

AUTHENTICITY, IDENTITY, AND FIDELITY TO SELF

Sigurður Kristinsson

Faculty of Law and Social Science, University of Akureyri

sigkr@unak.is

ABSTRACT: An ideal of authenticity is deeply imbedded in western culture. Following Harry Frankfurt, authenticity might be characterized as fidelity to the person's essential nature. This suggestion fails, no matter whether personal identity is construed in a narrow, metaphysical sense or in the broader sense involved in talk of individual self-conceptions and social groups. Another Frankfurtian approach would appeal to self-reflective attitudes, but this falls prey to counterexamples as well. A third approach sees personal identity as a complex empirical fact. This fits the phenomena better, but leaves no room for fidelity to self. I conclude that the modern ideal of authenticity does not lend itself to a coherent and rationally compelling account, and will more profitably be seen as a mixture of logically unrelated concepts.

1. Introduction

When people's behavior genuinely expresses what they are like, it may be described as authentic. The nature and value of authentic behavior has been explained by associating it with personal identity (being/remaining oneself) and fidelity to self (being true to oneself). Consider the following characterization from Charles Taylor's *Ethics of Authenticity*:

There is a certain way of being human that is *my* way. I am called upon to live my life in this way, and not in imitation of anyone else's. But this gives a new importance to being true to myself. If I am not, I miss the point of my life, I miss what being human is for *me*.

[...] Being true to myself means being true to my own originality, and that is something only I can articulate and discover. In articulating it, I am also defining myself. I am realizing a potentiality that is properly my own.¹

If authenticity involves being true to one's "originality", as Taylor suggests, it will presumably also be a matter of avoiding blind conformism and undue deference. This is readily apparent in André Gide's literary expression of the value of being true to oneself:

The borrowed truths are the ones to which one clings most tenaciously, and all the more so since they remain foreign to our intimate self. It takes much more precaution to deliver one's own message, much more boldness and prudence, than to sign up with and add one's voice to an already existing party [...] I believed that it is above all to oneself that it is important to remain faithful.²

Gide suggests that we behave inauthentically because it is easier than "delivering one's own message". The result may be what D.W. Winnicott calls a "false self", a condition that David Velleman describes as "the paradigm case of inauthenticity":

¹ Charles Taylor, *The Ethics of Authenticity* (Cambridge, Ma: Harvard University Press, 1991), 28-9.

² André Gide, *Journals* ii 248. Cited by Leslie Green, "Sexuality, Authenticity, and Modernity" *Canadian Journal of Law and Jurisprudence* 8,1 (1995).

This person laughs at what he thinks he is supposed to find amusing, shows concern for what he thinks he is supposed to care about, and in general conforms himself to the demands and expectations of others. The motives that his behavior is designed to simulate are motives that he doesn't genuinely have.³

Velleman notes that such inauthenticity consists not in lack of self-guidance, but rather in the aim by which the agent guides himself, viz. that of imitating others and fulfilling their expectations. In making this his aim, the agent apparently fails to consult and express "himself" in some further sense. In Gide's terms, his "intimate self"—what he believes, what moves or touches him, and what he cares about—is not expressed in his behavior or in what he professes to believe and want. This perceived mismatch between outward expression and actual self generates our sense that the person's behavior is inauthentic.⁴

Various modern philosophers have regarded authenticity as an important ideal. Rousseau linked inauthenticity with living "in the opinion of others", and romantic thinkers like Herder and Humboldt took

³ David Velleman, "Identification and Identity" in Sarah Buss and Lee Overton (eds.) *Contours of Agency: Essays on Themes from Harry Frankfurt* (Cambridge, Ma: MIT Press 2002) 91-123, p. 97.

⁴ Admittedly, Taylor's analysis of the self and human agency suggests that inauthenticity involves not a mismatch between outward expression and actual self, but rather a state where the self to be expressed exists only potentially until the individual attempts to express or "articulate" it. Taylor's views will be dealt with in sections 3 and 5 below.

authenticity to be a matter of expressing one's individuality.⁵ Mill notes his indebtedness to Humboldt in *On Liberty*, where he stresses the importance of individuals' forming their own purposes, as opposed to having their individuality stunted by fear of sanctions from public opinion.⁶ These diverse authors insist on the value of individuals' conducting themselves based on their own substantive beliefs, desires, emotions, and ideas, rather than dismissing or repressing them in favor of what they think is acceptable to others.

It is hard to disagree with Taylor's claim that an ideal of authenticity is deeply imbedded in modern (western) culture. The aim of this paper is to see whether this persistent ideal can be given a coherent and rationally compelling characterization. To do this, three theoretical approaches are explored, two of them directly inspired by Harry Frankfurt's moral psychology. As a first approximation, then, it might be suggested that authenticity is a matter of preserving or expressing the person's essential nature. I argue that this suggestion fails, no matter whether personal identity is construed in a narrow, metaphysical sense or in the broader sense involved in talk of individual self-conceptions and social groups. Another Frankfurtian approach would appeal to the person's self-reflective attitudes or "identifications", in order to arrive at the phenomenon hinted at by

⁵ See Kasper Lippert-Rasmussen, "Identification and Responsibility" *Ethical Theory and Moral Practice* 6 (2003) 349-376, p. 369.

⁶ Individual human beings realize their potential, according to Mill, through exercising their capacity for autonomous choice while "he who lets the world, or his own portion of it, choose his plan of life for him, has no need of any other faculty than the ape-like one of imitation." *On Liberty, The Collected Works of John Stuart Mill*, ed. John M. Robson (Toronto: University of Toronto Press, 1963-91), vol. 18, p. 262.

