Journal of Human Cognition
2022, VOL. 6, NO. 1, 24-32
DOI: 10.47297/wspjhcWSP2515-469903.20220601

Is Chalmers' Virtual Reality "Mirror Argument" Sound?

Shaohua Xue

School of Humanities and Social Sciences, Beijing Institute of Technology, Beijing 100081, P.R.China

ABSTRACT
Extended reality devices provide users with unprecedented immersive and hybrid perceptual experiences, and users will act their bodies according to the information perceived. This shows that visual perception plays a crucial role in the formation and shaping of self-perception and spatial position. Users have a strong perceptual experience of their physical presence and self-perception in the real world as a result of their avatar perspective based on visual perception in a virtual hybrid environment, as is issued by Chalmers in his "Mirror Argument". However, users can still clearly differentiate their self and spatial experience between virtual and reality. To refute Chalmers' argument regarding perceptual experience in virtual reality, this article proposes a thought experiment called "Mary's Room in the Virtual Reality World" based on the experimental evidence of neuroscience. Finally, a possible future solution to this dilemma is presented.

KEYWORDS
Virtual reality; Affordance; Perceptual experience; Thought experiment; Mary's room

1 Extended reality and its perception problems

Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR), Diminished Reality (DR) and other technologies are collectively regarded as Extended Reality (XR) by academia and industry. XR technology attracts extensive attention from researchers, industrialists and consumers for it provides people with interactive perceptual experience of mixing virtual and reality. In virtual reality, people project their bodily sensations into a brand new virtual space to obtain perceptual experiences that never occur in reality and experience virtual worlds that never really exist. The application prospect of XR is very wide even though the technologies are under development, currently in video games, education and teaching, medical practice, combat training, art design and even adult entertainment.

Anyone who has experienced XR technology may feel strongly astonished by the indistinguishable virtual and real environments created by relevant devices. When a user is in an XR program like roller coaster or simulated flight, the sight of the virtual setting he/she experiences from first-person perspective causes obvious physical responses including standing unsteadily, heart beating faster, and shrinking head in fear, which are undoubtedly the “presence” experiences. However, the sense of presence is usually generated from perceptual experiences in the real world. After removing XR devices and set their vision back in the real world, the users can still differentiate their perceptual experience generated in the virtual and real settings. This phenomenon raises three questions to be discussed:

(1) Why do the users come to real perceptual experiences even though they know the environment and objects seen are virtual?

(2) Although the perceptual experience in (1) is real, people still sense the difference of perceptual experience in the real world. What are the causes for such a huge gap between the two kinds of perceptual experiences?

(3) Based on the question in (2), this paper proposes the thought experiment "Mary's room in the virtual reality world": if a person wears a virtual reality device ever since birth, can his perceptual experience in the real world match his self-perception and his perception-action coupling ability match the real world after taking the device off for the first time?

2 The eco-psychological explanation for extended reality

Extended reality provides users with immersive on-site visual perceptual experiences. Most of the popular virtual reality devices such as Valve Index, Oculus Quest 2, HTC Vive Pro and SONY PSVR are equipped with high-resolution screens, head motion tracking technology, whole body motion capture positioning system, and sound field headset specially designed for VR audio positioning to realize or enhance human-computer interaction behavior in the virtual environment. Virtual reality devices generate a first-person perspective in the virtual world based on the user's motions and behavior, and provide visual, auditory and other sensory simulations for users to create a strong sense of immersion and presence. The question is, how can the visual perception experience of these virtual settings (like riding a roller coaster and skiing downhill) feel so real to the users?

Ecological optics may offer an answer to this question. Gibson's eco-psychology discusses the close relevance between visual perception and action, and answers a core question: how do people make proper actions after being stimulated by visual materials, and is there some robust coupling relevance between the two factors?

