Perception in VR 3D vs VR 360 video : How a key cognitive process in learning operates in virtual environments

Anne-Dominique Salamin
Christophe Hadorn
e-learning Center HES-SO Cyberlearn
University of Applied Sciences Western Switzerland (HES-SO)
Switzerland
adominique.salamin@hes-so.ch
christophe.hadorn@hevs.ch

Abstract : This paper reports on the findings of a study about perception in virtual contexts, conducted with 111 persons from 4 countries. Perception constitutes one of the key cognitive learning process. In this study, we aimed at assessing whether items displayed in a VR360 video context were more rapidly perceived compared to a VR 3D environment. Results showed that for the considered sample, items displayed were more rapidly perceived in a VR 3D environment than a VR 360 video environment. Women were slower than men, but with a shorter perception gap for the VR 360 video than the VR 3D. The elder the people, the less performing the perception becomes, both in 360 and in VR 3D, but more efficient in VR 3D than in 360. These results will help design visual environments adapted to virtualized experiential learning simulations.

Introduction

21,000 students (27% of the 8 Swiss UAS students) enroll every year in the different curricula proposed by the HES-SO. This 20-year-old university offers students strong references to the real professional world, either by linking the teaching laboratories with real experiments or by developing projects with professionals in action and has become a great source of skills, ideas, innovation, creativity and knowledge. With its six faculties (Design and Fine Arts, Business, Management and Services, Engineering and Architecture, Music and Performing Arts, Health and Social Work), HES-SO plays a preeminent role in the seven cantons of Western Switzerland where its faculties stand.

The HES-SO has created the e-learning Center Cyberlearn, in 2004. The Center is in charge of developing and conducting research in blended learning, comprising the pedagogical use and implementation of new innovative and disruptive technologies.

By 2018, virtual reality (VR) has become an innovative technology which will soon influence the methods for knowledge acquisition. Using immersive helmets, VR has come of age as far as hardware is concerned. Its relatively low cost enables considering using it in the context of education. The VR environment provides an extremely interesting learning frame indeed. It enables real, immersive and interactive experiments. Ranging from medecine to aeronautics, many fields rely on this technology to develop competencies, with low human, hardware and safety costs. The industrial sector increasingly uses this type of technology to train its employees (Durand, 2017). The army, the gaming business or marketing expand widely in this area.

Until now, such programs have seldom been applied to universities, and rarely developed for abstract learning fields (maths, physics, economics, etc.) or for more applied fields such as communication (language, welcoming etc)

Developing VR learning formats to acquire academic subjects seems relevant since this provides a means for giving abstract learning a touch of reality, it facilitates immersion in a context close to reality. Nevertheless, in spite of its relatively low cost, the actual number of students and learners potentially concerned hinders a stronger committment to this technology, for the primary, secondary or tertiary learning levels. Moreover, a number of issues remain questionable. Can experimenting in a virtual situation be comparable to a real experience ? Is virtual immersion as significant as real immersion ? How does virtual scenography impact on the actual sensation of reality ?

Before developing learning experiments adapted to the acquisition of academic knowledge, we wish to establish whether the use of VR can positively influence the learning process. As a means of measurement we identified five key cognitive processes involved in this acquisition : perception, attention, memorization, motivation and transfer. (Collective, 2014 ; Roulin, 2006). We assess the proceedings of these processes in virtual reality by developing measurable learning experiments.

This paper presents the results of a study conducted by the Center on item *perception* in a VR 3D environment compared to a VR 360° video. This study aims at determining which environment provides the best learning aid.

Experiential learning

As the main feature of VR experimentation consists in the ability to be immersed and to interact in a simulated world, we looked at the existing experiential learning contributions in the field of education. A number of writers (Freinet, 1964 ; Montessori, 1912 ; Dewey, 1963) claim that experiential learning offers a learning method particularly pertinent for developing competencies thanks to real life and rational experiments. In such a model, learning consists in three steps : integration into a meaningful context, concrete experiment (*hands-on*), then decentralization and analysis of the performed experiment. (Bruner, 1991 ; Kolb, 1984). This theory pertains to the contructivist theories by Piaget, demonstrating that the learner learns by performing activities as an actor of his own learning. (Piaget, 1936).