Gide's phrase "the intimate self". I argue that such attempts fall prey to counterexamples as well. Finally, some theorists consider a person's identity to be a complex fact about her physiology, character, social relations, and self-reflective attitudes. This account avoids many of the difficulties with Frankfurt's theories and can better explain the phenomena of authentic and inauthentic behavior, but it does so by sacrificing the possibility of fidelity to self as a reasonable ideal. I conclude that despite its persistent appeal, the modern ideal of authenticity does not lend itself to a unified and rationally compelling account, and will more profitably be seen as a mixture of logically unrelated concepts.

2. Authenticity as obedience to one's essential individual nature

As a first approximation, authenticity might be construed as being motivated by something that "genuinely belongs" to the agent.⁷ A desire for a cigarette might be truly attributed to an agent who also had the opposite desire not to smoke, and yet, one of these desires may be seen as belonging to her in a more genuine way than the other. If the desire not to smoke thus genuinely belongs to the agent, it provides her with a special kind of reason that the desire for a cigarette does not: By smoking, she would be going against the desire that represents her agency. Authenticity might then be thought of as acting according to elements that belong to the agent in this genuine way.

In virtue of what might certain desires "genuinely belong" to an individual agent? An especially definitive answer would be that some

⁷ This phrase is intended to be interchangeable with membership in the "intimate self", as referred to by Gide.

desires belong to the agent's essential nature, while others don't.⁸ This view has been argued for by Harry Frankfurt. He asks us to consider that a person's will may be structured in such a way that it is impossible for her to perform certain actions intentionally. For example, it may be impossible for someone to hurt the object of her love intentionally, or to sever her relationship with it. This may be impossible for that person even after she has judged that she has overwhelming reason to do so; she may even have formed the intention to sever the relationship only to discover, when the chips are down, that she cannot bring herself to do it.⁹ Frankfurt calls this phenomenon "volitional necessity", arguing that it is not mere compulsion because "the essence of a person...is a matter of the contingent volitional necessities by which the will of the person is as a matter of fact constrained."¹⁰ When the will is constrained by love, for example, "the authority for the lover of the claims that are made upon him by his love is

⁸ This status would seem to endow such desires with special authority as reasons for action for the particular individual in question. This answer is not to be confused with Kant's view that certain reasons are authoritative in virtue of being commanded by ourselves as rational agents. Kant did not think we can locate that authority by looking (or "feeling") among our empirical desires, but sought to establish it by deriving substantive practical principles from a self-legislation inherent in practical rationality as such. By contrast, authenticity may be thought to involve commands made by each person's contingent individual nature, even though such commands would certainly be heteronomous in Kant's sense.

⁹ "Identificaton and Wholeheartedness" *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press 1988), p. 175-176.

¹⁰ Ibid.

the authority of his own essential nature as a person.”¹¹ “In betraying the object of his love, he therefore betrays himself as well”.¹² On this account, volitional necessity is seen as the will’s defense against self-betrayal, while at the same time revealing the person’s essential nature.

Do we have personal essences that reveal themselves through volitional necessity? David Velleman suggests, debunkingly, that our belief in the special authority of desires that are experienced as authoritative may be the result of wishful thinking: “We’d like to have motivational essences, and so we’re happy to agree when someone says that we do.” Perhaps “Frankfurt’s conception of the self appeals to us only because its implicit ideal represents us as we wish we could be.”¹³ If Velleman is right, being true to oneself cannot be construed as fidelity to one’s essential nature because our belief in such a nature is unfounded. In order to assess the force of Velleman’s challenge, we need to consider what it might mean to speak of personal essences, and then see whether any coherent notion of personal essences could provide a plausible ground for an ideal of authenticity.

3. Narrow personal identity

“The essence of a person”, Frankfurt explains, “pertains to the purposes, the preferences, and the other personal characteristics that the individual cannot help having and that effectively determine the activities of his

¹¹ “Autonomy, Necessity, and Love” *Necessity, Volition, and Love* (Cambridge: Cambridge University Press 1999) 129-141, p. 138.

¹² *Ibid*, p. 139.

¹³ “Identification and Identity”, pp. 99-100.

will.”¹⁴ This claim would result in a very narrow criterion for inclusion in the “intimate self” to be expressed in authentic behavior. To qualify, a personal characteristic would need to be both unsheddable and determining of voluntary action. The second necessary condition is especially troublesome, because it would imply that inauthenticity is practically impossible. If it is a defining characteristic of elements of the intimate self that they “effectively determine the activities of the will”, then we will never succeed in bringing ourselves to act inauthentically, because success would mean that the action was not contrary to our intimate self after all.

Perhaps the unsheddability condition would fare better all by itself. The intimate self is then seen as consisting of personal characteristics that are essential to the person’s nature in virtue of being unsheddable. This is an odd way of talking about essences, however. One would expect that a personal characteristic belongs to a person’s essential nature just in case she would not be the same person without it. This is how talk of essential natures translates into talk about identity, which Frankfurt does not shy away from.¹⁵ But the mere fact that I cannot change a personal characteristic of mine does not mean that I would not be the same person without it. Whether I can change it and whether I would be the same if it were to change are simply two separate questions. The notion of unsheddability is therefore not promising.

¹⁴ “Autonomy, Necessity, and Love”, p. 138.

¹⁵ In the discussion referred to here, Frankfurt uses “the essential identity of an individual” interchangeably with “the essence of a person”. “Necessity, Volition, and Love” p. 138. See also his remark that “a person can have no essential nature or identity as an agent unless he is bound with respect to that very feature of himself—namely the will—whose shape most closely coincides with and reveals what he is.” *The Importance of What We Care About*, ix.

Although Frankfurt's own explanation of personal essence is thus not helpful, an alternative explanation might enable us to test the hypothesis that the "intimate self" consists in that which is essential to a person's identity. Let us consider, then, what it could mean to say that I would not be the same person if something about me were to change.

