement has more authority than any other.

Upon closer inspection, then, active accounts seem to require either, in the case of structural accounts, an infinite number of MPAs, or, in the case of historical accounts, an infinitely long personal history. Of course, it is unclear whether actually having infinitely many MPAs or infinitely long histories would actually help to terminate the regress, since we would still be without good reason for terminating it at any particular point. What should be clear, however, is that since we are finite beings, unable in any case to enjoy the benefits of such infinite resources, active accounts seem to set the bar for autonomous action far beyond our reach. Active accounts, then, united in their acceptance of AP, invite what is known as the *regress problem*.

The susceptibility of active accounts to the regress problem is the main motivation for *passive accounts* of autonomy, which holds that agents need not *do* anything at all to make it the case that they themselves produce their actions. That is, passive accounts reject the claim that agential sanction is *necessary* for autonomous action (though they can, but need not, allow that it

is *sufficient*). By rejecting AP, passive accounts hope to avoid the troublesome regress problem altogether.

There is, however, a potential problem with passive accounts, which is that of having to explain just how autonomy is possible without the characteristic sanctioning requirements of active accounts. Active accounts, despite being problematic, are at least intuitive; it is very reasonable to suppose that an agent governs himself only to the extent to which he endorses or sanctions his actions. In order to counter this intuition, passive accounts need to do more than simply object to the active approach to autonomy. Without a positive account of what autonomy requires if not agential sanction, it would be unclear why we should shift away from the intuitive active approach in favor of passive accounts, as opposed to simply attempting to repair or adjust the active approach (or, for that matter, to adopt a skeptical attitude regarding human autonomy). In short, without such a positive account of the requirements of autonomy, passive accounts, despite being amply motivated, are toothless. Call this the *incompleteness problem for passive accounts*, or just the incompleteness problem.

In the next section, I will discuss one passive account, recently defended by Sarah Buss (2012), and suggest a solution to the incompleteness problem. The key to a solution, I will argue, lies in having available a satisfactory account of how autonomy develops in normal human beings.

3 Buss on Manipulation

In this section, I aim to show that Buss's passive account is incomplete in the way that I've just described, though it provides important foundations for an effective passive account of autonomy that I will begin to develop in the following section. I will now provide some background on Buss's account in order to facilitate that discussion.

3.1 The Human Flourishing Condition

Buss's account is motivated partly by the aforementioned concerns about the vulnerability of active accounts to the regress problem. But her account is also motivated by what she sees as an important though almost entirely overlooked connection between the concepts of autonomy and human flourishing. We can approximate the nature of this connection by considering that some psycho-physiological conditions manage intuitively to undermine our autonomy, even when we deliberate carefully about our actions.

Consider Mary, for instance, who is my normally kind and pleasant coworker. Mary usually greets me in the morning politely and happily. On this particular day however, Mary could hardly be said to have "greeted" me at all; she responded to my 'hello' tersely and with a rude remark. What should I make of Mary's terse and rude remark? Suppose I think to myself: *something must be wrong with Mary, she doesn't usually act like this.*

This is a reasonable response. But insofar as I mean to be offering an *excuse* for her behavior, I can't mean simply that she acted out of character. After all, the simple fact that Mary acted uncharacteristically does not, by itself, indicate that she didn't do so autonomously. No, if I mean 'Mary doesn't usually act like this' as an excuse for her behavior, I must be pointing to something else about Mary, some deeper fact about her that does the work of explaining away her behavior.

As a proponent of the passive account, Buss rejects the idea that this deeper fact relates to the manner in which Mary exercised her agency in that instance. Rather, according to Buss, the deeper fact to which we draw attention when we say ‘Mary doesn’t usually act like this’ is that whatever caused her to act uncharacteristically did so in a way that was external to, or independent of, Mary’s own influence on her behavior.

Presumably, then, certain psycho-physiological conditions can constitute external influences on our behavior even when they are, in one obvious sense, internal (i.e., bodily) influences. In what sense, then, do those conditions operate externally to our own influence?

Buss writes:

To be a human agent is to be a representative member of a species. This means that there is an important respect in which even well-integrated, long-standing psychological and physiological conditions are external to a human agent’s identity insofar as they are causes or symptoms of human malfunctioning. And this means that there is an important sense in which when such conditions exercise a decisive causal influence on the formation of someone’s intentions, she does not produce her own actions – no matter how carefully she reviews her options and no matter how wholeheartedly she endorses her response to these options.

(Buss 2012: 660)

Some psycho-physiological conditions, then, are either causes or constituents of human malfunctioning, and operate externally to *our* influence over our behavior for just this reason. But what does it mean for a particular psycho-physiological condition to be a cause or constituent of human malfunctioning? Buss has something like the following in mind. We, as human beings, are a particular kind of creature with a particular way of functioning *well*. We have basic human needs, psycho-physiological in nature, the satisfaction of which enables us to live *minimally decent* lives. A psycho-physiological condition is a cause or constituent of human malfunctioning when and to the extent that it prevents its bearer from satisfying one or more of these basic needs.

