The Effect of VR Learning on Learning Outcome:
Understanding Visually Complex Concepts

Jarno E. Smit
STUDENT NUMBER: U1274850

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF SCIENCE IN COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE
DEPARTMENT OF COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE
SCHOOL OF HUMANITIES AND DIGITAL SCIENCES
TILBURG UNIVERSITY

Thesis committee:

Dr. Marie Postma
prof. Dr. Eric Postma

Tilburg University
School of Humanities and Digital Sciences
Department of Cognitive Science & Artificial Intelligence
Tilburg, The Netherlands
May 2019

Preface

I would like to express my gratitude to my thesis supervisor, Dr. Marie Postma, whose guidance throughout the development of this thesis has been very insightful. I could not have hoped for a better supervisor.

The creating of the virtual reality application and conducting of the experiments for this study have been done in collaboration with another student (Reinier Zwart). However, this thesis was written individually.

The Effect of VR Learning on Learning Outcome:

Jarno E. Smit

In this work the effect of learning about a visually complex concept in VR on learning outcome was investigated. An experiment was conducted where participants learned about Principal Component Analysis on paper or in VR. This study has not found any effects of learning in VR on learning outcome. However, participants did enjoy learning in VR more than learning on paper. In addition, the effects of flow on learning outcome were analyzed. The results did not show any effects of flow on learning outcome. A first step was made to supplement current research on VR learning and to create a concrete way to make visually complex concepts understandable for people that would otherwise struggle with this. Two limitations are discussed which, when removed, could provide other findings.

1. Introduction

In recent years, virtual reality (VR) systems have developed rapidly. As an effect of technological advances and decreasing costs, VR has gotten much more accessible. Consequently, VR has made its way into many fields of research and business. As is the case for education (Merchant et al. 2014). Research has been done on the use of virtual reality for educational purposes, and has proven VR to be an appropriate medium for education. In addition, studies have concluded that learning in VR can improve learning outcome, when compared to traditional learning media (Merchant et al. 2014; Mikropoulos and Natsis 2011; Psotka 1995; Winn 1993). In addition, VR has some advantages over traditional learning media and offers some possibilities that would otherwise not be possible. For example, VR can create both environments that would otherwise be too difficult or expensive to simulate and create environments that cannot be represented in the real world.

Another trend is the increase in interest in Artificial Intelligence and Data Science. The two fields are quite similar and have a shared factor: machine learning. Machine learning techniques often use complex data transformations to give new insights about the data. These transformations usually require spatial ability (i.e. visual thinking) to be understood, which makes them difficult to understand for many people. Since VR offers the possibility to visualize objects in three dimensions and represent data in a way that would not be possible in the real world, it is interesting to explore VR as a learning medium to explain these concepts in. Although much research has been done on learning and VR, research on learning about visually complex concepts, specifically machine learning techniques, in VR has yet to be done .

This study will explore the use of VR as a learning medium for visually complex concepts. The goal is to make a contribution to current research on VR learning and explore a concrete way to make visually complex concepts understandable for people that would otherwise struggle with this. Principal component analysis (PCA), a widely used machine learning technique, will be used as learning material in this study. The study aims to answer the following research question:

RQ_1: What is the effect of learning in virtual reality on understanding of Principal Component Analysis?

To answer this question, an experiment will be conducted with two conditions. In the first condition participants will learn about PCA on paper. In the second condition participants will learn about PCA in virtual reality. A VR application will be created for this study.

In addition to the effect of VR learning on learning outcome this study will investigate the effects of flow on learning outcome. Flow is a state of mind where people are completely involved and focused and is known to increase learning outcome. To measure flow, participants will fill in a self-evaluating questionnaire after finishing the explanation. The second research question is formulated as follows:

RQ_2: What is the effect of flow on the understanding of Principal Component Analysis?

This study did not find an effect of learning in VR on learning outcome. Furthermore, no effect was found of flow on the understanding of PCA. Participants did indicate they enjoyed learning in the virtual environment, which is in line with previous research. Cognitive overload, due to the amount of information given to the participants, and the difficulty level of the test on PCA are seen as limitations to this study.

This study has the following outline. Firstly, the related work on VR and learning, spatial ability, and flow are discussed. Secondly, the method of this study is stated, followed by the results. Then, the results are interpreted in the discussion. Finally, the conclusion chapter gives a summary of the main findings and discusses possibilities for future research.

2. Related Work

Virtual reality could prove itself to be a successful medium to explain visually complex concepts with, since it offers the ability to visualize data in three dimensions and represent abstract information. This study explores the use of virtual reality as a learning environment to learn about a widely used machine learning technique, Principal Component Analysis. The following section will cover the literature on the three main constructs of this study: virtual reality learning, spatial ability, and flow.

2.1 Virtual Reality and Learning

In this section two aspects of learning in virtual learning environments are discussed. Firstly, the possibilities and advantages that VR offers in comparison to the 'real world' are reviewed. Secondly, literature on the effects of virtual environments as a learning medium on learning outcome is covered.

Virtual reality consists of technologies that can create synthetic and highly interactive three-dimensional environments (Mikropoulos and Natsis 2011). Virtual reality (VR) knows a relatively short history, yet has come a long way since its birth in the late twentieth century. One of the first virtual reality systems was the 'Sword of Damocles' (Sutherland 1968), which was a head-mounted-display system for entertainment purposes. The head-mounted-display weighted so much that needed to be carried by a mechanical arm. In the early stage of VR development, VR systems were mainly used for entertainment and training. However, when costs of development decreased and the technology advanced and become more accessible, virtual reality made its way into education around 1990 (Merchant et al. 2014). Since the upcoming of VR, many researchers have studied virtual reality and its implementations for educational purposes. Constructivism is agreed upon by a majority of researchers to be the best theoretical framework to build virtual learning environments on (Winn 1993; Mikropoulos and Natsis 2011). The theory of constructivism regards how people learn and states that, in a nutshell, people construct their own knowledge and meaning while learning (Hein 1991). Constructivism is the theoretical framework on which the virtual environment used in this study will be based upon.

Virtual reality offers some unique educational possibilities compared to traditional educational techniques. In a study about educational applications of virtual reality, Winn (1993), states that immersive virtual reality environments can provide distinctive learning possibilities from those that students usually get in education. The research states that immersion is the key to applying VR in education, since it offers the possibility to experience a synthetic world in first-person and learn from that experience. VR is immersive when the virtual environment, to some extent, looks and feels like the real world (Psotka 1995, p. 406).

Another possibility virtual reality offers is that it can aid situated learning (Psotka 1995, p. 420). Situated learning states that learning in contextual situations can positively affect the learning process, compared to learning only from conceptual, and out of context, knowledge (Brown, Collins, and Duguid 1989). A situated learning environment in VR can be, for example, a simulation of a real world context that would otherwise be hard or expensive to replicate. Furthermore, VR learning environments can create simulations of situations that cannot be represented in the real world. Winn (1993) describes three kinds of learning applications that cannot be created in the real world: size, transduction and reification. Size refers to the ability to disregard real world limitations of size changes. Using virtual reality, a person could be shrunk down to the size of a molecule and learn about molecular biology. Transduction is described as hardware that modifies sensory information that we cannot process into a form which we can process. This allows the user to 'feel' things that could otherwise not be sensed. Reification is the process where normally imperceptible representations of information are made perceptible. The current study aims to implement situated learning by representing a context in a virtual learning environment which cannot be created in the real world.

In addition to size, transduction and reification, Mikropoulos and Natsis (2011) add first order experiences, natural semantics, autonomy and presence as aspects of VR that positively affect learning outcome in a study reviewing ten years of educational VR research. A first order experience arises from experiencing the world from a first person view and being able to maneuver at will. Natural semantics can contribute to learning since it is not dependent on using symbols. First order experience and natural semantics both reinforce the feeling of presence in the virtual environment. Presence, which Mikropoulos concludes is a key factor VR learning, is the feeling of being present in the virtual environment. There are indications that presence has a positive effect on learning outcome (Mikropoulos 2006; Winn et al. 1999). However, research is still inconclusive about the effects of presence. For example, Whitelock et al. (2000) concluded that presence could cause cognitive overload and could negatively affect understanding. Since current literature is inconclusive about the effects of presence on learning outcome, this study will explore an alternative construct, namely flow.

In a study regarding the effects of virtual reality learning environments on learning outcome, Merchant et al. (2014) analyzed studies in the categories: games, simulations and virtual worlds. The authors define simulations as environments where real-life situations or processes are replicated. Games are defined as a special kind of simulation where the focus lies on goals, achievement and rewards. Virtual worlds are defined as open-ended environments where users are autonomous, these environments can be used to interact with objects, avatars and other users. The study concludes that all three categories had a positive effect on learning outcome gains. In addition, the study concludes that simulation environments are better for assessing and acquiring knowledge than skill, since skill requires recurring practice. Furthermore, the results suggest that participants retained the information in long-term. The current study will implement a simulation in which Principal Component Analysis is explained to the user.

Research has shown the possible advantages and effects on learning outcome of virtual reality learning environments. There are many studies regarding learning in virtual reality and representing abstract information. However, understanding machine learning, which requires understanding of visually complex concepts, in virtual reality remains an unexplored field. Therefore, this study will investigate the effect of learning in virtual reality on the understanding of such a concept. Two other constructs are covered in this chapter. Firstly, spatial ability is linked to the understanding of visually complex concepts and implications of spatial ability for this study are discussed. Secondly, flow is explored as a construct that could positively influence learning outcome.

2.2 Spatial Ability

Spatial ability refers to the ability to understand and interpret the spatial relations of an object. Simply said, it is ones ability to think visually. Spatial ability is an important component of human intelligence and is a key component for multiple professions, such as surgeons and mathematicians (Dünser et al. 2006). VR offers the ability to represent objects in 3D and is therefore very suitable for working with spatial concepts. Studies have shown that virtual reality environments can be used to teach about spatial concepts and improve spatial ability. For example, Rizzo et al. (1998) investigated if VR could be used to improve spatial ability by learning about mental rotations. The results showed that participants in the VR condition had higher learning outcome gains than those in the control group.

A study measuring the effect of VR-based learning on high school students with different levels of spatial ability found that students learning in VR achieved better results than those that did not (Lee and Wong 2014). Furthermore, the results showed that performance of students with a low spatial ability were affected more positively than those with a high spatial ability. This suggests that learning about concepts that require spatial ability in VR might be more beneficial for people with a low spatial ability. Jang et al. (2017) add to this, in a study on the effect of two virtual environments (manipulation versus passive viewing) on learning outcome for medical students studying complex anatomical structures, that being embodied in the virtual environment benefits participants with low spatial ability more than those with high spatial ability. This implies that having the feeling of being physically present in the virtual environment is, mainly for participants with low spatial ability, important when being presented with visually complex concepts. However, literature is not conclusive on the effect of spatial ability on learning outcome when learning visually complex concepts. For example, Huk (2006) concluded that only participants with high spatial ability perform well when presented with three-dimensional visualizations and that participants with low spatial ability suffer from cognitive overload, causing them to perform worse.

