Techné: Research in Philosophy and Technology ISSN: 1091-8264
19:1 (Winter 2015): 92–112 DOI: 10.5840/techne20154229

The Realism and Ecology of Augmented Reality: An Ecological Way to Understand the Human-Computer Relationship

Giovanni Simonetta

Abstract: Unlike in the phrase "Virtual Reality," in the phrase "Augmented Reality" (AR) the stress is put on the word "reality." It seems, though, that we still lack a concept of reality which can fit the world of both humans and computers. In connection with this philosophical issue, this paper aims to provide the background for a better insight into the meaning of Augmented Reality and its impact on human behavior. My thesis is that an ecological version of direct perception's realism constitutes the most natural framework from which to start. The ecological approach to perception—namely, the Gibsonian theory of affordances—together with a non-dualistic, pragmatist and evolutionist notion of reality, perfectly fits this purpose. Thus, after a brief survey of the present state of AR technologies, it should become natural to interpret AR digital contents as implementations of affordances.

Key words: augmented reality, ecological psychology, animal-environment synergy, affordances, realism

1. What is Augmented Reality?

Knowing what hides behind that building's corner; being well acquainted with the streets of a city never visited before; wandering around an archeological site visualizing how those ruins appeared two thousand years ago; assembling an airplane starting from five million metal pieces; fixing a failure in the water network without blocking traffic for a week; doing minimally-invasive surgery; having at your disposal the profile of an unknown person, his preferences, his expectations, his diseases or his criminal record. All those activities presume for the successful

Giovanni Simonetta, University of Florence, Via Bolognese n. 52, 50139 Firenze, Italy; +39 334 3561095; giovannisimonetta84@gmail.com

performer a knowledge of the world so extraordinary we could call her a magician. On the contrary, this is all reality, or better, Augmented Reality (AR).

In this section I will give some useful definitions and make a brief report on the actual state-of-the-art in AR. Following Paul Milgram and Fumio Kishino's lead (1994), an initial taxonomic definition of AR can be given in terms of AR's location along the "Reality-Virtuality Continuum," which runs from the "real environment" to the "virtual environment." A real environment is reality as common sense sees it. A virtual environment is a totally synthetic world, fully modeled, digitally created from scratch. Between these two poles lies the environment of "Mixed Reality," composed of two subclasses. On one side—going from the center to the virtual environment pole—lies "Augmented Virtuality" (AV), on the other side—from the center to the real environment pole—is "Augmented Reality." AV is Virtual Reality with the addition of real data, while AR is reality augmented with digital objects. Notice, however, that there are no sharp boundaries between these two things, only fuzzy ones. Indeed, the more we venture towards the center of the continuum, the more difficult it is to say whether we are in AV or in AR. Thus, unlike Virtual Reality, AR constitutes a supplement to, and not a replacement of, reality *tout-court*.

A more explicit definition was given by Ronald T. Azuma (1997) according to whom AR is whatever system satisfies the three following characteristics: 1) combining the real and the virtual (as seen above); 2) interacting in real time: that is, it follows the observer's movements in the augmented environment; 3) being registered in 3-D. The latter is the most challenging task for engineers, as it implies that virtual 3-D objects must respect the same laws as real objects: the same occlusion relationships, having a shadow, and so on.

Within this general definition, keeping in mind the various applications and the possible developments of this kind of technologies, it is possible to distinguish three different meanings of the expression "Augmented Reality" (Simonetta 2012). The first, and most common, account of AR identifies it with certain technological devices with specific practical purposes. Connected to the first, although distinct, is AR as communication strategies, mainly carried out in advertising to promote services, events or products by means of digital apps augmenting (or replacing) traditional advertisements. Finally, looking towards the future, AR can be interpreted as a scenario in which the contamination and the full mimesis of digital contents will make useless the need to distinguish between Augmented Reality and common reality.

Let us turn now to AR devices and the ways AR contents are being used. Focusing on the visibility of the device and on the camouflage of the contents, it is possible to classify all of them into three categories: "video devices," "optical devices," and "spatial devices." In the first case we are dealing with a display (a PC monitor, a smartphone, a tablet screen or some kind of closed-view Head-Mounted Display) on which digital contents are superimposed over images of the real world captured by a video camera. These are the most common devices. Systems such as "1st and Ten" (the dynamic super-impression of the first down line during a football game on television), "paper markers" (printed onto paper magazines, schoolbooks or on the packaging of various products which, once captured by a video camera, instruct the computer to overlay 3-D graphics onto the real world reproduced on the computer screen), smartphone apps like "Google Sky Maps" (allowing us to visualize the constellations situated in a certain portion of the sky by simply pointing at it with the phone camera, even when it's cloudy), and the so-called "mirror apps" (permitting us to visualize an image of ourselves and to superimpose onto it such digital contents as a pair of glasses or a shirt) fall into this category. Here the awareness of the device is high. Moreover, by simply taking the eye off the display we can look at reality, eliminating the computer-generated part which is, thus, minimally camouflaged. In the case of see-through "optical devices" (e.g., certain types of Head-Mounted Displays, special glasses or, more futuristically, special contact lenses), the mediation of the camera is disposed of, and reality is seen in the traditional way, but through a semi-transparent surface upon which the digital contents are visualized. Devices of this kind are the most promising for the near future. The images on the semi-transparent display appear in an appropriate portion of the visual field allowing us, for example, to see road signs where they are absent; to look at some ruins and see over them the amphitheater as seen by an observer in the same position two thousand years ago; to repair our washing machines without calling the technician; or, again, although not yet, to do micro-invasive surgical operations where the surgeon can see inside the patient's body by means of digital mock-ups superimposed on the patient's skin. Incidentally, also falling into this category is the device invented by Tom Caudell to help Boeing workers put together the five million pieces constituting an airplane. It was this same Caudell who coined the expression "Augmented Reality" in the early 1990s (Caudell and Mizell 1992). Of course, Caudell's device was much bigger than today's glasses. Optical devices can be more or less visible and their contents more or less camouflaged, however the awareness of wearing a device is always present and one is always able to distinguish the objects traditionally seen from the digital ones. Lastly, we come to what I call “spatial devices,” borrowing “spatial” from the locution “Spatial Augmented Reality” (SAR). This type of device enriches the real environment with digital contents by projecting them. In this way every real environment surface becomes (more) interactive (than before). Furthermore, there is no need for wearable devices as the observer meets AR contents by moving and touching both ordinary and computer-created objects. Observers are immersed in a totally new environment, they have no direct awareness of the device, and the contents’ camouflage is maximum. What would make traditional objects different from augmented ones would be only a prior acquaintance with the former.