This is an important move for Buss's account, as it grounds a normative distinction between sick and healthy human beings that, Buss argues, is inseparable from the distinction between autonomous and nonautonomous human beings. In one straightforward sense, sick human beings are no less *human* than healthy ones; all human beings, sick or healthy, are members of the same species. However, if Buss is right, then sick human beings, despite being no less human, are *less perfect* representatives of the species than their healthy counterparts; since their basic human needs aren't met, *they fail to function minimally well as human beings* and, as such, are prevented from expressing their identity as *healthy* human beings. And it is in this sense that psycho-physiological conditions, when they constitute causes or symptoms of human malfunctioning, operate externally to the influence of their bearers; they prevent, rather than facilitate, their minimal well-functioning as human beings.

Now, not all causes or symptoms of human malfunctioning undermine autonomy. Such conditions might prevent someone from being perfect representatives of the human species without thereby undermining her control over her behavior. It is quite possible, for instance, to imagine people with striking health issues – for example, amputees, blind people, and those with terminal but well-managed diseases – who nonetheless manage to have decent lives and to act autonomously in doing so. It is only when such conditions *do* manage to play a decisive role in determining one's behavior that they undermine autonomy. Thus, we arrive at the following principle, which serves as the basis for Buss's passive account of autonomy:

Human Flourishing Condition (HFC): [A] human being fails to act autonomously if and only if she forms her intentions under the decisive [...] influence of conditions that are elements or symptoms of human malfunctioning.

(Buss 2012: 660)

One might worry that 'human *flourishing* condition' is a misleading name for this principle, since it runs the risk of obscuring the purported relationship between human well-functioning,

human flourishing, and personal autonomy. On Buss's account, it is not human flourishing that is required for autonomy; indeed, to require that a human agent *flourish* as a human being before that agent could be considered autonomous would be to set the bar for autonomy far too high. Rather, it is minimal human well-functioning that is required for autonomy, and autonomy in turn is but one constituent of human flourishing. Perhaps, then, a better name for this principle is the 'human *functioning* condition', since its satisfaction requires only that the agent functions minimally well *qua* human being. However, for the sake of consistency, I will continue to refer to this principle as the 'human flourishing condition'.

The HFC is Buss's way of overcoming the incompleteness problem that I set out in the opening. That is, the HFC describes how it is that agents manage to act autonomously without having to rely on agential sanction; agents control their actions autonomously when they do so in ways that aren't detrimental to their minimal well-functioning as a human being. There are many interesting facets of the HFC that, for lack of space, I will not be able to discuss here.² The remaining discussion will focus solely on how Buss makes use of the HFC in responding to a kind of counterexample that has been very difficult for autonomy theorists to overcome.

3.2 Buss's Passive Treatment of Manipulation

For the remainder of this section, I will evaluate Buss's response to manipulation cases – cases wherein an agent who shows the outward (and even inward) signs of acting autonomously is, nonetheless, nonautonomous – and argue that she still has more work to do in providing a complete account of the passive requirements for autonomy. Consider now the following manipulation case, which I will refer to as *Ann/Beth*:

² One might wonder, for example, exactly how to fill out the list of basic human goods that might bear on the question of whether any agent at any particular time satisfies the HFC. For one compelling attempt to complete such a list, see Martha Nussbaum, *Women and Human Development: The Capabilities Approach* (New York: Cambridge, 2000).

Ann is an autonomous agent and an exceptionally industrious philosopher. She puts in twelve solid hours a day, seven days a week; and she enjoys almost every minute of it. Beth, an equally talented colleague, values a great many things above philosophy, for reasons that she refined and endorsed on the basis of careful critical reflection over many years. She identifies with and enjoys her own way of life – one which, she is confident, has a breadth, depth, and richness that long days in the office would destroy. Their dean (who will remain nameless) wants Beth to be like Ann. Normal modes of persuasion having failed, he decides to circumvent Beth’s agency. Without the knowledge of either philosopher, he hires a team of psychologists to determine what makes Ann tick and a team of new-wave brainwashers to make Beth like Ann. [...] Beth is now, in the relevant respects, a “psychological twin” of Ann. She is an industrious philosopher who enjoys and highly values her philosophical work. [...] Her critical reflection, like Ann’s, fully supports her new style.

(Mele 1995: 145)

The intuition that we are supposed to have is that Beth is not autonomous, neither with respect to her critical reflection post-manipulation, nor with respect to the rededication inspired by that reflection. The task, now, is to identify the difference between Ann and Beth that explains why the former is autonomous but the latter is not.

For proponents of structural accounts, there seems to be no satisfactory way to complete this task. Since Ann and Beth are “psychological twins” following Beth’s manipulation, it seems hopeless to look for a difference between them in their psychological or motivational structures. A structural approach to manipulation, then, isn’t likely to succeed.

Historical accounts fare somewhat better. The dean, by his intervention, denied Beth the opportunity to deliberate or reflect on the change before he hired the team that instituted the change; indeed, her own deliberations pre-manipulation indicate that, given the choice, she would not accept the lifestyle that the dean wants her to have. In other words, the dean’s intervention prevented Beth from completing the appropriate historical process regarding her sudden lifestyle change, thereby undermining her autonomy with respect to having made that change. Historical accounts, then, can at least initially accommodate the intuition that Ann is autonomous while Beth is not.