Research in experiential learning shows **that a more significant experience and closest to reality for the learner, increases the effectiveness of the learning process.** (Kohonen)

VR 3D and VR 360 video environments

Virtual reality provides a « computerized model, based on digital images in VR 3D, of a fictive or real environment, where a user can move and interact using a specific helmet. » (Glossary, 2018)

Compared with a 2D simulation on a computer or mobile tablet, VR provides an immersive experience. The subject is actually « inside » the scene and can move around by interacting with the items available in the scene. The brain is « fooled » and the subjects take in the fictive universe in which they are immersed, physically avoiding a virtual obstacle, leaning down to collect an item, part of this virtual world.

VR uses two types of visual displays *: three dimensional computer generated images* (VR 3D) and the VR *360 video scenes*. The VR 3D is « drawn » and can, depending on the means invested, reach an extremely realistic display, as it is the case for video games. In such a frame, the human brain « understands » that it is not the actual reality, but still avoids the virtual obstacles.

The virtual scenography VR 360 video, on the other hand, lacks three dimensional display, but provides a 360 sideways movement. Movement in depth is not available (moving forwards). Several interactions are available (staring at a given point, voice interaction, etc.), but the items displayed cannot be grabbed. However, the scene, the items, such as people, are consistent with reality.

Figure 1 : Comparison between the VR 360 video experience and the VR 3D universes

The Study

In this experiment on perception, we developed a recognition experiment with items placed on a desk. We predicated that the scenes filmed in VR 360 video and proposed with a VR helmet would be perceived more efficiently (rapidity) than those in VR 3D because of : 1) the realistic display (the desk is « real »), 2) the familiarity (the items placed on the desk look alike items found on any desk), and 3) the sensations (the book is open, the mouse ready to function etc.).

In order to perform this experiment, we built two scenes representing a desk in the *first-person*[1]. Each scene comprised 20 everyday items (computer, cup, photo, sheet of paper, figurine, mouse, etc.). This experiment was proposed to 111 subjects from 4 different countries, filed into five different age groups. In the end results, we left out 10 persons from this sample, as some personal information was missing (age/gender/country/position).

	20-30 years	**%**	**31-45 years**	**%**	**46-55+ years**	**%**
Women	30	38%	6	55%	5	45%
Men	49	62%	5	45%	6	55%
	79	100%	11	100%	11	100%
Total	**% 20-30 y**		**% 31-45 y**		**% 45-55+ y**	
103	77%		11%		11%	

Table 1 : Breakdown gender/age of participants

Protocol for data collection

The experiment was proposed to second year Bachelor students of the HES-SO, as well as people participating in various Swiss and international conferences in which we also took part. The experiment procedure was explained to participants before it started. If needed, they were given a short explanation on how to use the joystick for interaction. Seated in an armchair, the subjects were wearing a VR helmet (HTC Vive), a voice indicating the name of the items to be found, at which they would aim with the joystick. The system recorded the time span needed by the subject to recognize an item, as well as the total time spent by the subject to complete the experiment.

The conditions in which the study was conducted were identical for both environments (VR 3D and VR 360 video). A number of factors may have influenced the experiment, namely : the subjects' visual quality (wearing optical glasses was allowed), how familiar they would be with video games (use of the joystick), and the perception threshold of the items, which varies according to individuals (Fraisse, 1980). These authors have demonstrated that the perception threshold (the shortest time to identify an item) decreases with age, that geometrical shapes are perceived more rapidly (10ms) than stylized drawings (21ms) and words (25ms), and that shapes seen are a « virtual » generated image produced by the brain.

Results

The results invalidate the starting hypothesis. All data taken into consideration, by age group, gender, country, position), indicates that the perception is faster in VR 3D than VR 360 video. On average, items are recognized in 00 :01 :07 sec. in the VR 3D environment, against 00 :01 :20 sec in the VR 360 video environment (+ 13 sec). The result collected per item in VR 3D shows that the longest item to be recognized is the sheet of paper (10 sec.), while the fastest identified item is the keyboard (<1 sec). For the 360 video, the most rapidly perceived items are the keyboard and the screen (<1 sec) and the longest is the briefcase. The probability of finding the item on the desk, as well as the semantic proximity (the keyboard for instance) may explain the results obtained by the keyboard for both environments.