Wanzel et al. (2002), comparable to the current research, studied the effect of spatial ability on learning of spatially-complex surgical skills. Spatial ability was measured using six spatial ability tests, among which the mental rotations test used in the current study. Participants, junior surgical residents, performed two procedures differing in difficulty after receiving an instruction on the procedure. The study concluded that participants with better spatial ability scored significantly better than those with lower scores. This suggests spatial ability has a positive effect on learning outcome in the understanding of complex visual concepts. Lufler et al. (2012) conducted a similar research, studying the effect of spatial ability on performance of medical students in a gross anatomy course. The authors concluded that students that scored high on the spatial ability test were more likely to score 90 percent or more on the anatomy exam, compared to students that scored low on the spatial ability test. These findings reinforce the conclusions made by Wanzel et al. (2002). The authors add that students who scored highest on the spatial ability test not only performed better on questions requiring visual thinking, yet also performed better on non-spatial questions.

Many techniques and tests to measure spatial ability exist, ranging in difficulty. Some examples of these tests are the Snowy Pictures Test, Gestalt Completion Test, Shape Memory Test, Cube Comparison Test, Card Rotations Test, Form Board test, and the Mental Rotations Test (Ekstrom, Dermen, and Harman 1976, p. 4). Each of these tests focuses on a different component of spatial ability. The current study focuses the understanding of visually complex concepts, specifically Principal Component Analysis. In order to understand how PCA works, a person has to be able to rotate three-dimensional figures (data) in their mind. Therefore, the Mental Rotations Tests (MRTA) (Peters et al. 1995) is used in this study, since this test is widely used (Jang et al. 2017; Lufler et al. 2012; Rizzo et al. 1998; Wanzel et al. 2002) and requires participants to mentally rotate three-dimensional figures.

This study aims to supplement current research on learning in virtual reality by investigating the effect of learning about a visually complex concept in VR on learning outcome. More specifically, this study will use Principal Component Analysis, a technique often used in machine learning, as an example to learn about. PCA was selected since it makes use of visually complex data transformations, which require visual thinking to understand. The current research will attempt to answer the following research question:

RQ_1: What is the effect of learning in virtual reality on understanding of Principal Component Analysis?

Research has shown that learning in VR can have positive effects on learning outcome. More specifically, it can increase learning outcome when learning about concepts requiring visual thinking in VR (Lee and Wong 2014; Rizzo et al. 1998). Therefore, the following hypothesis is formulated:

H_1: Participants who learn about PCA in VR will score significantly better on understanding than participants learning about PCA on paper.

PCA is a machine learning technique that is visually complex. Based on previous studies it can be expected that participants with a higher spatial ability will score better overall on understanding of PCA in comparison to participants with a lower spatial ability, based on the visual nature of the technique (Lufler et al. 2012; Wanzel et al. 2002). However, based on the literature, it could also be that participants with a low spatial ability will have higher learning outcome gains than those with a high spatial ability (Jang et al. 2017; Lee and Wong 2014). The effects described above could interfere with each other: participants with a high spatial ability might naturally understand PCA better, yet participants with a low spatial ability could be more positively affected by learning in VR. This study will analyze the effect of spatial ability and learning medium (paper versus VR) on learning outcome, where Spatial ability will be seen as a moderator in this study.

2.3 Flow

As defined by Csikszentmihalyi and Larson (2014, p. 136), "flow denotes the holistic sensation present when we act with total involvement". The authors state that the flow state only arises when a person is actively engaged in an interaction with the environment and uses some skill. A so-called 'flow activity' is needed to enter a state of flow. Csikszentmihalyi and Larson (2014, p. 240) define 9 dimensions of flow. The first dimension is Challenge-Skill Balance, which states that the challenge of the situation should be fitting to one's skills. The second dimension is Clear Goals, which means that one's goals for the situation should be clear. The third dimension is Concentration on Task at Hand and states that one should have intense and focused concentration in the situation. The fourth dimension, Merging of Action and Awareness, consists of intense involvement in the situation which merges ones actions and awareness, as if it were to go automatic. The fifth dimension is Sense of Control, which is defined as the sense one gets that all actions are under their control. The sixth dimension, Transformation of Time, states that one experienced a sense of altered time, usually that time passed faster than normal. The seventh dimension is Loss of Self-Consciousness, which states that during the situation one loses the sense of self-consciousness as an actor in the situation. The eighth dimension, Autotelic Experience regards if the situation is intrinsically rewarding. The final dimension is Unambiguous Feedback, which states that one gets immediate and clear feedback in the situation.

The effect of flow on learning has been studied in all kind of fields and being 'in the flow' has been linked to having positive effects on learning (Choi, Kim, and Kim 2007; Skadberg and Kimmel 2004; Webster, Trevino, and Ryan 1993). Custodero (2002), for example, studied the effect of flow in music education. The study concluded that flow helped construct musical meaning and motivated children to stick with music. In a study testing an e-learning model based on the flow theory as a training method, Choi, Kim, and Kim (2007) concluded that flow experience significantly and positively affected learning outcome. Furthermore, the authors add that flow, besides having a direct effect on learning outcome, also had an indirect effect by inducing a positive attitude toward what was being learned. There are multiple other studies that showed flow to have indirect effects. For example, Ghani and Deshpande (1994) concluded that flow resulted in exploratory behavior user, which in turn resulted in more extensive use and enjoyment. In another study regarding human-computer interactions, the authors concluded that flow triggers playfulness and can positively affect work-related outcome (Webster, Trevino, and Ryan 1993). Another possible effect of flow on learning was found by Guo et al. (2007), who stated that flow had a positive effect on participants' perceived learning of the subject, skill development, and learning satisfaction. These findings are supported by conclusions of a similar study by Rossin et al. (2009).

While some studies find a direct link between flow and learning outcome, others do not. Guo et al. (2007), for example, concluded that flow only had indirect effects on learning. The authors state two main reasons as to why flow might not have had an effect on learning outcome. Firstly, it is concluded that the multiple-choice questions were quite challenging for the students and might not be able to properly assess flow. The authors suggest future studies investigate if the effects of flow can be measured with other performance measures. Secondly, the authors add that the study was conducted in one course and future research could examine if the effects might differ in other courses or academic disciplines. In a similar study, Rossin et al. (2009) also conclude multiple-choice questions might not properly assess the effect of flow on learning outcome, and other performance measures should be tried. The current study will use open-ended questions to measure learning performance and will explore if this measure is able to assess the effect of flow on learning outcome.

Ibáñez et al. (2014) compared an augmented reality (AR)-based application with a web-based application and studied the effects on learning outcome and enjoyment. The study showed that participants in the AR condition experienced significantly more flow overall and had better learning outcomes. In particular, participants in the AR condition scored higher on the dimensions 'Concentration on Task at Hand' and 'Transformation of Time'. The author concludes that this might be the cause of the increase in performance from participants in the AR condition. Furthermore, participants in the AR condition scored higher on sense of control, unambiguous feedback, and autotelic experience. The current study focuses on virtual reality, which shares many features with augmented reality. In addition, Reid (2002) states that VR meets the requirements to induce flow. These findings suggest that participants in the virtual reality condition could experience more overall flow than those in the paper condition.

Ibáñez et al. (2014) additionally stated that not all dimensions had an effect on the experienced flow, only Concentration on Task at Hand, Transformation of Time, Sense of Control, Unambiguous Feedback, and Autotelic Experience contributed to the flow experience. Likewise, Rossin et al. (2009) concluded only the dimensions Clear Goals, Unambiguous Feedback, Challenge-Skill balance, and especially Autotelic Experience contributed to learning. To account for the possibility that some dimensions of flow have a different effect on learning than others, the effect of each dimension and the effect of flow as a construct on learning outcome will be investigated in this study.

Measuring flow is difficult due to its "inherently unstable, un-self-conscious, and subjective" nature (Csikszentmihalyi and Larson 2014, p. 245). There are three main types of measures for flow: interviews, questionnaires, and the experience sampling method. Interviews can be used to determine and give in-depth insights into when flow experiences arise, what causes them or what disrupts them. Questionnaires are often used when the goal to measure the dimensions and measure differences between subjects, instead of determining if a flow experienced has occurred. The experience sampling method requires participants to fill in an questionnaire at certain moments throughout the day for a certain amount of days. The goal is to recognize in which situations flow experiences arise. In the current study, the goal is to compare levels of flow among participants and assess its effect on learning outcome. Therefore, a questionnaire will be used to measure flow. Specifically, an adapted version of the Flow State Scale (FSS) by Jackson and Marsh (1996) will be used.

Literature has shown that flow has the ability to positively and directly affect learning outcome (Choi, Kim, and Kim 2007; Ibáñez et al. 2014; Webster, Trevino, and Ryan 1993), yet some studies concluded flow only indirectly affects learning (Ghani and Deshpande 1994; Guo et al. 2007; Rossin et al. 2009). This research aims to examine the relationship between flow and learning outcome. The following research question is formulated:

RQ_2: What is the effect of flow on the understanding of Principal Component Analysis?

To answer the research question, the following hypothesis is formulated:

H_2: A higher level of flow experience will result in significantly higher learning outcome.

Based on the literature discussed, a theoretical model was made (figure 1). To answer the research questions an experiment has been conducted. In the experiment, three constructs were measured: understanding of Principal Component Analysis, flow and spatial ability. In the following chapter, the methodology of the study is described.

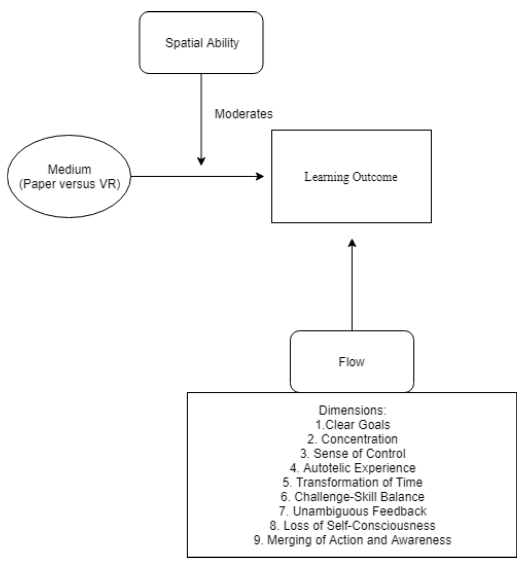

Figure 1
Theoretical model

3. Experimental Setup

In this chapter, the methodology of the current study will be described. A between-subjects design was used for this study, where the virtual reality condition was the experimental group and the paper condition was the control group. Firstly, the development of the two experimental conditions will be described in section 3.1. Secondly, demographics regarding the participants of the study are given in section 3.2. Then, experimental design and procedure will be discussed in section 3.3 and 3.4. Section 3.5 covers the construction and some descriptive statistics of the measurements. The chapter is concluded with a brief description of the soft- and hardware used in this study.