But we can imagine a slightly different case wherein Beth was manipulated in a more distant manner. Perhaps the dean's team rearranged Beth's motivational structure such that she would make the change only after a year or so of critical reflection. In this case, one could argue that Beth completed the relevant historical process by, for instance, engaging with her new MPAs for a full year. Yet the intuition persists that she is not autonomous with respect to making the lifestyle change, even after a year of critical reflection.

A historicist could reply that though Beth completed the right historical process, the impetus for completing that process came from a source external to Beth, rather than from Beth herself. That is, though Beth did engage with her MPAs in the right kind of way, the motivation for beginning the historical process in the first place came not from her, but from the dean and his team of scientists. And so, despite appearances, Beth is actually not autonomous.

To press this reply, however, would have the unfortunate consequence of also disqualifying Ann as autonomous, since Ann's own motivational structure also originated from external sources. After all, no one creates their MPAs entirely from scratch; Ann, like everyone else, is the result of the confluence of numerous influences during her formative years. And it is at best unclear why we should privilege Ann's external influences over Beth's.

Active accounts, then, fail to successfully explain manipulation cases. Structural accounts fail to specify the relevant psychological difference between Ann and Beth, while historical accounts fail to do so isolate the relevant historical difference. Let us, then, move on to evaluation of Buss's passive response to the case, and see if it fares any better.

Some recent discussions of manipulation cases have focused, not on how manipulation prevents agents from sanctioning their actions, but rather on their problematic use of what can be

called *effective background agency* (EBA).³ EBA occurs when the manipulator not only wants the featured agent to do a certain thing, but is also able to *ensure* that the featured agent does that thing. The idea is that in cases involving EBA, the featured agent either (1) has her extant MPAs replaced by the ones the manipulator wants her to have, which then motivate the action that the manipulator wants her to perform; or (2) has new MPAs implanted in her which, while they don't replace her own MPAs, nonetheless do all the work of motivating her to perform the relevant action. If (1) is true, then there is little sense in which the agent *herself* is in control over her behavior; she may not fail to sanction or endorse her actions, but she does fail to do so with respect to MPAs that are *her own*. If (2) is true, then though she retains her own set of MPAs, it isn't *those* MPAs that motivate the relevant action. The motivational work is done entirely by the MPAs that the manipulator implanted. In cases involving EBA, then, with respect to the action that the manipulator wants the featured agent to perform, the agent's motivational structure is either eliminated or bypassed entirely.

Buss's own response to manipulation cases builds upon this idea, offering an explanation as to *why* EBA reliably undermines autonomy. According to Buss, having a *psychic identity*, by which she means the set of "relatively stable psychological dispositions [which] are constituents of [one's] identity" (687), is an essential component of what it means to function minimally well as a human being. Further, it is important that one's psychic identity be allowed to exert causal influence on one's actions, since "having no character, or personality, or psychic identity of any kind (or, what amounts to the same thing, no psychic identity that can directly influence the formation of [one's] intentions) is incompatible with being a minimally well-functioning adult human being" (688).

³ See, for example, Robyn Repko Waller, "The Threat of Effective Intentions to Moral Responsibility in the Zygote Argument," *Philosophia*, Vol. 42 (2014): 209-222. See also Kristin Demetriou, "The Soft-Line Solution to Pereboom's Four-Case Argument," *Australasian Journal of Philosophy*, Vol. 88 (2010): 595-617.

If Buss is right about this, then the HFC helps to explain cases like *Ann/Beth* in a way that is importantly different from similar attempts by active accounts. The dean, simply by intervening in the production of Beth's actions, prevents her from performing those actions autonomously. But this is not because Beth was prevented from reflecting fully enough or in the right kind of way about her actions, nor is it because she failed to take the right sanctioning attitudes toward her actions. Rather, Beth fails to be autonomous because the dean, by disrupting Beth's psychic identity, prevented her from functioning minimally well as a human being. And, presumably, all EBA-style manipulations prevent agents from acting autonomously in just this way.

What, then, should we say about those that *don't* involve EBA? We can imagine for example that Beth's manipulators, rather than instituting a change in her motivational structure that would *guarantee* her rededication, instead made her rededication only *more likely*. Perhaps they could accomplish this by somehow allowing Beth's own MPAs to interact with the implanted ones in a way that leaves it open whether Beth will actually make the change they want her to make. Buss has the following to say about this kind of case:

On my account, agents act autonomously as long as their intentions reflect the decisive influence of their character or personality; and this is true no matter how they acquired these psychological traits. When someone is moved by a desire that has just been forced upon her in the manner envisioned in the thought experiments, she does not act autonomously (since, again – at least if one is older than a toddler – having no functioning stable psychological traits is not a way of functioning-minimally well as a human being). As time goes by, however (and it may not be much time at all), a person who has been manipulated in this way may come to have a new identity. And as soon as this happens, the mad scientist's prior intervention cannot prevent her from governing herself. [...] What matters is whether the [manipulation] left her in a condition that impedes her ability to function minimally well as a human being.

(Buss 2012: 688-689)

This strikes me as the right response. That is, I see no conclusive reason that Beth, even if she is not autonomous immediately following the manipulation, could *never* become autonomous again. It seems plausible to me that there is some mechanism or process that allows for people to

“recover” from undue influences on their motivational structures. If not, we would have to say that people who have had troubling or overbearing childhoods are, for that reason, prevented henceforth from being autonomous. Of course, in extreme cases, they *might* actually be permanently unable to become autonomous. But I don’t see why this should be true *necessarily* or *in every case*.