Rough Data **Average for**	**VR 360** **Perception Time**	**VR 3D** **Perception Time**	**VR 3D**
Paper	00:00:04	00:00:10	+6
Figurine	00:00:10	00:00:05	-5
Book	00:00:03	00:00:03	0
Plant	00:00:02	00:00:01	-1
Briefcase	00:00:12	00:00:06	-6

---

[1] refers to a graphical perspective rendered from the viewpoint of the user's character

Rough Data	**VR 360**	**VR 3D**	
Car	00:00:08	00:00:02	-6
Candle	00:00:05	00:00:03	-2
Headphones	00:00:03	00:00:01	-2
Picture	00:00:02	00:00:05	+3
Screen	00:00:00	00:00:04	+4
USB key	00:00:05	00:00:05	0
Pencil	00:00:01	00:00:02	+1
Mug	00:00:02	00:00:02	0
Apple	00:00:01	00:00:01	0
Coin	00:00:02	00:00:02	0
Keyboard	00:00:00	00:00:00	0
Computer Tower	00:00:03	00:00:01	-2
Computer Mouse	00:00:01	00:00:02	+1
Coffee	00:00:03	00:00:04	+1
Smartphone	00:00:02	00:00:01	-1
Total	00:01:20	00:01:07	-0:14

Table 2 : Results in all

Detailed results

- *By age group*

The results by age group give some indication on the relevance of providing either of the two environments for basic trainning (20-30 year old) or for continuous education (>31 year old). Without surprise, the younger ones idendify the items more rapidly, regardless of the environment. The slowest group in recognizing the items is the 46-55+ group. The lesser improvement rate for both universes is found among the 31-45 (-16 sec), while the 46-55+ group (-117 sec) improves its performance in both environments, but nevertheless remaining the slowest group of all.

Age groups (years old)	**participants**	**VR 360**	**VR 3D**	**difference**
20-30	78	00:01:13	00:00:48	-25 sec
31-45	13	00:01:43	00:01:27	-16 sec
46-55+	11	00:03:31	00:01 :34	-117 sec

Table 3 : Results grouped by age

- *By gender*

Results by gender show that women are slightly slower than men in identifying items displayed in either environment. The gap « won » when switching from VR 360 video to VR 3D is less significant (-19 sec) for women than for men (-21 sec). The difference between women and men is weaker for video 360 (-8 sec), and higher for VR 3D (- 10 sec).

Gender	**Number**	**VR 360**	**VR 3D**	**Difference**
Men	58	00:01:15	00:00:54	-21 sec
Women	44	00:01:23	00:01:04	-19 sec
Différence M/W		-8 sec	-10 sec	2 sec

Table 4 : Results grouped by gender

- *By position*

The results prove coherent when comparing administrative staff with teaching staff. Age is higher in the first two categories compared to students, so the perception time measured is slower in the first two categories. It can be argued that the students group is trained to follow instructions, which could influence the quickness of the perception task. The category « administrative staff » gains 29 sec switching from one system to the other, while the category « teaching staff » gains 27 sec and the students gain 11 sec. In the VR 360 video environment, students are 51 sec quicker than the « administrative staff », while in the VR 3D they are 33 sec faster. When comparing the students with the « teaching staff », the time gap amounts to 70 sec for VR 360 video, and only 54 sec in VR 3D.

So students are the fastest in VR 3D, faster than the administrative staff and the teaching staff, on average (28 sec against 11 sec), as well as in the position category.
This result is promising because VR training may be aimed firstly at students. Administrative and teaching staff, who may also get training in such an environment, is less efficient concerning perception.

Position category	**VR 360**	**VR 3D**	**Difference**
Administrative staff	00:01:52	00:01:23	-29 sec
Teaching/research staff	00:02:11	00:01:44	-27 sec
Students	00:01:01	00:00:50	-11 sec
Students vs AS	-51 sec	-33 sec	-18 sec
Students vs TRS	-70 sec	-54 sec	-16 sec

Table 5 : Results grouped by position category

Findings

The obtained results demonstrate that perception is more efficient in a VR 3D environment. All participants, regardless of age, position and gender perceive items faster in this environment. Several reasons can be put forward to explain this convergence of the results.

