
Augmented Reality: Beyond Interaction

Anton Nijholt

University of Twente, Enschede, The Netherlands

ABSTRACT

In Postphenomenology, various human-technology-world relations have been investigated. They provide an understanding of how humans and technological artifacts shape each other and how human intentionality is mediated by technology. As discussed in this paper digital technology as it has developed over the last twenty years has not received the attention it deserves. The reflections on smart environments and augmented reality technology show a lack of insight into this research and are based on (premature) commercial applications that are not representative of this research. We take a closer look at this research from a human-computer interaction perspective and use this premise to comment on these relations and reflect on their values.

Keywords: Augmented reality, Human-computer interaction, Smart environments, Postphenomenology, Google glass, Optical see-through

INTRODUCTION

Don Ihde (Ihde, 1990) introduced four human – technology – world relations in what has become known as a Postphenomenology philosophy. How we experience the world and in the world is analyzed in Phenomenology. Postphenomenology focuses on technology and the mediating role technology has in this examination. Humans interact with the world, or rather their environment, with and through technology. It is argued that the relation between the human subject and technology is the result of mutually shaping each other (Verbeek, 2015). In this paper, we focus on digital technology as we encounter it in smart environments and augmented reality. It is not the case that after the initial years of Postphenomenology research digital technology has been neglected. Rather we think that the observations on digital technology have been wrongly guided by how it has been presented in commercial applications instead of how in research communities longer-term research developments are envisioned.

This paper focuses on Ihde’s human – (digital) technology – world relations and not on the many other aspects of the philosophy. We often see in investigations of these relations that humans are assumed to be users of digital technology. This is not untrue, but it ignores situations where the user is not or hardly aware of the presence of the digital technology and that rather than the human user, the technology guides the user, decides about the user’s behavior, and enforces this behavior. From a “use” perspective, such smart environment technology is quite different from more traditional situations where a human decides to use a hammer, a tablet, or to take his bicycle for a

ride. And it is different from situations where computer-generated interactive content is added to a user's perception of an environment in which he has to perform a task, wants to be entertained, or just happens to roam around.

In the next section, we introduce Ihde's fourfold relationships, followed (section 3) by a discussion on extensions of these relations. Section 4 contains observations on smart environments and augmented reality (AR). We distinguish between "annotated augmented reality" and AR where we introduce virtual objects that need to be aligned with and embedded in the user's perceived environment and that we can interact with. In section 5 we comment on Ihde's relations and their extensions, emphasizing their shortcomings and associated challenges. Conclusions can be found in section 6.

POSTPHENOMENOLOGICAL HUMAN-TECHNOLOGY-WORLD RELATIONS

In this section, we introduce the original human – technology – world relations and some views on their possible extensions. Moreover, we have a look at the notational schemes that have been introduced to give some formalizations to these relations. These relations are the "Embodiment", the "Hermeneutic", the "Alterity", and the "Background" relation.

In the "Embodiment" relation users relate through technology. That is, perhaps after a habituation period, humans are not aware anymore of the technology they are using to perceive their environment or act upon it. Think of eyeglasses, oven gloves, a remote control, or even a bicycle or a car. In Ihde's notation, we have (*human – technology*) → *world*. The mediating technology withdraws to the background.

In the "Hermeneutic" relation humans access the world with technology. The technology makes it possible to "read" the environment. A wristwatch allows one to read the time. Ultrasound and infrared scans can be used to provide visualizations of the invisible. For example, objects in the environment emit infrared light that makes them visible with night vision goggles. Images can be constructed from telescope observations that make galaxies in the early universe visible. We "read" the display interface of the mediating technology. We have *human* → (*technology – world*).

This hermeneutic "reading" is not present in the "Alterity" relation where the technology is experienced as a quasi-other. We have this experience when interacting with an ATM or a household robot. It certainly helps if we are nudged into thinking that in some way the device has some human characteristics (anthropomorphism), for example, playing chess and trying to be smarter than the algorithm. In a scheme, *human* → *technology* (– *world*). The world withdraws to the background.

