Est. 1841 | YORK ST JOHN UNIVERSITY

Carruth, Alex and Hill, David (2015) Identity and distinctness in online interaction: encountering a problem for narrative accounts of self. Ethics and Information Technology, 7 (2). pp. 103-112.

Downloaded from: https://ray.yorksj.ac.uk/id/eprint/3557/

The version presented here may differ from the published version or version of record. If you intend to cite from the work you are advised to consult the publisher's version: https://link.springer.com/article/10.1007/s10676-015-9364-y

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. Institutional Repositories Policy Statement

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorksj.ac.uk

Identity and distinctness in online interaction: Encountering a problem for narrative accounts of self

It is widely accepted that online interactions involve an encounter between users. Notably, the assumption of encounters between users is a direct result of the observation of users identifying with the proxy that stands in for them online – or what is referred to as an *avatar*. Sherry Turkle, having established that the participants in her study understand themselves as (re)writing their identity online, could then remark that when this self enters an online environment 'other people are there too' (1997, p. 9). Similarly, and more recently, Tom Boellstorff's anthropological study of Second Life established that avatars were understood as simply 'the human online' (2008, p. 25), after which it is possible for him to conceive of avatar interaction as a kind of mediated social encounter. Whilst it is true that Turkle has recently taken a dimmer view to life online than that offered in her more affirmative *Life on the Screen* (1997), the somewhat exclamatory title of her most up-to-date meditation on digital connectivity – *Alone Together* (2011) – suggests an insurmountable solitude not present in the book itself. That is, whilst she observes a proliferation of instrumental relations online, of emotional distance between users, and, most strikingly, of hiding behind communication to avoid others, it is important to note that these are user-generated dispositions, imposed on technology because, according to Turkle, we dislike intimacy with others. Encounters are still assumed to be possible online and indeed to take place; the argument from Turkle now is that we want to cheapen interactions to make sure that they are not too demanding. There are numerous other empirical researches across the years that make the same move, from observing users establishing their avatar as an expression of themselves to the claim that meetings or encounters take place (see for instance Markham 1998; Ben-Ze'ev 2004; Hillis 2009).

In this paper we examine the assumption that encounters between users occur online via avatar interaction. In Section One we argue that users and their avatars are distinct entities. We do this by examining a number of accounts of identity and distinctness, and demonstrating that under each of these accounts, either the user and avatar come out as distinct, or else that holding them to be identical is problematic. It is important to adopt a systematic approach to tackling the question of whether users and their avatars might be held to be identical, for, as shall be seen below, there are a variety of concepts of identity which might be employed in discussions of the relationship between users and their avatars. We delineate three such notions – *absolute identity*, *relative identity* and *identity-as-selfhood* – exploring whether each of these can be appealed to in order to ground the claim that users encounter one another via the medium of their avatars. In Section Two, working from the conclusions

reached regarding the ontology of users and avatars in Section One, we show that it cannot be taken for granted that users encounter *each other* when they interact online through their avatars. As such, we begin to explore the viability of grounding the encounter between users in narrativity, alongside a variant grounded in trace. Whilst these can be shown to give a sufficient account of how the user relates to their own avatar first-personally, we argue that they are not sufficient to ground encounters *between* users. This ought not to suggest that such accounts of the user-avatar relationship should be rejected; rather, we conclude that supplementary work is required in order to provide a grounding for the claim that when users interact online via an avatar, an encounter between these users takes place – without this, any number of social, ethical and political issues around online interaction will remain unsettled.

1. The distinctness of user and avatar

The claim that encounters between users take place online is widespread in the literature. How are such encounters possible? One very simple way for this to be possible would be if users were identical, in the sense of being-the-same-entity-as, their avatars – if this were the case, users would literally be encountering each other online. We shall call such a conception of identity *absolute identity*. None of the authors discussed above explicitly endorse the claim that the user and their avatars are absolutely identical. Nevertheless, the claim that they could be deserves attention, since, as shall be seen below, paying close attention to exactly what distinguishes users and their avatars in this absolute sense can inform us as to how plausible other sorts of identity claims which such authors are more likely to endorse are: namely, claims of *relative identity* and *identity-as-selfhood*.

1.1 Absolute identity

A reasonable starting point for any discussion of identity is the principle of the identity of indiscernibles and its converse, as formulated by J.M.E. McTaggart (1921), the dissimilarity of the diverse. These claim that no two distinct entities can share all their properties and thus that two putatively distinct entities, if it turns out that they do in fact share all their properties, are one and the same. Two concepts putatively applied to a single entity, if they entail two non-shared sets of properties, must in fact apply to two entities. It should be clear that the user

and the avatar do not share all their properties (for example, the user has a determinate height not shared by the avatar), and therefore, if we accept these twin-principles, the user and avatar are distinct entities. Brey (2003) draws on this aspect of online representations – their inability to instantiate some of the physical properties of the entities they are representations of – to distinguish between 'mere simulations' and 'genuine ontological reproductions' (Brey 2003, p. 282); the latter being representations of entities all of whose properties *can* be virtually reproduced. However, it is not clear this distinction can be neatly carried over to the case of avatars, for it is often the case that, unlike, for instance, the representation of a building or forest in a computer game, the online representation of the user doesn't even *attempt* to represent the properties that the user instantiates in real life.