From the above distinctions it should be clear that “Augmented Reality enhances a user’s *perception* of and *interaction* with the *real* world” (Azuma 1997, 3; italics mine). Unfortunately for now we lack an analysis of the notions of perception, interaction, and reality with specific reference to Augmented Reality. This is a task for psychologists and philosophers. The present work represents a step in this direction. Its main thesis is that ecological theory of perception, as formulated by James Gibson between the 1930s and the 1970s, is the best theoretical framework to understand this emerging technology. In the following I will try to bring out some problematic issues raising from a philosophical consideration of AR, of its contents and, most of all, of what constitutes its biggest innovation, namely its modality of fruition. How are the digital contents augmenting reality perceived? What is their status? Are they virtual or are they real? And if we could find some reasons to argue they are real, to what extent they are and what allows us to distinguish them from ordinary reality? Such questions are well known in the field of philosophical consideration of virtual reality, but in the case of AR they come to the fore because here the fruition of computer-created contents—the so-called virtual environment—takes place along the very same perceptual modality which is active in the perception of the natural environment. The main feature of AR systems is that, going back to Azuma’s definition, they are “registered in 3D.” That is, at least in the ideal case of invisible devices with totally camouflaged contents, they are such that one cannot distinguish them from the reality to which they have been added. Experiencing an AR content is different from looking at a picture or watching a movie, insofar as perception of the latter carries along with it the perception of its support: the picture or the screen, while the former’s support is just air. In the case of an AR object I can go around it and look at its sides and back. Its edges are such that the occlusion relationships with real objects are consistent with its position among them. AR systems are completely different from other media as we can listen to the radio, read a newspaper, and watch television, while an AR system is experienced, with the whole body and all the senses, just as the real world is experienced.

Before going on to introduce ecological theory, a clarification is in order. This paper will focus on visual perception both because Gibson's work has been mostly concentrated on visual systems, and because current AR devices mainly convey these kinds of contents. Nevertheless, the argument could be easily extended: the ecological theory of visual perception is part of a theory of all perceptual systems (see Gibson 1966), and technological innovation is making progress in view of the transmission of olfactory, gustatory and even tactile contents. So, I ask the reader to try to visualize the ideal scenario described above, where "augmented" reality and "ordinary" reality are indistinguishable: where there would be full tri-dimensional alignment, real time interactivity, and no awareness of the device. At the current state of technological development, the above-mentioned AR devices cannot realize this technological dream but perhaps such realization is not too far away (for some developments in the creation of AR contents for all the senses see, e.g., Craig 2013). Keeping in mind this scenario, the ecological interpretation of AR will reveal itself with all its pregnancy.

2. The Ecological Theory of Perception and Ecological Realism

Walking along an uneven road without stumbling; halting before falling into the ravine; grabbing objects; sitting on a chair; recognizing places, situations and people; and other similar activities presuppose in the doers a certain knowledge of their surroundings. And, it will be admitted that an adequate theory of perception should account for all these facts.

The inquiry into human perception has a long-standing tradition both in philosophy and psychology and, at least in the Western world, has mostly served the task of justifying supposed true beliefs about the so-called "external world." The ways of justification have been of the most varied sorts but, both in rationalism and empiricism, they have dealt with demonstrating a correspondence between something inside the mind, perceptions and representations, and something outside the mind: objects and events in the external world. In short, they have always followed along the lines of Plato's soul-body dualism by translating the problem of perception into the Cartesian one of the mind-body relationship. According to some ecological psychologists we could go even further by saying that there is a more fundamental dualism from which all representational theories of mind and perception arise: animal-environment dualism (Turvey and Shaw 1979). Experimental psychology also encourages this kind of dualism to the extent that it assumes that incoming sensory stimuli are insufficient and ambiguous, and that an adequate perception of the environment requires some form of organization and correction of these very stimuli on the part of the mind. Cognitivism, one of the most recent formulations of this type of theory, holds that the mind has a computational nature: it processes sensory inputs (*per se* meaningless) in an algorithmic way (coding them according to its own language, so-called 'mentalese,' and in compliance with its own syntactic rules), thereby giving knowledge about the external world in the form of mental representations. Following Norman (2002) the labels "constructivism" and "constructivist approach" will be used to refer to all those theories of perception which, although highly varied in nature, do share some basic tenets, namely: the stimulation coming through the senses is impoverished and needs to be processed by an intelligent cognitive structure; the senses are channels transmitting, in the form of electric impulses, the messages received from the physical world of atoms, molecules and forces to a central processing unit called "mind" and whose seat is often supposed to be the brain; sensation is different from perception insofar as the former is an immediate consequence of a stimulus while the latter is the product of a mediate cognitive process; memory has a major role to play in this process of mediation between sensations and perceptions; researchers' attention is largely focused on the process of the elaboration of the stimulus rather than on the stimulus itself. The problem faced by constructivism is to explain how we can construct a perceptual environment, made out of meaningful objects and events, starting from physical and *per se* meaningless stimuli. Put in another way, the problem is how to justify the belief that sensory stimuli carry information about their external sources, i.e., about the environment's features and properties. How to distinguish, in the percept, what is the information coming in from the external world and what is a mere product of the mind?