Lastly, we have the "Background" relation where technology operates in the background of the user's environment. The sounds of a fridge or an air-conditioning system, but we can add technology in urban environments that we are hardly aware of, such as traffic lights, announcement boards, or billboard animations.

We conclude this section by mentioning that in most postphenomenological papers there are no or hardly attempts to relate these traditional

human-technology-world relations to 21st-century research into smart technology designed with an agency. That is, smart devices and environments that not only monitor users but also take the initiative to support users or enforce particular human behavior that is foremost in the interest of the smart environment (its owners, its managers, and its virtual or physical agents that perform tasks in the environment). Before entering a discussion about how to fit smart and AR technology and smart environments into the relation framework, we will look at how Ihde's relations have been interpreted and extended.

TECHNOLOGICALLY MEDIATED INTENTIONALITIES

In (Verbeek, 2008) attention is drawn to technologically mediated intentionalities to augment Ihde's four human-technology-world relations. In Ihde's relations, except the alterity relation, human intentionality is mediated by technology. Verbeek distinguishes three types of "posthuman" intentionalities that involve "artifactual" or technological, rather than human intentionality. They are the augmented, constructive, and reflexive intentionalities. We focus on the constructive intentionality where the technology is assumed to have intentionality directed towards the world, constructing a new reality, and the involved human intentionality is directed towards the result of this technological intentionality. The assumption is that any technology is directed at specific aspects of reality. The usual example to show the difference between technological and human directedness and intentionality is the distinction that can be made between a sound recorder that picks up many different sounds that are not perceived by human beings and the way humans will focus on the sounds that are important for them. This technological directedness is also involved in the "technology" of the Hermeneutic human-technology-world relation. With this technological directedness and intentionality in mind, we can rewrite the Hermeneutic relation as a composite relation *human* → (*technology* → *world*).

The "Hybrid" or "Fusion" relation (Verbeek, 2008), addresses the situation where humans have implants (for example, cochlear or brain implants). Technology is merged with the human body. We have a (*human/technology*) → *world* intentionality and relation. This intentionality has an integrated human and technological form.

In these observations, the emphasis is on perceiving rather than on acting and creating changes to the real world. Perception with technology is an active process. From proximity sensors to radio telescopes, signals are sent out to perceive activity and objects, including otherwise imperceptible objects. Perception technology is also active from the point of view of enhancing and transforming observations and making them accessible and experienceable. Perception technology turns into pervasive or ubiquitous computing technology when this technology starts interpreting and acting on the results of the perception. Agency becomes part of the technology. In ubiquitous computing, the technology perceives the environment and its users and decides how it will interact with its users, support, guide, persuade and control them.

SMART AND AUGMENTED REALITY TECHNOLOGY

Smart technology embedded in our environments, whether at home, shops, offices, and public spaces, knows about us, our behavior, and preferences, and it is pro-active, meaning that it can predict and anticipate activities and perform acts, without the user being necessarily aware of. In these environments, we can have multiple users, including social robots and virtual assistants that show autonomous behavior. Users have wearables (smartphones, smartwatches, health monitors, fitness trackers, smart glasses) that connect them to their environments and the Internet, making them nodes in an Internet of Things (IoT) with limited freedom of behavior. In addition to the users, there are other stakeholders, the suppliers, the owners, managers, and there are regulations that determine the behavior of smart environments. To summarize, a ‘user’ of smart environments can not be compared with a hammer user. The practical dealing with things disappears in the background (Kiran, 2012).