Buss, however, offers little guidance as to what such a mechanism or process might consist in or entail. Of course, since she is merely trying to establish a conceptual link between autonomy and human flourishing, perhaps she needn’t provide such an explanation in detail. I think, however, that such an explanation would help to bolster her account for two main reasons. First, and most obviously, such an explanation would complete her response to manipulation cases by helping to make sense of the reasonable intuition that Beth could overcome the influence of a non-EBA-style manipulation, and thereby strengthen her account overall. Second, such an explanation would also help us to make sense of Ann who, again, like everyone else, was surely heavily influenced during childhood by her formative environment. Presumably, we want to be able to explain another reasonable intuition, which is that people are not necessarily doomed by their childhoods to lives of nonautonomous action simply by having been influenced in certain undesirable ways (and we also want to be able to explain why, in the extreme cases, agents *could* be prevented from ever being autonomous again).

It is likely that Buss would again make use of the notion of psychic identity and its importance for human well-functioning. Indeed, it is the requirement that an autonomous agent have a non-tampered-with psychic identity that allows her to offer an explanation as to why manipulated agents are not autonomous. However, while I think that this is a good point to press,

I think that it requires some elaboration. More specifically, I think that Buss needs to be more precise in her definition of psychic identity.

To see the problem more clearly, consider a case slightly different from *Ann/Beth*, in which Beth rededicates her life to philosophy not because she was covertly manipulated by the Dean, but instead because Ann visits her at home one night and convinces her that she is no longer realizing her full potential as a philosopher. At first, Beth disagrees with Ann's reasoning; she is, after all, very satisfied with the balance that her current lifestyle affords. But suppose that after a relatively short period of deliberation – perhaps just one night, the same amount of time that it takes the Dean's team of manipulators to do their work in *Ann/Beth* – Beth decides that Ann is right after all. Beth has had a change of heart.

I take it that we should say that this version of Beth, the one that had the change of heart, is autonomous with respect to her rededication. And presumably this is because the relevant change to her MPAs occurred in a way that was more organic and 'normal' than covert psychological manipulation. But without a more precise conception of 'psychic identity', as well as of how it relates to (and is constitutive of) autonomy and human flourishing, it becomes difficult to differentiate between manipulated Beth and change of heart Beth. After all, in both cases, Beth's rededication was inspired by another agent, and took place in roughly the same amount of time. What we need is a fuller account of what 'psychic identity' amounts to, as well as of the principled difference between those changes in psychic identity that undermine autonomy (*e.g.*, covert manipulation) and those that do not (*e.g.*, a change of heart inspired by a friendly intervention).

In other words, an explanation of how autonomy might re-emerge after cases of manipulation (*e.g.*, Beth), as well as of how it develops in ordinary cases (*e.g.*, Ann), would help

Buss to better handle the incompleteness problem that I set out in the opening which, again, is the problem of having to describe just how it is possible for agents to act autonomously if not by recourse to agential sanction. To this end, I turn now to the task of evaluating one promising account of the development of autonomy, recently put forward by Ishtiyaque Haji and Stefaan Cuypers (2007) and seeing how well it fits with Buss's account.

4 The Development of Autonomy

I've been assuming thus far that the intuition that Beth is not autonomous with respect to her rededication to philosophy is the right one to have. Michael McKenna (2004) has challenged this intuition by using a concept that we can call *instant agency*.⁴ Haji and Cuypers (hereafter H&C; 2007) offer their account of the development of autonomy as a response to McKenna's challenge.

To see how instant agency is supposed to challenge the aforementioned intuition, we can imagine that Beth, rather than living a normal life up until the point when she was made into Ann's psychological twin, was instead created instantaneously by a powerful deity with a full suite of MPAs and memories that make her, essentially, Ann's psychological twin. Furthermore, we can imagine that Beth comes into existence at precisely the time that Ann performs some action *A* on the basis of her MPAs, and that Beth, being Ann's psychological twin, performs the same action at the same time and for the same reasons.

McKenna goes on to argue that in this case, an agent like Instant Beth would be autonomous with respect to *A* if Ann would be. This is because Instant Beth would not be relevantly different from an agent created in the way envisioned in Mele's (2006) Zygote Argument (i.e., the god intentionally designed the agent's zygote in such a way that she would grow up to become an autonomous adult and perform some action *A* at a particular time), and also an intentionally designed agent would not be relevantly different from an agent, like Ann, who develops her autonomy in the normal way. Consequently, since the intentionally designed agent is autonomous (owing to her similarity to the normal agent), and since the instant agent is

⁴ Another, more counterintuitive strategy is to simply bite the bullet and insist that Beth is autonomous despite the dean's manipulation. For an excellent discussion of this sort of reply to manipulation arguments, see Michael McKenna, "A Hard-line Reply to Pereboom's Four-Case Manipulation Argument," *Philosophy and Phenomenological Research*, Vol. 28 (2008): 142-159.

autonomous (owing to her similarity to the intentionally designed agent), then if the manipulated agent is relevantly similar to the instant agent, then the manipulated agent is also autonomous. McKenna then supposes that manipulated agents are indeed similar to instant agents. And if McKenna were right about this, then our initial intuition regarding Beth's non-autonomy would turn out to be mistaken.