3.1 Development

Two experimental condition have been created for this study. For the first condition, a paper explanation of Principal Component Analysis was written. For the second condition, a virtual reality application was made to explain PCA with. The following subsections describe the construction of these ’teaching media’. The creation of both conditions and conducting of the experiments has been done in collaboration with another researcher. All data analyses and interpretations were made individually.

3.1.1 Paper condition. A paper explanation of Principal Component Analysis was created to teach participants about the subject. The main objective for this explanation was to cover the basics of PCA in such a manner that every participant (university students) would be able to understand the concept. The explanation was inspired by George (2013). This article was chosen as a basis of the explanation since it offers very clear and basic information and visualizations. All visualizations were color coded to add extra clarification. In addition, an example about an iris data-set was added to the explanation. The same example was used in the VR condition.

The explanation was made iteratively. Firstly, one of the researchers made a draft, which was then revised by the second researcher. Afterwards, the explanation was proofread by multiple the thesis supervisor and other students, and then revised again. The final and complete version of the explanation can be found in appendix A.

3.1.2 Virtual Reality Condition. A VR application was made in Unity (version 2018.3.2) to explain PCA to the participants. The information given in the application coincides with the paper explanation, to ensure that participants get the same information, only presented differently. During the development, some pilot versions were tested by other students and the thesis supervisor, on the basis of which the application was then adapted.

The explanation in the application was given by a voice over, which was recorded by a native English speaker. It was chosen not to embody the voice, since this might divert participants’ attention to the avatar from where it should be. The environment was designed to minimize distraction, yet still arouse interest. The environment existed of a big surface surrounded by four walls and a star-filled sky above, making it a surreal environment. Figure 2 and 3 can give somewhat of an impression of the environment. During the experiment, there were some moments where participants could walk freely in the environment to investigate the data. However, for most of the explanation, participants were in a fixed position to retain their attention.

The final application can be divided into five scenes. The first scene gave a general introduction to the VR application and the controls. In the second scene, participants were introduced to the data-set that was used in the application (figure 2), which is the same data-set as that in the the paper condition. In addition, the problem that is tackled in the explanation, trying to find the view were most information about the data can be seen, was introduced. In the third scene, a very basic and abstract introduction to PCA was given. The introduction asked participants to find the angle where most of a teapot could be seen and used this as an example to apply PCA on. The third scene was inspired by an instructional video on PCA (J. 2009). In the fourth scene, the teapot example was swapped for the data-set from earlier, on which PCA is then applied (figure 3). In the final scene, the concept of dimensionality reduction was explained and applied to the data and a summary of what was explained was given. After each scene participants had the option to either replay the scene or continue to the next one, giving participants that might struggle with the information the chance to revise it.

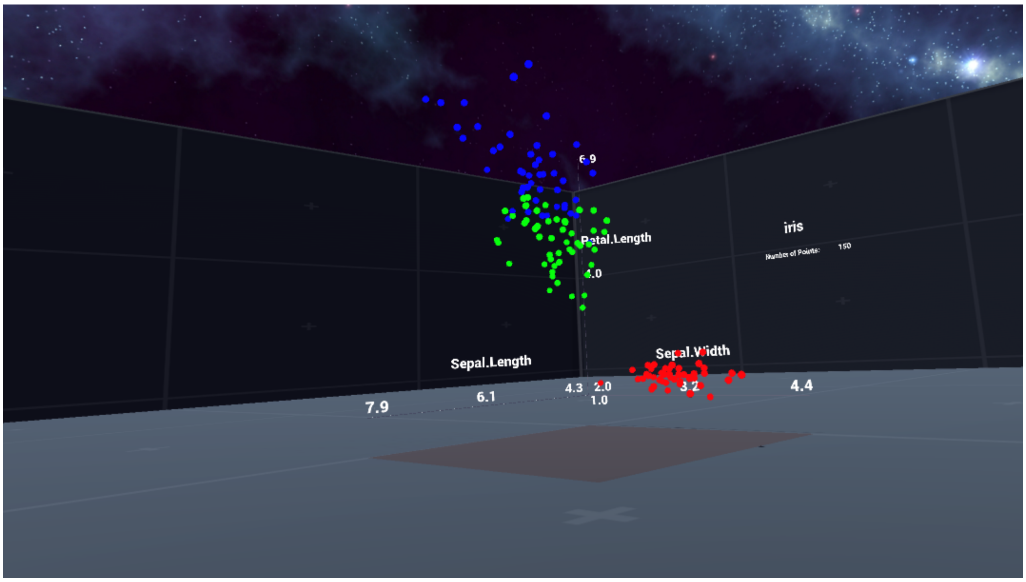

Figure 2
The data-set in the VR environment

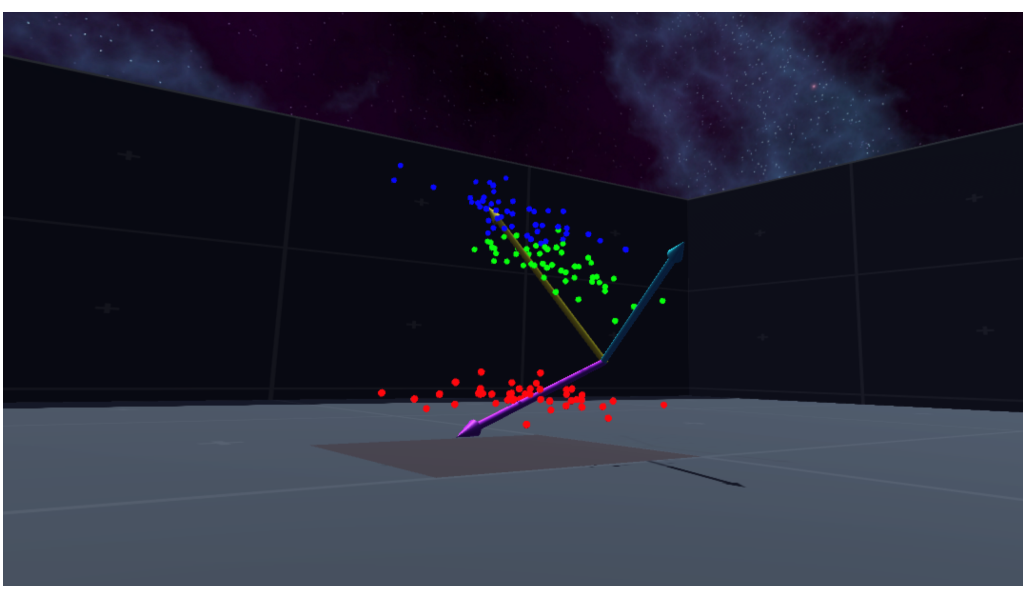

Figure 3
The principal components fitted to the data in the VR environment

3.2 Participants

Participants for this study were recruited using the Human Subject Pool from Tilburg University. In total, eighty people participated in the experiment. The average age of participants was 22 (*M*=21.88, *SD*=4.08), the participants were between 18 and 42. There was one outlier somewhat skewing the average, as one participant was considerably older than the rest. Forty-six percent of all participants was male (*N*=37) and fifty-four percent was female (*N*=43). There were an equal amount of males and females in the VR condition, in the paper condition there were seventeen males and twenty-three females. People with a background of epilepsy or migraine were not accepted to the experiment, since virtual reality might bring risk to them. Participants were asked before the experiment whether they: 1. Did not know what PCA was, 2. Recognized the term, but could not explain it, or 3. Recognized the term, and could explain the workings of PCA. Eighty-nine percent of participants (*N*=71) had never heard of PCA, eight and a half percent (*N*=7) recognized the term but could not explain it, and two and a half percent (*N*=2) knew the term and could explain the workings of PCA. Participants in the VR condition were asked to rate their prior experience with virtual reality ranging from: 1. No experience, 2. Once or twice, to 3. More than twice. Overall, most participants had experienced some form of VR. Twenty percent (*N*=8) had never experienced VR, Sixty-five and half (*N*=27) had experienced VR once or twice, and twelve and a half percent (*N*=5) had experienced VR more than twice.

3.3 Design

To answer the research questions of this study, an experiment was conducted. In the experiment, participants got a basic introduction to PCA in either a virtual reality learning environment or on paper. In advance to the explanation, participants' spatial ability was measured using a mental rotations test. Afterwards, a self-evaluating questionnaire was made to measure perceived flow experience. Furthermore, participants made an open question test to measure their understanding of PCA. The dependent variable of this study was learning outcome, which was represented as a score based on the understanding of Principal Component Analysis. There were two independent variables for this study. The first is the condition in which participants learned about PCA (VR versus paper). The second independent variable was the level of experienced flow. In addition, spatial ability was used as a moderator in this study. A more detailed description of how the concepts were measured can be found the subsection 3.5.

3.4 Procedure

Before starting the experiment, each participant first read an information letter about the study and then signed a consent forum if they agreed with the presented information. Afterwards, participants filled out a short questionnaire with personal information. Then, participants' spatial ability was measured using a mental rotations test. The mental rotations test is further described in section 3.5.3.

In the VR condition, the experiment continued with an explanation about the application and how to use the equipment (Oculus Rift). Before entering the VR environment, participants were told that they would make a open-question test about PCA after the explanation. In the VR application, participants got a basic introduction to PCA divided into five scenes. In the other condition, participants were presented a paper explanation of PCA. Before reading the text, participants were told that they would make an open-question test about PCA after the reading the explanation. Participants had fifteen minutes to study the text.

After the explanation in either the VR environment or on paper, participants made a questionnaire to evaluate their flow experience. Afterwards, a test containing eight open questions about PCA was made by the participant, to measure their understanding. All participants had ten minutes to finish the test.

3.5 Data

Three concepts were measured in the experiment: Flow experience, understanding of PCA, and spatial ability. Flow was measured using a self-evaluating questionnaire, understanding was measured using a test the researchers developed, and spatial ability was measured using a mental rotations test. The measures and some descriptive statistics are thoroughly discussed in the following subsections.

3.5.1 Flow. Flow was measured using an adapted version of the Flow State Scale (FSS) by Jackson and Marsh (1996). The questionnaire is based on the nine dimensions of Flow (Csikszentmihalyi and Larson 2014, p. 240). One of the dimensions was omitted, namely Merging of Action and Awareness. This was done because this dimension focused on performing actions, which was not relevant for this study. The remaining questionnaire contained thirty-two questions divided equally into eight dimensions. All questions were measured on a five-point scale. The full questionnaire is listed in appendix B.

Reliability was measured using Cronbach's alpha, which determined reliability to be acceptable (32 items; α=0.92). In appendix C, an overview of average score per question and condition is stated. The average Flow Score was calculated and used in further analysis (*M*=3.47, *SD*=0.57). In appendix D the distribution of Flow score per dimension and condition is visualized. In table 1, the mean and standard deviation of Flow score per dimension and condition is stated. There were some missing values, one participant did not fill in the Flow questionnaire.