We now zoom in on AR to find out how we should place this technology in the whole of Ihde’s and Verbeek’s human-technology relations. Post-phenomenology researchers are not always clear about their definition of AR. We can read the opinion that the world has been augmented since we have languages to describe it (Feyles, 2020) or that bringing a smart speaker into your home makes your home environment an AR reality environment (Wellner, 2020). There is a well-known definition of AR that has guided AR research for 25 years and that nevertheless has not been discussed in this philosophical approach to AR. In short, following AR pioneer Robert T. Azuma (Azuma, 1997), AR combines real and virtual objects in a real environment (1), registers (aligns) real and virtual objects (2), and runs interactively, in three dimensions, and in real-time (3). More recently Azuma predicted that AR, in particular optical see-through glasses, will be the dominant platform and interface, supplanting the smartphone. These glasses are seen as (Azuma, 2019) “... the best chance of achieving the long-term vision of ubiquitous consumer AR displays.”

Rather than following AR researchers we see the philosophy of technology researchers basing their philosophy on the still rather limited possibilities of commercial AR technology such as the early version of Google Glass, a game such as Pokémon GO, or more recently the IKEA Place app, and assuming they are representative for AR. From (Wellner, 2020) one may understand that only in 2015 (Liberati and Nagataki, 2015) it has been suggested that a virtual layer can include objects rather than information that improves a user’s perception. In contrast to this, it should be noted that in the first AR demonstration in 1968 graphics (3D wireframes) were superimposed on reality. Doing so in such a way that virtual 3D objects and possible dynamic behavior become indistinguishable from real objects and their behavior has been the main aim of AR research since the mid-nineties of the previous century. Indeed, in this post-phenomenology literature, AR is mainly seen as adding verbal and symbolic information to a user’s perception of the real world, rather than looking at AR as a creator of new experiences, not only at home or work but also in an outdoor environment (Nijholt, 2021a).

Hence, our starting points are Azuma's AR definition and the research aims of AR researchers as they are reported in the yearly ISMAR conferences since the late 1990s. Moreover, we focus on optical see-through (OST) AR with head-mounted devices (HMDs) such as Microsoft's HoloLens or the Magic Leap HMD. Capabilities of these bulky and heavy devices will become available for smart glasses or even smart contact lenses and retina implants in the future. One example is Facebook's research on smart glasses in the ARIA project (Abrash, 2020).

Some basics of AR need to be mentioned. HMDs have position and orientation sensors. They provide a user with the right perspective of virtual objects embedded in and aligned with the perceived real world. HMDs have cameras and vision software that tries to understand the real world as far as needed for taking care of this alignment. This includes taking care that, depending on the user's perspective, parts of real and virtual objects may be hidden (the occlusion problem). The digitally created objects may address senses other than the sight sense. Perhaps we can hear, smell, touch, or even taste them, in a multi-sensorial approach to AR (Karunanayaka et al. 2021). But it should be clear that AR devices allow us to manipulate these objects and the objects themselves can be dynamic. To position AR in the human-technology framework we focus on visual-oriented AR, but we should be aware that we can interact in many different ways with our HMD and through this HMD with the virtual, computer-generated objects. In addition, using an HMD does not prevent us from interacting with real-world objects. One aim of vision-oriented AR is to make virtual objects indistinguishable from real objects (Itoh et al. 2022). We can walk in the real world around a virtual building to see how it fits in the environment where it is supposed to be built. While driving our car, our smart glasses can show traffic information and we can use speech commands to search for alternative routes while being attentive to the traffic in the real world.

We mention two AR examples that will play a role in our next section on extensions of Ihde's and Verbeek's human-technology relations framework. We do not consider AR applications where the user is 'only' provided with textual or symbolic information about the environment or an activity that is perceived in a real environment. In this case, the real world is annotated with useful information about buildings, streets, shops, restaurants, and public spaces, but there are no changes to the look of the world. We rather call it "annotated reality". Notice that with the original Google Glass you can have a small rectangle window transparently displayed on the glasses' view that can provide textual information about what you see. In that case, you can say that the glasses 'read' the environment, not that different from a thermometer reading the temperature. But mainly it has the function of a speech-controlled control panel that lets you make a video call, take pictures, make recordings, show directions, et cetera. That is, there is not always (need of) alignment and there is not necessarily a relation to what is happening in this window with what you see of your environment through the glasses. In Pokémon Go, the AR aspects come down to (re-)annotate locations with different Pokémon creatures. Hence, these are not very good examples to discuss how augmented

reality technology fits in a human-technology-world relation à la Ihde and Verbeek.