Identity can be either numerical or qualitative, and there are two corresponding interpretations of the statement that a person *P* would not be the same person without the characteristic *c*. I will argue that while the numerical interpretation would make better sense of why self-betrayal is not rational, it is only the qualitative one that has a chance of fitting the constellation of beliefs and values associated with authenticity. The qualitative interpretation, however, does not help us draw a principled distinction between characteristics that do and do not belong to the "intimate self". If a notion of identity is to be helpful at all, it must be construed as something other than numerical or qualitative identity.

A person's numerical identity is the fact that she is one person. Synchronic personal identity is the fact that an entity (e.g. a human body) constitutes one person at a given time, as opposed to many or none, while diachronic personal identity is the fact that an entity constitutes one person at two different times. What is essential to a person's nature, in this sense, is that without which she would not exist as one person. On this interpretation, the survival of the person is logically incompatible with the alteration or removal of that which is essential to her nature, even though it

might mean that another person (or perhaps more than one) comes to occupy what used to be her body.¹⁶

We have an obvious interest in the preservation of our numerical identity, both synchronic and diachronic. Being one person is very valuable to us, and so is our continued existence. If the value of authenticity were grounded in this interest, it would therefore be easy to explain its normative force. But this interpretation is inconsistent with what is typically at stake when the value of authenticity is invoked. It should be clear from our initial characterization of inauthenticity as living according to “borrowed truths” that whatever the authentic person is supposed to be true to, it is not her numerical identity. Admittedly, there is room for debate concerning whether the preservation of numerical identity requires the persistence of substantive qualities, for example those that are fundamental to the person’s self-conception. One view holds that alteration in certain highly distinctive and personally valued characteristics would amount to the end of the person, while another view holds that a person could in principle survive the alteration of any of her substantive characteristics.¹⁷ Even on the former, more demanding view, however, living according to borrowed truths is not a threat to continued existence unless the person already happens to be deeply attached to, or distinctively characterized by, the value of authenticity. This would make nonsense out of the idea of urging people to care more about their originality in the name of authenticity,

¹⁶ This is of course the sense of personal identity addressed by Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984) and the papers collected in John Perry (ed.) *Personal Identity* (Berkeley: University of California Press, 1975).

¹⁷ Velleman attributes the former view to Parfit and the latter to Locke, and comes down on Locke’s side himself, in “Identification and Identity”, pp. 96-97.

understood as fidelity to their essential nature. The intended audience of such exhortations consists precisely of those who have no reason to think of inauthenticity as a threat to their survival, nor as non-expressive of their essential characteristics. Living according to borrowed truths is therefore not a threat to numerical personal identity, nor does it involve a failure to express the characteristics, without which the inauthentic conformist would no longer be the same person.

While this argument works in relation to diachronic identity, inauthenticity might still be seen as a threat to synchronic identity. Synchronic personal identity is the fact that the entity in question is one person and not none or many. It is threatened, then, by any prospect of the disintegration or multiplication of the person. Leaving prospects of multiplication to one side, we may consider whether inauthenticity threatens to eliminate any substantive or structural characteristics necessary for personhood. To be sure, pleas for authenticity tend to emphasize, as in Mill, that those who live according to “borrowed truths” neglect their important capacities, such as the capacity for discrimination, reasoning, and choice, for example. Insofar as the invoked capacities are considered necessary for personhood, it might be thought that inauthenticity threatens to eliminate the person. However, personhood hardly requires the robust exercise and development of such capacities, as opposed to their mere possession. Even thoroughgoing neglect, through habitual deference and self-effacement, would hardly turn the conformist into a non-person. This holds also on Frankfurt’s view that personhood essentially requires second-order desires; a person might reflect on her first order desires based on her

second-order desire to want whatever is in line with the “opinion of others”.¹⁸

Taylor may be fishing in these same waters when he claims:

Our identity is [...] defined by certain evaluations which are inseparable from ourselves as agents. Shorn of these we would cease to be ourselves, by which we [...]mean that] we would lose the very possibility of being an agent who evaluates; that our existence as persons, and hence our ability to adhere as persons to certain evaluations, would be impossible outside the horizon of these essential evaluations, that we would break down as persons, be incapable of being persons in the full sense.¹⁹

This statement is in fact ambiguous between the claim that the actualized capacity for making a certain type of evaluation is a necessary condition for personhood, and the further claim that if the content of (some or all of) a person’s evaluations were to change, she would no longer be the same person. The latter claim is about diachronic identity, and is irrelevant to an assessment of the predicament of the inauthentically behaving conformist; his problem is not that of a fundamental shift in evaluations but rather of either not having or not abiding by his own evaluations. The former claim, however, about synchronic identity, is relevant because it might be claimed that the conformist does not make the requisite sort of

¹⁸ See more on this possibility in the section on “Identification theory” below.

¹⁹ “What is Human Agency?”, *Human Agency and Language: Philosophical Papers*, vol. 1 (Cambridge: Cambridge University Press, 1985) 15-44, pp. 34-35. [Originally published in T. Mischel (ed.), *The Self: Psychological and Philosophical Issues* (Oxford: Basil Blackwell, 1977)].

evaluation and therefore is not a person “in the full sense”. I have already questioned whether personhood should be taken to require the actual exercise of whatever capacities are thought to be crucial, as opposed to their mere possession. But even if we were to accept Taylor’s demanding view of personhood, why think the inauthentically behaving conformist necessarily fails to make evaluations? It seems quite possible to imagine that he values the role of the obedient conformist, articulating it through a “vocabulary of worth”,²⁰ even if it means that he will often be trying to follow prescriptions that run counter to spontaneous motives, furnished by emotions and other mental states that are actually embedded in his psychology.²¹ The result is inauthentic behavior by a Taylorian person who has clearly not lost “the very possibility of being an agent who evaluates.”