Optic flow is one of the core concepts in ecological optics. It is used to describe the perceptual information an agent's visual perception system receives providing its action when it moves in an environment, such as air pressure, the density of optic flow, and the size and distance of the objects. The concept was initiated with Gibson's theory of ecological psychology, later often applied in extended reality technologies such as optical scene design, automatic driving program algorithm, scene recognition of natural road conditions and other relevant fields, showing robust theoretical adaptability. Gibson believes that perception is not a result of mechanisms such as representational symbols or mental processing, but an act of "picking up" information directly from the environment. Optic flow is the pattern of environmental movement received by an agent's visual perception during its motion, carrying various and vital motion information like speed and direction. The vision can better perceive the perception-action information carried in the optic flow when the agent is moving towards a specific target at a certain speed.

The massive sense of immersion and presence brought by virtual reality lies in the fact that it creates a virtual environment matching the biological acquisition of visual perception information, like roller coasters and tracks, ski slopes and snow-capped mountains, flying sky and fields. When a user picks up the perception-action information from the environment, the body will carry out responding actions matching visual perception.

Taking a user in a downhill-skiing programme seeing everything from a first-person perspective as an example, although he/she understands that all visual perception content is virtual and those visual optic flow motions do not occur in the real world, the body still automatically adjusts its postures and changes its centre of gravity according to the direction of the visual actions. According to the theory of ecological psychology, this is actually the user's adaptive perception-action based on the change in optic flow after picking up the affordance information provided by the virtual environment, which means the user is experiencing the affordance effect carried by the optic flow.

Affordance can be briefly comprehended as the chance or possibility of action provided to animals by a certain environment or object, such as a large flat rock providing “sitable” affordance and a strawberry providing “edible” affordance.

In the above skiing-downhill case, although the agent’s body remains stationary and only interacts with the virtual environment through visual perception, it still makes the agent feel it is in real motion. And this “virtual” real bodily sensation is caused by the change of optic flow provided by virtual reality equipment.

We can imagine the situation of driving a car on a highway. Seeing a car driving from the opposite side of you gives you the feeling that the closer the car is to you, the faster the speed is. This is because the optic flow in the opposite direction travels twice as fast as usual, and its information carrying high-speed affordance can attract individual’s attention more than other optic flow information on the side of the highway like highway guardrails and isolation belts on both sides. Since our visual perception system always pays attention to the high-speed targets around us, someone will directly obtain the meaning from the high-speed optic flow visual perception, that is, “may be dangerous”. In this way, your body will unconsciously produce evasive tilting movements even if you are driving.

Virtual reality devices can provide brand new visual perception experiences. The devices adjust and enhance the user's depth perception experience according to the density and texture of the optical flow. Some of the virtual objects are far and some are near, the users can see a large new space on a screen only a few centimetres from the eyes, which brings very realistic visual perception. The authenticity of this spatial experience in virtual reality can actually be illustrated by the "visual cliff" experiment (Gibson & Walker, 1960).

Eleanor J. Gibson and her colleagues used the concept of optic flow from ecological optics and the affordance theory to explain why infants are born with the ability to directly perceive deep space at the edge of a plane and “pick up” the information carried by the deep space perception to guide its follow-up actions under the condition of avoiding falling from a height. This experiment is to verify whether infants have deep space perception in an instant environment. The researchers asked the infants’ mothers to stand opposite the visual cliff device and call or guide their children to crawl over it. However, most infants refused calls from their mothers or researchers and crawled in the opposite direction away to the safety zone of the visual cliff device.

Based on the visual cliff experiment, Eleanor J. Gibson and her colleagues came to the conclusion that infants who have not been socialized can “pick up” the information of optic flow cues in a visual perception environment only through their visual perception, so as to directly obtain affordance about depth and space.

Infants do not have any perceptual experiences via social acquisition, but all small animals and infants in this experiment showed reactions relating to fear and caution when they are on the edge of the cliff. This proves that both infants and small animals can directly obtain information from the environmental surface and structural texture according to the instant surrounding light and optic flow, that there is a large height difference at the cliff providing a “fall-able” affordance while the other side of the cliff is safe providing “support-able” plane.

We can see that this environmental optic flow information directly obtained by visual perception made infants have an affordance effect on environmental features and take adaptive actions. Therefore, if someone puts on virtual reality glasses to experience a virtual environment for the first time, he can take adaptive actions like infants based on the perceptual information about depth and space formed by optic flow information in this virtual environment even if he has never used the device.