However, these principles are not universally accepted (see for instance Black 1952, which discusses the possibility of a universe containing only two qualitatively identical spheres at some distance from each other, or Lowe 2002, p 62). In the context of the current discussion, the principles discussed above may seem to be concerned with a kind of distinctness that lacks the significance required for the other aims of this paper – two copies of the same book are distinct under this conception but not in any way that holds significance for a discussion as to whether two people reading different copies are reading the same book or not. Likewise, it might be agreed that the principles discussed above will lead us to consider user and avatar as different *entities* per se but not different *persons* (or perhaps versions of persons). This is a thought to which we will return later in this section.

Another approach to determining absolute identity is to consider persistence conditions. Persistence conditions determine the sorts of change a given entity can undergo and still be considered to be the very same entity that existed prior to the change.[1] It seems plausible that one of the persistence conditions of users is spatial continuity – it would be reasonable at least to ask, if the case were to arise that a user ceased to be in one spatial location and a seemingly qualitatively similar user began to be in another spatial location, without any continuous movement from the first to the second, whether this second user really was the same user.[2] It is not clear that

[1] For an introduction to discussions of this sort see Lowe 2002, pp. 23-74; for more detail see Van Inwagen 1990.

[2] Questions about tele-transportation and personal identity (for example Parfit 1984) are contentious. Whilst Parfit holds that it is clear that a person could survive tele-transportation, those who hold that no reductive

users are the sorts of thing that survive spatial discontinuity. Conversely, if an avatar does indeed have a spatial location (say the physical media on which the data comprising the avatar is stored), it seems intuitively plausible to suppose that if that media were destroyed, and then exactly similar data placed onto spatially distant media, this avatar would be the very same avatar. If we suppose that avatars do not have a spatial location then any prospect of identifying them with users – in the sense of their being-the-same-entity-as – which quite clearly *do* have a spatial location, seems vanishingly small.

Similar thought can be applied to temporal continuity – if a user stopped existing for a period of time, and then a qualitatively similar user began to exist, it would be reasonable to question whether these two users were identical or not. Intuitively this does not seem to be the case with the avatar; avatars seem to be the sort of things that can survive temporal discontinuity. If questions of identity and distinctness are to be settled by considering persistence conditions then it appears that user and avatar are distinct entities, as constraints on what sorts of changes they can survive seem to differ quite starkly.

Alongside the differences in properties instantiated (and thus differing identity conditions) and differing persistence conditions, we contend that the following situation adds credence to the claim that there is an ontological divide between user and avatar. It does not appear that avatars can be mapped to users (and *vice versa*) in an isomorphic fashion. If they could be so mapped, then for each and every user there would be exactly one avatar, and *vice versa*. However, there does not seem, *prima facie*, to be any reason to exclude the possibility of one avatar being related to several users, either simultaneously or consecutively, in the manner in which we normally take an avatar to be related to a single user. It is not hard to imagine such cases: a group of people get together and create an avatar as a joint project; or one person hands over control of an avatar to another – indeed, such things are fairly common online (Klang 2004; for example the selling of characters in *World of Warcraft*).

This invites the question of whether there is a difference in kind between multi-user avatars and avatars which have only a single user. Whilst there is quite clearly a difference between the overall situations (one involves several users, the other does not), there does not seem to be any reason to distinguish between the single- and

---

criterion of personal identity can be given may well disagree with this intuition (see for instance Lowe 1996 or Merricks 1998). It is not the aim of this paper to settle this controversy here; it suffices for current purposes to establish that such cases do *raise questions* about the persistence conditions of persons.

multi-user avatar as *different kinds* of avatar – or at least no reason to do so that does not assume that user-avatar isomorphism is definitive of or essential to one such kind. If single- and multi-user avatars are not fundamentally different in kind then our theoretical treatment of them ought to respect that fact. It is reasonable to suppose that the latter may be qualitatively indistinguishable from the former. Thus a single avatar may be related to (simultaneously or in series) a number of different users, in just the same way that we normally take a single avatar to be related to a single user. If this situation is plausible, as we contend it is, there seems little prospect in forging *any* strict and necessary relation between user and avatar, let alone that of identity. It should also be noted that a single user may have several online proxies (and, indeed, almost all users of the internet do have a variety of avatars) and if these avatars are taken to be genuinely distinct then by transitivity of identity they cannot be considered identical (in the sense of absolute identity elaborated above) to the user;[3] it cannot be the case that $U=A_1$ and that $U=A_2$ if it is the case that $A_1 \neq A_2$.

Two scenarios have been sketched – a single avatar created and controlled by a group of users working together (what we shall call the 'Synchronic' case) and avatars who, during their existence, shift from being controlled by one user to another (what we shall call the 'Diachronic' case).[4] In the Synchronic case it might be claimed that rather than this example speaking against user-avatar identification, the avatar in question is identical to the group of users, just as a single-user avatar is identical to its sole user. Firstly, it should be noted that an ontological commitment to group-entities existing over and above their members may be considered controversial. Secondly, avatars do not *prima facie* seem to be group-like entities and so it is not immediately apparent how they could be identified with some group-entity. Furthermore, if the avatar in this situation is taken to be of a different kind to avatars in the single-user case (in this case, as a group-like entity), there ought to be some grounds independent of the fact of the avatar being related to multiple-users to motivate this – otherwise

[3] Transitivity is a formal property of the identity relation. Transitive relations are relations such that if something *x* bears a transitive relation to something *y*, and *y* also bears the same relation to something else, *z*, then *x* must also bear that relation to *z*. Being-taller-than, for instance, is a transitive relation, as is identity. Loving, for instance, is intransitive – just because John loves Jane and Jane loves Janice, we cannot infer that John loves Janice (although, of course, he might – but not just in virtue of his loving Jane and her loving Janice).