Frankfurt has offered a different argument that might support the idea that a person’s conformistic, inauthentic behavior sacrifices her synchronic self-identity. He argues that the absence of volitional limits “gives rise to a diminution, or even to a dissolution, of the reality of the self.”²² An agent who would be capable of intentionally doing just anything, given that she thought she had a reason to, would not have a will with a “determinate character”, according to Frankfurt, and this would weaken the sense in which she had a self at all. Now, the conformist may be seen as a person

²⁰ Ibid., p. 24. This is “the language of higher and lower, noble and base, courageous and cowardly, integrated and fragmented, and so on.” (24). Such articulation is the hallmark of the type of evaluation Taylor hints at in the previous quote, and usually refers to as “strong evaluation”.

²¹ An example of this sort, the “moviegoer example”, is discussed in some detail two sections below.

²² “Rationality and the Unthinkable”, *The Importance of What We Care About*, 177-190, p. 179

who is prepared to do anything, given that the “opinion of others” demands it; her own substantive beliefs and desires won’t stop her. So Frankfurt’s argument may suggest that those who fail to tend to their own originality are in fact undermining their selfhood. Even if their conformism doesn’t turn them into non-persons, it turns them into persons without determinate selves.

In order for this argument to be convincing, it would need an explanation of what it could mean to be a person without a self. In the absence of such an explanation, we must stand by the arguments already given for the view that conformism doesn’t threaten either diachronic or synchronic personal identity. Still, it tends to affect them in a way that arguably makes them less valuable. A paragon of Millian individuality may enjoy his personhood more than the conformist, or make better use of it as a source of meaning in his life. An emotionally robust and willful exemplar of Frankfurtian caring will also perhaps make more of his selfhood than the malleable self-effacer. Such claims need to be argued for, however, and the arguments can not be based on the claim that the conformist is betraying his essential nature; for this claim is not supported by any reasonable understanding of what it would mean to have an essential personal nature.

So much for numerical identity. Now, what about qualitative identity? Are the boundaries of the “intimate self” drawn by the person’s qualitative identity? Two things are qualitatively identical if they are exactly similar, i.e. if they have exactly the same qualities. It is not as clear what it would mean to speak of the qualitative identity of one thing only. It could be the entire collection of qualities possessed by that thing at a given time or period, or it could be some subset thereof. How do we draw that line? One “thing” can fall under different concepts and we may group the qualities it possesses accordingly. For example, a living human being has

certain characteristics *qua* material object, others *qua* living organism, still others *qua* member of the human species, *qua* person, and so on. The qualitative identity of a person may then perhaps be thought of as the collection of qualities she possesses *qua* person. These will presumably not include her weight or skin color, but instead qualities like her beliefs, desires, emotions, and character traits. While this way of drawing the boundary seems to provide an acceptable explanation of a person's qualitative identity, it will not do as a demarcation of the intimate self. Characteristics external to the intimate self are presumably still qualities the person possesses *qua* person, but by hypothesis, the current explanation of qualitative identity takes it to include all such qualities.

4. Broad personal identity

A person possesses various qualities *qua* member of a group, including nationality, race, gender, sexual orientation, age, occupation, etc. This brings us into an arena where the word 'identity' is widely used, but clearly in a different and broader sense than what we have been considering so far. In this broader sense, a person may be described as having several different identities, understood as membership in different groups and the characteristics that constitute such membership. This also brings us into the arena of discussions of identity as those aspects of the person that are basic to either her self-conception or to her ideals for herself.

As an example of this broader use of 'identity', consider the following definition of authenticity, offered by Michele S. Moses:

Authenticity [...] is defined here as a state of being within which one has the ability to act in keeping with one's true self, that is, to make uncoerced

choices and to feel public affirmation of one's personal identity, of which one's cultural identity is a central part.²³

The reason why public affirmation is so important, according to Moses, is that otherwise, the person's self-definition becomes false or skewed:

We must define ourselves not just within the sometimes-damaging categories society might have assigned to us but within a secure personal and cultural structure.²⁴

Moses argues that in order to develop what she calls an "authentic identity", one needs a social context that affirms the value of the the relevant identity or identities. Speaking of identity adjustments of immigrants, she states:

What leads to inauthentic identity is a social context within which only the dominant culture is affirmed as worthwhile. In order to define one's authentic identity, one always must take into account a personal history and background—what one has come to believe is significant, especially distinct from others.²⁵

²³ Michele S. Moses, "The Relationship between Self-Determination, the Social Context of Choice, and Authenticity" *Philosophy of Education* 2000: 294-302; p. 297.

²⁴ Ibid.

²⁵ Ibid. p. 298. Joel Anderson and Axel Honneth similarly emphasize the importance of social conditions that allow for the recognition of people's different (cultural) identities ("Autonomy, Vulnerability, Recognition, and Justice" in to John Christman and Joel Anderson (eds), *Autonomy and the Challenges to*

When we speak of identity in this sense, we are not referring to numerical identity, and not exactly to qualitative identity either, although that is clearly involved. Instead, we seem to be referring to a relationship between the person's qualitative identity, existing collective identities, and her own self-conception or self-definition. Authenticity here tends to be associated with having a self-conception that does justice to what one is really like—an accurate self-conception—while inauthenticity is associated with a skewed or false self-conception.