Ecological theory, as formulated by Gibson in *The Ecological Approach to Visual Perception* (1979) and in other works from the 1970s, stems from the very critique of the constructivist concept of stimulus for perception. Gibson denies that what the perceiver meets in the world is just rough material, isolated and meaningless stimuli summing up to electric impulses. He refuses to identify the senses with the sense organs, the receptors, thus rejecting the idea that they are mere channels for the transmission of impulses to the brain. Since his early works on perception Gibson cut short the above-mentioned problem of constructivism, claiming that the sensation-perception dichotomy is a false one (Gibson 1937): on the one hand, if sensation is meant in term of electric impulses triggered when the receptors of a single sense organ come in contact with certain composition of atoms in the physical world, then we must conclude that sensation and perception belong to two absolutely separate and incommensurable universes of description; on the other hand, if sensation is the information encountered by the whole perceiving agent in his contact with the environment, then it is indistinguishable from perception as the information received through the senses does not need any processing, being already structured and meaningful. This last claim summarizes Gibson's contribution to psychology. Its radicality is a sign of the deep divide between the ecological approach and the theories of perception which preceded it. Its justification has been the aim of over fifty years of work that led Gibson to a profound redefinition of what is the information we encounter when we are in contact with the world and of how this contact happens. It is a radically new conception of both stimuli and the senses. We turn now to a presentation of Gibson's proposal.

The Gibsonian solution to the problem of how to convert meaningless sensory data into perceptions of meaningful objects and events lies, as we already pointed out, in the rejection of the stimulus as rough, simple, isolated material. The ground that ensures solidity to this solution is the distinction between two different and incommensurable levels of description: that of psychological description and that of physical description. The ecological notion of the perceptual environment, which is assumed to be the only relevant for psychology, has nothing to do with the world as described by physics. While the latter is made out of atoms, molecules and forces—with animals too being viewed as complex configurations of these elements—the former, being made out of meaningful objects and events, implies the presence of a living animal perceiving them. Animal and environment constitute an indivisible pair; an animal implies an environment to act and perceive in, and an environment implies an animal to surround. Thus, in the framework of ecological psychology it is impossible to study the perceptual process independently from the environmental context in which it is performed; a strong animal-environment synergy is here presupposed.

In the psychological description of the environment one must take into account what matters to the animal, namely its "layout" or structure of the terrestrial environment necessary to grasp persistence under change, its "variants" and "invariants." The layout of the environment is described by Gibson in terms of three elements. First, a "medium" which allows the animal to move and perceive substances. In the case of humans this is a homogeneous medium like air, which permits vision being completely illuminated. Secondly, "substances," more or less solid, which are not homogeneous. Lastly, "surfaces" that separate substances from the medium (every substance has a surface), each of which provided with its unique combination of resistance, texture, form and light reflection in compliance with the "ecological laws of surfaces." Objects are thus defined as persistent substances with closed surfaces and with a particular layout, color, resistance, texture, shape and reflectance.

A key role in ecological theory is played by light. Gibson distinguishes "ambient light" or "illumination" from the light of physics considered as mere electromagnetic radiation. Unlike the latter, ambient light is light reflected many times off the surfaces in a medium before reaching a certain point located within the medium itself: the percipient. By virtue of these reflections ambient light is structured and so is capable of being a means of information, itself structured, about the environment. Ambient light, as it is structured relative to a certain point of observation in the medium, is called an "ambient optic array" and can be described in terms of "solid visual angles," a sort of point of view. With a changing point of observation, that is, as a consequence of the observer's movements the optic array becomes an optic flow and solid visual angles and perspective change too. It is during this very change that it is possible to detect some "invariant structures" constituted by certain characteristics of the environment showing a tendency to persistence. Changes in the optic flow, amounting mainly to variations of the surfaces' texture gradient and modifications of the occlusion relationships between them, can be determined by movements of the observer, by motions of the objects, and by qualitative changes of the latter like solidification, evaporation, or breakage. However, the perceiver is always able to discriminate among them and this skill's explanation is to be found in proprioception and in the close link between action and perception. The traditional explanation in terms of successive retinal images conceived as point-to-point projection of the external object upon the retina and integrated by inferential mental activities directly leads to the paradox of the little-man-in-the-brain. Notice that here the organ of vision is viewed as a passive receiver of point sensory stimuli. On the contrary, in Gibson's explanation vision refers to a whole perceptual system constituted by a nested hierarchy of organs. The perceiver's awareness of the movements of his eyes, of his head, and of his whole body allows him to establish whether a surface's accretion is due to the growth of an object, to its approaching the perceiver, or to the perceiver's approaching it. Like perception, proprioception too is grounded on the picking up of propriospecific information through all the perceptual systems "and the two-eyes-with-head-and-body-system, in cooperation with postural equilibrium and locomotion, can get around in the world and look at everything" (Gibson 1966, 42). Information contained in the optic flow univocally specifies whether the changes occurring in it are due to movements of the observer, to motions of the objects or to some qualitative change in the latters. Ambiguity is mostly due to conditions of impoverished stimuli. Ambiguity often arises in some laboratory experiments where a static observer is asked to look at a frozen, two-dimensional optic array but, in general, under normal conditions, exploration of the natural environment is sufficient to clear it up. Thus, the perceiver's explorative activity is essential to the detection of the invariants of structure in the optic flow. Perception and action, where the latter coincides with the observer's movement, are both purposive behaviors and, although distinct, are also in close relation: perception investigates the environment, detects meanings and values, and therefore is a guide for action; action performs movements, gets results, and looks for further information for perception. The information in the ambient optic array is potentially infinite and the perceiving animal is engaged in a task of goal-oriented selection; perception is a process of purposive attuning, of coordination, between the animal and relevant aspects of his environment. The animal does not add anything to the environment, limiting himself to the pickup of available information; nevertheless, one cannot say what has been said by earlier critics of Gibson's theory (see, e.g., Richards 1976; Epstein and Park 1964; Gregory 1972): that his is a passive theory of perception. In fact, what Gibson points out against his critics is that the only activity they are willing to admit in perception is a kind of mental activity, while