[4] It might be argued that there are no genuine cases of Synchronic multi-user avatars, and seeming cases of this can actually be reduced to very complex patterns of the Diachronic case.

the claim is either *ad hoc* or else begs the question against the claim that user-avatar isomorphism is not an essential feature of single-user avatars.

In the Diachronic case it could be argued that the avatar is identical to one user at one time, and later is identical to another user. Such a claim would be inconsistent with endurantist accounts of identity over time (which hold that an entity is wholly present at all times it exists; see for instance Lewis 1986, pp. 202-204), on pain of violating the transitivity of identity; it cannot be the case that $A=U_1$ at t_1 and then $A=U_2$ at t_2 if $U_1 \neq U_2$. It is however consistent with perdurantist (sometimes called four-dimensionalist) accounts which hold that entities persist by being composed of not only spatial but also temporal parts – ultimately of momentary space-time slices or the like (Lewis 1986, pp. 202-204). What we ordinarily think of as a single entity persisting over time is really a set of temporal parts (one might think of this as analogous to a complex object being a set of spatial parts, bottoming out in fundamental particles) which together make up the whole. An entity persists 'through' a period of time just in case it has a temporal part which exists at each moment of that period, just as an entity occupies a region of space just by having some part which occupies part of that region.

According to perdurantism, it would be the case that $A\text{-at-}t_1=U_1\text{-at-}t_1$ and $A\text{-at-}t_2=U_2\text{-at-}t_2$, where the entity referred to as A is a space-time worm – that is, a collection of all these momentary slices tracing a particular trajectory through space and time, each of which is a distinct proper part of the worm – composed of $\{A\text{-at-}t_1, A\text{-at-}t_2,...,A\text{-at-}t_n\}$. Note that the worm need not be considered as an additional entity over and above the parts that make it up (just as, for instance, a collection of stamps need not be considered to be an additional entity to the individual stamps of which it is composed).[5] As $A\text{-at-}t_1$ is not absolutely identical with $A\text{-at-}t_2$, as these are two different temporal parts of A, there is no violation of the transitivity of identity due to the fact that $U_1\text{-at-}t_1 \neq U_2\text{-at-}t_2$. That is, it is perfectly consistent to claim that two *different* (temporal) parts of the avatar might be identical to two different (temporal) parts of two different users; sometimes (some part of) the avatar is identical with (some part of) User One, sometimes with (some part of) User Two. The claim here is, strictly speaking, that there can be no case of a single avatar having different users at different times, because strictly speaking, no

[5] This view could be disputed, but follows trivially for anyone who accepts the thesis that composition is identity. For a canonical statement of this thesis, see Lewis 1991, section 3.6. For a critique of the thesis see van Inwagen 1994.

avatar (nor any other entity) can exist at more than one time; absolute identity is always synchronic and momentary.

However, it still seems to be a live option that every temporal part of A is identical to some temporal part of some user or another. However, such a strict view of identity would not aid the proponent of user-avatar identification, as to claim that each spatio-temporal part of an avatar is identical with some spatio-temporal part of a user is clearly indefensible – for users and their avatars do not occupy the same spaces. It seems that multi-user avatars, whether Synchronic or Diachronic, pose a problem for holding that users and avatars can be identified one with the other – at least in the sense of absolute identity, of being-the-very-same-entity.

1.2 Relative identity

Appeals to relative identity hold that it is possible for two (absolutely) distinct entities to still be (non-absolutely, relatively) the same thing, that is, some *x*s may be the same F but different Gs. For instance, if Jane has bought two tickets for the same flight, it might be argued that whilst she is just one person, she is two different passengers. A response to the arguments presented above might be the claim that whilst it has been established that avatar and user are, as we have argued, distinct in some sense – they are not absolutely identical, not the very same entity – they are nonetheless identical relative to some sortal term, the most plausible candidate being *person*.

To plausibly maintain such a position some criterion of identity for persons will need to be appealed to. Whilst this clearly raises a lot of substantial questions for the metaphysics of personhood which cannot be settled here, below we address a number of popular positions with regard to the focus of this paper specifically:[6] psychological accounts; bodily accounts; and narrative accounts. The first two accounts have primarily been proposed as answers to the Persistence Question – when is a person at one time identical to some future/past

[6] It should be noted that we do not take the relative successes or failures of these positions to unify user and avatar as identical relative to the sortal *person* to speak to the wider debate regarding these positions – the question of just what is the correct account of personal identity is a question beyond the scope of this paper.

person? – that is, to answer questions of diachronic identity. What we seek here is a principle of synchronic identity, that is, criteria to answer the question of what unites diverse phenomena in a single person. We hold that from both of these diachronic accounts a synchronic principle can be derived. The narrative account focuses, rather, on the grounds we can have for attributing various experiences, psychological states and characteristics, actions etc. to some particular person.