Dimension	Paper		VR	
	M	*SD*	*M*	*SD*
Challenge Skill	3.67	1.02	3.60	0.84
Clear Goals	3.74	0.92	3.8	1.06
Unambiguous Feedback	3.15	1.08	3.41	0.95
Concentration	4.10	1.13	3.85	0.99
Sense of Control	3.82	1.00	3.85	0.99
Loss of Self-Consciousness	3.54	1.24	3.65	1.22
Transformation of Time	3.38	0.90	3.2	1.06
Autotelic Experience	3.62	1.09	4.25	0.95

Table 1
Descriptive statistics of the flow dimensions across conditions.

A Principal Component Analysis was conducted on the flow questionnaire to investigate if the construct consisted of any (unknown) sub-constructs. Initial extraction of the components resulted in thirty-two principal components. Upon assessing the scree-plot 4 and variance explained by the principal components, two principal components were retained (accounting for 43 percent of total variance). The first principal component (eigenvalue=10.51) accounted for 33 percent of the total variance. The second principal component (eigenvalue=2.87) accounted for 9 percent of the total variance The rotated factor loadings (Varimax rotation), stated in appendix H, were too low to distinct underlying constructs.

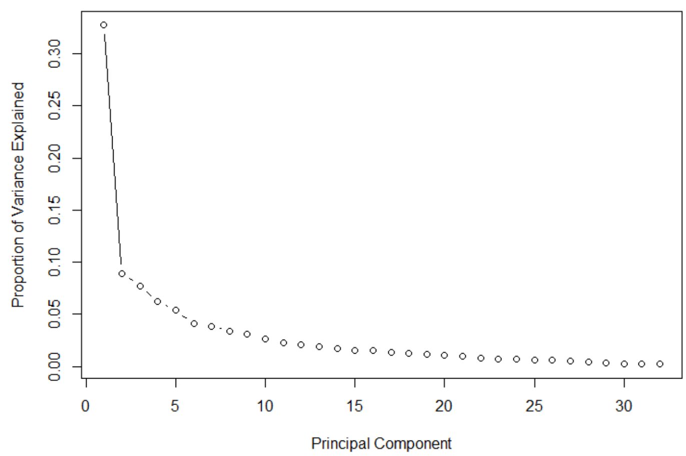

Figure 4
Scree plot of PCA flow questionnaire

3.5.2 Understanding. To measure the understanding of Principal Component Analysis, a test was created by the researchers and approved by the thesis supervisor. The test consisted of eight open questions and can be found in appendix E. All question were made to measure understanding, there were five 'textual' questions and three 'visual' questions. In order to grade the tests, an answer key and point distribution was made.

All tests were checked by the researchers individually. An interrater reliability analysis, using the Kappa statistic, was performed to determine consistency among raters. Cohen's Kappa showed outstanding overall interrater reliability, with 96.6 percent agreement (K=0.95, p=0). All test were graded on a scale from zero to ten, the final scores were used for analysis (M=4.35, SD=2.13). Distribution of overall PCA understanding score and PCA understanding per condition can be found in appendix F. There were no missing values.

3.5.3 Spatial Ability. A mental rotations test was used to measure spatial ability. This study used a revised version of Vandenberg and Kuse's Mental Rotations Test by Peters et al. (1995) (version MRT-A). Participants were presented an example figure with four stimuli to the right of it (figure 5). Two of these four were a rotated version of the left figure. Participants had to cross out these two rotations and only received a point if both rotations were correctly crossed out.

The original test consisted of twenty-four figures. Due to time limitations, the current study only used twelve figures per participant. The original questions where split in halve and both halves where used an equal amount of times, to minimize possible bias. Before making the test, the participants received a verbal and textual explanation and practiced four figures without time limitations. Participants had three minutes to make the twelve-figure test.

A spatial ability score was calculated for each participant, this is the amount of points obtained divided by the total amount of points that could be scored (M=0.40, SD=0.19). The distribution of spatial ability score overall and per condition can be found in H. There were no missing values.

To test if the shortened version of the mental rotations test was reliable, scores and distribution of the shortened test were compared to those of the original test (Peters et al. 1995). The average score, standard deviation and score distribution of the shortened test were very similar to those of the original test. Therefore, it can be concluded that the shortened test was an acceptable measure of spatial ability.

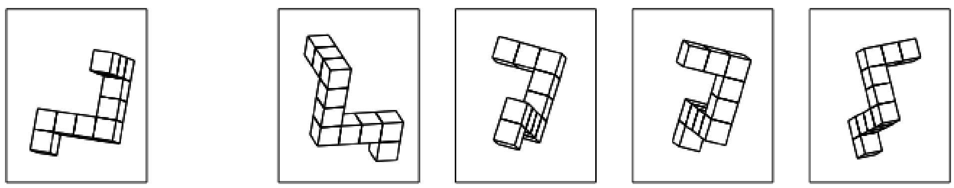

Figure 5
Example from Mental Rotations Test

3.6 Models

The hardware used for the virtual reality application was an Oculus rift. The controls in the application used the Oculus Touch controllers. The application was made with Unity, version 2018.3.2. All analyses were done in RStudio (version 3.5.1). Table 2 shows all the packages that were used.

Package	Abbreviation
Procedures for Psychological, Psychometric, and Personality Research	psych
Tools for Splitting, Applying and Combining Data	plyr
Create Elegant Data Visualisations Using the Grammar of Graphics	ggplot2
Tools for Tall Distributed Matrices	kazaam
Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files	xlsx
Various Coefficients of Interrater Reliability and Agreement	irr
Extension of 'data.frame'	data.table
Support Functions and Datasets for Venables and Ripley's MASS	MASS
STAT: Interactive Document for Working with Basic Statistical Analysis	stats
S3 Infrastructure for Regular and Irregular Time Series	zoo

Table 2
R-packages.

4. Results

In this section, the main findings of this study will be presented. The section will start of with some descriptive statistics of the measurements. Afterwards, multiple analyses of the data are conducted.

Three constructs were measured in this study across two conditions. In table 3, the descriptive statistics of all measures are listed per condition. The mean, standard deviation, and median are given to give insight on the distribution of the measures. In addition, the minimum and maximum values are given together with the range they could fall in. The skewness and kurtosis values of all measures are within the acceptable range (between -1.96 and 1.96), thus normality of distribution can be assumed.

Cond.	Measure	*M*	*SD*	Med.	Min.	Max.	Range	Skew.	Kurt.
Paper	Flow	3.43	0.62	3.4	2.22	4.66	1-5	-0.23	-0.42
	Spatial Ability	0.38	0.19	0.42	0	0.75	0-1	0.10	-0.59
	Understanding	4.25	2.37	4.00	0.5	9	0-10	0.29	-1.01
VR	Flow	3.51	0.53	3.48	2.50	4.47	1-5	0.01	-0.77
	Spatial Ability	0.43	0.19	0.42	0	0.75	0-1	-0.20	-0.39
	Understanding	4.45	1.89	4.5	1	9.5	1-10	0.42	-0.07

Table 3
Descriptive statistics of Flow, spatial ability and understanding across condition.

The first research question aimed to measure the effect of learning in virtual reality on understanding of Principal Component Analysis. An independent t-test was conducted to compare understanding score between conditions. There was no significant effect between the paper condition (M=4.27, SD=2.40) and the VR condition (M=4.45, SD=1.89) on understanding score, conditions; $t(74.26)$=-0.42, p=.678.

Previous research suggested that participants in the VR condition could experience more overall flow than those in the paper condition. Therefore, an independent t-test was conducted to compare Flow scores between conditions. There was not a significant difference in the Flow scores for participants in the paper condition (M=3.43, SD=0.62) and the VR condition (M=3.51, SD=0.53), conditions; $t(74.53)$=-0.60, p=.551.

A correlation matrix was made to visualize the relationship between Flow, Spatial Ability and Understanding in table 4. Overall correlation and correlation per condition is stated. Overall, both Flow and spatial ability have a weak positive correlation with understanding and with each other. However, in the paper condition Flow has a moderate positive correlation with understanding and the correlation between spatial ability and understanding also positively increased. In contrast to the paper condition, Flow has a weak negative correlation with understanding in the VR condition. Furthermore, spatial ability has an almost nonexistent correlation with understanding in the VR condition.

The correlation matrix shows a weak positive correlation between spatial ability and understanding, yet does show difference in correlation between conditions. To investigate the difference in correlation of understanding and spatial ability between conditions, spatial ability was split into participants with high (>0.5) and participants with low (<0.5) spatial ability. In the paper condition participants with a high spatial ability (N=6, M=5.00, SD=2.52) scored higher than those with low spatial ability (N=34, M=4.12, SD=2.34). However, in the VR condition participants with a low spatial ability

		Flow	p	Signif.*	Spatial Ability	p	Signif.*
Overall	Flow	-			-		
	Spatial Ability	0.09	.415		-		
	Understanding	0.16	.169		0.13	.260	
Paper	Flow	-			-		
	Spatial Ability	0.09	.573		-		
	Understanding	0.35	.031	*	0.22	.184	
VR	Flow	-			-		
	Spatial Ability	0.08	.637		-		
	Understanding	-0.13	.439		0.01	.951	

*Signif. codes: 0 '∗ ∗ ∗' 0.001 '∗∗' 0.01 '∗' 0.05 '.' 0.1 ' ' 1

Table 4
Correlation Matrix of Flow, Spatial Ability and Understanding per condition and overall.

(*N*=23, *M*=4.63, *SD*=1.85) scored higher than those with a high spatial ability (*N*=9, *M*=4.17, *SD*=1.89).

Another correlation matrix was made to analyze the relation between each dimension of Flow and understanding to see if correlation differs greatly between dimensions. Only the dimension Challenge-Skill Balance had a moderate correlation with understanding. Therefore, two sets of regressions will be run, one with the complete Flow construct and one with only the Challenge-Skill Balance dimension.

Dimension	Understanding	p	Signif.*
Challenge-Skill	0.30	.008	**
Clear Goals	0.21	.067	.
Unambiguous Feedback	0.12	.302	
Concentration	0.11	.327	
Paradox of Control	0.09	.410	
Loss of Self-Consciousness	0.01	.952	
Transformation of Time	-0.08	.481	
Autotelic Experience	0.13	.256	

*Signif. codes: 0 '∗ ∗ ∗' 0.001 '∗∗' 0.01 '∗' 0.05 '.' 0.1 ' ' 1

Table 5
Correlation Matrix of Flow score per dimension and understanding.

As stated above two sets of regressions will be conducted to analyze the effects of Flow/challenge-skill balance, spatial ability, and condition on learning outcome (PCA understanding). Spatial ability, and specifically mental rotation ability, has shown to differ across gender, in favor of males (Peters et al. 1995; Rizzo et al. 1998). Therefore, the difference in spatial ability score between males and females was analyzed to determine if the regression model should account for gender. An independent t-test was conducted to compare spatial ability scores for males an females. There was not a significant difference in the scores for males (M=0.39, SD=0.16) and females (M=0.41, SD=0.21), conditions; t(77.12)=-0.35, p=.725. Therefore, gender was not accounted for in both regression sets.