> the kind of activity . . . that seems to me important is the looking, listening, touching, tasting, and sniffing that goes on when the perceptual systems are at work. These acts involve adjustments of organs, not mere stimulation of receptors. They are not mental. Neither are they physical, for that matter, but functional. My notion of the pickup of information by extracting invariants over time involves the optimizing activity of a system . . . an activity that orients the organs of perception, explores the ambient array, and seeks an equilibrium. (Gibson 1982b, 397–98)

Thus, since information univocally specifies its source and the perceptual systems evolved to detect it, there is no need of any mental process to elaborate it. The only activity required here is exploration.

The animal-environment synergy, the indissoluble link between perception and action, the theory of perception as information pickup, the idea of psychology as engaged in the study of the behavior of purposive animals inside their own environment, all of this finds a definitive systematization in Gibson's concept of "affordance." The ways the surfaces' layouts appear have an intrinsic meaning relative only to the animal to which they offer. Affordances are, specifically, “values” and “meanings” that the environment affords an animal in terms of possibilities of use and action, benefits and dangers, and behaviors in general. Affordances are to be found in the environment and arise, like all perceptions, from the surfaces’ layouts and then are immediately detectable in the optic array, although the ways these are detected are species-specific. A certain layout will offer different affordances to different animals, perfectly in accordance with the ecological principle of inseparability between animal and environment. To paraphrase the words of the philosopher Henri Bergson, whose ideas about perception have striking affinities with the theory of affordances, perception is a virtual action (Bergson 1911).

After Gibson, other ecological psychologists felt the need to intertwine the notion of affordance with that of “effectivity”: that is to say, the ability of the acting animal to recognize and use the environment’s affordances as demonstrated by successfully performing goal-oriented behaviors. This insistence on the animal side of the animal-environment duo emphasizes the need for an ecological definition of the animal. As Gibson defined an environment as a set of affordances or an ecological niche, in the very same way an animal is definable as a set of effectivities, a specific combination of functions determined by the animal’s morphology in relation to its environment (Shaw, Turvey, and Mace 1982). Some of these psychologists conclude that action is virtual perception (Shaw, Kadar, and Kinsella-Shaw 1994).

The idea that the phenomenon of perception from the ecological standpoint concerns not two distinct entities but the animal-environment system in its fundamental unity, and the theory according to which information for perceiving consists of invariants across transformations that are specific to their source, implies a philosophical position that is difficult to collocate among the traditional categories of epistemology. Nevertheless, it is no doubt a form of realism. Besides, if for “realism” we mean the philosophical position that an objective reality exists, with perceivable objects and events the existence of which is fully independent from their being perceived, and which is, at least in part, knowable—i.e., true perceptions are possible—then the ecological theory of perception is sure to be committed to realism. In fact, it asserts the existence of an environment: perceivable objects and events, whose properties or affordances, although existing in their own right, are directly perceivable from animals acting in such a way that it is impossible to deny their having a true knowledge of them.

The keystone of this realism, which has been called “ecological realism,” is obviously the notion of affordance, particularly in its pointing to the direct and indissoluble link between animal and environment. Of course, every attempt to justify the theory of affordance is condemned to failure if one keeps looking at the environment, and so at reality, from the standpoint of physics. Wondering how our knowledge of the world, as described by psychology in terms of values and meanings, could correspond to the physical properties of the world, described by physicists in terms of atoms, molecules and forces, lays down a strong precondition for realism by asking the unanswerable question of how the mind can build a world of meanings out of meaningless electric impulses and chemical reactions. Actually, this formulation of the problem excludes realism *a priori*, insomuch as it states an absolute alterity between the world of stimuli and that of knowledge. Dualism comes up yet again. One could go further in affirming that most of the objections against realism are rooted in the animal-environment dualism. These could be refuted by keeping the psychological description apart from the physical description and by looking at the environment as it is, in its relationship with the animals that live in it, *ergo*, doing without that useless dualism (for some of the most frequent criticisms of ecological realism and the responses to them see Shaw, Turvey, and Mace 1982 and Michaels and Carello 1981). At this point it should be clear that the knowledge to be justified is more pragmatic, a knowledge expressed by animals' goal-oriented actions. Here the only valid criteria of truth are the appropriateness of behaviors and the compatibility between an animal's effectivities and the environment's affordances. At the species level, evolution by means of natural selection guarantees compatibility in the form of a genetic attuning of the animal to the environment. The survival of a species is an indisputable sign that such a compatibility is actually extant. At the individual level, the very same compatibility is achieved by means of perceptual learning intended by Gibson as an education of attention in the selective detection of the invariants, the environmental properties, to the extent that they motivate and/or constrain action. Education of attention means that "the information-pickup will improve with practice. This is perceptual learning. But this does not imply that information is stored in memory. The information continues to be available outside the skin, i.e., the invariants that specify the world. Perception is a skill, not a constructing of the mental world out of psychic components" (Gibson 1982c, 372).