There is still little indication of what an accurate self-conception is supposed to track, and this leaves us with the very same problem we started with, namely that of explaining the difference between our “intimate selves” and other aspects of ourselves. I have argued that this distinction is not explained by an appeal to personal essences or personal identity in the narrow sense. We have also seen, however, that there is a broader sense of ‘identity’ that might be appealed to. I now want to consider two theoretical alternatives for trying to understand authenticity, and what it means to be true to oneself, along the lines of identity in this broader sense. The first is “identification theory”, which emphasizes the authority of the person's reflexive attitudes, and the second is “trait theory”, which emphasizes the individual's objective and (in principle) empirically discernable features. I will argue that while each alternative illuminates, neither yields a plausible account of the modern ideal of authenticity.

Liberalism, pp. 127-149). They view such recognition as a necessary support for autonomy, but more precisely, their concern is presumably with the authenticity condition for autonomy rather than with the capacity condition, to borrow Christman's terminology (see note 6 above).

5. Identification and hierarchy

According to hierarchical theories of the self, a first-order desire belongs to the self in an especially strong sense if it is endorsed by a higher-order desire.²⁶ Can a hierarchical approach yield a plausible account of the “intimate self” that is supposedly tracked by an accurate self-conception?

Following Frankfurt, a hierarchical account of authenticity could state that the intimate self consists of desires (in the broad sense) with which the agent identifies. The agent identifies with desires he wants to have,²⁷ and he does not identify with desires he wants not to have. He may also have desires he has no reflexive attitude toward (e.g. because he is unaware of them), and desires he has mixed attitudes toward. Frankfurt calls the former “wanton” desires, and implies that they do not belong to us *qua* persons—hence presumably not to our intimate selves. The latter are

²⁶ See Harry Frankfurt, “Freedom of Will and the Concept of a Person”, *Journal of Philosophy* 68 (1971) 5-20; Gerald Dworkin, “Acting Freely,” *Nous* 4 (1970) 367-83, “Autonomy and Behavior Control”, *Hastings Center Report* 6 (1976): 23-28, and *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press 1988); and Wright Neely, “Freedom and desire” *Philosophical Review* 83 (1974) 32-54. There are important differences between these versions, and Frankfurt and Dworkin have amended their theories over time. For present purposes, however, I take Frankfurt’s version as representative of the hierarchical approach.

²⁷ We can ignore the distinction between wanting to have a desire and wanting the desire to be effective in bringing the person to act. This distinction matters for the topic of agency, but not when the question is which desires are part of the “true self”.

more problematic, for the higher-order attitudes that are supposed to determine their rank are themselves divided. With respect to such desires, the agent lacks what Frankfurt calls ‘wholeheartedness’ and defines as satisfaction or the absence of higher-order ambivalence.²⁸ Frankfurt now argues that such wholeheartedness is precisely what identification consists in.²⁹ The desires the agent is ambivalent toward should therefore not be viewed as part of the intimate self, on Frankfurt’s view. Accordingly, Frankfurt’s account implies that the intimate self is comprised only of the desires the agent is satisfied in wanting to have.

If identification demarcated the “intimate self” neglected by Winnicott’s person with a “false self”, it should not turn out to be possible for him to identify with his desire to conform himself to the expectations of others. But this does in fact seem quite possible. Imagine a thoroughgoing conformist, who wholeheartedly identifies with his desire (or tendency) to imitate a certain group of people, his idols. Now suppose this person is brought to laughter, tears, and other genuine emotional responses by watching movie A, and is utterly unmoved by movie B. However, he is also

²⁸ “Consider a person who believes something wholeheartedly, who is wholehearted in some feeling or attitude, or who intends wholeheartedly to perform a certain action. In what does his wholeheartedness with respect to these psychic elements consist? It consists in his being fully satisfied that they, rather than others that inherently (i.e., non-contingently) conflict with them, should be among the causes and considerations that determine his cognitive, affective, attitudinal, and behavioral processes” (“The Fainted Passion”, *Necessity, Volition, and Love*, p. 103).

²⁹ “Identification is constituted neatly by an endorsing higher-order desire with which the person is satisfied” (“The Faintest Passion”, p. 105). I take this to imply that wholeheartedness is a necessary condition for identification.

convinced that his idols would have exactly the opposite experiences. Furthermore, movie A (unlike movie B) appeals to our conformist at a more cognitive level, giving him a satisfying sense of a new perspective on something he cares about, but at the same time, he is quite sure that his idols would have no such experiences upon watching A. Now, if this wholehearted conformist just went by the strength of his first-order desires, he would praise movie A, see it again, or see other similar movies. Being a wholehearted conformist, however, volitionally bent upon imitating his idols, he reflexively hates his desire to see this sort of movie (and the fact that he is frequently assaulted by thoughts about it) and loves his desire to see movies like B, however weak that desire may be on its own. His desire to see movies like B may indeed be fuelled only by his desire to imitate his idols, and not at all by the quality of his experience while watching the movie. Nevertheless, his higher-order desires unequivocally favor it over the desire to see movies like A. If he praised movie B, went to see it again, tried to laugh at its jokes, and sought out similar movies, he would be following his wholehearted desire and yet, his behavior would be stereotypically inauthentic. Conversely, he would behave authentically if he praised movie A and went to see similar movies, allowed himself to savor thoughts about it and discuss it with others and yet, he would not be following his wholehearted desire and thus not acting on a motive with which he identifies, given Frankfurt's account. Wholeheartedness of motivating desire therefore appears to be neither necessary nor sufficient for authenticity of behavior.

Another reason to reject the identification approach (as exemplified by Frankfurt) is that even if we allowed that wholehearted identifications created membership in the intimate self, the account would not be relevant to the person who was trying to establish an "authentic identity" in Moses'

sense. It would presumably tell such a person to define herself not based on what is considered acceptable by available ideologies or social categories, but rather based on what it is about herself that she is satisfied with. She needs to remember, for example, that when there is a conflict between the qualities in her that are valued by her surrounding culture and the qualities that she herself is satisfied with, it is the latter that she needs to favor, in order to be true to herself.