First, a multiple regression analysis was performed to examine if understanding score could be predicted based on Flow score, spatial ability and condition. The regression was performed stepwise, using backward selection. The model started of with Flow, Spatial ability and Condition as predictors and with condition and spatial ability as an interaction term. Before fitting the model, all values were centered and scaled. None of the steps of the regression model could significantly predict understanding score. The first step was the full regression model and included Flow score, condition, spatial ability score, and an interaction term between condition and spatial ability as predictors of understanding score, $F(4,74)=0.99$, $p=.418$, $R^2=0.05$. The second step excluded the interaction variable and used Flow score, condition, and spatial ability score as predictors, $F(3,75)=1.00$, $p=.397$, $R^2=0.04$. The third step used Flow score and spatial ability score as predictors of understanding score, $F(2,76)=1.48$, $p=.234$, $R^2=0.04$. The fourth and final step used only Flow score to predict understanding score, $F(1,77)=1.93$, $p=.169$, $R^2=0.02$. The first set of regression analyses are visualized in table 6.

Predictors	p	Signif.*	R^2	F	DF
Flow, Condition, SA*, Condition-SA Interaction	.418		0.05	0.99	4, 74
Flow, Condition, SA	.397		0.04	1.00	3, 75
Flow, SA	.234		0.04	1.48	2, 76
Flow	.169		0.02	1.93	1, 77

∗SA is Spatial Ability
*Signif. codes: 0 '∗∗∗' 0.001 '∗∗' 0.01 '∗' 0.05 '.' 0.1 ' ' 1

Table 6
First Regression Set.

A second multiple regression analysis was conducted to examine if understanding score could be predicted based on the Flow dimension: challenge-skill balance, in combination with spatial ability and condition. The first step was the full regression model and included challenge-skill balance, condition, spatial ability score and an interaction term between condition and spatial ability as predictors of understanding score. The model did not significantly predict understanding score, $F(4,74)=2.41$, $p=.057$, $R^2=0.12$. The second step used spatial ability score, challenge-skill balance and the interaction term as predictors. The model did significantly predict understanding score, $F(3,75)=2.77$, $p=.048$, $R^2=0.10$. The third step used spatial ability score and challenge-skill balance as predictors. The model significantly predicted understanding score, $F(2,76)=3.96$, $p=.023$, $R^2=0.09$. The fourth and final step used only challenge-skill balance to predict understanding score. The model significantly predicted understanding score,

$F(1,77)$=7.48, p=.008, R^2=0.09.The second set of regression analyses are visualized in table 7. [1]

Predictors	p	Signif*	R^2	F	DF
CS*, Condition, SA, Condition-SA Interaction	.057	.	0.12	2.41	4, 74
CS, Condition, SA	.048	*	0.10	2.77	3, 75
CS, SA	.023	*	0.09	3.96	2, 76
CS	.008	**	0.09	7.48	1, 77

∗CS is Challenge-Skill balance
*Signif. codes: 0 '∗∗∗' 0.001 '∗∗' 0.01 '∗' 0.05 '.' 0.1 ' ' 1

Table 7
Second Regression Set.

5. Discussion

The aim of this study was to investigate the effect of learning in VR on learning outcome when learning about Principal Component Analysis, a concept that requires visual thinking to understand. An experiment was conducted where participants either learned about PCA from a paper explanation or from a VR application, which was designed for this study. This section will discuss the findings of this study, as presented in the results

The main research question of this study was to investigate the effect of learning in virtual reality on understanding of Principal Component Analysis. Previous research has shown that learning in VR can have a positive effect on learning outcome (Psotka 1995; Mikropoulos and Natsis 2011; Winn et al. 1999; Merchant et al. 2014) and that this effect also occurs when learning about concepts that require visual thinking (Rizzo et al. 1998; Lee and Wong 2014). Based on these findings, a hypothesis for the main research question was formulated: *Participants who learn about PCA in VR will score significantly better on understanding than participants learning about PCA on paper*. The first analysis, an independent-test, showed that although understanding score was higher in the VR condition, the difference was not significant. The findings do not support the first hypothesis. This finding contradicts the literature as presented above. Most participants already had some experience with VR and concentration score indicates most participants did not have a problem focusing on the information, so there are no indications that the environment might have been to distracting. A possible explanation of the findings is cognitive overloading. Whitelock et al. (2000) concluded that the feeling of 'being there' in a virtual world works motivating, yet takes a toll on working memory and can cause cognitive overload. In addition, the participants were presented a lot of new information in a relatively short time, which could have cause cognitive overload.

1 There were two participants that indicated to have prior knowledge of PCA. An additional regression model was made where these participants were excluded to test if this affected the outcome. However, excluding the participants had no significant effect on the model.

A second t-test was conducted to analyze if participants in the VR condition experienced higher levels of Flow than those in the paper condition. The t-test showed that although the average Flow score was higher in the VR condition, the difference was not significant. This implies that learning in VR does not result in higher levels of experienced Flow compared to learning from paper. This finding is in contradiction with the study by Ibáñez et al. (2014). A possible explanation for these findings lies in the conditions for experiencing Flow. Flow is state of mind that occurs when people are completely involved in and focused on the activity they are undergoing. Csikszentmihalyi and Larson (2014, p. 232) states that there are three key conditions for Flow experience: clear goals, a balance between challenge and skill, and the presence of clear and immediate feedback. In the experiment, these three condition did not vary across condition. All participants got a clear goal, to understand the concept of PCA and make a test about it. The challenge of the task was the same same in both conditions, since they were presented the same information. Additionally, the level of feedback participants got was the same across conditions. In conclusion, the conditions for inducing Flow were similar in the experimental conditions which can explain why participants' Flow score was not significantly different between conditions.

Correlation analysis was performed on Flow, spatial ability and understanding. Overall, both spatial ability and Flow had a weak positive correlation with understanding score. The results showed that spatial ability had a weak positive correlation with understanding in the paper condition and an almost nonexistent correlation with understanding in the VR condition. To give further insight in this result, the participants were split into high and low spatial ability groups and conditions were compared. The group-sizes were extremely small. Therefore, no statistical conclusions can be made. However, it is interesting to note that in the paper condition participants with high spatial ability scored higher on understanding than those with low spatial ability, yet in the VR the opposite is observed. This suggests that while people with high spatial ability naturally understand visually complex concepts better, people with low spatial ability are more positively affected by learning in VR. The latter is in line with the studies by Jang et al. (2017) and Lee and Wong (2014).

The correlation analysis showed that in the paper condition Flow had a moderate positive correlation with understanding score, yet in the VR condition Flow had a weak negative correlation with understanding score. This implies that Flow as a construct might not have a fixed effect on learning outcome, since the Flow score did not differ much across conditions. To further analyze the Flow construct, another correlation analysis was conducted between each dimension of Flow and understanding score. The results showed that only the dimension challenge-skill balance (moderately) correlated with understanding score. This result reinforces the suggestion that Flow did not significantly contribute to learning outcome. The results from the correlation analyses were further explored with a set of regression analyses.

The second research question aimed to measure the effect of flow on the understanding of Principal Component Analysis. Based on the literature, a second hypothesis was formulated: *A higher level of Flow will result in significantly higher learning outcome*. The results from the first set of regression analyses did not support the hypothesis. This is contradictory to previous research, which showed that being 'in the flow' can have positive effects on learning outcome (Choi, Kim, and Kim 2007; Custodero 2002; Webster, Trevino, and Ryan 1993). A possible explanation why no effect is found could lie in the understanding test. Studies have suggested multiple choice questions might not properly asses the effect of flow on learning (Guo et al. 2007; Rossin et al. 2009). It is possible that the performance of the current study, open-ended questions, has not been capable of properly measuring the effects of flow on learning outcome.

The results of the second regression analysis did show three models that could significantly predict learning outcome, using spatial ability, challenge-skill balance and an interaction term between spatial ability and condition. However, the models explained only ten percent of the variance at best. Furthermore, spatial ability and the interaction term only account for around one percent. Nine percent of the total explained variance was accounted for by challenge-skill balance alone. Therefore, the models including spatial ability and the interaction term were, although significant, not very insightful. In addition, the challenge-skill balance dimension measured the extent to which people found their own skills to be in balance with the challenge. Therefore, it is not surprising that the regression model is able to predict a lower understanding score for participants that indicated that the task was too difficult for their skill, and the other way around.

The regression analysis indicated that condition was unable to significantly predict learning outcome. This finding corresponds to the results from the t-test. Therefore, it can be concluded that no effects of learning in VR on learning outcome were found in this study. As stated above, one explanation for this finding could be cognitive overload. In addition, the average understanding score in both conditions is quite low, at around four and a half out of ten points. This indicates that the test measuring understanding was too difficult for participants.

The findings of this study have not found a significant effect of learning in VR on learning outcome, nor for experienced Flow on learning outcome. However, it should be stated that many of the participants indicated they enjoyed learning in the virtual environment. The scores for the autotelic dimension of Flow, which measures if a situation is intrinsically rewarding, confirm this observation. This indicates that although learning in the VR condition does not have a significant effect on learning outcome, participants did have more learning satisfaction.

As stated above, a limitation of this study was the difficulty of the test that measured understanding of PCA. Participants in both conditions scored low on the test, which makes it more difficult to differentiate between the groups. Creating a test which is more in balance with the level of understanding could provide a more insightful measure to compare groups and could results in significant results. Another possible limitation was the amount of information participants had to process in a relatively short time, which might have caused cognitive overload. This effect is more likely to have occurred in the VR condition, since being in the virtual environment takes its toll on working memory and can cause cognitive overload. Decreasing the amount of information participants need to process or increasing the amount of time participants have to process the information could result in different learning outcomes.

The current study has investigated the effect of learning visually complex concepts in virtual reality on learning outcome. This study contributes to existing research on VR learning by focusing on the understanding of concepts that require visual thinking. No effects of learning in VR on learning outcome were found. In addition, the effect of Flow on learning outcome was studied. No effects of Flow on learning outcome were found. Participants did enjoy learning in VR more than learning on paper and the virtual environment was not found to be distracting. Cognitive overload as a result of the amount of information to process and the impressions of the virtual environment could explain why no effects were found by this study. Furthermore, the test on understanding proved to be too difficult. Decreasing the amount of information given in the explanation and developing a test that better matches the acquired knowledge could alter the results and provide different findings.

6. Conclusion

The current research has tried to answer the following research questions: 1. *What is the effect of learning in virtual reality on understanding of Principal Component Analysis?*, and 2. *What is the effect of Flow on the understanding of Principal Component Analysis?* To answer these questions, an experiment was conducted where participant either learned about PCA on paper or in a virtual environment. This study, in contrast to previous research, did not find any effects of learning in virtual reality on understanding. However, participants indicated they enjoyed learning in VR more than learning on paper. There were two possible limitations to the study that could explain why no effects were found. Participants might have experienced cognitive overload, due to the amount of information they needed to process. In addition, the test that measured understanding proved to be too difficult. This study did not find any effects of Flow on the understanding of PCA, which contradicts previous research. There are no clear indications as to why no effects were found.