Psychological accounts come in a variety of forms – appealing, for instance, to continuity of memory (see for example discussions in Parfit 1984; Shoemaker 1970) or causal dependencies between sets of psychological states (see Shoemaker 1984). According to such accounts, generally speaking what it takes for a person to persist is for there to be some unbroken chain of appropriately related psychological states. From this diachronic principle a synchronic one can be derived: what it is for there to be a person at any given moment is for there to be some unity of psychological states. This sort of account cannot be appealed to in order to maintain that avatars are identical to their users, either absolutely or relative to some sortal, as avatars do not have psychological states of any sort and so cannot partake in the required psychological unity.[7]

Bodily approaches hold that what it takes for a person to persist is for there to be some particular sort of physical continuity (see for instance Olson 1997), generally the continuity of a body, or some part of a body – a brain perhaps. The synchronic principle that can be derived from this is that what it is for there to be a person at any given moment is for there to be some particular kind of physical (bodily) unity. Again, this sort of account cannot be appealed to in order to maintain that avatars are identical to their users, either absolutely or relative to some sortal, as avatars do not form a bodily unity with users.

Narrative accounts hold that to correctly attribute something to some person, that thing (be it a past action, a belief, an experience, whatever) must fit into a story that person tells about their own life in the right sort of way (see Schechtman 1996). Such an account, if it is to provide the basis for a relative identity claim between user and avatar, would hold that both feature as part of a consistent narrative of a single person – insofar as separate narratives can be told regarding just the user or just the avatar, the two are distinct, but these would be sub-

[7] The claim that avatars lack beliefs, desires, memories and the like at least seems to enjoy a very high degree of *prima facie* plausibility. We note that perhaps this claim could be disputed, but confess that we find any literal ascription of psychological states to avatars hard to motivate.

narratives of a super-narrative that constitutes the *person*. Taking a narrative approach to personal identity to substantiate the claim that, relatively speaking, the user and avatar can be identified qua 'person', shows more promise than either the psychological or bodily approaches but it is not without problems. First, the non-isomorphism of users and avatars (as discussed in section 1.1) seems to undermine the viability of this claim. Whilst under relative identity transitivity does not always apply, it should be maintained relative to the relevant sortal, in this case 'person'; that is to say, if an avatar is the same person as User One and the same person as User Two, then Users One and Two must also be the same person as each other. But if an avatar is held to be the same person as multiple users, who are *ex hypothesi* not the same people as each other (on the assumption that all users are persons in their own right, that users would still count as people even if they never had any avatars), then transitivity is violated.

To avoid this violation, it would seem that in these cases it would have to be argued that there is some super-narrative comprising the person of which there are sub-narratives of User One, User Two, the avatar, User-One-and-the-avatar, User-Two-and-the-avatar, none of which themselves qualify as persons. However, the last two of these are formally similar to what the proponent of the narrative account of personal identity would take to be the 'person' in the single-user case – that is, the narrative that interweaves one user and one avatar – and so it would be *ad hoc* to claim they do not constitute persons in this case.

As we have seen, none of the major accounts of personal identity – Psychological, Bodily or Narrative – can be unproblematically coupled with a relative identity thesis in order to identify the user and avatar qua 'person'.[8] Without such substantiation, no informative appeal can be made to a relative identity thesis in the case of user and avatar, and so we must still regard user and avatar as distinct, leaving the question of how users encounter one another online unanswered.

1.3 Identity-as-selfhood

There is another sense in which we might identify users and their avatars. Users express themselves through their avatars. As such, users may take their avatars to instrumentally encode aspects or features that are central to that

[8] It is not our intention to argue against these positions themselves – simply to note that they cannot substantiate a relative identity thesis in this case.

user's notion of their own identity, of what makes them the very person they are. The identification of user and avatar in terms of this sense of 'identity' differs significantly from their identification in terms of either absolute or relative identity. Absolute and relative identity are formal relations grounded in the natures of the entities they involve; identity-as-selfhood seems, rather, to be grounded in an attitude or stance which the user takes towards their avatar. David W. Shoemaker (2010) discusses the strengths and weaknesses of a variety of potential accounts of identity-as-selfhood – however, his starting point assumes some relatively circumscribed and united set of properties or entities from which the self is composed. The key question at hand isn't *which* of an already defined set of properties or entities do in fact compose the self but rather *whether* the avatar and the user *can* enter into such a unity.

It is not immediately clear that, if User One and her avatar are identified in this sense – which might be expressed by saying that the User One *identifies with* her avatar – we can infer from the fact that User Two has had an online encounter with User One's avatar, that User Two has had an encounter with User One. Note that this is not so with either of the previously discussed senses of identity. If User One and her avatar are absolutely identical, or identical relative to the sortal 'person', then an encounter with User One's avatar is straightforwardly an encounter with User One herself. This is because both absolute and relative identity, being grounded in the nature of the entities in question, carry the requisite ontological weight to straightforwardly license the inference. Conversely, as identity-as-selfhood is grounded primarily in the psychological attitudes of a user, it is not *prima facie* clear that a user identifying with her avatar has any ontological consequence, and so there is no straightforward implication from a user's identifying with her avatar to encounters with that avatar being encounters with that user.

Users and their avatars ought not to be held to be identical: they differ in the properties that they each instantiate; they have radically different conditions applying to the types of change they can survive and, as a consequence of the transitivity of identity, the possibility of multi-user avatars excludes the identification of the avatar with any of its several distinct users and so by extension the single-user avatar with its sole user. However, there is a sense of identity – identity-as-selfhood – which it is plausible obtains between users and their avatars. As we have seen above, identity-as-selfhood does not provide a straightforward route for considering online encounters with the avatars of others as encounters with those others themselves. In Section Two we examine whether, on closer inspection, identity-as-selfhood can provide a more circumspect route to this same outcome.