I wish at present to avoid questions regarding whether intentionally designed agents are autonomous, as well as those regarding whether there are relevant differences between intentionally designed agents and instant agents. Here, however, I will focus instead on whether a manipulated agent like Beth is, as McKenna supposes, not relevantly different from an instant agent. It is important for our purposes that we be able to resist this supposition, since accepting it entails the failure of the account that I mean to use to bolster Buss's account. To that end, I proceed now to a brief discussion of the details of H&C's account, and of what that account has to say about the difference between manipulated agents and instant agents.

4.1 A 'Two-Stage' Account of the Development of Autonomy

In a discussion about the sense in which "self-creation" is possible for beings like us, Joel Feinberg writes:

The extent of [a] child's role in his own shaping is [...] a process of continuous growth already begun at birth. From the very beginning that process is given its own distinctive slant by the influences of heredity and early environment. [...] After that, the child can contribute towards the making of his own self and circumstances in ever increasing degrees. These contributions are significant even though the child is in large part (especially in the earliest years) the product of external influences over which he has no control, and his original motivational structure is something he just finds himself with, not something he consciously creates. [...] At every subsequent stage the immature child plays a greater role in the creation of his own life, until at the arbitrarily fixed point of full maturity, he is at last fully in charge of himself, his more or less finished character the product of a complicated interaction of external influences and ever-increasing contributions from his own earlier self. At least that is how growth proceeds when parents and other authorities raise a child with a maximal regard for the autonomy of the adult he will one day be. [...]

Perhaps we are all self-made in the way just described, except those who have been severely manipulated, indoctrinated, or coerced throughout childhood. But the self we have created in this way for ourselves will not be an authentic self unless the habit of critical self-revision was implanted in us early by parents, educators, or peers, and strengthened by our own constant exercise of it.

(Feinberg 1986: 34-35)

There are several key ideas in this passage that serve as the foundations upon which H&C build their account. The first of these ideas regards the “habit of critical self-revision.” It is important that autonomous agents have a relatively stable and coherent set of MPAs, as well as the ability to act in accordance with those MPAs. But it is perhaps just as important that autonomous agents have the ability to step back and evaluate those MPAs. That is, it is important that an autonomous agent be able to reflect on his own motivational structure and decide for himself which of its components he should accept and keep, which he should modify, which he should remove, and whether any new components need to be added.

H&C capture this notion of critical self-revision with a concept that they call the *evaluative scheme*, which is composed of “[n]ormative standards the agent believes (though not necessarily consciously) ought to be invoked in assessing reasons for action, or beliefs about how the agent should go about making choices, [...] the agent’s long-term ends or goals he deems worthwhile or valuable, [...] deliberative principles the agent utilizes to arrive at practical judgments about what to do or how to act,” as well as the motivation to act on his accepted normative standards in pursuit of his accepted goals at least partly on the basis of his accepted deliberative and practical principles (350). It is an agent’s evaluative scheme that ensures a harmonious relationship among the agents MPAs, as well as between those MPAs and the decisions that the agent makes.

Feinberg’s second major idea is that there are two distinct stages in the development of autonomy. During the first stage, the person’s ability to govern his behavior is unrefined and

necessarily relies on certain external influences, which play a crucial role in facilitating the transition to the second stage. Once the person reaches the second stage, we say that he has matured to the point the he can now rightly be said to be “in control of himself.” Moreover, many of the interferences that were necessary during the first stage are now strictly prohibited, as they would influence the agent’s behavior in a way that controverts his now well-established authority over himself.

H&C adapt this idea by distinguishing between two kinds of evaluative scheme: *initial evaluative schemes*, which correspond to the first stage of development, and *evolved evaluative schemes*, which correspond to the second stage. The goal of development is to ensure a successful change from an initial scheme into an evolved scheme, by way of acceptable interferences during the initial stage of development.

The third and final idea regards determining what exactly counts as an acceptable interference. According to H&C, the acceptability of any particular interference depends on which stage of development the interfered-with agent is in. During the initial stage, interferences are acceptable to the extent that they facilitate the transition of the agent into a psychologically healthy and autonomous human being. Interferences during the evolved stage, however, are acceptable only if they were initiated by the agent’s own evaluative scheme or, after the right kind of deliberative activity, deemed allowable by the agent’s own evaluative scheme.

One important thing to note is that though agents in both stages can be autonomous, they cannot be autonomous in exactly the same sense. Agents in the initial stage of development act under the influence not only of their rudimentary motivational structures, but also of external forces that, if they are acceptable, promise to increase the agent’s autonomy or, at the very least, not do it damage. To be ‘autonomous’ during the initial stage, then, is to be something like

‘autonomous-insofar-as-future-autonomy-is-not-subverted’. This means that autonomy at this stage is possible only in a restricted and forward-looking sense. In the evolved stage, however, the assumption is that the agent governs himself in a fuller manner than does his initial-stage counterparts. ‘Autonomy’ in the evolved stage, then, is closer in meaning to ‘full autonomy’ or ‘autonomy *simpliciter*’, unrestricted by forward-looking considerations.