Our first example is about having virtual 3D blocks added to the real world, for example in a parc, such that they constitute an obstacle course. In the past, traditional video games have been given similar real-world implementations (Johnson, 2017). Our HMD shows the obstacles and we obediently jump over these virtual obstacles or jog around them. The alignment according to Azuma's definition should take care that the obstacles fit naturally in the environment and the course can be transported from one environment to another. Moreover, the user can be given tools that allow them to adapt the course, introduce more or other (virtual) obstacles. This requires the automatic modeling of a particular environment by the AR device, which is a topic in current AR research. Our second example is about introducing humanoids or animoids in a virtual layer where they are superimposed on reality. They can play the role of an assistant that has particular knowledge of the real world or educational activity that will take place in the AR environment. But they can as well act as a companion of the user. AR allows us to walk a virtual dog in the real world. The virtual dog should display real dog-like behavior, be aware of approaching real pedestrians, avoid them, and show awareness of the real world, for example by being startled by a real dog barking and answering. This requires a correlation between the real world and the content in the layers that are superimposed on the real world.

HUMAN – (SMART-AND-AR) TECHNOLOGY RELATIONS

As mentioned in (Michelfelder, 2015) and (Luan, 2020) the postphenomenological framework should be extended to account for contemporary technologies. In this section, we discuss some of the proposed extensions. We will first look at smart technology as we explained above, and then zoom in on AR. In (Verbeek, 2015), and some other papers, an "Immersion" relation is suggested. While in the original Ihde's relations we can speak of a user of technology we now have a more balanced situation where we can say as well that the technology has the intention to use the user. It is not only that the technology is directed to particulars of the world, but it is also that, from the point of view of the "user", the technology displays autonomous behavior, not only perceiving human beings but also acting upon them. As mentioned in (Rosenberger and Verbeek, 2015), human beings are directed towards technologies that are also directed towards them and perceive their users and act upon them. Therefore, we now have bidirectional intentionality: *human* \longleftrightarrow *technology/world*.

In (Vindines and Watson, 2021) it is mentioned that the Immersion view can be understood as an active version of the Background relation, where the environment is aware of human beings, actively interacts with them, and is pro-active, anticipating a user's actions and being able to respond to them smartly. The environment and technology become merged, just as in the fusion relation humans and technology become merged. Looking at our description of smart environments we conclude that this immersion relation acts at a very global level of human-technology world relations and the other

relations that have been mentioned above can be present as well for a human moving around in a smart domestic or urban environment.

We now turn our attention to augmented reality. Above we discussed augmented reality according to Google Glass and we gave two other AR examples. In various papers, the “Augmentation” relation is introduced. In these papers, Google Glass is seen as representative of AR and that is certainly not in line with how the scientific research AR world views AR. And that applies not only to Google Glass but also to a different kind of AR such as the handheld non-OST Pokémon GO game.¹ In (Rosenberger and Verbeek, 2015), the Augmentation relation consists of a parallel Embodiment relation, that is, $(human - technology) \rightarrow world$, and a Hermeneutic relation, that is, $human \rightarrow (technology - world)$ relation. Before adding details, it should be mentioned that despite augmentations that can be perceived, an AR user remains in the real world and therefore the previously mentioned relations remain. But we can look at specific characteristics of AR technology and see how they relate to these human-technology-world relations. In (Wellner, 2020) it is argued that each of the four original relations covers a certain aspect of the AR experience. We have to make a caveat here. There are various kinds of AR (spatial, optical, and video see-through, and mirror-AR) and therefore various kinds of AR devices with different capabilities. As mentioned before, we focus on OST HMDs such as HoloLens or Magic Leap and their future miniaturized versions such as smart glasses with similar capabilities. Google Glass is not representative of what the AR research community considers to be AR, but it nevertheless shows some of its characteristics so we can adapt to some views on AR expressed in (Rosenberger and Verbeek, 2015), (Wellner, 2020) rather than starting from scratch.