2. A problem for encounters between users

Whilst in the previous section we raised problems with a narrative identity thesis for establishing strict identity conditions between user and avatar, if we take narrativity to ground identity-as-selfhood rather than relative identity, then we take it to be a potentially useful way of understanding how users conceive of their relationship with their avatars. In this section we will argue that narrativity is not sufficient, however, to ground encounters between users mediated by these avatars. It should be apparent, if we observe Paul Ricoeur's (1994, p. 116) distinction between identity as *sameness* (*idem*) and identity as *selfhood* (*ipse*), that our claim that the narrative thesis plausibly offers a sense of personal identity online is not paradoxical to the argument in the first section of this paper that the user and the avatar are not, strictly speaking, identical.

If the user and the avatar are not identical then a problem arises: we cannot say that a user encounters another user, all we can say is that a user encounters another avatar. If we could have concluded that the user and their avatar were identical, then we could unproblematically consider that the online interaction is directly between two users – as it literally would be so. Were this the case, online interaction would not differ in any significant dimension from other cases of intersubjectivity and so there would be no special question regarding the status of the online encounter. However, given the systematic explorations of the preceding section, we are in a position to pose the question of whether online interaction, involving users and avatars, has the nature of an inter-personal encounter. Without strict identity the encounter would need to be cashed out in terms of a four-place relation, between User One, User One's avatar, User Two's avatar and User Two. It is the presence of *both* users in the relation described that is key to our license to consider interaction with an avatar as an encounter with its user. Yet, from the perspective of either user, it is the presence of the other user that is the hardest term in the relation to motivate. Whilst it is clear that User One is related to their own avatar, and in their dealings online comes across other avatars (for example, User Two's), given the non-identity between user and avatar, it is yet to be demonstrated that in encountering User Two's avatar, User One also encounters User Two. Likewise, although User Two is related to *their own* avatar, and encounters User One's avatar, it is yet to be demonstrated that in doing so User Two also encounters User One. Where we would hope unproblematically to find a four-place

relation, we can thus far maintain only a three-place one, given the work in Section One: that between just one of the users and both of the avatars.[9]

Unless we can reinstate the fourth term in the relation we cannot hold that two users encounter each other online when they interact via avatars. What is needed is a way to conceive of the avatar such that to encounter it is at the same time to encounter the user that controls it.

2.1 Autobiography

One option[10] would be to understand the avatar as part of one's narratival sense of self, a position offered by Turkle (1997) and, more recently, Marya Schechtman (2012). Galen Strawson has noted a 'widespread

[9] We thank an anonymous reviewer for bringing our attention to the following case: certain instances of technologically mediated communication, such as those conducted via telephone or VoIP applications such as Skype, do not seem to fit the four-place relation model suggested above. Rather, such instance might be better conceived as involving two three-place relations, one by which User One encounters Avatar Two and through it, User Two; the other by which User Two encounters Avatar One, and through it, User One. Furthermore, it seems fairly uncontentious that we encounter others when we talk to them on the telephone or via Skype. If what has just been said is correct, then one might wonder whether this model can be extended to all online interactions: whatever the case, we simply have two three place relations as described above.

It does not seem to us, however, that this model can be so extended. The cases mentioned above—the telephone; VoIP—share the following feature: the acts of communication they involve are merely recorded and broadcast (as voice or voice plus image) from one user to another. But a broad and interesting range of cases are not like this (think about interaction via social networks, MMORPGs and the like). Rather, in these cases, both users' actions are mediated by and through their avatars, and so it is simply not clear that it can be uncontentiously asserted about these latter cases that they involve encounters between users.

[10] We are concerned here with narrativity since it presents a particular problem for understanding encounters online. As suggested by one of the anonymous reviewers, the Levinasian notion of the encounter-with-the-face-of-the-Other might go some way towards the settling the question of how encounters can be taken to occur between users in avatar environments. For example, Richard Cohen (2000) employs the work of Emmanuel Levinas to advance his claim that information technologies can ground face-to-face encounters in a morally

agreement that human beings typically see or live or experience their lives as a narrative or story of some sort, or at least as a collection of stories' (2004, p. 428). For example, Schechtman's narrative self-constitution view holds that 'a person creates his identity by forming an autobiographical narrative – a story of his life' (1996, p. 93). That is, a person's identity is precisely the content of their self-narrative. Daniel Dennett's account of personal identity similarly has us all as self-made novelists (1984; 1988): 'We try to make all of our material [the episodes of our lives] cohere into a single good story. And that story is our autobiography' (1988). The autobiography, for Dennett, *is* our self. Alasdair MacIntyre (2006) argues that it is narrative that allows us to understand ourselves as coherent through seemingly distinct partitions of life – childhood, adulthood, old age, say – as well as allowing us to situate ourselves within a role in society and to make decisions accordingly. He puts forward 'a concept of a self whose unity resides in the unity of a narrative which links birth to life to death as narrative beginning to middle to end' (2006, p. 205).[11]

---

significant way. The main thrust of his argument is beyond the scope of this paper but Cohen's claim that we cannot hide behind screens, refusing the encounter on the basis that it is indirect and mediated, offers a potential solution to the problem at hand. It consists of the primacy of the moral seriousness of being in an encounter with another person over the always deferred epistemological question of that encounter, and that information technologies are simply tools such that whether we directly see the other person or not we are nonetheless brought into a relationship with their face. Such a Levinasian account might be supplemented with a radical conception of semiotics such that the avatar becomes a visceral sign of the user (see Hill 2013) - in order to more fully explain how this indirect communication constitutes an encounter. Nonetheless, we set such solutions aside for now in order to highlight the problem of encounters in narrative accounts of identity online since such accounts are currently influential and since the setting out of this problem – too rarely recognised – might help frame future solutions.