On H&C’s account, then, an autonomous agent is one who is either (1) in possession of an initial evaluative scheme and is being externally influenced, if at all, in ways that are acceptable in the sense described above; or (2) in possession of an evolved evaluative scheme that is the result of acceptable interferences upon the initial scheme.

Such are the details of H&C’s account. But what does it have to say about the difference between manipulated agents and instant agents? Manipulated agents such as Beth act on the basis of an evolved scheme that is the direct result of an unacceptable interference. In Beth’s case, the dean, in acting as he did, bypassed Beth’s evaluative scheme and, thereby, her chance to deliberate at all about the ensuing change.

Instant agents, on the other hand, have no such checkered past – indeed, they have no past at all. As such, it can hardly be said that they act on the basis of unacceptable influences on their earlier evaluative scheme. They simply come into existence with a full complement of MPAs and mature deliberative capacities, albeit without having spent any time actually exercising those capacities for the purpose of evaluating their motivational structure. Interestingly, this means that whereas a child might take years to get to the evolved stage (owing to their nascent psychological capacities), an instant agent may be able to make the transition in a matter of days or even hours. At any rate, we should say that instant agents, at the moment of

their creation, are in possession of an initial evaluative scheme and are at least possibly autonomous (again, in the restricted, forward-looking sense).

H&C's account, then, seems able to preserve the intuition that manipulated agents are not autonomous, and also to demonstrate along the way they are relevantly different from instant agents. Returning now to the main inquiry, my next task will be to highlight the various ways in which H&C's account and Buss's account are mutually supportive.

4.2 Bringing Together Buss and Haji & Cuypers

We've already seen why Buss would say that a manipulated agent like Beth is not autonomous: her already established psychic identity was disrupted by the Dean's team of scientists, and she was thereby prevented from functioning minimally well as a human being. This response is compatible with H&C's response which, as we've seen, is that Beth's own evaluative scheme (analogous to Buss's 'psychic identity') was either bypassed or replaced altogether, making the resulting evaluative scheme a nonautonomous one. And, if H&C accept Buss's HFC, then they could say that such bypassing or replacement undermines our autonomy *precisely because* it prevents us from functioning in the way that we need to in order to develop and maintain autonomy.

Buss and H&C also have compatible appraisals of instant agents. H&C are prepared to allow that instant agents are autonomous, but only in the restricted, forward-looking sense (since, once created, they possess an initial evaluative scheme that isn't obviously the product of unacceptable interference with an existing scheme). While Buss says nothing at all about instant agents, I believe a plausible case can be made on her behalf that would align with H&C's verdict.

Recall that the HFC states that psycho-physiological conditions undermine an agent's autonomy just in case they decisively influence the production of that agent's behavior and are themselves causes or symptoms of human malfunctioning. As explained above, it is easy enough to see how a manipulation could fall into the category of such undermining influences. But could the same be said about the instantaneous creation that is the cause of an instant agent's actions? It isn't obvious that to me that it could. After all, in the case of an instant agent, there is no 'self' or MPAs that exist prior to the creation event that could have been affected in any way at all; there is no psychic identity that could have been disrupted. Indeed, one might wonder whether cases of instant agency could pass as *manipulation* cases at all. At any rate, on the basis of what has been said, it is unclear whether creation *ex nihilo* could be considered an undermining influence in the sense relevant to Buss's account, or, if it could, whether it would be undermining in the same way that more paradigmatic forms of manipulation are. A response that Buss could make, then, is that instant agents are at least possibly autonomous not because it is so obvious that creation *ex nihilo* isn't in any way contributory to human malfunctioning, but because it isn't so obvious that it is. And this response would align well with what H&C have said on the matter.

Buss's HFC supplements H&C's account in that it explains not only *that* certain kinds of interference are unacceptable, but also more precisely *why* they are so. What remains to be seen, however, is whether H&C's account can, as I have suggested, complete Buss's response to manipulation cases by helping to explain how it might be possible for manipulated agents to "recover" from their manipulations. If this sort of explanation is possible, the combination of Buss's and H&C's accounts allows for a more comprehensive treatment of manipulation and instant agency than either of the accounts seem to have been able to provide separately. And at that point, I will have accomplished the goal that I set out in the opening.

How, then, can manipulated agents become autonomous again? Consider that the kinds of cases that we are trying to explain – non-EBA manipulated Beth, unacceptable child rearing, etc. – are cases in which the nonautonomous agent is in possession of a nonautonomous evolved scheme. An answer to our question, then, will likely begin with what has already been said about how evolved schemes can be acceptably changed in the first place. As a brief reminder, evolved schemes can be acceptably changed so long as such changes occur under the supervision of the agent's own evaluative scheme.

Now, given that the manipulation cases currently under investigation are of the non-EBA variety, we can safely assume that at least some of the agent's own MPAs will remain intact following the manipulation. Further, if the agent is to act autonomously again, then it will have to be under the influence of *those* MPAs. More specifically, it is those MPAs – those that are left unaffected by the manipulation – that will have to do the work of evaluating, revising, and otherwise interacting with the implanted MPAs. Otherwise, those new MPAs, as well as the actions they might produce, cannot be considered autonomous.