However, if the designer of the virtual scene environment does not consider optic flow and surrounding light information when designing, then users will not experience the sense of immersion even if the resolution of the virtual devices is higher, the screen scene refresh delay rate is lower, and the motion capture sensor is more sensitive. Therefore, getting users to obtain affordance information in the environment by simulating optic flow and surrounding light information is the key to obvious immersive perceptual experiences.

While the sense of immersion and realism do exist, we can still distinguish between virtual reality and real environments. In fact, besides the limitations of screen display, motion capture sensing and technical computing power, another important reason is that the changes in optic flow prompt changes in the user's spatial position while the body remains still. This means that visually you are starting to move forward while your body remains stationary. In this way, the perceptual information of the vestibular system to maintain the balance of the body cannot completely match with the visual information of the retina, and the user will inevitably feel dizzy after using the virtual reality devices for a period of time, which is called "motion sickness".

3 Extended reality and grid cells

According to ecological psychology, users can obtain deep space perception in virtual environments similar to that in real environments based on the environmental information carried by the optic flow, therefore producing adaptive behaviours in virtual environments. This deep space perception comes from the user's ability to instantly judge the relevance between space and distance. In order to achieve this ability, the relevant brain areas of the user's cerebral cortex play a very important role.

It can be seen that the affordance effect of deep space perception produced by the user's visual perception system is closely related to the spatial positioning information received by its visual perception. Many researchers have worked on this, the most famous of which is the work of John O O'Keeffe. For example, O'Keefe and Dostrovsky (1971) studied cells in the hippocampus of the brain that are closely related to the function of processing spatial locations, which were later called place cells by O'Keefe (1976). They can send out the positional information of where the observer is, and make the brain have spatial memory. The hippocampus, where these special neurons are located, was later regarded as the main place for the formation of cognitive maps, and a series of scholars carried out key research.

Hafting and Moser et al. (2005) went further in O'Keefe's research and discovered grid cells, which can provide coordinates to the observer like a navigation system. The functions of these two types of cells can cooperate with each other to jointly perform spatial cognition and location navigation for the cognitive subject.

So, if the user's visual perception is in a virtual environment, what will happen to the neural activity of the user's brain? Is the perceptual experience obtained from the virtual environment consistent with the neural activity of the brain obtained from the perceptual experience in the real environment? Perhaps on the basis of answering these questions, we will find biological evidence and find the key to why users can tell the difference between the virtual environment and the real world.

It should be noted that Ravasssard et al. (2013) experimented and monitored the visual perception of mouse, and focused on the hippocampus of the mouse brain. There are place cells and grid cells in this brain area, and their function is to establish and control the observer's cognitive map and position navigation system. When the mouse enters a new physical environment, these localization cells activate and begin to build a cognitive map inside the mouse's brain. When this cognitive map is constructed, the mouse can judge its spatial position and direction area according to the cognitive map in its brain, which can guide its subsequent adaptive behaviors in the environment. Next, the researchers gave two different environments (that is, virtual reality and real world), and then monitored the neural activity of place cells and grid cells in the hippocampus of mouse in these two different environments respectively. There are two possible outcomes in this experiment:

(1) The activities of related cells in the hippocampus in these two environments were the same, with no significant difference;

(2) There were significant differences in the activity of related cells in the hippocampus under these two environments.

First, to test whether creating a spatial cognitive map of the hippocampus requires only the perceptual ability to "see", the researchers set up a virtual environment to observe the mouse's actions in the device, and simultaneously monitored the neuronal activities in the hippocampus of the mouse. In this virtual space setup, mouse were placed on a spherical device whose field of view was covered by a video screen, thus forming a virtual environment of full-view immersion for the tested mouse. The purpose of this is to put the mouse in a virtual reality environment, which can create an immersive sense of presence. Then, the tested mouse performed activities in this virtual environment device, and the researchers began to record the immediate neural electrical signals of the relevant cells in the cerebral cortex and the hippocampus of the mouse.