[11] It is important to note that although Schechtman and MacIntyre see narrativity as inseparable from moral life – Schechtman argues that 'weaving' stories about our lives is what makes us persons and that to enter moral commerce with other persons one must share the same (i.e. narratival) sense of self as others; MacIntyre argues that to consider oneself as a narrative is 'normal' and that the unity of life consists in the narrative of a quest for the Good – such positions are by no means uncontroversial. See Strawson 2004 for an argument against (1) the universality of the narrative notion of self and (2) the moral claims that are made of it.

How can we conceive of the avatar in terms of the narrative identity thesis? Let us take the latter account. MacIntyre notes that the narrative unity of the self is often obscured by the sorts of partitions forced upon our lives by society – not just the seemingly separate states of 'childhood' and 'old age' but also the distinctions between 'work' and 'leisure', 'public' and 'private', and so on (2006, p. 204). We might add also 'real' and 'virtual', the partition that seems to exist between lives lived offline and lived online – between user and avatar. What if what we are missing when we conceive of online environments as 'cyberspace', as an 'other place' to the 'real' world, is the unity of the self that occupies both? For MacIntyre (2006, pp. 208-209), an intelligible action is one that takes place within a sequence in a given context; that is, human actions are characterised by their place in a narrative and a setting. Imagine the following sequence of actions: a user sits at their computer; they tap away at their keyboard; their avatar waves at another avatar (in *Second Life*, say); they watch excitedly for a response from the other avatar. Here we have a coherent sequence of actions of one person that does not lose unity by virtue of replacing the term 'user' with 'avatar.' Consider the intentions at play here. MacIntyre (2006, pp. 207-208) observes that the question 'What is he doing?', asked of, say, an academic working, offers a set of possible answers that refer to different lengths of time. His example is the following set of answers: writing a sentence; completing a book; adding to a philosophical debate; trying to get tenure. The writer's action – writing – is the result of intentions that stretch into the future. Similarly we can imagine our waving avatar being part of a story about temporally projected intentions: the user is typing; the user is flirting through their avatar; the user is looking to initiate a relationship; the user is trying to settle down with someone. Here the actions of the user *and* the avatar – its waving as flirting – are made sense of as part of a story. There is no disconnection between the intentions of the user and the actions taken by that user and the avatar; they are all part of the same unitary story. If we accept MacIntyre's account of narrative self then we can plausibly include the avatar as part of the user's narrative.

However, such an account would not be sufficient to ground an encounter between users. If another user encounters the waving avatar (through their own avatar) then they experience only one part of the sequence of actions and are unaware of the full set of possible intentions (from flirting to settling down). The narrative of that encounter, then, is only a story about encountering a waving avatar; it offers no place in the narrative to the user that initiated the wave. This holds also if Dennett's (1988) 'virtuoso novelist' includes their avatar amongst the 'material' that they attempt to make cohere into an autobiography, or if Schechtman's autobiographer weaves the avatar into their life story. When we encounter an avatar we would only encounter the chapter about that avatar –

those that precede or follow it, those chapters that would identify the avatar as part of a story about its user, are inaccessible. MacIntyre (2006, p. 218) argues that the narratives we spin are 'interlocked'; we might play the lead role in our own story but in the stories of others we have only a supporting role. The point is that the story is different in an encounter between people depending on whose perspective it comes from. However, online the other avatar may appear in the narrative of the user that encounters it (through their own avatar) but the narrative of the other user that stands behind that avatar is not 'interlocked' with that of the user. In other words, there has been no encounter between users.

2.2 Autograph

A variation on the writing theme would see the avatar as a *mark* made by an author, intelligible to other users as a product of that author – which would seem to hold more promise for grounding encounters between users at the same time as establishing identity between user and avatar. For example, Ken Hillis (2009) offers an account of the relationship between user and avatar in terms of trace. Hillis's account would seek to add substance to narrativity by establishing a material connection between user and avatar. A trace is a remainder or residue, 'a mark in the here and now of something absent' (Hillis 2009, p. 115). 'Trace', as a verb, is also to follow. Hillis understands the avatar, then, as a material remainder of the user that can be traced back to that user. The user is able to see their avatar on the screen and act in a way consistent with the idea that that avatar is themselves – a trace of themselves – that stands-in for them in the environments it occupies online. Hillis writes: 'The digital *representamen*/trace of the operator, when viewed by the *interpretent*/embodied operator, appears as a stand-in for or body-double replica of the operator' (2009, p. 129). In other words, the user can interpret their representation in the form of their avatar as possessing a 'performative link' (Hillis 2009, p. 130) with themselves in that the avatar can be viewed as a vehicle or surrogate for their online actions. The user 'traces himself as a trace' allowing him (or her) 'to produce himself' (or herself) online (Hillis 2009, p. 130). The problem with such an account is that it begins with the user's experience of their own avatar and then extrapolates from this an encounter.