We have explicit commands to the AR device. For example, telling the device to take a picture. We have explicit interaction with the perceived virtual content. For example, using a data glove to grasp a virtual object or telling a virtual dog to stop barking. These commands are mediated through the AR device or other connected wearables, but not directed to these devices. If after a habituation period, an AR user is focused on the task at hand we can indeed assume an Embodiment relation with the device. But to realize this Embodiment relation in a situation where an AR user may change his position and orientation, or just makes head movements, a lot of computation has to be done in the background, hidden for the user, to take care that the virtual content remains aligned with the real content despite the changes in the user’s view of its AR environment. This is not an issue for the original version of Google Glass, nor is this managed in AR games such as Pokémon GO. The Embodiment requires computing, it requires activity from the device. AR devices provide us with a “reading” of the augmented world, so we can also include the Hermeneutic relation here. Again, a caveat here. A wayfinding application on Google Glass can have a route, for example by generating colored lines, on a real-world view of streets or a map representation of the

¹A ‘proof of concept’ implementation of Pokémon GO for smart glasses was introduced by Niantic in 2021.

real world. We “read” descriptions of characteristics of the real world. However, rather than annotations, we have virtual layers that can contain virtual objects that do not annotate real objects, but that add (dynamic) virtual content that is not necessarily distinguishable from the real-world content in the view presented by the AR device.

Not being able to distinguish (multisensorial) artificially generated stimuli from natural stimuli from a real-world environment introduces the “AR Immersion” relation. It requires the pose-change computations and the continuous alignment of the virtual with the real content. In the AR environment, we can also have the Alterity relation with quasi-others, either real or virtual. For example, the virtual dog in our AR example is a quasi-other. More generally, we can have virtual humans as quasi-others in the virtual layer (Nijholt, 2021b) with whom we can interact.

AR devices are active. They update the alignment of the virtual with the real content whenever there are pose changes. In the case of vision-oriented AR, every pose change requires new computer-generated imagery to compute a new AR view. An AR device’s depth camera emits infrared light to construct a 3D model of the real environment that allows the handling of occlusion problems. For these reasons, we suggest changing the Hermeneutic relation in AR to the Composite relation $human \rightarrow (technology \rightarrow world)$. Making changes in a virtual layer can affect the real world. For example, sensors in the real world can be triggered by changes in the AR world and allow a virtual human to enter a physical location by opening an automatic real door. This can be interpreted as Ihde’s Background relation. Earlier we mentioned the Fusion relation. There is an ongoing miniaturization process where bulky AR HMDs are transformed into less obtrusive smart glasses, contact lenses, and retina implants. Hence, when that is happening, we also have, from an AR’s point of view, a Fusion relation with a merging of human and technological intentionalities.

Smart and AR environments have in common that they can be directed to the users. The user can be monitored, guided, and controlled by the smartness that is present in the smart environment or the virtual AR layer. For that reason, in the context of AR, (Wellner, 2020) introduced the “Relegation” relation $human \leftarrow (technology \rightarrow world)$, where, using her words, “... the human intentionality “withdraws,” and the technological intentionality “takes over.””

CONCLUSION

We surveyed the different points of view philosophers have adopted for incorporating AR and smart environments into their human-technology-world relations and added some of our interpretations of these relations to make clear that the “technology” in the human-technology-world relations is not always fully understood or is oversimplified. We also mention the ambiguity and the ‘black boxing’ in the relations and the schemes. The postphenomenological framework does not take into account technology that makes the so-called quasi-others agents that take the initiative and make changes to a “user’s” daily (digitally enhanced) environment. We can also conclude

that the views mainly focus on smartphone AR which has different capabilities than a head-mounted AR device such as the HoloLens and its future miniaturizations. A question that pops up is whether we should extend the framework of relations with a productive Hermeneutic relation as suggested in (Wellner, 2017) where technology adds objects (and leaves traces) to a perceived environment. The important question remains: does the postphenomenological approach to how users experience technology have implications for interaction design by human-computer interaction researchers? At the moment that question cannot be answered in the affirmative.