Summarizing what has been said in this section, the ecological theory of perception provides us with the following definitions. First off, "perception" is the detecting of properties of objects and events in the environment intended as a picking up of invariants in the observer's optic flow. Secondly, the "interaction" between an animal and his environment is viewed as kind of osmotic trade between the two, where perception provides information for action in the guise of affordances—meanings, possibilities of action, goals—and action executes goal-oriented movements while providing new perspectives for the detection of new information. Lastly, "reality" is the potentially endless supply of information with which animals interact according the above-mentioned modalities. Real objects and events exist in their own right, their properties belong to them and are not imposed by the perceiver. However, these properties are not to be described in the terms of physics, referring to them as particular configurations of atoms, molecules and forces, but rather in the terms of ecology, as opportunities for action, as affordances. Consequently, the truth of a perception is a matter of the appropriateness and successfulness of the purposive behaviors it suggests to the animal.

3. Augmented Reality as Implementation of Affordances

In Section One of this paper we said that AR systems are systems mixing virtual and real elements, the former being registered in 3D and interactive in real time with the latter, and aiming to "enhance a user's perception of and interaction with the real world" (Azuma 1997, 3). We can point out that "an augmented perception of reality is meant to provide useful information that will allow for a better understanding of our surroundings and improve our decisions and actions" (Kipper and Rampolla 2013, 30). From the technological point of view, this means that "AR can offer interfaces which propose either, more explicitly, information, or, more explicitly, a better mastery of our actions with regard to real events" (Hugues, Fuchs, and Nannipieri 2011, 48). Given this characterization and keeping in mind what we said about AR applications, it becomes evident that the fruition of AR contents is essentially action-oriented. We also claimed that the main feature of AR contents is their being "registered in 3D," that is, their being merged with ordinary objects in such a way that one cannot distinguish the former from the latter, and such that they are experienced in the very same manner, as they were part of the natural environment.

In Section Two we showed that ecological theory, unlike constructivist theories of perception, can account for the very close link between perception and action by means of the notion of affordances coming from the environment. Put in this way it immediately seems to me that AR contents can be interpreted as sources of affordances for the user and that AR systems can be considered as devices for implementing the environment's affordances. AR emphasizes information relating to the environment otherwise not detectable but after a great deal of exploration. Like the affordances of natural objects, AR offers the user what the environment "*offers* the animal, what it *provides* or *furnishes*, either for good or ill" (Gibson 1979, 127; emphasis in original). Let us try to go deeper in this comparison so to see to what extent it works.

One could reasonably ask whether AR augments reality itself or just the user's perception of it. Hugues, Fuchs, and Nannipieri (2011) hold that the question raised by AR environments concerning perception is better understood starting from the sensory-motor approach as opposed to the representational one, insofar as in AR environments perception always aims at a possible action. They ascribe this approach to the theory of Gibson (1966), to the philosophy of Henri Bergson (1911), and to William James (1907). In discussing how AR modifies perception they suggest a distinction between two functionalities of AR environments: augmenting the user's perception of reality on the one hand, and creating an artificial environment, i.e., a new perceptual space, on the other. The former functionality provides the user with a better understanding and mastery of the real environment; this augmented environment "only aims to unveil parts of that space, . . . does not modify anything, it only reveals what already exists" and "presents us reality (what I can see, what I can do)" (Hugues, Fuchs, and Nannipieri 2011, 61). As an example, the authors describe an AR application allowing one to locate and reach the nearest subway station. The latter functionality concerns the creation of an artificial environment "which does not correspond to anything which exists," (49) which "modifies the spatial structure by adding objects, beings or relations which do not belong to it" (61) and "presents us what is imaginary: what I cannot really see—since it is not actually real—but which I can, paradoxically, see thanks to an AR interface" (61). The example suggested here is that of an empty room filled with virtual furniture. This distinction, they say, does not affect the essence of AR for, whether we have to reach the subway station or to furnish a room, the aim is always a practical one. Regarding the question of what is actually augmented in AR environments, the authors admit that it would be hasty to say that reality cannot be augmented because it already contains all that exists and that what is augmented is just perception. Although leaving this question unresolved on purpose, they hold that "the issue concerns perception, i.e., the phenomenon—reality as it is perceived, not the noumen—reality in itself, to use Kant's word" (Hugues, Fuchs, and Nannipieri 2011, 48).

I think that the relation of an affordance to its source is different than that of a phenomenon to a noumen, just as it is different from a relation of a mental representation to an object made out of atoms. Affordances are invariant combinations of features of their sources (objects and events made out of surfaces with their layouts, their textures, changing arrangements in the optic flow, etc.); their relation is univocal and direct, so we can say that to perceive an object is to perceive its affordances. On the contrary, phenomena and noumena are on distinct levels of reality, and representations and the objects of physics belong to different levels of description. In short, the latter distinctions pose the animal-environment dualism again, and starting from the above Kantian distinction would raise all the difficulties of traditional approaches. Moreover, I think that the augmented environment is not artificial in the sense of being not real, i.e., simply virtual, as it seems to be meant by Hugues, Fuchs, and Nannipieri, in the same manner in which an artificial environment like a city cannot be said to be unreal as compared to a forest. Let's see if Gibson's theory can help us solve this question.