Now, this is a perfectly workable conception of how a user might understand the relationship they have with their own avatar. This avatar can be seen by the user to perform their actions and so there is coherence between the user's intentions and the avatar's actions. Such an account allows the user to take possession of a distinct

entity in a way that would allow themselves to understand avatar actions as their own intended actions. The user sees the performative link between themselves and their avatar and then traces this back to themselves in a circular fashion; this allows the user to understand their avatar as a trace of themselves through a sort of feedback loop. However, it is quite a leap from such a 'from the inside' account of avatar action to an account of interaction between two users mediated by avatars. Yet this is precisely the move that Hillis makes when he argues that this notion of trace is sufficient to ground encounters online. 'Appearance,' he claims, 'in the form of the trace, stands in for presence' (2009, p. 119). So an account of trace that is initiated from one's own tracing from avatar back to oneself is posited as a form of encounter between users. Since this tracing is reliant upon the observation of a performative link between user and avatar – an observation that is foreclosed to others who are elsewhere – this tracing cannot be open also to the other user. As such when users encounter avatars online they cannot experience them as a material residue of other users; what we get are two users who relate to their avatars in a circular fashion (the feedback loop) and two avatars that interact with each other – but the two users remain absent to each other.

Instead of autobiography – or a chapter in an autobiography – what we have now is an autograph. In Hillis's account 'the avatar [...] is a kind of moving autograph' that, apparently, allows us to leave a trace of ourselves in online environments (2009, p. 126). The thing about autographs (or signatures) is that they are frequently indecipherable to others. A scrawl on a piece of paper offers no clues to the identity of whoever scrawled it. If it is a perfectly legible rendition of a name then the question is: to which 'Jane Smith' does it belong? When we collect autographs of the famous, it is not the scrawl but the encounter with that famous person that gives the autograph its meaning. If an avatar is a trace as described above then, firstly, it may be unidentifiable from the outside as the person that controls it if, say, it has a radically different appearance or personality, something not uncommon given what Mitch Parsell calls the 'freedom of presentation' or 'the ability to construct our own online identity as we see fit' (2008, p. 47); secondly, it may be a perfect rendition of the user but without the experience of that user it is impossible from the outside to have awareness of this; and, thirdly, it is the encounter with the user that would give the avatar-as-autograph its relation to that user – and yet this is not given unless there has been a prior meeting. The avatar as an autographical trace does not give us the encounter between users.

When we conceive of the user-avatar relationship as one of narrative identity, or a variant of this such as autographical trace, it is difficult to establish how encounters between users can be conceptualised. Where such

accounts are successful in explaining the psychological coherence of, in Boellstorff's words, being a 'human online' (2008, p. 25), it is precisely this appeal to psychological identification that causes difficultly for grounding encounters, since it is not apparent how this might migrate from the user to other users. As Shoemaker observes, of Schechtman's account, 'the narrative view is a purely subjective, first-person account of self-identity – my self is constructed by the story *I* tell' (2010, p. 8); whilst we do not question users' 'avatar attachment' (Wolfendale 2007, p. 112), we nonetheless have cause to question how this internally motivated identification would allow us to ground an encounter between users in an interaction mediated by avatars.

Concluding remarks

It ought not to be concluded from this that encounters between users do not in fact take place – we take it to be clear that they do – and we do not question the efficacy of any of these accounts for understanding personal identity online (as opposed to establishing strict identity conditions), finding the narrative account as articulated by Schechtman (2012) to be highly promising in this respect. Nor, it should be stressed, are we claiming that such accounts are in anyway problematic when it comes to understanding offline encounters. Precisely, the problem is that the strict non-identity of avatar and user – their being distinct entities in terms of their material embodiment – means that accounts of identity that are essentially psychological in nature, whilst allowing a user to bridge the ontological gap between themselves and their avatar, do not permit us to understand how another user can cross the same bridge, from the avatar they encounter to the user that identifies with it. This was an argument made in two stages.

In Section One we examined a number of accounts of identity and distinctness, considering what each has to say about the case of users and their avatars. As users and their avatars instantiate distinct sets of properties then, according to the principle of the identity of indiscernibles, they should be considered to be distinct. Differing persistence conditions apply to users and avatars; they can each survive (or not survive) different kinds of change so, insofar as one takes this as supporting evidence for a distinctness claim, they should again be considered distinct. The relation between users and avatars appears to be many-many – one avatar may be related to multiple users and vice versa, in just the same way a single avatar and user are – which presents, in the form of a violation of transitivity, a challenge to any straightforward identity claim. Appeals to relative identity – that whilst the user and avatar may be distinct in some sense, they might nevertheless be the same *person* – have been

considered. Three manners of substantiating this claim – the Personal, Bodily and Narrative accounts of personal identity – have all been shown to be of no help in supporting the strict identity of user and avatar qua person: adopting either of the first two entails that they are not in fact identical, and the third faces the same problems regarding the possibility of multi-user avatars and transitivity. The prospects for maintaining that users and their avatars can in some sense be identified – and thus for an easy answer as to how it is that users might encounter each other online – seem slim.