Then, the researchers placed the tested mouse in a physical real space equivalent to the virtual environment and continued to record the relevant nerve cell activity data of the mouse. After comparing the available data, the researchers found that the activity of place cells was significantly different. In the real environment, 45% of the place cells were activated and working, while only 22% of the place cells were activated under the virtual reality conditions, which is less than half of the real environment showing an obvious comparison. This shows that the activity of the hippocampus of the mouse brain in the virtual environment is completely different from that in the real environment, and there are different levels of activity. In the virtual environment, the activation of neurons in the hippocampus was random, just like the neurons could not locate the location of the tested mouse. Although the mouse had normal appropriate behaviour in the virtual environment, it appeared that the mouse knew exactly where it was in the virtual space. In this way, the cognitive map seems to disappear in the virtual environment. The neurons of most of the mouse tested that were activated in the real environment also stopped processing neurotransmitter information at that time. In an immersive extended reality experience, "place" cells are not as active as they would be in a real environment. This undoubtedly shows that the cognitive map and spatial location navigation system in the brains of the tested mouse cannot work normally in the virtual environment.

In fact, it is not a surprising conclusion for us. Because previous researchers such as P. Martin (2001) have shown that the activation of place cells requires the synthesis and processing of various effective information, such as visual information, self-displacement information and proximal stimuli to smell, hearing, touch and other perceptual information. Therefore in order to form spatial perception, people must rely on information from multiple sources and comprehensively process perceptual information, and the hippocampus is the brain area that forms a complete and clear spatial representation. When the tested mouse made any exploratory spatial perception in the real environment, the relevant nerve cells in the hippocampus immediately produced a cognitive map. However, in the virtual environment, the tested mouse only produced visual information. There was no self-displacement information, let alone some proximal olfactory or somatosensory perception information, which were very important parts of proximal stimulation, so the brain could not generate an effective cognitive map.

In summary, multiple sources of information are required to form a normal spatial experience within the brain. Although extended reality devices create realistic effects visually, they cannot bring us the normal perception like the real environment. Therefore to create real immersive experiences, extended reality devices have to include some near-end perceptual experiences besides visual information, otherwise, the formation of the brain's cognitive map will be disrupted.

Vision alone is not enough for the generation of real environment perception, perceptual information from multiple channels needs to get involved. So how does this conclusion from neuroscientific data affect the philosophical considerations of extended reality?

4 Mary’s room in the virtual reality world

Extended reality raises many philosophical questions that have never been discussed. In the paradigm of philosophy of mind, D. Chalmers (2015) discussed virtual reality together with issues like visual illusion and sensory quality. There are two of his main concerns:

(1) Is the perceptual experience in virtual reality a kind of illusion, and is the spatial experience in virtual reality real?

(2) Is the perceptual experience obtained in the real world still valid in virtual reality? Conversely, can the perceptual experience obtained in the virtual environment be effectively applied in the real world?

For the first question, Chalmers believes that there is no hallucination phenomenon in the experience of virtual reality, all experiences are real and they do exist.

First of all, in his opinion, the perceptual experience in virtual reality is completely different from the hallucination experience in the traditional sense. The so-called hallucination is actually a special visual perception experience, that is, when the observer perceives a specific scene or object, these scenes or objects will make the observer feel that they exist in a certain way or appearance. However, those scenes or things actually exist in another way or form, and the observers will obviously misjudge the relationship between the size, shading, colour, length, shape, and distance of the observed object. In psychological or philosophical experiments, more than hundreds of illusion patterns have been found in the phenomenon of hallucinations.

For example, the Müller-Lyer illusion can be explained like this: it refers to the fact that the lines between the two symbols "<>" and "><" are exactly the same length, expressed as: "<——>" and ">——<". When we observe these two figures, we will feel that the line in the symbol ">——<" is longer than "<——>". However, the visual information received by the eyes is actually significantly inconsistent with the results measured by physical instruments, and it can be seen that there is a cognitive deviation between the observer's experience and the way things themselves exist.

However, virtual reality is a visual environment simulated by a computer, and this concept is distinct from the concept of hallucination. Unlike this article, which argues for the authenticity of virtual reality experiences based on Gibson's eco-optics and functional visibility effects, Chalmers offers another interesting line of argument. While both arguments are valid, Chalmers proposed the "Mirror Argument". In his view, the relationship between virtual reality and hallucinations can be analogised to the mirror. So, is what we see in the mirror an illusion to us?