Consider the above examples of AR environments, that of the subway station and that of the virtually-furnished room. The computer-created objects are, in the first case, labels or arrows superimposed over and/or interposed among the natural scenario of a city; in the second case they are virtual furniture displaced inside a room; in both cases we take it for granted that they are interactive in real time and perfectly aligned in 3D with the ordinary objects. Now, are these computer-created objects real or virtual? Will one say that just the city and the room are real while the digital arrows and the virtual furniture are not? If the exploration of the digitally-furnished room would provide exactly the same affordances as would the same room with real furniture, what would allow us to distinguish from one another? Maybe the fact that affordances come from atoms and molecules in one case and from bits in the other case? But, as we already saw, this would amount to confuse two incommensurable levels of description, the physical one of atoms and bits and the psychological one of the affordances, and one would find oneself facing the traditional problem of justifying the correspondence between them. On the contrary, keeping the discussion on the psychological level, one would have to deal, on the one hand, with the affordances coming from both the digital furniture and the room's layout, and, on the other hand, with the decision of buying certain furniture rather than other, i.e., with a behavior induced by those very affordances. In this case, it would be possible to evaluate the truth, as it were, of a perception by evaluating the appropriateness and successfulness of the corresponding behavior. If the furniture I bought following an AR catalogue fit the room, then I can conclude that my perception of the AR environment was trustworthy.

From what we said about the two examples of AR environments it seems reasonable to conclude, at the level of ecological description, that digital objects and natural objects enjoy the same reality status to the extent that they are sources of the very same affordances. Therefore, to the question of what is augmented in augmented reality, one would have to answer that it is neither perception as intended only on the user's side, nor reality in itself intended as what already and always exists, but rather reality at the psychological/ecological level of description, the environment intended as a set of affordances. Gibson himself, although in passing, noted that mankind has always intervened in the environment to modify its affordances with the only purpose of making "more available what benefits him and less pressing what injures him," adding that "this is not a *new* environment—an artificial environment distinct from the natural environment—but the same old environment modified by man" (Gibson 1979, 130; emphasis in original). Of course Gibson did not have in mind an environment modified by adding digital elements; nevertheless the analysis of the latter in terms of their affordances does not permit distinguishing them from any other human artifact. Thus, AR constitutes an implementation of the environment's affordances and, since ecological reality is a set of affordances, the digital objects of an AR environment are as real as the ordinary objects and, exactly like them, are perceived directly.

To complete this comparison is advisable to examine a subject which occupied Gibson throughout his career: the perception of pictures. Gibson's ideas about this topic were the main source of his criticisms regarding laboratory experimentation in optics and led him to the idea of the senses as perceptual systems. According to him,

> a picture is a surface so treated that a *delimited* optic array to a point of observation is made available that contains the *same* kind of information of an ordinary environment. . . . In short the optic array from a picture and the optic array from a world can provide the same information without providing the same stimulation. (Gibson 1982a, 277)

But it is not exactly the same information, in that "in any picture, chirographic or photographic, information to specify the surface on which it is made is unavoidable and can be noticed by an exploring observer" (Gibson 1982d, 291). Hence, the perception of pictures, by analogy with that of environmental perception, is based on the pickup of information from the picture optic array; the perception of a pictured object provides information about the pictured object. The only difference with respect to the direct perception of the object is that this perception is a mediated or indirect one, as it were, a second-hand perception where the information about the pictured object includes the information that it is pictured, i.e., not directly present. This awareness of both the pictured object and the surface on which it is pictured is unavoidable, therefore "the notion of an image that is literally and actually undistinguishable from the reality is a myth. . . . No matter how faithful, how lifelike, how realistic a picture becomes, it does not become the object pictured" (Gibson 1982a, 280). Now, what are the implications of this argument for AR? The answer is simple: Gibson's argument, ascribing the detection of the difference between the perception of pictures (which, incidentally, applies also to motion pictures) and that of the ambient to the explorative (read: locomotive) activity of the perceiver, does not applies to AR contents because the latter are interactive in real time and aligned in 3D with the objects of the environment.

To conclude this section, it seems to me to have shown that the ecological theory of perception, although being the most appropriate framework to study perception in AR environments, does not have the theoretical tools, at least in Gibson's formulation, to distinguish between computer-created objects and ordinary ones, thus forcing the ecological theorist to grant reality to AR contents.

4. Summary and Conclusions

This paper's main aim was to highlight the close relationship between AR and ecological psychology, having candidated the latter as the best and most natural framework for explaining the impact of this technology on humans.

In Section One AR was defined, following Azuma (1997), as a mixed reality system merging computer-created objects, interactive in real time and registered in 3D, with real world objects toward the aim of enhancing the user's perception, i.e., allowing him to better understand and interact with his own environment.

Section Two consisted in an exposition of the main features and concepts of Gibson's ecological theory of perception. Ecological perception is direct, it does not require any cognitive process, it does not hypothesize representations in the mind nor homunculi looking at them. The information for perception is found already structured in the ambient optic array, in the form of invariants of structure univocally specifying properties of the environment. The perceiver, being an essential component of its own environment, explores the latter carrying out his main activity: the pickup and selection of the information relevant for his own purposes. The notion of affordance encapsulates the idea that environment information is the object of perception to the extent that it consists in what the environment has to offer the animal in terms of possibility of action. Affordances are real insofar as they exist independently from the perceiving animal but, nevertheless, they have an essentially relational character in virtue of their practical orientation.