The problem is that if the person is in fact struggling with an identity question, it is because she already feels torn or ambivalent at the level of higher-order desires. Telling her to follow the desires she endorses wholeheartedly will therefore not be of much help. We do not typically start to worry about “who we are” until we experience ambivalence in our self-conception, and this is the very opposite of wholeheartedness. As questions of identity emerge, it seems, Frankfurt’s formula for answering them becomes decreasingly relevant, because we ask such questions precisely with regard to those parts of ourselves that we do not already endorse (or reject) wholeheartedly. When we are wholehearted (as the wholehearted conformist in the example above), we are not in doubt about who we are or what we stand for, even if we may have to engage in some heavy internal fighting against our unruly first-order desires.

We should reject the identification approach, then, both because conformistic desires might be wholehearted (and thus identified with), and also because the question of identity arises just when it seems least likely that the agent has already identified with a relevant desire. This conclusion is not due to the peculiarities of Frankfurt’s account of identification. On Michael Bratman’s competing account, for example, identification with a desire is a decision, with which the agent is satisfied, to treat the desire as a

reason for action.³⁰ Nothing rules out that this might apply to conformistic desires. Also, the person who is struggling to find his “authentic identity” will not find this answer any more helpful than Frankfurt’s, because his problem is precisely that he has not made such a decision concerning which desires, highlighted by different aspects of his self-conception, he should treat as a reason for action. The self, to which he is struggling to be true, is not constituted by his identifications. If identification enters the picture at all it is because he is concerned to be true to himself in his identifications.

There are variants of the hierarchical approach that do not take identification to be essential to membership in the “intimate self”, and might therefore avoid the current problem of demandingness.³¹ These variants will also run into problems, however, when asked for a criterion for the “intimate self” that might be of help to the person struggling with an identity question. They replace identification with some higher-order attitude that is typically involved in self-reflection. It may be the agent’s evaluations, as in Gary Watson’s account, or “strong evaluations”, as in the account offered by Taylor. David Brink similarly presupposes in a recent article that “authenticity requires acting on the ideals that the agent reflectively and sincerely accepts at the time of action.”³² The problem with all these proposals is that reflectively accepted ideals and evaluations can

³⁰ Michael Bratman, “Identification, Decision, and Treating as a Reason” *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), 185-206.

³¹ Gerald Dworkin, *Theory and Practice of Autonomy*, Charles Taylor “What is Human Agency?”, Gary Watson “Free Agency” in *Free Will*, ed. Gary Watson (Oxford: Oxford University Press, 1982) 96-110.

³² “Prudence and Authenticity: Intrapersonal Conflicts of Value” *Philosophical Review* 112 (2003) 215-245, p. 215.

be deeply opposed to features of the individual that intuitively have at least as much claim to membership in his intimate self. Therefore, the agent who manages to act according to the sincere deliverances of his self-reflection need not behave authentically.

Let me substantiate these claims a bit further. The person struggling with an identity question is typically torn between socially available categories and ideologies on the one hand, and her own experiences and emotions on the other. The threat to authenticity Moses identified was “a social context within which only the dominant culture is affirmed as worthwhile.”³³ If the dominant culture resonates badly with the person’s own emotions and experiences, the ideal of authenticity demands that she should not simply follow or adopt the value-judgments implicit in the dominant culture, but rather give her own emotions and experiences their due. Definitions of authenticity as fidelity to the agent’s evaluations and ideals are here faced with the possibility that an agent might make evaluative judgments and reflectively accept ideals that represent the dominant culture and do not sit well with her own emotions and experiences. It is indeed in cases of this sort that questions of identity arise most acutely, because the self-alienated agent is divided not only against society but also against himself.

Consider a gay teenager who has been raised in an exclusively homophobic social environment. The norms and conceptualizations he has been exposed to, concerning homosexuality, are thoroughly inconsistent with any positive evaluation. As this teenager comes to experience his own sexuality, it will inevitably be in conflict with the conceptual and social

³³ “The Relationship between Self-Determination and the Social Context of Choice”, p. 298.

resources available to him for self-reflection. We can imagine that he genuinely grows to believe that homosexuality is evil and comes to experience his own homosexual leanings as a mere collection of urges that are not “really him”. The more he reflects, the more secure he feels in the judgment that his way of experiencing sexuality is contaminated by evil, and constitutes an ailment that needs to be suppressed or eradicated. To be true to the ideals he sincerely and reflectively accepts, he must make sure he does not express his homosexuality through his behavior; instead he must fight it as hard as he can and seek all the help he can get in doing so.

This line of behavior would hardly amount to authenticity. Recall Velleman’s description of the inauthentic person who “laughs at what he thinks he is supposed to find amusing, shows concern for what he thinks he is supposed to care about.” The gay homophobic will similarly try to enjoy heterosexual relations, to be aroused when he thinks he should, and to show contempt for what he believes is evil. But these patterns of behavior will not be appropriately supported by his feelings. In other words, “the motives that his behavior is designed to simulate are motives that he doesn’t genuinely have.” He will have a “false self” in Winnicott’s sense, and his self-definition will be a prime example of “inauthentic identity” in Moses’ sense.