Chalmers distinguished between two situations. In one case, mirror experiences can be explained as hallucinations under certain conditions. For example, when someone does not know the mirror is in front of him, and he sees and thinks that there are some equidistant objects on the opposite side with exactly the same shape, arrangement and size which do not actually exist, then this mirror experience is an illusion.

In the other case, however, mirror experiences are not hallucinations at all. For example, when you know that there is a mirror in front of you, and you see the scene in the mirror when you look in the mirror, the mirror experience in this case is not an illusion but a real visual perception experience. The experience of virtual reality is actually almost identical to our mirror experience. Once we know the specific environment we are in, then we can be sure that our visual perception experience is real, no matter if we are in virtual reality or looking into the mirror.

Chalmers uses the example of a rearview mirror to further explain the reality of this mirror experience. When we are driving a car, the rearview mirror installed in front of the windshield of the car is a must-have device, which allows us to see the road and the vehicle behind. So, if someone is driving and sees a car with double flashing lights in the rearview mirror in front of the car's windshield, will he think the car is in front of his car? Hallucinationists would say that this is of course an illusion because we judge that the car is behind us, but we see that car in front of our vision. Here the problem arises that the visual information does not correspond to reality, hence it is a hallucination. However, Chalmers refuted that we use mirrors a lot and know the existence and function of the rearview mirror, so when we obtain the visual information that the rearview mirror provides us when it performs its function, we generate the correct judgment that the car is indeed behind us. This is because we know that the purpose of the rearview mirror is to see the road behind us. Under the condition of virtual reality equipment, this argument still holds. In VR, we perceive virtual things in a virtual space, and virtual things are also real things, although they are data structures simulated by computing. Furthermore, Chalmers proposed a hypothesis trying to solve the aforementioned problem (2), that is, the generalizability of perceptual experience in two environments. He divided virtual reality experiences into two categories:

The first is a brief virtual reality experience. For example, if a person usually lives in a real environment, once he briefly experiences a virtual reality device for dozens of minutes, can his perceptual experience in real life guide his perception in the virtual reality environment? The second is a long-term virtual reality experience. For example, if someone wears a virtual reality device for a long time, will he confuse the spatial experience of the virtual reality with the spatial experience of the real environment?

The first condition is very common, and most people have this experience, where their perceptual experience can switch between virtual reality and real environments without confusion. This situation is very similar to the example of the rearview mirror: you know the function of the device, and then you get the perceptual experience that the device has on you. The spatial experiences in both virtual reality and real environments are real. Thus, there is no hallucinatory confusion in the spatial experience between the two environments.

Regarding the second case, Chalmers believes that spatial experience in virtual reality also produces real mental representations. If someone perceives virtual things in virtual reality for a long time, he may adapt to the virtual reality environment and take virtual reality as an independent thing and space. The longer he uses the virtual device, the less likely he can tell the difference between the virtual and real.

However, Chalmers's hypothetical argument is not convincing enough. We can assume an example that is more extreme than the second case. If Chalmers's conclusion does not hold in this extreme case, then in the second case Chalmers's conclusion is less likely to hold.

This example is somewhat similar to Jackson's (1970) thought experiment “Mary's room argument”: Mary grows up in a black and white room with all kinds of scientific knowledge, and she reads all kinds of knowledge about physics and the world. She knows that there are many colours in the world, but there are only two colours in her living environment: black and white. One day she walks out of the room and sees all kinds of real colours. So, will she gain knowledge or experience that she doesn’t have before because of this new experience?

This extreme case in virtual reality might be called "Mary’s room in the virtual reality world". If someone wears a virtual reality device ever since birth and he doesn't know that his visual perception is virtual reality, then all the visual perception he has is the information of the virtual world scene simulated by the computers, and all the perceptual information and knowledge representation about the real world that he sees come from virtual reality. If one day he takes off the virtual reality device and sees the real world, can his previous visual perception experience and knowledge guide his perceptual actions in the real environment? How is the perceptual experience in virtual reality different from the perceptual experience in the real environment?