This study investigated the effect of learning about a visually complex concept in virtual reality on learning outcome. A first step was made to supplement current research on VR learning and to create a concrete way for making visually complex concepts understandable for people that would otherwise struggle with this. Future research could use the limitations of this study to investigate if removing these would have an effect on learning outcome.

References

Brown, John Seely, Allan Collins, and Paul Duguid. 1989. Situated cognition and the culture of learning. *Educational researcher*, 18(1):32–42.

Choi, Duke Hyun, Jeoungkun Kim, and Soung Hie Kim. 2007. Erp training with a web-based electronic learning system: The flow theory perspective. *International Journal of Human-Computer Studies*, 65(3):223–243.

Csikszentmihalyi, Mihalyi and R Larson. 2014. Flow and the foundations of positive psychology. *Recuperado de http://www. springer. com/us/book/9789401790871*.

Custodero, Lori A. 2002. Seeking challenge, finding skill: Flow experience and music education. *Arts education policy review*, 103(3):3–9.

Dünser, Andreas, Karin Steinbügl, Hannes Kaufmann, and Judith Glück. 2006. Virtual and augmented reality as spatial ability training tools. In *Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI*, pages 125–132, ACM.

Ekstrom, Ruth B, Diran Dermen, and Harry Horace Harman. 1976. *Manual for kit of factor-referenced cognitive tests*, volume 102. Educational testing service Princeton, NJ.

George, M. Dallas. 2013. Principal component analysis 4 dummies: Eigenvectors, eigenvalues and dimension reduction.

Ghani, Jawaid A and Satish P Deshpande. 1994. Task characteristics and the experience of optimal flow in humanâĂŤcomputer interaction. *The Journal of psychology*, 128(4):381–391.

Guo, Yi, Barbara Klein, Young Ro, and Donald Rossin. 2007. The impact of flow on learning outcomes in a graduate-level information management course. *Journal of Global Business Issues*, 1(2):31–39.

Hein, George. 1991. Constructivist learning theory. *Institute for Inquiry. Available at:/http://www. exploratorium. edu/ifi/resources/constructivistlearning. htmlS*.

Huk, Thomas. 2006. Who benefits from learning with 3d models? the case of spatial ability. *Journal of computer assisted learning*, 22(6):392–404.

Ibáñez, María Blanca, Ángela Di Serio, Diego Villarán, and Carlos Delgado Kloos. 2014. Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education*, 71:1–13.

J., Xinzhi Li. 2009. A layman's introduction to principal component analysis.

Jackson, Susan A and Herbert W Marsh. 1996. Development and validation of a scale to measure optimal experience: The flow state scale. *Journal of sport and exercise psychology*, 18(1):17–35.

Jang, Susan, Jonathan M Vitale, Robert W Jyung, and John B Black. 2017. Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computers & Education*, 106:150–165.

Lee, Elinda Ai-Lim and Kok Wai Wong. 2014. Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education*, 79:49–58.

Lufler, Rebecca S, Ann C Zumwalt, Carla A Romney, and Todd M Hoagland. 2012. Effect of visual–spatial ability on medical students' performance in a gross anatomy course. *Anatomical sciences education*, 5(1):3–9.

Merchant, Zahira, Ernest T Goetz, Lauren Cifuentes, Wendy Keeney-Kennicutt, and Trina J Davis. 2014. Effectiveness of virtual reality-based instruction on students' learning outcomes in k-12 and higher education: A meta-analysis. *Computers & Education*, 70:29–40.

Mikropoulos, Tassos A. 2006. Presence: a unique characteristic in educational virtual environments. *Virtual Reality*, 10(3-4):197–206.

Mikropoulos, Tassos A and Antonis Natsis. 2011. Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education*, 56(3):769–780.

Peters, Michael, Bruno Laeng, Kerry Latham, Marla Jackson, Raghad Zaiyouna, and Chris Richardson. 1995. A redrawn vandenberg and kuse mental rotations test-different versions and factors that affect performance. *Brain and cognition*, 28(1):39–58.

Psotka, Joseph. 1995. Immersive training systems: Virtual reality and education and training. *Instructional science*, 23(5-6):405–431.

Reid, Denise. 2002. Virtual reality and the person–environment experience. *CyberPsychology & Behavior*, 5(6):559–564.

Rizzo, Albert A, J Galen Buckwalter, Ulrich Neumann, Carl Kesselman, Marcus Thiébaux, Peter Larson, and Andre van Rooyen. 1998. The virtual reality mental rotation spatial skills project. *CyberPsychology & Behavior*, 1(2):113–119.

Rossin, Don, Young K Ro, Barbara D Klein, and Yi Maggie Guo. 2009. The effects of flow on learning outcomes in an online information management course. *Journal of Information Systems Education*, 20(1).

Skadberg, Yongxia Xia and James R Kimmel. 2004. VisitorsâĂŹ flow experience while browsing a web site: its measurement, contributing factors and consequences. *Computers in human behavior*, 20(3):403–422.

Sutherland, Ivan E. 1968. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pages 757–764, ACM.

Wanzel, Kyle R, Stanley J Hamstra, Dimitri J Anastakis, Edward D Matsumoto, and Michael D Cusimano. 2002. Effect of visual-spatial ability on learning of spatially-complex surgical skills. *The lancet*, 359(9302):230–231.

Webster, Jane, Linda Klebe Trevino, and Lisa Ryan. 1993. The dimensionality and correlates of flow in human-computer interactions. *Computers in human behavior*, 9(4):411–426.

Whitelock, Denise, Daniela Romano, Anne Jelfs, and Paul Brna. 2000. Perfect presence: What does this mean for the design of virtual learning environments? *Education and information technologies*, 5(4):277–289.

Winn, William. 1993. A conceptual basis for educational applications of virtual reality. *Technical Publication R-93-9, Human Interface Technology Laboratory of the Washington Technology Center, Seattle: University of Washington*.

Winn, William, Hunter Hoffman, Ari Hollander, Kimberley Osberg, Howard Rose, and Patti Char. 1999. Student-built virtual environments. *Presence: Teleoperators & virtual environments*, 8(3):283–292.

Appendices

A. PCA Paper Explanation

Imagine a scatter-plot of a large set of data, containing the sepal and petal length and width of 150 flowers, divided over 3 species. In this scatter-plot, each dot represents one flower. The dimensions of your scatter-plot are equal to the amount of features you want represented in it. For example, in the scatter-plot in figure 1, the data is plotted with sepal width and length as the two features. It is consequently 2-dimensional.

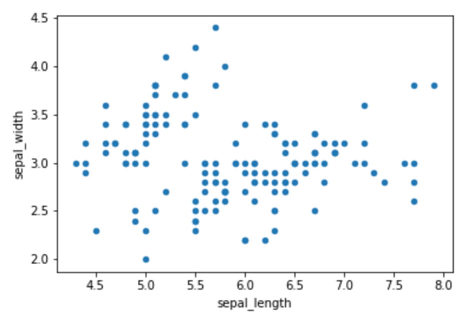

Figure 6

Now imagine you want to take another feature, such as petal length or petal width into consideration. In order to do this, you will have to add a dimension to your scatter-plot, which will make the scatter-plot go from 2-dimensional to 3-dimensional, see figure 2.

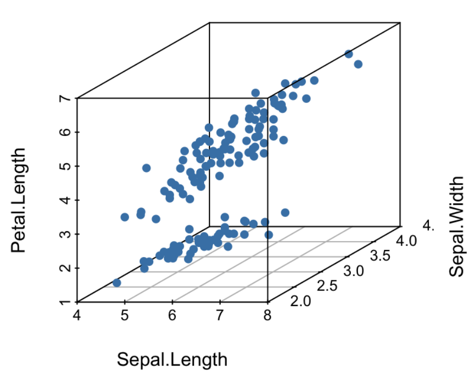

Figure 7

Suddenly your data becomes a whole lot less easy to interpret. Data points are in front and on top of each other, and their position is not very clear. So, is there a way we can get a more clear view of our data, with minimal information loss?

Enter Principal Component Analysis (PCA). Principal components are the underlying structure in our data, and they will help us perform dimensionality reduction with minimal information loss. This might seem a little confusing, so let's explain.

Imagine our iris data-set only had eight entries, that look like figure 3 when they are plotted in a 2-dimensional graph, with sepal length as the x-axis and sepal width as the y-axis.

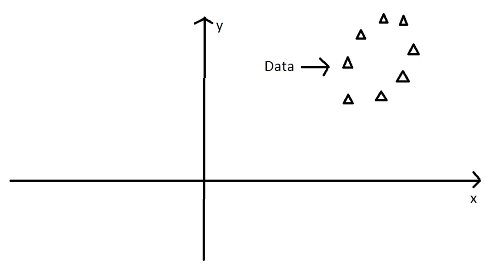

Figure 8

Principal components are lines that are drawn through our data, along which the variance of our data is greatest. The variance is a mathematical term for how spread out the data points are when they are projected on the line you have drawn. Each line has its own corresponding amount of variance. Such a line is called a eigenvector, while the amount of variance corresponding with the line is called the eigenvalue.

In order to find the first principal component, we need to keep drawing these lines (eigenvectors) in our data until we find an eigenvector that has the highest corresponding variance (eigenvalues). Note that not every eigenvector in a data-set is automatically a principal component; only the eigenvectors with high enough eigenvalues to be interesting (so the lines along which most information is found) are the principal components of the data.

The amount of eigenvectors/values found in a data-set is always the same as the amount of dimensions: 2-dimensional data has 2 eigenvectors/values, while 3-dimensional data has 3 eigenvectors/values etc. To keep it short, let's assume we have already found our 2 eigenvectors for our smaller iris data-set example.

Let's draw our first eigenvector in blue (figure 4).

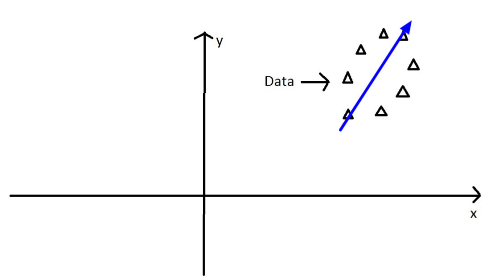

Figure 9

Now let's draw the second eigenvector, which is the eigenvector that is perpendicular to our first, and that has the highest eigenvalue. We'll draw the line in red (figure 5).

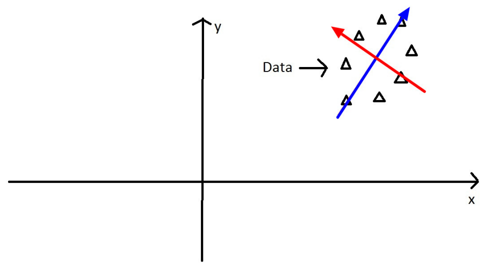

Figure 10

Let's say we have also calculated the eigenvalue of each of these eigenvectors, and both were high enough to be considered interesting: we can call these eigenvectors the Principal Components of our data. They represent the underlying structure in the data. Using these principal components, we we can create a new graph. By taking our principal components and rotating them, we can create a new x- and y-axis and plot our data points around those axes. The following graph (figure 6) would be the result.