In Section Three, the emphasis placed on the practical character of AR systems permitted the beginning of a comparison between the digital contents of AR and the notion of affordance. It was pointed out that the definitions and aims of AR systems are consistent with the Gibsonian characterization of the affordances of the natural environment, thus leading to the main thesis of this paper: the interpretation of AR systems as implementation of the environment's affordances. Continuing the comparison, the question of the status of AR contents was raised. The discussion, undertaken within the framework of the ecological realism implied by the theory of affordances, led to some arguments in favor of the reality of AR contents. The main argument was that these are indistinguishable from real objects insofar as their affordances are exactly the same as those of real objects, i.e., they motivate the very same users' behaviors that would have been motivated by usual perceptions.

Of course, the question of realism posed by AR is not as straightforward as it may seem from the above discussion. On the one hand, I am convinced that the ecological approach and the notion of affordance is the best background for a study of perception in AR environments. On the other hand, the issue of the realism of AR is hard to address. In part this is due to the lack of a complete theorization of AR which, besides, is a consequence of its novelty. But it is also true that Gibsonian theory needs to be renewed, at least in what concerns the conception of learning and memory, the animal's side of the animal-environment pair, and its implications in terms of realism (for some suggestions see Ben-Zeev 1984; Michaels and Carello 1981; Norman 2002). I think that this issue is of major interest for both philosophy and technology-related disciplines in that, as we saw earlier, Augmented Reality brings us face-to-face with new problems which are rather different from the well-known ones posed by other media and technologies.

Appendix: A Note on Indirect Augmented Reality

During the review process preceding the publication of this paper, the reviewers brought to my attention that I did not mention the version of AR named Indirect Augmented Reality, developed at the *Nokia Research Center* by Azuma and others. This appendix aims to fill this gap and to explain the reasons why I left Indirect AR out of my picture of AR. Indirect AR has been characterized by its developers in a paper titled *Indirect Augmented Reality* (Wither, Tsai, and Azuma 2011). In this paper the authors introduce a variant of AR, designed mostly for portable devices, which does not suffer from the biggest problem affecting mobile AR devices, that is, the registration problem following from errors and/or lags in the tracking of the real world scenario. The idea is that of substituting the real world scenario, usually captured by a camera on traditional AR devices, with "pre-captured panoramas" so to generate a completely virtual scenario where no registration error is detectable by the user. As an example imagine your smartphone's screen displaying not what its camera "sees," but a previously captured image representing (almost) the same scenario in front of you filled with digital contents. This kind of AR in which registration errors are eliminated offers to the user a more comfortable experience (no effort needs to be made to pass over possible registration mismatches or lags) without compromising information. Now, can this be said to be a "variant" of AR? And, if so, to what extent? In my opinion here we are dealing with AR (although "indirectly") only in a very broad sense.

If we consider only what is displayed on the smartphone's screen it becomes evident that it is not AR, for it doesn't meet the requirement of interactivity in real time. Moreover, Indirect AR only applies to what I called "video devices" and not to "optical" and "spatial devices." If we broaden our point of view to consider how the smartphone's screen inserts into the whole real scenario in front of us, then we get genuine AR, but the registration problem arises again. Here the problem is the registration between what is displayed on the screen and what the user sees directly in front of him. At the very beginning of the paper, the authors recognize that Indirect AR is a shift, not a true solution, of the registration problem, but minimize it by saying that people are already used to confining part of their world within the boundaries of a screen and are not disturbed if the outside doesn't match exactly with what is displayed.

As an engineer, as an AR developer, and even as a man in the street I would feel forced to agree with this. It is better to "see" directly the boundaries of my window-on-the-world rather than seeing a mess when I look out the window. And it is also true that this doesn't affect the practical purposes pursued by the users (e.g., detecting a certain building among many others). But as a philosopher I can't help but notice that this is the very same problem rising again, that is, the link between augmentations and ordinary objects. Neglecting the real time interactivity requirement exposes Indirect AR to the argument used by Gibson for the perception of pictures (see Section Three above). The pre-captured images of the world, although being sources of reliable affordances, are far from having the same status of genuine augmentations; cutting a little world out of the whole world we live in implies an unavoidable loss of meaning thus broadening the gap between the (Indirect) AR experience and the ordinary one. Indirect AR maintains its usefulness and can be considered "better" than traditional AR only under certain respects—

i.e., as support for specific tasks—and under certain conditions—where "Good conditions in this case meant users stood at the same location the panorama were captured from, and the real world environment was predominantly static" (Wither, Tsai, and Azuma 2011, 816). But, as I pointed out earlier, these are not "normal" conditions because neither the world nor the user of a technological device is ever at rest, and common experience is experience of an ever-changing world. Azuma himself in his 1997 *Survey*, discussing some perceptual issues in AR concerning the connection between a user's performance and an allowable registration error, asked if the allowable registration error can be "larger while the user moves her head versus when she stands still" (Azuma 1997, 33). Here again, couldn't this be considered the technological counterpart of Gibson's insistence on the perceptual advantages deriving from having a mobile point of view? Finally, as the authors themselves claim in the introduction, a way for Indirect AR to become more powerful would be to increase the complexity of the virtual representation of the world and "the extreme conclusion of this increase in complexity might be the case of a real environment filled with a large array of video cameras and other sensors that would capture, *in real time*, the real environment and permit a perfect reconstruction of that environment, *in real time*, as seen from an arbitrary viewpoint" (Wither, Tsai, and Azuma 2011, 811; italics mine). Wouldn't this, though, amount to a return to the good old traditional AR?

For these reasons I consider Indirect AR as something different from genuine AR. At most it can be considered a sub-product of AR, useful for certain specific purposes, but always as a kind of fall-back solution, an expedient, a surrogate of AR, just as pictures can be considered as surrogates of reality.

References

Azuma, Ronald T. 1997. "A Survey of Augmented Reality." *Presence: Teleoperators and Virtual Environments* 6(4): 355–85. Available at http://www.cs.unc.edu/~azuma/ARpresence.pdf. Page citations throughout the article are to the PDF version.