REFERENCES

- Abrash, M. (2020). Facebook Connect Keynote. Project ARIA. Available from www.youtube.com/watch?v=woXmJMw2lTM [accessed 9 February 2022].
- Azuma, R.T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6:4, pp. 355–385.
- Azuma, R.T. (2019). The road to ubiquitous consumer augmented reality systems. *Hum Behavior & Emerging Technologies* 1, pp. 26–32.
- Feyles, M. (2020). Phenomenology of Augmented Environments. In: Iannilli, G.L. (Ed.) *Aesthetic Environments*, Aesthetica Edizioni No. 114, pp. 99–111.
- Ihde, D. (1990). *Technology and the lifeworld: from garden to earth*. Bloomington: Indiana University Press.
- Itoh, Y., Langlotz, T., Sutton, J. and Plopski, A. (2022). Towards Indistinguishable Augmented Reality: A Survey on Optical See-through Head-mounted Displays. *ACM Comput. Surv.* 54, 6, Article 120, 36 pages.
- Johnson, L. (2017). Watch This Guy Become Super Mario in a HoloLens AR Game. *Motherboard*. Available from www.vice.com/en/article/bj84w3/watch-this-guy-become-super-mario-in-a-hololens-ar-game [accessed 9 February 2022].
- Karunanayaka, K. et al. (2021). Multisensory Augmented Reality. In: Ardito, C. et al. (Eds.) *INTERACT 2021*. LNCS, vol 12936. Cham: Springer, pp. 558–563.
- Kiran, A.H. (2012). Technological Presence: Actuality and Potentiality in Subject Constitution. *Hum Stud* 35, pp. 77–93.
- Liberati, N. and Nagataki, S. (2015). The AR glasses’ “non-neutrality”: their knock-on effects on the subject and on the givenness of the object. *Ethics and Inf. Technol.* 17, 2, pp. 125–137.
- Luan, S.T. (2020). The Hidden Dimensions of Human-Technology Relations. *Philosophy and Technology* 33 (1), pp. 141–165.
- Michelfelder, D. P. (2015). Postphenomenology with an eye to the future. In *Postphenomenological investigations*. Lexington Books, pp. 237–246.
- Nijholt, A. (2021a). Experiencing Social Augmented Reality in Public Spaces. *UbiComp-ISWC ‘21 Adjunct proceedings*, New York, ACM, 6 Pages.
- Nijholt A. (2021b) Augmented Reality Humans: Towards Multisensorial Awareness. In *Digital Economy. Emerging Technologies and Business Innovation. ICDEc 2021*. LNBIP, vol. 431, Cham: Springer, pp. 237–250.
- Rosenberger, R. and Verbeek, P.-P. (2015). A field guide to postphenomenology. In *Postphenomenological Investigations*. Lexington Books, pp. 9–41.
- Verbeek, P.-P. (2008). Cyborg intentionality: Rethinking the phenomenology of human-technology relations. *Phenom Cogn Sci* 7, pp. 387–395.
- Verbeek, P.-P. (2015). Beyond interaction. *Interactions* 22, pp. 26–31.
- Vindenes, J. and Wasson, B. (2021). A Postphenomenological Framework for Studying User Experience of Immersive VR. *Front. Virtual Real.* 2:656423.
- Wellner, G. (2017). I-Media-World: The Algorithmic Shift from Hermeneutic Relations to Writing Relations. Chapter 11 in *Postphenomenology and media*. van den Eede, Y. et al. (Eds.), Lexington Books, pp. 207–227.
- Wellner, G. (2020). Postphenomenology of Augmented Reality. In *Relating to Things: Design, Technology and the Artificial*. Bloomsbury, pp. 173–188.