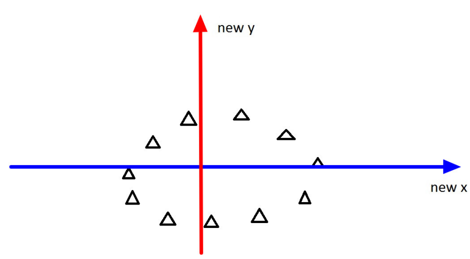

Figure 11

As you can see, the principal components have just been turned so they become the new axes, and the data points have turned with it. The data itself has not been changed, we just changed our point of view of it. The relative distance between the data points remains the same as before.

Mind you, these new axes do not mean anything. Where previously the axes were very clearly defined (the x-axis was sepal length and the y-axis was sepal width), these new axes do not have such an explicit meaning. There's often a good reason why these axes represent the data better, but that is for you to work out.

So how does this help our problem of making our 3-dimensional flower graph become more readable? Well, PCA can be used to perform so-called 'dimensionality reduction' in data. Dimensionality reduction is the process of removing dimensions from the data, so it becomes less complex. Imagine you would plot the data points of the example in a graph with 3 dimensions, namely sepal length, sepal width and petal length. Since we've added an extra dimension, there's also a third eigenvector to be found apart from the two we already found in our 2-dimensional example. We find this principal component by finding the eigenvector with the highest eigenvalue, that is perpendicular to both the first and second principal component. The first two eigenvectors are drawn in the same color as in our 2-dimensional example, the third is drawn in green, see figure 7.

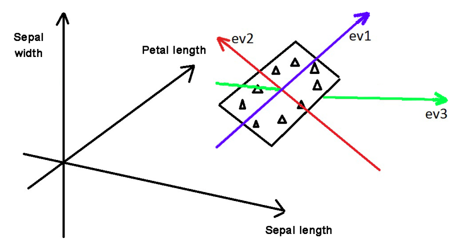

Figure 12

Let's assume the eigenvalue of EV1 is the biggest of the three, the eigenvalue of EV2 is a non-zero value and the eigenvalue of EV3 is 0. An eigenvalue of 0 tells us that eigenvector is meaningless; it does not provide us with any information about the data. We can consequently ignore EV3 when we are looking for principal components. After all, a principal component is an eigenvector with a high eigenvalue, and EV3 has an eigenvalue of 0. This leaves us with only EV1 and EV2. We can use those to plot our data in 2 dimensions again, see figure 8.

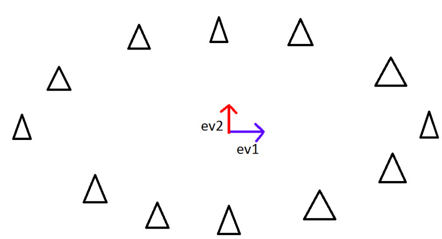

Figure 13

Our data has suddenly become a whole lot easier to interpret now that it has gone from 3-dimensional to 2-dimensional , don't you agree? This process of removing dimensions is called dimensionality reduction, and is the main reason why PCA is such a neat trick. PCA can be performed on data with relatively few dimensions, like our example. However, where PCA truly shines is when it is used for data sets with many more dimensions, like in the following example:

In a report, 2000 people were asked questions about their internet use. Let's say every person was asked 50 questions. 50 questions will equal 50 variables in the data, and will consequently give our data 50 dimensions. Let's say the eigenvalues in the data were (in descending order): 50, 29, 17, 10, 2, 1, 1, 0.4, 0.2 etc. Only 4 of these eigenvalues are large enough to consider, since their eigenvalue indicates a lot of information is found in that direction. Using PCA, the data can consequently be reduced from 50 dimensions to only 4 (which is a whole lot easier to work with). In other words, PCA helped us simplify the data-set by finding the dominant dimensions within it.

B. Flow State Scale

1. Challenge-skill balance
    - I was challenged, but I believed my skills would allow me to meet the challenge.
    - My abilities matched the challenge of the situation.
    - I felt I was competent enough to meet the demands of the situation.
    - The challenge and my skills were at an equally high level.
2. Clear goals
    - I knew clearly what I wanted to do.
    - I had a strong sense of what I wanted to do.
    - I knew what I wanted to achieve.
    - My goals were clearly defined.
3. Unambiguous feedback
    - It was really clear to me that I was doing well.
    - I was aware of how well I was performing.
    - I had a good idea while I was performing about how well I was doing.
    - I could tell by the way I was performing how well I was doing.
4. Concentration on task at hand
    - My attention was focused entirely on what I was doing.
    - It was no effort to keep my mind on what was happening.
    - I had total concentration.
    - I was completely focused on the task at hand.
5. Sense of control
    - I felt in total control of what I was doing.
    - I felt like I could control what I was doing.
    - I had a feeling of total control.
    - I felt in total control of my body.
6. Loss of self-consciousness
    - I was not concerned with what others may have been thinking of me.

- I was not worried about my performance during the event.
- I was not concerned with how I was presenting myself.
- I was not worried about what others may have been thinking of me.

7. Transformation of time

- Time seemed to alter (either slowed down or sped up).
- The way time passed seemed to be different from normal.
- It felt like time stopped while I was performing.
- At times, it almost seemed like things were happening in slow motion.

8. Autotelic experience

- I really enjoyed the experience.
- I loved the feeling of that performance and want to capture it again.
- The experience left me feeling great.
- I found the experience extremely rewarding.

C. Flow Question Scores

Statements	Paper		VR	
	M	*SD*	*M*	*SD*
1. I was challenged, but I believed my skills would allow me to meet the challenge.	3.67	0.90	3.6	0.78
2. I knew clearly what I wanted to do.	3.74	0.88	3.80	0.99
3. It was really clear to me that I was doing well.	3.15	1.04	3.41	0.85
4. My attention was focused entirely on what I was doing.	4.1	0.99	3.85	1.01
5. I felt in total control of what I was doing.	3.82	1.00	3.85	1.00
6. I was not concerned with what others may have been thinking of me.	3.54	1.41	3.65	1.42
7. Time seemed to alter (either slowed down or sped up).	3.38	0.85	3.2	0.99
8. I really enjoyed the experience.	3.62	1.09	4.25	0.74
9. My abilities matched the challenge of the situation.	3.31	1.03	3.73	0.88
10. I had a strong sense of what I wanted to do.	3.59	0.97	3.45	1.04
11. I was aware of how well I was performing.	3.13	1.08	3.13	0.99
12. It was no effort to keep my mind on what was happening.	3.49	1.32	3.70	1.11
13. I felt like I could control what I was doing.	3.71	1.04	3.97	0.84
14. I was not worried about my performance during the event.	3.21	1.13	3.58	1.08
15. The way time passed seemed to be different from normal.	3.21	0.83	3.25	1.08
16. I loved the feeling of that performance and want to capture it again.	2.69	0.98	3.48	1.09
17. I felt I was competent enough to meet the demands of the situation.	3.51	1.14	3.82	0.79
18. I knew what I wanted to achieve.	3.82	1.00	3.48	1.06
19. I had a good idea while I was performing about how well I was doing.	3.28	1.12	3.03	1.00
20. I had total concentration.	3.85	1.16	3.45	1.15
21. I had a feeling of total control.	3.38	1.09	3.4	1.03
22. I was not concerned with how I was presenting myself.	3.49	1.30	3.6	1.13
23. It felt like time stopped while I was performing.	2.56	1.02	2.95	1.08
24. The experience left me feeling great.	3.00	1.05	3.48	0.88
25. The challenge and my skills were at an equally high level.	3.15	0.99	3.38	0.91
26. My goals were clearly defined.	3.87	0.83	3.45	1.13
27. I could tell by the way I was performing how well I was doing.	3.18	1.06	3.00	0.96
28. I was completely focused on the task at hand.	4.08	1.04	3.9	0.67
29. I felt in total control of my body.	4.05	0.86	3.65	1.10
30. I was not worried about what others may have been thinking of me.	3.87	1.13	3.68	1.23
31. At times, it almost seemed like things were happening in slow motion.	2.26	0.88	2.98	1.10
32. I found the experience extremely rewarding.	2.97	1.22	3.13	1.11