Ben-Zeev, Aaron. 1984. "The Kantian Revolution in Perception." *Journal for the Theory of Social Behaviour* 14(1): 69–84. http://dx.doi.org/10.1111/j.1468-5914.1984.tb00488.x

Bergson, Henri. 1911. *Matière et Mémoire. Essai sur la relation du corps à l'esprit.* Paris: F. Alcan.

Caudell, Thomas P., and David W. Mizell. 1992. "Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes."

Proceedings of the 25th Hawaii International Conference on System Sciences, 659–69. http://dx.doi.org/10.1109/HICSS.1992.183317

Craig, Alan B. 2013. *Understanding Augmented Reality. Concepts and Applications*. Waltham, MA: Morgan Kaufmann-Elsevier.

Epstein, William, and John Park. 1964. "An Examination of Gibson's Psychophysical Hypothesis." *Psychological Bulletin* 62: 180–96. http://dx.doi.org/10.1037/h0044097

Gibson, James J. 1937. "Adaptation with Negative After-Effect." *Psychological Review* 44: 222–44. http://dx.doi.org/10.1037/h0061358

Gibson, James J. 1950. *The Perception of the Visual World*. Boston: Houghton Mifflin.

Gibson, James J. 1966. *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Gibson, James J. (1971) 1982a. "The Information Available in Pictures." In *Reasons for Realism: Selected Essays of James J. Gibson*, ed. Edward Reed and Rebecca Jones, 269–83. London: Lawrence Erlbaum Associates Publishers.

Gibson, James J. (1976) 1982b. "The Myth of Passive Perception: A Reply to Richards." In *Reasons for Realism: Selected Essays of James J. Gibson*, ed. Edward Reed and Rebecca Jones, 397–400. London: Lawrence Erlbaum Associates Publishers.

Gibson, James J. (1974) 1982c. "A Note on Current Theories of Perception." In *Reasons for Realism: Selected Essays of James J. Gibson*, ed. Edward Reed and Rebecca Jones, 370–73. London: Lawrence Erlbaum Associates Publishers.

Gibson, James J. (1974) 1982d. "Notes on Direct Perception and Indirect Apprehension." In *Reasons for Realism: Selected Essays of James J. Gibson*, ed. Edward Reed and Rebecca Jones, 289–92. London: Lawrence Erlbaum Associates Publishers.

Gregory, Richard 1972. "Seeing as Thinking." *Times Literary Supplement*, June 23, 707–08.

Hugues, Olivier, Philippe Fuchs, and Olivier Nannipieri. 2011. "New Augmented Reality Taxonomy: Technologies and Features of Augmented Environment." In *Handbook of Augmented Reality,* ed. Borko Furth, 47–63. Dordrecht: Springer Press. http://dx.doi.org/10.1007/978-1-4614-0064-6_2

James, William. 1907. *Pragmatism: A New Name for Some Old Ways of Thinking*. New York: Longman Green and Co. http://dx.doi.org/10.1037/10851-000

Kipper, Greg, and Joseph Rampolla. 2013. *Augmented Reality: An Emerging Technologies Guide to AR*. Waltham, MA: Syngress-Elsevier. http://dx.doi.org/10.1016/B978-1-59-749733-6.00001-2

Michaels, Claire F., and Claudia Carello. 1981. *Direct Perception*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Milgram, Paul, and Fumio Kishino. 1994. "A Taxonomy of Mixed Reality Visual Displays." *IEICE Transactions on Information Systems* E77-D (12): 1321–29.

Norman, Joel. 2002. "Two Visual Systems and Two Theories of Perception: An Attempt to Reconcile the Constructivist and Ecological Approaches." *Behavioural and Brain Sciences* 25: 73–144.

Richards, Robert J. 1976. "Gibson's Passive Theory of Perception: A Criticism of Müller's Specific Energies Hypothesis." *Philosophy and Phenomenological Research* 37: 221–34. http://dx.doi.org/10.2307/2107193

Shaw, Robert E., Endre E. Kadar, and Jeffrey Kinsella-Shaw. 1994. "Modeling Systems with Intentional Dynamics: A Lesson from Quantum Mechanics." In *Origins: Brain and Self-organization*, ed. Karl Pribram, 54–101. Hillsdale, NJ: Lawrence Erlbaum & Associates Inc.

Shaw, Robert E., Michael T. Turvey, and William M. Mace. 1982. "Ecological Psychology: The Consequence of a Commitment to Realism." In *Cognition and the Symbolic Processes*, vol. 2, ed. Walter Weimer and David Palermo, 159–226. Hillsdale, NJ: Lawrence Erlbaum & Associates Inc.

Simonetta, Gianluca. 2012. "Contenuti e Authoring". In *Communication Strategies Lab—University of Florence, Realtà Aumentate. Esperienze, strategie e contenuti per l'Augmented Reality*, ed. Gianluca Simonetta, 1–64. Milano: Apogeo.

Turvey, Michael T., and Robert E. Shaw. 1979. "The Primacy of Perceiving: An Ecological Reformulation of Perception for Understanding Memory." In *Perspective on Memory Research: Essay in Honor of Uppsala University's 500th Anniversary*, ed. Lars-Göran Nillsson, 167–222. Hillsdale, NJ: Lawrence Erlbaum & Associates Inc.

Wither, Jason, Yun-Ta Tsai, and Ronald Azuma. 2011. "Mobile Augmented Reality: Indirect Augmented Reality." *Computer & Graphics* 35: 810–22. http://dx.doi.org/10.1016/j.cag.2011.04.010