

The Moral Status of Virtual Artefacts In Computer Games

Edward H. Spence

Department of Philosophy

University of Twente

The Philosophy of Computer Games 2009

University of Norway, Oslo

1. Introduction

In this paper I examine the question concerning the Moral Status of Virtual Artefacts in Computer Games: What is, if any, the moral status of virtual artefacts such as virtual swords (v-swords), v-wands, v-buildings, avatars, v-apparel, v-cars, etc., in computer games? Following Floridi (2007), I argue initially and in agreement with Floridi that artefacts generally and in my case virtual artefacts (Varts) existing in computer games can be correctly conceived as *informational objects*. As such, such virtual artefacts can be correctly conceived as *moral patients* deserving of some prima facie minimal moral consideration in their own right and independently of the moral and legal rights-claims, if any, that the creators and/or owners of those Varts might have to them. In this paper I will be primarily concerned with the moral status of Varts as such and not with the question concerning the moral or legal rights that computer-game players or other stakeholders may or may not have to Varts.

However, in a radical departure from Floridi's metaphysical thesis of information ethics that attributes to all informational objects, sentient as well as not-sentient, a moral status based exclusively on their existence qua *Being* (according to Floridi everything that *exists* qua Being as an informational object has an intrinsic moral status that renders it worthy of a minimal moral respect) I argue that Varts as informational objects have a moral status not in virtue of their existence qua Being, which is problematic (Spence, forthcoming) but rather because of their *Designed-in-Purposive Agency (DiPA)*.

By contrast to existence or Being, purposive or goal-orientated behaviour can confer value in the manner demonstrated by Alan Gewirth's argument for the PGC (Gewirth, 1978). Namely, the necessary conditions for purposive agency, freedom and wellbeing, which are also necessary for a meaningful and worthwhile life, provide the basis for having rights to freedom and wellbeing and hence provide the universal foundation for the moral status of all purposive agents or patients, be they human or non-human, such as animals and possibly in the future, androids.

One way to extend the moral status to non-sentient informational objects can be accomplished by showing how non-sentient informational objects possess in some sense

and to some degree a form of purposive agency or some other teleological property that is value conferring.

2. The Argument from Designed-in-Purposive Agency

Using an argument that I refer to as the *Argument from Designed-in-Purposive Agency (A-DiPA)* (Spence, forthcoming) I will argue that artefacts and other non-sentient informational objects, both real and virtual, have a functional instrumentality. They are designed to perform a certain specific functional and instrumental role, whether in a virtual or in a real environment. Take a knife, for example. The functional role of a knife is to cut materials of a certain kind. It has been designed with that functional purpose in mind. This functional role or purpose is inherently designed in the knife and as such *inheres* in the knife unless removed. All things being equal the knife when used as intended will cut perfectly well according to the purpose for which it was designed – its designed-in-purpose. Now let us suppose that someone *for no good reason* and merely on a whim destroys the teleological (its design-in-purpose) and functional capacity of the knife to cut. Let us also assume that this someone, call him Mack, is the owner of the knife. The knife is now blunt and has lost its functional purpose of cutting. No doubt the knife has been *damaged* (harmed) instrumentally as it can no longer fulfil the instrumental role or the purpose for which it was designed and created. But has any moral harm being committed and if so to whom and by whom?

To answer this question let us first ask a different question: Would it have been better if Mack had not and for no good reason, destroyed the capacity of a perfectly good knife to cut? If the answer to that question is yes, as it is likely to be, we can then proceed and ask what kind of damage or harm has been committed. I think we can allow that an instrumental harm has taken place which would have been better had it not occurred. What about a moral harm? Has the knife suffered a moral harm by it being made blunt? Clearly not as an agent, since the knife lacks the capacity for agency¹. However, even if the knife lacks the capacity for agency in the full-blooded and traditional sense, could we not argue that the knife because of its inherent or *designed-in-purposiveness* or *designed-in-teleology* has some other type of *distributed* agency (Floridi and Sanders, 2004, p. 351) or *contributive* agency (Korsgaard, 1983, p. 172) which affords it some minimal moral role? After all, a knife can be used to murder, a typical immoral action. Let us assume that if the murderer had not possessed a knife they would not have been able to commit the murder and thus an immoral act would not have taken place. Under this assumption, the knife can be said to have contributed to the murder in virtue of its inherent teleology or *designed-in-purposive-agency* (DiPA), or that the immoral act of the murder can be defined as *morally distributed* across a *moral-field* or *moral-network* that at least includes the murderer (the prime moral agent), the teleological instrument (the knife as a morally contributing *and* instrumental agent) and the victim (the moral patient).