According to the third part of this article, we know that the main difference in spatial experience between virtual reality and the real environment is: according to the experimental data of neuroscience, the shaping of the spatial cognitive map of the real environment requires information obtained from multiple sensing channels to be comprehensively organized and processed so that a complete cognitive map can be formed within the cognitive system of the brain.

In this way, if Mary in virtual reality only receives one-way stimulation of visual perception information, then her cochlear vestibular system will never match visual perception, and the resulting vertigo will always company all her spatial experiences. Meanwhile, the activity of place cells and nearby grid cells in the hippocampus will remain at a low level of activity, so the cognitive map inside the brain will not really be formed.

When she takes off the virtual reality device for the first time and uses her eyes to perceive the real world, her cochlear vestibular system and visual perception will truly match for the first time, and the vertigo she was accustomed to before will also disappear. This is when the cognitive map of the brain is then truly formed. On the basis of these obvious physiological changes, her spatial experience will inevitably change significantly, so she can definitely feel the huge difference between the spatial experience of virtual reality and the real environment.

5 Conclusion

It should be seen that Chalmers and Gibson have strongly argued from different perspectives on whether the spatial experience of virtual reality is real, that even though virtual reality is a kind of visual perception information simulated by computers, its spatial experience is undoubtedly real. Regrettably, although such visual perception experiences are real, virtual reality is limited by the current technology. Even with the gradual rise of wearable devices (such as backpack-type portable computing units, wearable virtual reality gloves, etc.), these technologies cannot effectively provide multi-channel perceptual information, and there is still a considerable experiential gap between virtual reality and real environment spatial experience that needs to be bridged.

If the matching between the cochlear vestibular system and the visual perception information can be solved at the same time, then the virtual reality users can be provided with the same spatial experience as the current real environment. Moreover, in the extreme case "Mary′s room argument in the virtual reality world", Mary can still feel the huge difference between the spatial experience of the virtual reality and the real environment, then Chalmers′ conclusion in the second case—observers wearing virtual reality devices for extended periods of time may not be able to tell the difference between virtual reality and the real environment—would be self-defeating.

If Chalmers' assertions and ideals are realized in the foreseeable future, then virtual reality devices should provide as many channels and contents of multi-channel perception information as possible. For example, the virtual reality device "Virtuix Omni" that can realize location tracking provides a solution for the observer to locomotion or run in the virtual space. This indicates that the future virtual reality devices may be able to meet Chalmers′ theoretical assumptions, which is very worth anticipating.

References

[1] Gibson E J, Walk R D. The" visual cliff"[M]. WH Freeman Company, 1960.

[2] Walk R D, Gibson E J. A comparative and analytical study of visual depth perception[J]. Psychological Monographs: General and Applied, 1961, 75(15): 1.

[3] O'Keefe J, Dostrovsky J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat[J]. Brain research, 1971, 34(1), 171-175.

[4] O'Keefe J. Place units in the hippocampus of the freely moving rat[J]. Experimental neurology, 1976, 51(1): 78-109.

[5] Hafting T, Fyhn M, Molden S, et al. Microstructure of a spatial map in the entorhinal cortex[J]. Nature, 2005, 436 (7052): 801-806.

[6] Ravassard P, Kees A, Willers B, et al. Multisensory control of hippocampal spatiotemporal selectivity[J]. Science, 2013, 340(6138): 1342-1346.

[7] Martin P D. Locomotion towards a goal alters the synchronous firing of neurons recorded simultaneously in the subiculum and nucleus accumbens of rats[J]. Behavioural brain research, 2001, 124(1): 19-28.

[8] Chalmers D J. The virtual and the real[J]. Disputatio, 2017, 9(46): 309-352.

[9] Jackson F. Epiphenomenal qualia[J]. The Philosophical Quarterly, 1982, 32(127): 127-136.

About the Author

Shaohua Xue, Ph.D, associate professor at institute of science technology and society, school of humanities and social sciences, beijing institute of technology, research directions: cognitive science, philosophy of mind.