- *Quality of the 360 scene display*
  The VR 360 video scene, captured by video, produces a mediocre rendering. Moreover, inside shooting induces a loss of quality compared to outdoor shooting. Nevertheless, the items were easily identified. Perception time of the items varies considerably : the first item (paper) is identified in 00 : 00 :04 sec, while the second item (figurine) is identified in 00 : 00 : 10 sec. The last item (smartphone) is identified in 00 :00 :02 sec. This indicates that the items were recognized with no interference in the identification process due to possible visual or cognitive fatigue, as the last item obtains a better perception score than the first one.
  As for the VR 3D scene, the first item obtained 00 :00 :10, while the second scored 00 :00 :05, and the last 00 :00 :01.
- *Fuzziness and flickering*
  The VR 360 video images are slightly fuzzier and more flickering than the VR 3D images, which can hinder perception. However, the brevity of the experiment (less than 3 minutes) make this feature rather irrelevant.
- *Geometrical simplicity*
  Perceiving shapes is a paradox, as the brain cannot apprehend images as a whole. The shapes of images are combined by the brain from visual information captured by each visual stare. The brain creates in its memory, a representational prototype of a shape or an item by combining each pattern of shape by category (Bonnet, 1989). Biederman demonstrated that « item recognition is based on the perception of basic geometrical elements from which an item can be build » (Biederman, 1987). Such elements are called « geons » (contraction of geometrical ion), whose perception enables association between the item and its prototype as build by the brain. The image is identified by the brain by associating it with fundamental geometrical shapes. Less than 40 basic shapes (round, square, rectangle, etc.) enable perception and identification of any image, regardless of lighting direction and surface characteristics. Faces and words, however, obey other decoding rules.
  These theories indicate that the simpler, the more symetrical and the more regular the shape is, the easier the

reconstruction process by the brain becomes. This could explain the better performance of the test in VR 3D. The drawn scene is simple, the items more schematic than those displayed in the VR 360 video.

The VR 3D images generated for this study are close to reality, except for the computer mouse. This image is hardly identifiable, as it consists of a large dark grey rectangle, and may represent any other item. In spite of being placed very close to the keyboard, thus providing a good perceptive clue, the computer mouse in VR 3D was only identified after 00 :00 :02 sec, while it was identified after 00 :00 :01 in VR 360 video. It can be deduced that the simplicity of an item in VR 3D must go hand in hand with a similitude with the real item, in order to guarantee a short perception time.

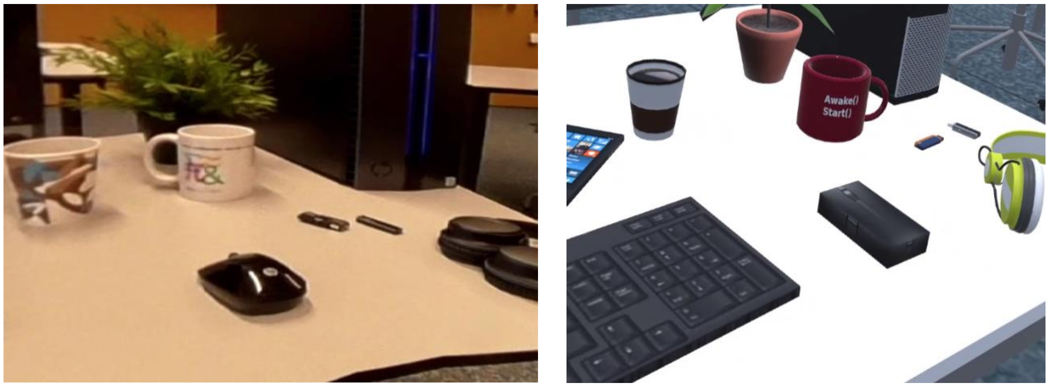

Figure 1 : Comparison of a computer mouse in VR 360 video and VR 3D

Conclusions

The study conducted on a random sample of 111 persons lead to the conclusion that the cognitive process of perception was more efficiently called for, during immersion in a VR 3D scene, rather than a scene based on VR 360 video.

This result can probably be explained in that the brain constructs the perceived images, by working on basic geometrical shapes (geons). The simpler the shapes, the less complex is the reconstruction model and thus the shortest the perception time.