Following Floridi and Sanders (2004, pp. 366-369) I will argue that although the knife can of course not be held in anyway morally *responsible* for the murder it can nevertheless be held *accountable* in virtue of its contributed role to the murder via its designed-in-purposive-agency or DiPA. There is, as Floridi and Sanders rightly claim a conceptual difference between moral responsibility and moral accountability. Although

an earthquake can be held accountable for the moral harm of its victims as the primary *cause* of that harm it cannot, because it lacks the relevant full-blooded agency, be held morally responsible.

Adapting and extending Alan Gewirth's argument from the Principle of Generic Consistency (Spence, 2006 and Gewirth 1978) on the basis of which it is shown that purposive agents have rights to freedom and wellbeing for the sufficient reason that they are purposive agents (that is, they possess the natural property of purposive agency) can we not *reasonably* say that artefacts such as knives with a designed-in-purposive-agency (the designed-in goal or purpose to do x, in the case of the knife, x= to cut) have to some minimal degree prima facie rights to (Art)freedom (artificial freedom) and (Art)wellbeing (artificial wellbeing) *as patients* if not as agents? That is to say, can we not reasonably say that such artefacts have the right not to have their (Art)freedom in exercising their designed-in-purposive-agency thwarted or interfered with for no good reason, or their (Art)wellbeing violated by having their DiPA, within which their (Art)well-being can be defined and understood in terms of *what they are good for* (their designed-in "functional goodness" or "designed-in-capacity" to do x) reduced or eliminated for no good reason?

Can we not say following this line of thinking that Mack's knife that was rendered useless by being made blunt for no good reason had its (Art)freedom and (Art)wellbeing unjustifiably violated and thus suffered not only an instrumental harm by having its instrumental functional role damaged, but also a moral harm qua artifact worthy of some minimal respect owed to it by virtue of its DiPA? Although the instrumental role of the knife can be replaced by the replacement of the damaged knife by a new one, the knife itself that was made blunt for no good reason has not only lost its replaceable instrumental functionality but also its irreplaceable particular inherent capacity to do what it was designed to do best, namely, cut well. That inherent capacity is something that the knife possessed as *a thing-in-itself* and as such it is something that can be valued for its own sake and not merely instrumentally for the sake of being able to cut well for some human agent. Following Christine Korsgaard's distinction between objective intrinsic and unconditional value on the one hand and objective but extrinsic conditional value on the other (1983), I wish to adapt and extend that distinction to claim that the knife has suffered moral harm by being damaged: that is, by having its DiPA to cut well, rendered useless, for no good reason.

According to Korsgaard (1983) something X has an objective extrinsic but conditional value if X meets the relevant conditions under which it is held to be valuable and X is also something that is valued for its own sake or as an end, and *in addition* to its instrumentality as a means (p. 84ff). Going along with Korsgaard we can then say that a knife or other relevant informational object is valued or can be valued partly for its own sake as an end in addition to its instrumental use as a means for human ends, provided certain relevant conditions are met. Having this dual value, both instrumental as a means and extrinsic or inherent value² as an end, the instrumental disvalue of a knife or other object that is being used to commit a moral wrong, overrides its inherent value as an end. This follows from the fact that the knife and other objects of this ontological type, only

have conditional value so that it would be justified to destroy a perfectly good knife if that were the only way to prevent a murder, for example.