There is a perfectly straightforward sense of valuing that permits that we do not care deeply about what we value. The gay homophobic may persistently believe heterosexuality is better than homosexuality, have some desire to promote heterosexuality, and consider sexual orientation to be a matter of importance. These assumptions strongly suggest that he values

heterosexuality.³⁴ At the same time, homosexuality may be deeply imbedded in his psyche. It may reveal itself in his fantasies, emotive responses, and feelings of satisfaction and dissatisfaction in various situations. More ominously, it may perhaps reveal itself in the psychological consequences of repression as, for example, his repressed desires provide fuel for his hatred toward homosexuals (including the aspects of himself he is trying to expel). It is therefore not plausible to object that the rift between the gay homophobic's values and his emotions means that these aren't his "genuine values", that he cannot really believe homosexuality is evil if he also feels it residing in his body, that to be true to himself, he would need to find what values or ideals he "genuinely accepts", and then be guided by them. Each of the actual accounts mentioned above permits that the individual might experience a certain amount of alienation toward his own evaluations. Watson notes that it is possible to recoil from one's own sense of sin,³⁵ and Taylor remarks that actual (strong) evaluations are typically tentative, mere yardsticks that fail to capture our deepest sense of what is important.³⁶ Each author thus acknowledges that a person might to some extent feel dissatisfied with the evaluations he sincerely makes. They should therefore allow that the gay homophobic might genuinely condemn homosexuality while also feeling somewhat dissatisfied with that condemnation. Consider also, however, that strengthening the account of evaluation so as to exclude the possibility of alienation would not only result in an implausible and *ad hoc* account of

³⁴ See Insoo Hyun's characterization of values in "Authentic Values and Individual Autonomy", p. 197.

³⁵ "Free Action and Free Will", *Mind* 96 (1987) 145-72.

³⁶ "What is Human Agency?" p. 42.

evaluation, but also in fact fail to solve the problem. For such a move would invite the difficulty already noted with Frankfurt's account of identification, that the result would not be of use to an ambivalent person struggling with an identity question. The self-alienated gay homophobic would have no intimate self to be true to if the intimate self could only contain elements endorsed by non-alienated value judgments. I conclude that hierarchical accounts of the self, with or without an identification requirement, will not provide a satisfactory explanation of what it is that the person who acts authentically is supposedly true to in his or her behavior.

6. Empirical identity: The trait theory

The examples of the gay homophobic and the wholeheartedly conformistic moviegoer indicate that there is more to a person's identity than what she herself endorses or values. This realization is central to an account of identity offered by Amelie Rorty and David Wong, according to which "a person's identity is constituted by a configuration of central traits."³⁷ Traits, according to Rorty and Wong, are "dispositions to beliefs, desires, habits, attitudes, and actions".³⁸ Traits can become central to a person's identity in various ways, including, but not confined to, its "subjective appropriation".³⁹ The other forms of a trait's centrality are: Objective ramification, contextual or regional ramification, resistance to the person's

³⁷ Amélie Rorty and David Wong, "Aspects of Identity and Agency", in Owen Flanagan and Amelie Rorty (eds.) *Identity, Character, and Morality* (Cambridge, Ma: MIT Press, 1990), 19-36; 19.

³⁸ *Ibid.*, p. 20.

³⁹ *Ibid.*

own efforts to change, social ramification, dominance in situations of coping with stress or conflict, and domination during conflicts with other traits.⁴⁰ Rorty and Wong further distinguish between different kinds and features of traits that all can be central to a person's identity. Traits can be somatic, temperamental/psychological, acquired through social role casting or membership in socially defined groups, and influenced by the person's ideals for herself.⁴¹ Rorty and Wong suggest that while philosophical arguments can be offered to persuade us to value certain aspects of our identity rather than others, our actual identity is always a complex configuration of central traits that often conflict with each other.⁴²

Suppose we adopt Rorty and Wong's view of identity. Would that provide a plausible understanding of authenticity and fidelity to self? The understanding would be that the person who acts inauthentically fails to express his central traits. For instance, he may try to act contrary to his temperament and hide his feelings. His bodily movements may also be contrived and contrary to his somatic dispositions. These would seem to be instances of inauthentic behavior. A person may also value, and think himself characterized by, a trait that is hardly measurable along any of the objective dimensions of centrality. Such mismatch between what is subjectively appropriated and objectively instantiated is indeed present in our examples of the gay homophobic and the wholehearted conformist, and it generally seems to be part of what happens when individuals or groups have traits that are not socially valued, and as a result risk developing what

⁴⁰ Ibid.

⁴¹ Ibid., pp. 21-26.

⁴² Ibid., pp. 32-34.

Moses called “inauthentic identity”. This understanding of identity therefore seems to succeed where the hierarchical accounts failed.

An apparent problem with the trait theory is that it does not provide a principled way of telling what fidelity to self consists in when central traits conflict. What a person values may conflict with what she idealizes unselfconsciously, which in turn may conflict with traits that are reinforced through her role casting, which in turn may conflict with her temperamental traits, and so on. While trait theory may help us understand the nature of such conflicts, and even point toward ways of alleviating them, it does not offer any particular analysis of what it means to be true to oneself in such situations. Are all aspects of identity equally important? Consider Rorty and Wong’s example of a person who is “unselfconsciously guided by a John Wayne ideal of masculine identity but [does] not accept the moral values that support such an ideal.”⁴³ Which is more indicative of this person’s intimate self, his moral values or his unselfconscious ideal? Suppose his somatic traits are in fact quite unmasculine. Does this make the ideal less authentic, and the moral values more authentic? What if the person also has temperamental traits that are in fact John Wayne-like. Does this tip the balance in the other direction? The current analysis of identity is useful in that it allows us to pose such questions, but it does not tell us how to answer them.