In Section Two it was argued that this non-identity of user and avatar posed a problem for grounding the idea – not sufficiently justified by, for instance, Turkle (1997), Boellstorff (2008), Schechtman (2012) and Hillis (2009) – that encounters between users take place when they are mediated by avatars. We then argued that, whilst narrative accounts of self online are convincing for conceptualising how a user might understand their continuity with their avatar, such accounts are not sufficient to move from establishing personal identification – identity-as-selfhood – to identification of another user so as to constitute an encounter and so do not, as they stand, take us out of the problem. Whilst we are not suggesting that the accounts of the user-avatar relationship discussed in Section Two should be rejected, we do contend that supplementary work is required in order to provide grounds for the claim that when users interact online via an avatar, an encounter between these users takes place. It is the work of another paper to provide this supplementary work. We take such work to be important for the discussion of a number of philosophical issues online. For example, one might appeal to users' avatar attachment to afford moral significance to virtual harms, since that attachment is morally significant to the user; however, if we cannot make claims about users encountering users then we may not have the licence to lay moral responsibility with a user who harms another user's avatar and in so doing that user, since the attachment is not visible to them. Similarly, one might appeal to various accounts of moral evaluation in order to reject the Asymmetry Thesis – that is, the idea that actions that can be deemed morally wrong offline can be simultaneously understood as permissible online (for a fuller account of the Asymmetry Thesis see Dunn 2012) – but whilst an account of online encounters is lacking, then the very process of moral evaluation in this context is undermined, for such evaluation concerns interactions between *users*, and not between a user and what they encounter as merely an unattached avatar. Questions such as these may only be settled once user-avatar identification can also provide an account of encounters between users.

References

Ben-Ze'ev, A. (2004). *Love Online: Emotions on the Internet.* Cambridge: Cambridge University Press.

Black, M. (1952). The identity of indiscernibles. *Mind*, 61(242), 153–164.

Brey, P. (2003). The social ontology of virtual worlds. *American Journal of Economics and Sociology*, 62(1), 269–282.

Boellstorff, T. (2008). *Coming of Age in Second Life: An Anthropologist Explores the Virtually Human.* Oxford: Princeton University Press.

Cohen, R. A. (2000). Ethics and cybernetics: Levinasian reflections. *Ethics and Information Technology*, 2(1), 27-35.

Dennett, D.C. (1984). *Elbow Room: The Varieties of Free Will Worth Wanting.* Oxford: Oxford University Press.

Dennett, D.C. (1988). Why everyone is a novelist. *Times Literary Supplement*, 16-22 September.

Dunn, J. (2012). Virtual worlds and moral evaluation. *Ethics and Information Technology,* 14(4), 255-265.

Hill, D. W. (2013). Avatar Ethics: Beyond Images and Signs. *Journal for Cultural Research*, 17 (1). 69-84.

Hillis, K. (2009). *Online a Lot of the Time: Ritual, Fetish, Sign.* London: Duke University Press.

Klang, M. (2004). Avatar: from deity to corporate entity: a philosophical inquiry into digital property in online games. *Information, Communication & Society*, 7(3), 389–402.

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

Lewis, D. (1991). *Parts of Classes*. Oxford: Blackwell.

Lowe, E.J. (1996). *Subjects of Experience.* Cambridge: Cambridge University Press.

Lowe, E.J. (2002). *A Survey of Metaphysics.* Oxford: Oxford University Press.

MacIntyre, A. (2006). *After Virtue: A Study in Moral Theory*. London: Duckworth.

Markham, A. (1998). *Life Online: Researching Real Experience in Virtual Space.* Oxford: AltaMira Press.

McTaggart, J.M.E. (1921). *The Nature of Existence.* Cambridge: Cambridge University Press.

Merricks, T. (1998). There are no criteria of identity over time. *Noûs*, 32(1), 106–124

Olson, E.T. (1997). *The Human Animal: Personal Identity Without Psychology*. Oxford: Oxford University Press.

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

Parsell, M. (2008). Pernicious virtual communities: identity, polarisation and the Web 2.0. *Ethics and Information Technology*, 10(1), 41-56.

Ricoeur, P. (1994). *Oneself as Another*, trans. K. Blamey. London: The University of Chicago Press.

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

Schechtman, M. (2012). The story of my (Second) Life: virtual worlds and narrative identity. *Philosophy and Technology*, 25(3), 329–343.

Shoemaker, D. W. (2010). Self-exposure and exposure of the self: informational privacy and the presentation of identity. *Ethics and Information Technology,* 12(1), 3-15.

Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7(4),269–285.

Shoemaker, S. (1984). Personal identity: a materialist's account. In S. Shoemaker & R. Swinburne (Eds.), *Personal Identity* (pp. 67–132). Oxford: Blackwell.

Strawson, G. (2004). Against narrativity. *Ratio,* 17(4), 428–452.

Turkle, S. (1997). *Life on the Screen: Identity in the Age of the Internet.* New York, NY: Touchstone.

Turkle, S. (2011). *Alone Together: Why We Expect More from Technology and Less from Each Other.* New York, NY: Basic Books.

Van Inwagen, P. (1990). *Material Beings.* Ithaca, NY: Cornell University Press.

Van Inwagen, P. (1994). Composition as Identity. *Philosophical Perspectives* 8, 207-220.

Wolfendale, J. (2007). My avatar, my self: virtual harm and attachment. *Ethics and Information Technology,* 9(2), 111-119.