In the case of Mack's knife, by contrast, both the extrinsic and instrumental value of the knife have been diminished, eliminated in fact, *for no good reason*; that is to say, the relevant conditions under which the knife is considered or can be considered valuable have been violated by the blunting of the knife, *for no good reason*. The qualification for *no good reason* is crucial and seems to point in the opposite direction in which Floridi's argument for assigning moral value to informational objects seems to go. For I am partly in agreement with Korsgaard although for Gewirthian reasons rather than Kantian as in her case, that the objective and inherent value (or for Korsgaard extrinsic value) of an object, or informational object as in Floridi's case, is not just a matter of the ontological status of the object qua informational object but of practical reason as well (Korsgaard, 1983, pp. 183-84).

I said I am only *partly* in agreement with Korsgaard because her claim is that the extrinsic value or in my case, inherent value, of an object is only a matter of practical reason and not one of ontology. Orientating my own position somewhere between that of Korsgaard and Floridi, I want to argue that the value of an object and in particular an informational object is determined partly by its ontology by virtue of its designed-in-purposive-agency (DiPA) – the artificial equivalent of the natural property of purposive agency inherent in human beings and some other animals – and partly by the reasons we have for holding that artefact valuable, principally, in virtue of the reasons for which we hold artefacts of a certain kind to be good for doing x; that is, by virtue of possessing the capacity to fulfil certain designed-in goals or purposes for doing x.

What drives us to attribute objective but conditional value to an informational object as a thing valued for its own sake and not merely as an instrument for advancing our own ends, such as a knife, for example, are partly the reasons themselves for designing such objects. The value or goodness of those reasons is transferred through the designing and creation of those objects into the objects themselves. Through this transference of *reasonable* value into the objects on the basis of the functional excellence and efficacy of their designed-in-agency or functional teleology, the value transferred through the design of the objects persists to inhere in the objects until the relevant conditions under which those reasons hold valuable and good are diminished or eliminated over time.

In conclusion of this section we can say this much: Insofar as a knife, or generally any other artefact, can be said to have an inherent value or what Korsgaard defines as an objective extrinsic but conditional value and insofar as Mack knife's value has been eliminated for no good reason (the relevant condition in this case) the elimination or diminution of the value of the knife or of any other teleological object can be said to be a moral harm; for the unreasonable elimination or diminution of an objective inherent or objective extrinsic conditional value is unjustified (because no good objective reason can be given for it) and hence morally wrong as it diminishes value overall. In the case of Mack's knife it diminished both instrumental and inherent value as the knife in its prime condition possesses both. It has the instrumental value of being used as a perfectly good

knife to cut, an apple for example, but it also possesses an inherent designed-in-purposive-capacity to cut whether or not it is ever used in that way. A good knife that lay dormant and was not used to cut would retain that inherent value regardless of whether its designed-in-purposive-capacity was put to instrumental use or not. And it is this conceptual distinction just made between the knife's *in-use-instrumental-value* exercised in cutting things and its inherent value, which it has by virtue of its *designed-in-purposive-value* (that affords it the capacity to cut) which allows us to ascribe to the knife and other objects or artefacts of the type that possess a designed-in-purposive-agency (DiPA), two inter-related values: one instrumental and one inherent.

3. Applying the Argument from Designed–in-Purposive Agency to Virtual Artefacts

However, even if it is granted that the argument from *Designed–in-Purposive Agency (A-DiPA)* can be applied to *real artefacts* as in the example of a real knife above, does it necessarily follow that the same argument can be applied to virtual artefacts such as virtual swords and virtual buildings in the virtual environments of computer games, such as the *World of Warcraft* and *Second Life*, for example? I argue in the affirmative. Insofar as the moral status of artefacts is exclusively based on their DiPA, it makes no ontological or moral difference whether the artifact is real or virtual. All things being equal, the only *relevant difference* between real and virtual artefacts is their capacity for purposive agency qua artefacts, whether real or virtual. Hence, since there is a *relevant similarity* and no *relevant dissimilarity* between real and virtual artefacts by virtue of their DiPA, this provides a sufficient condition for both real and virtual artefacts having a minimal moral status as patients. Therefore, it follows that virtual artefacts in computer games have generally a minimal moral status as patients in their own right.