One might now worry that we have simply reverted back to the active strategy that we've been arguing against all along. What has just been proposed is that these new MPAs can only be considered autonomous if they were in some sense legitimized or validated by the agent who now possesses them. But this seems dangerously close to accepting again the problematic slogan of active accounts, which is that autonomous actions require agential sanction. The objection, then, is that perhaps passive accounts are just more convoluted versions of active accounts.

Remember that active accounts ground autonomy in certain kinds of actions (*e.g.*, reflective endorsements). That is, on active accounts, certain kinds of actions confer autonomy while other kinds of actions do not. In this way, actions play a direct role in establishing an

agent's autonomy. By contrast, passive accounts ground autonomy in some feature of the agent herself. On passive accounts, an agent's actions are obviously important, but they play only an ancillary role. They can help to establish an agent's autonomy only if that agent already satisfies the relevant passive requirement – which, in this case, is that the agent must be functioning minimally well as a human being. That is, if an agent's actions are autonomous, it will not be because they are sanctioning actions, but rather because they express her nature as a human being that is functioning minimally well as such.

Our objector might then reply that if this is what the passive strategy entails, then there is no reason why we couldn't keep our active accounts intact and supplement them with the notion that agential sanction is important not *as* a sanctioning action, but because it is one way in which we express our identity as, say, rational beings of a certain sort, rather than as human beings.⁵

This, however, would be an ineffective reply for a proponent of active accounts to make for at least two reasons. First, this reply concedes perhaps the main motivation for passive accounts, which is that active accounts provide a description of the requirements of autonomy that is at best incomplete and at worst incorrect. Second, it is unclear to me why we shouldn't regard the suggested account – one that emphasizes our nature as rational beings – as a passive account itself. If we should regard it as such, then it would be simply one among perhaps many passive alternatives to Buss's account, requiring its own motivations, supporting evidence, and solution to the incompleteness problem. And even if such an account could be filled out and adequately defended, it would not demonstrate that Buss's account is *wrong*. Again, it would simply present another viable option. A proponent of the active account who presses this reply,

⁵ Steven Weimer has argued in favor of a strategy very similar to this. See Steven Weimer, "Autonomy, Regress, and Manipulation," *Philosophia*, DOI: 10.1007/s11406-014-9552-8 (2014): 1-28.

then, either admits that his active strategy is incomplete or mistaken, or switches to defending a passive account himself.

Let us return now to the main inquiry. It seems that enough has been said to support the idea that becoming autonomous after manipulation is possible, as well as about what might be involved in that process. We might still wonder, though, about the manipulative techniques themselves, particularly about the strength or effectiveness of those techniques. It seems intuitive to say, for example, that a manipulation that targets 75% of an agent's MPAs is, all else being held equal, *stronger* or *more effective* than one that targets only 50%. And so perhaps the strength of a manipulation is a quantitative notion, dependent upon how much of an agent's motivational structure is affected. In contrast, it also seems intuitive to say that a manipulation that targets a small number of the agent's deeply-held and widely influential MPAs (e.g., moral commitments or rules) would be more effective than one that targets even a large number of less fundamental MPAs (e.g., beliefs regarding the agent's favorite football team). If so, then perhaps which particular MPAs are affected matters more for the effectiveness of a manipulative technique than does the ration of affected to non-affected MPAs. That is, perhaps there are qualitative differences among an agent's MPAs that bear on the effectiveness of any particular manipulative technique that might target them.⁶

What I'm driving at is that perhaps the effectiveness of any manipulative technique, or of any particular use of such a technique, can be defined as a complex interaction among the various changes made to the targeted agent's motivational structure, and also that effectiveness in this sense could be quantified and given something like a 'manipulative effectiveness' (ME)

⁶ Readers familiar with the free will debate will notice that this qualitative suggestion lines up with several prominent compatibilist theories of free action. For example, Alfred Mele has discussed the importance of what he calls "practically unsheddable values"; see his *Autonomous Agents*. Similarly, Susan Wolf has discussed the importance of the alignment between an agent's behavior and her long-standing commitments; see Susan Wolf, *Freedom Within Reason* (New York: Oxford, 1990).

value. Maybe *Ann/Beth*-style global manipulation has an ME value of 100, while an unpleasant and overbearing childhood might have an ME value of 40.

The implication of these musings is that a systematic way of talking about different manipulative techniques might enable us to, for instance, establish a threshold on the ME scale, such that agent's manipulated by techniques that fall above the ME threshold cannot recover, but those manipulated by techniques that fall below the ME threshold could. If so, then questions about how and whether an agent can recover from being manipulated may be inseparable from questions about each particular manipulation and the manner in which it affects the targeted agent.

On one final and interesting note, consider that in cases that would fall above the proposed ME threshold, it would be impossible for the featured agents to become autonomous again, but not *because* they were manipulated. Rather, it would be because their MPAs were compromised to such an extent that further attempts at becoming autonomous were precluded (owing to a severely disrupted psychic identity, or to an irreparably compromised evaluative scheme), which, in principle, could happen even if those agents were not manipulated. In cases that would fall below the proposed ME threshold, it would still be open to the agent to become autonomous again *despite* having been manipulated (given, of course, that enough of their MPAs are left over to facilitate the right kind of engagement with the implanted MPAs), meaning that the manipulation does not automatically preclude the eventual possibility of autonomy. In any case, it seems that there is nothing about manipulation *as such* that makes it especially threatening to personal autonomy. Manipulation represents just one of the many ways in which agents can fail to act autonomously.