The studies we are planning to conduct next, will aim at determining if rote memorization is improved in a VR 3D or VR 360 video environment. We have already demonstrated (Salamin, 2017) that rote memorization was less efficient via a smartwatch rather than on paper, but that the items memorized with the smartwatch lasted longer in memory, three months after the end of the experiment. Thus, a second important cognitive process can be analysed to provide relevant indicators for the development of simulations in VR 3D in VR.

In the field of education, professors and staff responsible for developing VR 3D simulations are sometimes reluctant to fully engage, because the level of graphic images of the scenes will never reach those developed by the gaming business. Therefore, VR 3D item databases lack VR 3D artefacts available for educational purposes.

Our study on perception shows that this must not be so. Simple and relevant scene displays suffice as the pedagogical quality relies less on full-fledged graphical displays, but rather on the embedding in a context which makes sense for the learner. The theories of experiential learning identify four key steps at work in this type of learning process : Active experimentation, concrete experimentation, reflexive observation and abstract conceptualization (Kolb, 1984). Anchorage in learning occurs during the last two steps, independently from the graphics and quality of VR 3D scenes. Therefore, focus must be made on context relevancy, on creating activities suitable to this context and on developing reflexive observation activities, for instance metacognitive ones. (Fogarty, 1994).

Learning in the 21st century aims at training learners for critical thinking and decision making, which are key elements of experiental learning. Simulations in virtual reality VR 3D, thus, provide an experimental and learning frame useful for developing and anchoring these new competencies.

References

Biederman, I. (1987). *Recognition-by-components : A theory of human image understanding*. Psychological Review

94 (2), 115–47

Bonnet, C. (1989). *La perception visuelle des formes*. Traité de psychologie cognitive, vol.1, Perception, langage. Paris. Dunod. p.17

Bruner J. (1991). *The narrative construction of reality*. Critical Enquiry 18(1), 1–21. Retrieved from : http://www.univie.ac.at/constructivism/archive//4039

Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Collective (2014). Introduction to Concepts and Issues. Estes.W.K. Psychology Press, vol. 1. Psychologie et pédagogie, Gonthiers Denoël. coll. Médiations, Paris.

Dewey, J. (1963). *Experience and education*. Colliers Books, NY

Dokic. J. (2004). *Qu'est-ce que la perception ?* Vrin, coll. Chemins philosophiques.

Durand, C. (2017). *Comment la RV/RA vont-elles influencer le monde des affaires en 2017*. Retrieved from : https://skilled.co/fr/ressources/comment-la-rv-ra-vont-elles-influencer-le-monde-des-affaires/

Fogarty, R. (1994). *How to teach for metacognition*. Palatine, IL: IRI/Skylight Publishing

Fraisse, P. Mc Murray, G. (1980). *Etude génétique du seuil visuel de perception pour quatre catégories de stimuli*. L'année psychologique.

Freinet, C. (1964). Tome 2 : *Les invariants pédagogiques*.

Glossary. (2018). Retrieved from : La revue européenne des médias et du numérique. http://la-rem.eu/glossary/

Biederman I. and Gerhardstein P. C. (1993). *Recognizing Depth-Rotated Objects: Evidence and Conditions for Three-Dimensional Viewpoint in variance*. Journal of Experimental Psychology: Human Perception and Performance 19 (6), 1162-1182

Kohonen, V. *Learning to learn through reflection – an experiential learning perspective*. University of Tampere. Retrieved from : http://archive.ecml.at/mtp2/Elp_tt/Results/DM_layout/00_10/05/Supplementary%20text%20E.pdf

Kolb, D.A. (1984). *Experiential learning*. ISBN-13: 978-013295261

Montessori, M. (1912). *The Montessori Method*. New York: Frederick A. Stokes Company. Retrieved from : http://digital.library.upenn.edu/women/montessori/method/method.html

Piaget J. (1936). *La naissance de l'intelligence chez l'enfant*. Paris : Delachaux et Niestlé. Retrieved from : http://www.fondationjeanpiaget.ch/fjp/site/textes/VE/JP36_NdI_avpropos_intro.pdf

Roulin, J.L., Auclair. L. (2006). *Psychologie cognitive*. Rosny : Bréal. Coll. Grand Amphi, psychologie

Salamin, A.D. Hadorn, C. (2017). *e-MEMENTO : a smartwatch experiment to investigate rote memorization in the connected age*. Site conference Proceedings. Retrieved from : https://www.academicexperts.org/conf/site/2017/papers/50269/