4. Some Possible Objections

1. The Purpose in Some Computer Games is To Destroy Virtual Artefacts

The whole purpose in some computer games is precisely to destroy virtual artefacts. The argument for the intrinsic moral status of Varts would make this impossible. Thus the argument results in absurdity.

Insofar as the designed-in- purposive agency (DiPA) of such computer games allows for the destruction of individual Varts within those games then the DiPA-L1 of those games has intrinsic normative priority over the DiPA-L2 of individual Varts, within those games, that allows the destruction of Varts that possess only DiPA-L2.

2. The Argument from DiPA Gives Equal Moral Standing to All Entities

One implication of the Argument from DiPA is that non-sentient entities, including virtual non-sentient entities, have an intrinsic moral status just as sentient entities have including human beings and other higher animals. Clearly, this is absurd and the argument is false.

The Argument from DiPA allows for variations in the level of the moral value of different entities based on their complexity that determines the level of their intrinsic capacity for purposive behaviour both as agents and patients. Thus I am in agreement with Floridi's claim that,

It seems reasonable to assume that different entities may have different degrees of relative value that can constrain a's [the agent's] behaviour without necessarily having an instrumental value, i.e., a value relative to human feelings, impulses or inclinations, as Kant would phrase it (2002, 293).

Although the capacity for purposive agency both naturally in the case of sentient entities and artificially in the case of non-sentient entities, including Varts, creates a continuum of moral worthiness and moral consideration across a wide network of informational objects, that continuum is separated by qualitative divisions between those entities that affords them various differentiated degrees of moral value in terms of the complexity of their capacity for purposive agency. Using the metaphor of canal or river locks we can say that because the moral continuum of informational objects is porous, the capacity of purposive agency sips through the various qualitative moral divisions like water through the locks in a canal or river. However, the transitions from one qualitative moral division to another requires, as in the case of the raising of the water level in a lock to allow a ship to transit from one level of the canal to another, the raising of the level of complexity of an entity's capacity for purposive agency so as to enable its transition from a lower to a higher qualitative moral division. Thus, a software agent's capacity for purposive agency would have to be raised to that of an intelligent android that meet the conditions of full agency before it can proceed to a higher moral division close to that of human beings.

The conceptual distinctions between on the one hand responsibility and agenthood and on the other accountability and patienthood help explain the relative moral value of different entities. Thus although we could only hold a software agent accountable but not responsible for the destruction of valuable information, we could by contrast hold an android or human agent both accountable and responsible due largely to their higher moral status. Similarly, although we ought to morally avoid killing a tiger unless in self-defence we cannot reasonably expect a tiger to morally reciprocate in the same moral way. This is because although a moral patient worthy of moral respect the tiger does not possess sufficient moral agency to warrant us holding the tiger bound to reciprocal moral obligations with regard to human agents. Thus the four conceptual distinctions of responsibility/accountability and agenthood/patienthood go some way in explaining the relative moral value of different informational objects in relation to the moral relevance and significance of those conceptual categories in specific contexts, including Varts in computer games and other virtual environments.

5. Conclusion

The paper has argued that virtual artefacts generally and those located in computer games specifically have, as in the case of real artefacts (Spence, forthcoming), a minimal intrinsic moral status as patients on the basis of their designed-in-purposive agency (DiPA). Hence, it is morally wrong to damage or destroy artefacts, either real or virtual,

for no good reason. In the event that an artefact either real or virtual has to be destroyed for a good reason, for example, destroying a knife as the only way of preventing a murder, the intrinsic moral status of the knife is overridden in order to prevent a greater harm; namely, the murder of a human being or other sentient creature that possesses a higher moral status than the artefact by virtue of its higher capacity for purposive agency.