Now, of course, is the wrong time to pursue these interesting topics in any more detail than we already have. In sum, then, I will say simply that based on what has been said, a plausible case can be built in favor of the idea that agents can, on some occasions, become autonomous again after being manipulated.

5 Conclusion: Toward a *Flourishing Account of Autonomy*

At the outset, I defined active accounts as those that adhere to the activity principle, according to which an agent is autonomous with respect to a given action only if the agent has appropriately sanctioned that action. I then demonstrated that this adherence to the activity principle makes active accounts susceptible to the regress problem. Passive accounts avoid the regress problem by eschewing the activity principle, but face in its place what I called the incompleteness problem. I then presented one recently defended passive account – Buss’s account based on the human flourishing condition – and suggested that it could avoid the incompleteness problem, though only if supplemented by a satisfactory account of the development of autonomy. In a discussion focused on manipulation cases, I argued that the development account put forward by Haji and Cuypers supplements Buss’s account especially well. My conclusion is that the two accounts together provide for a comprehensive treatment of manipulation cases, and thereby helps Buss’s account to avoid incompleteness problem.

Indeed, it seems that a conceptual link between autonomy and human flourishing can be defended, and that the theoretical groundwork for what could be called a *flourishing account of autonomy* has been laid. On such an account, minimal well-functioning is a necessary condition for the autonomy of human agents. While I haven’t offered a complete account of what minimal well-functioning consists in, we have seen that it will likely involve, at the very least, having a suitable psychic identity that comes about as the result of proper psychological development (or, perhaps, a psychic identity that doesn’t come about as the result of improper psychological development, as suggested by the case of instant agency).

Of course, there is much that has been left unsaid. For example, one philosophical task that is still incomplete is that of filling out what exactly is meant by a term like ‘human

flourishing’, as well as what it means to ‘function minimally well as a human being’ in respects relevant to autonomy. Buss is largely agnostic about specific answers to these kinds of questions (aside from her discussion of ‘psychic identity’). The next step for a proponent of the flourishing account would be to engage in a thorough exploration of this topic, which might include empirical inquiries into the traits and characteristics of autonomous agents, as well as into what kinds of circumstances are conducive to producing such agents.

Indeed, psychologists Edward Deci and Richard Ryan have been working earnestly on this very topic for years, and have developed a view that they refer to as Self Determination Theory (Deci & Ryan, 1985). On this view, personal autonomy is vitally important for human psychological well-being, and can be importantly affected by the myriad influences and environments that we experience, especially as we develop from children into adults. Many of the studies that investigate Self Determination Theory support several of the ideas that have been discussed here, such as the importance of autonomy to psychological well-being (Ryan, Deci, Grolnick, & La Guardia, 2006), as well as that certain environments seem more conducive to fostering autonomy than others (Ryan & Deci, 2001). Psychological work of this sort leads me to believe that a flourishing account of autonomy would have considerable empirical support. The flourishing account has much to recommend itself and is worthy of further research along the lines just mentioned.

REFERENCES

- Buss, Sarah. 2012. "Autonomous Action: Self-Determination in the Passive Mode," *Ethics*, Vol. 122: 647-691.
- Deci, Edward and Ryan, Richard. 1985. *Intrinsic Motivation and Self-Determination in Human Behavior* (New York: Plenum).
- Feinberg, Joel. 1986. *Harm to Self* (New York: Oxford).
- Frankfurt, Harry. 1971. "Freedom of the Will and the Concept of a Person," *The Journal of Moral Philosophy*, Vol. 68: 5-20.
- Haji, Ishtiyaque and Cuypers, Stefaan E. 2007. "Magical Agents, Global Induction, and the Internalism/Externalism Debate," *Australasian Journal of Philosophy*, Vol. 85: 343-371.
- McKenna, Michael. 2004. "Responsibility and Globally Manipulated Agents," *Philosophical Topics*, Vol. 32: 169-192.
- Mele, Alfred. 1995. *Autonomous Agents: From Self-Control to Autonomy* (New York: Oxford).
- Mele, Alfred. 2006. *Free Will and Luck* (New York: Oxford).
- Ryan, Richard and Deci, Edward. 2001. "On Happiness and Human Potentials: A Review of Research on Hedonic and Eudaimonic Well-Being," *Annual Review of Psychology*, Vol. 52: 141-166.
- Ryan, Richard, Deci, Edward, Grolnick, Wendy and La Guardia, Jennifer. 2006. "The Significance of Autonomy and Autonomy Support in Psychological Development and Psychopathology," in *Developmental Psychopathology: Vol. 1: Theory and Methods*, eds. Dante Cicchetti & Donald Cohen (New York: Wiley): 795-849.
- Valdman, Mikhail. 2011. "Autonomy, History, and the Origins of Our Desires," *Journal of Moral Philosophy*, Vol. 8: 415-434.