7. Separating authentic behavior from fidelity to self

Is this really a problem? Once we have a thick notion of identity in place, the door is open to questioning a fundamental assumption behind the notion

⁴³ Ibid., p. 24.

of being true to oneself, namely that in each situation, there is but one way of being true to oneself. If our identity can have (and perhaps typically has) various conflicting aspects, and the agent finds himself in a situation where one can only be expressed at the expense of another, one possible assessment is that he will be true to himself no matter which aspect he expresses. This assessment results in the somewhat counterintuitive judgment that the gay homophobic is true to himself when he follows his evaluations and tries to act as if he were heterosexual. But if we allow, as I think we should, that his homophobic evaluations are part of his identity, then we must accept that in following them he expresses one part of who he is, even though he does so at the expense of another. His problem is that given the conflict between his evaluations and his deeply embedded traits, it will be very difficult for him to be true to who he is without also betraying who he is. His self is fragmented. As Bennett Helm has argued, the self may be characterized by a holistic rational pattern of will and judgment on the one hand and desires and emotions on the other, and these two sources of direction and meaning need not converge on the same ranking of values.⁴⁴ When they don't, the person is true to herself, and betrays herself at the same time, no matter which pattern she follows.

This conclusion does not imply that the person with a fragmented self, or internally conflicting identity, will *act authentically* no matter which pattern she follows. The examples we have considered strongly suggest that authentic behavior essentially involves a continuity between agency and embodied self, a continuity that is not available to us insofar as our selves are fragmented. I cannot offer a full-blown account of such

⁴⁴ "Integration and Fragmentation of the Self" *Southern Journal of Philosophy* 34.1 (1996):43-63.

continuity here, but I find promising the idea that behavior is authentic to the extent that it is motivationally supported by states that are embodied in the individual, and inauthentic to the extent that it lacks such support.⁴⁵ This analysis seems plausible in light of the examples of the wholehearted conformist and the gay homophobic, whose embodied states revolt against their own will and judgment. As self-reflective, rational agents, we are capable of guiding ourselves according to ends, principles, and concerns that do not necessarily recommend the actions that would come naturally to us in every situation, actions that would give us the sense that our “whole being” has been satisfied, or actions that reflect what we actually care about at the level of strongly embodied emotions.⁴⁶ The habitual exercise of rational self-control may of course gradually affect what comes naturally to us, what we experience as genuinely satisfying, and how we see and respond to our circumstances, but it does so by affecting what conative and cognitive states are actually embodied in us. Surely, there are contingent limits to what kinds of practical principles can thus become the second nature of any given individual; efforts of habituation can vary in the degree to which they go against the grain of the person’s original psychological and physiological tendencies. If, as I have suggested, authenticity in behavior involves continuity between agency and embodied self, it will be sacrificed by efforts of habituation, while at the same time being part of their goal. To the extent that authenticity is valuable—and clearly it has at

⁴⁵ This idea is developed by Olav Elgvin in “Autonomy and Authenticity”, an unpublished manuscript.

⁴⁶ Our reasoned decisions may even recommend actions that we cannot perform successfully because such performance would require an absence of self-conscious effort.

least instrumental value for the person's well-being—it therefore needs to be taken into account when efforts of self-control or projects of habituation and character training are evaluated. It is also relevant to issues like cultural adjustment, immigration, and education in a multicultural society. Social identities (e.g. race, gender, sexual orientation, nationality, and vocation) provide norms and “loose scripts” that help us make narrative sense of our own lives and provide us with meaningful purposes.⁴⁷ These goods will not be realized insofar as the norms or purposes they prescribe go against the grain of the robust tendencies of the individual.

8. Conclusion

This paper started with a quote from Charles Taylor. His view may be seen as an attempt to reconcile the romantic view that authenticity requires each person to discover what is fundamental to her own identity, and the existentialist view that authenticity requires each person to create her identity from scratch.⁴⁸

These basic notions are tested in the various theoretical alternatives that have been considered in this paper. Frankfurt's assertions about the relation between volitional necessity and personal essences would support the romantic view, but as we have seen, these assertions turn out to be false. Moreover, the very notion of personal essence turned out to be a non-starter even when Frankfurt's own interpretation was omitted. We then considered the other side of the romanticism-existentialism divide. Identification

⁴⁷ See Kwame Anthony Appiah, *The Ethics of Identity*, pp. 21-3.

⁴⁸ *Ibid.*, p. 17. On Taylor's view, we create our identities by defining ourselves with reference to “objective horizons of significance” that we can only discover.

theory is not existentialism, of course, but it involves an element of creation in that the agent is taken to have the ability to stand back from her actual desires and make an independent decision as to which of them (if any) to endorse. Like the existentialist agent, the agent modelled by hierarchical theories of the self decides what she should stand for and thereby constitutes herself as an agent. This approach also turned out to be unable to do justice to the phenomena of authenticity and inauthenticity. We also considered the type of view Taylor himself defends, i.e. views that define authenticity as acting according to one's evaluations or reflectively accepted ideals. Such views turned out to be insufficient because they cannot account for cases where the evaluations and ideals are not supported by the person's embodied desires and emotions.

Authentic behavior, it turned out, must involve acting according to what is objectively instantiated in one's body, be it traits or specific mental states. More specifically, it involves a continuity between agency and embodied self; behavior is authentic to the extent that it is motivationally supported by states that are embodied in the individual, and inauthentic to the extent that it lacks such support. Authenticity does not constitute fidelity to self, however. Talk of fidelity to self is deeply ambiguous; once we have abandoned the ideas of personal essences and self-constituting identifications, we see that a person's identity has many dimensions that may have conflicting normative implications.⁴⁹

⁴⁹ An early version of this paper was presented at the Philosophy Department at Lund University in November 2005. I am grateful to members of the audience for a helpful discussion, and to Giorgio Baruchello for written comments on a later version. I also thank the University of Akureyri for granting me a sabbatical for the academic year of 2005-2006, during which this paper was written.