The conclusion of this paper has at least two general implications:

5.1 The world comprises a moral continuum consisting of different entities both sentient and non-sentient of varying degrees of moral status based on the level of their capacity, either natural as in the case of sentient entities or designed-in as in the case of non-sentient entities, for purposive agency.

5.2 The moral difference between real and virtual entities, including virtual artefacts in computer games, is one of *degree*. As such there is in principle no moral difference *in kind* between the real and the virtual.

References

- Beyleveld, Deryck (1991) *The Dialectical Necessity of Morality: An Analysis and Defence of Alan Gewirth's Argument to the Principle of Generic Consistency*. Chicago: University of Chicago Press.
- Floridi, Luciano (2008) "Informational Structural Realism" *Synthese*, 161.2, 219-253.
- _____ (2008a) "Information ethics: a reappraisal" *Ethics and Information Technology*, 10: 189-204.
- _____ (2008b) "Understanding Information Ethics: Replies to Comments" *APA Newsletter (?)*
- _____ (2007) "Understanding Information Ethics" *APA Newsletter on Philosophy and Computers*, Fall 2007, Volume 07, Number 1, 3-12.
- _____ (2007a) "Global Information Ethics: The Importance of Being Environmentally Earnest" *International Journal of Technology and Human Interaction*, 3.3, 1-11.
- _____ (2005) "Is Semantic Information Meaningful Data?" *Philosophy and Phenomenological Research*, Vol. LXX, No. 2.
- _____ (2004) "On the Morality of Artificial Agents" *Minds and Machine* 14: 349-379.
- _____ (2002) "What is the Philosophy of Information?" *Metaphilosophy*, Vol.33, pp. 123-145.
- _____ (2002) "On the intrinsic value of information objects and the infosphere" *Ethics and Information Technology* 4: 287-304.
- Gewirth, A. (1978) *Reason and Morality*, Chicago: University of Chicago Press.

_____ (1996) *The Community of Rights*, Chicago: University of Chicago Press.

_____ (1998) *Self-fulfillment*, NJ: Princeton University Press.

Hongladarom, Soraj (2008) “Floridi and Spinoza on global information ethics”, *Ethics and information Technology*, 10: 175-187.

Korsgaard, Christine M. (1983) “Two Distinctions in Goodness” *The Philosophical Review*, Vol. 92, No. 2, 169-195.

Spence, E. (forthcoming). “Understanding Luciano Floridi’s Metaphysical Theory of Information Ethics: A Critical Appraisal and an Alternative Neo-Gewirthian Information Ethics”. *International Journal of Technology and Human Interaction*.

_____ (2007) “What’s Right and Good about Internet Information? A Universal Model for Evaluating the Cultural Quality of Digital Information”. In Larry Hinman, Philip Brey, Luciano Floridi, Frances Grodzinsky and Lucas Introna (eds.), *Proceedings of CEPE 2007, The 7th International Conference of Computer Ethics: Philosophical Enquiry*, University of San Diego, USA, July 12-14 2007, ISSN 0929-0672.

_____ (2006) *Ethics Within Reason: A Neo-Gewirthian Approach*, Lanham: Lexington Books (a division of Rowman and Littlefield).

¹ Following Floridi and Sanders (2004, 349) the knife can be said to lack agency because it lacks its three essential features of interactivity (*response to stimulus by change of state*), autonomy (*ability to change state without stimulus*) and adaptability (*ability to change the ‘transitions rules’ by which state is changed*).

² I prefer to use the term *inherent* rather than Korsgaard’s *extrinsic* term because the value an artefact has by virtue of its DiPA inheres in the artefact and so it is not exclusively determined by the external reasons for which human beings hold it to be valuable. I should add, however, and perhaps this is in keeping with Korsgaard’s position, that in the event that an artefact was no longer held to be valuable its inherent value by virtue of its DiPA could be revoked. For what can be designed in can also be designed-out. This is in keeping with the correct thought that values are to a large degree determined by the underlying reasons for considering those values “valuable”.