Table 8
Flow State Scale Questionnaire

D. Flow Score Distribution per Dimension and Condition

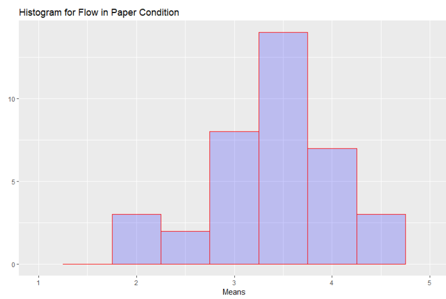

Figure 14
Histogram of Flow in paper condition

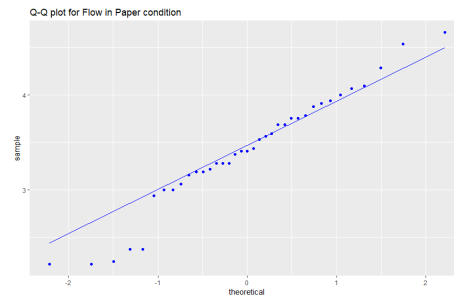

Figure 15
Q-Q plot of Flow in paper condition

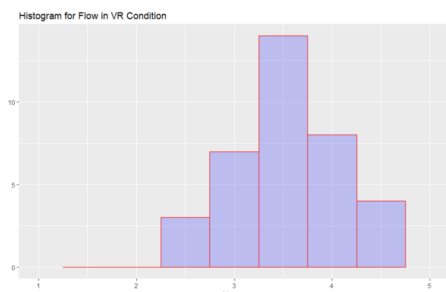

Figure 16
Histogram of Flow in VR condition

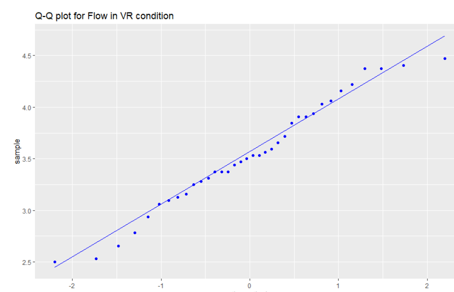

Figure 17
Q-Q plot of Flow in VR condition

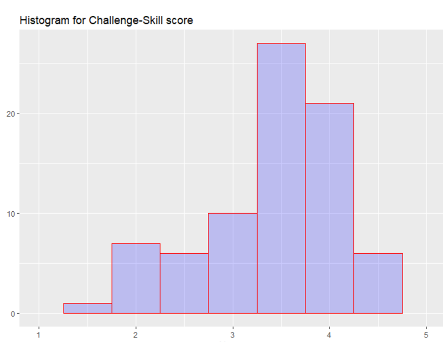

Figure 18
Histogram of Challenge-Skill Balance

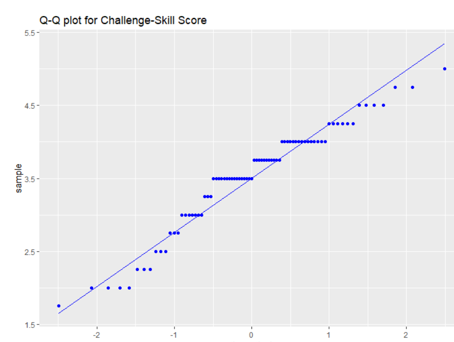

Figure 19
Q-Q plot of Challenge-Skill Balance

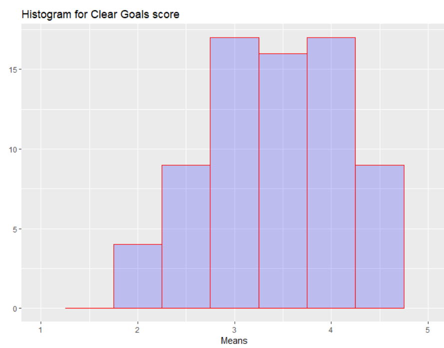

Figure 20
Histogram of Cler Goals

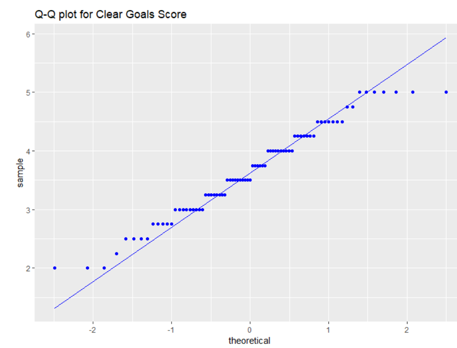

Figure 21
Q-Q plot of Clear Goals

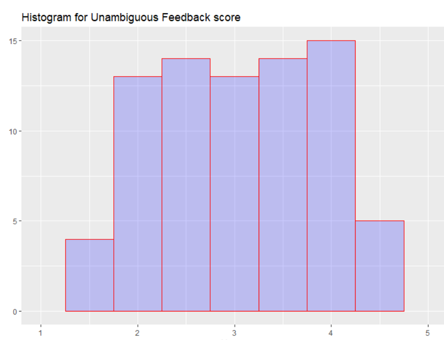

Figure 22
Histogram of Unambiguous Feedback

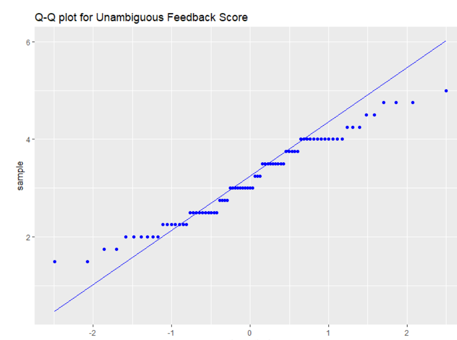

Figure 23
Q-Q plot of Unambiguous Feedback

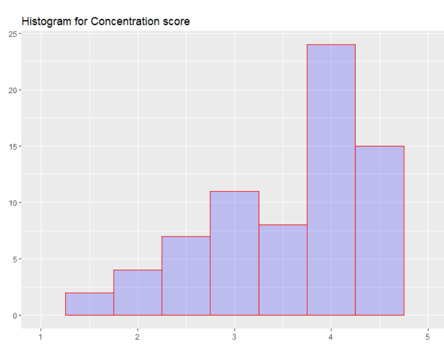

Figure 24
Histogram of Concentration

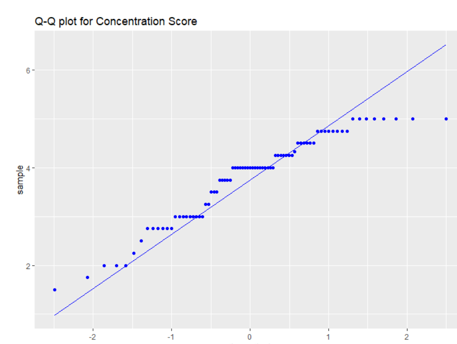

Figure 25
Q-Q plot of Concentration

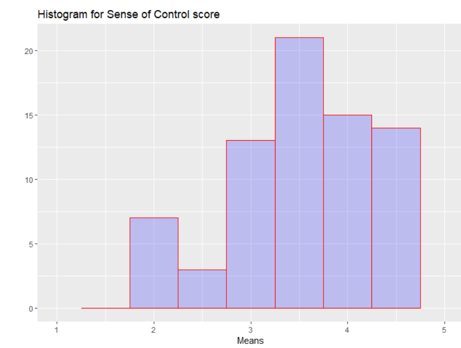

Figure 26
Histogram of Sense of Control

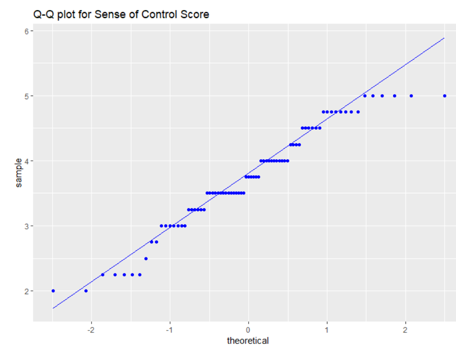

Figure 27
Q-Q plot of Sense of Control

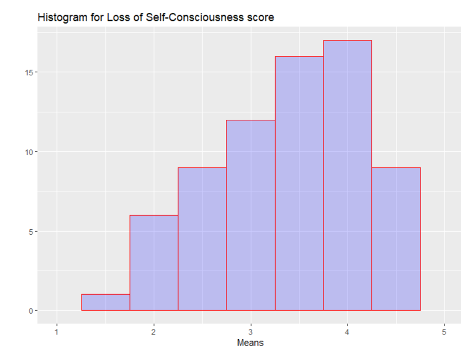

Figure 28
Histogram of Loss of Self-Consciousness

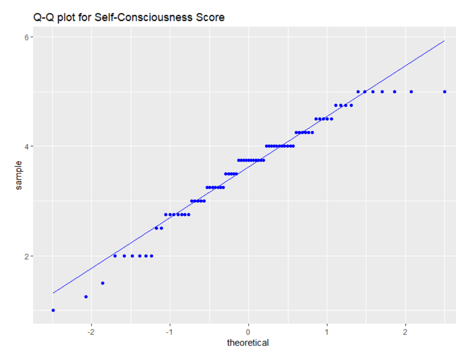

Figure 29
Q-Q plot of Loss of Self-Consciousness

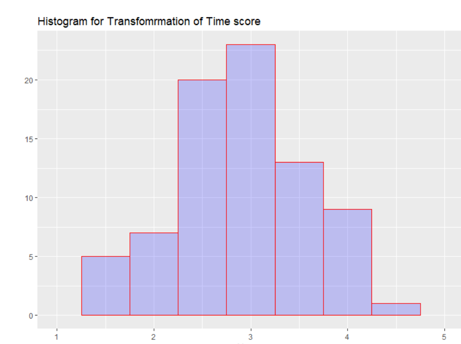

Figure 30
Histogram of Transformation of Time

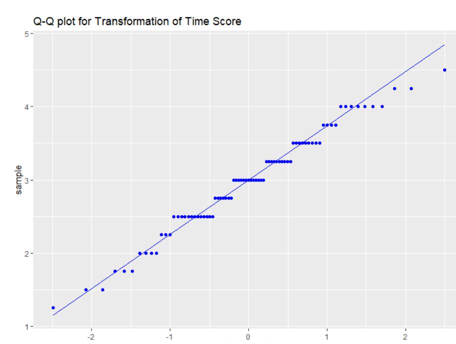

Figure 31
Q-Q plot of Transformation of Time

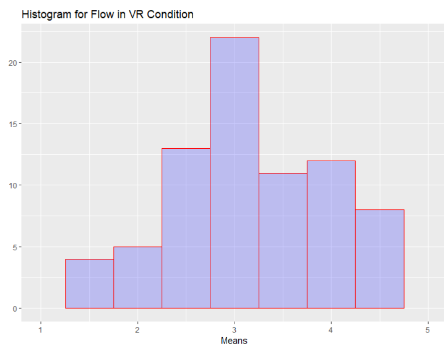

Figure 32
Histogram of Autotelic Experience

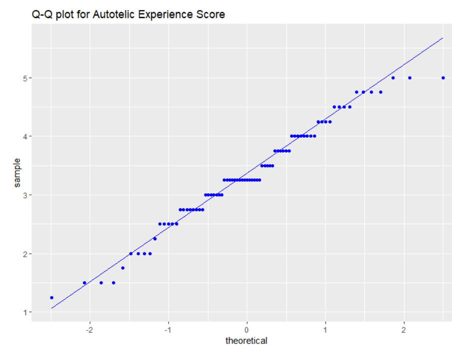

Figure 33
Q-Q plot of Autotelic Experience

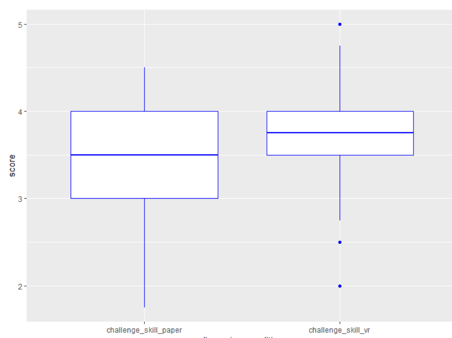

Figure 34
Box-plot of Challenge-Skill Balance

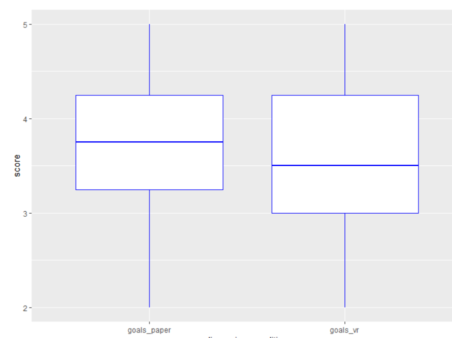

Figure 35
Box-plot of Clear Goals

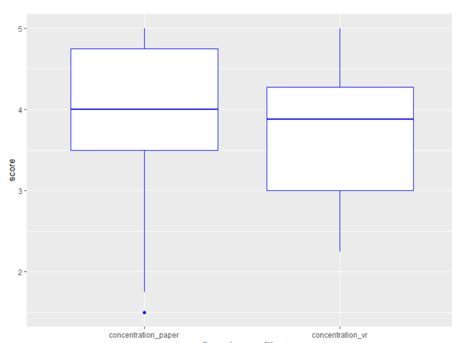

Figure 36
Box-plot of Concentration

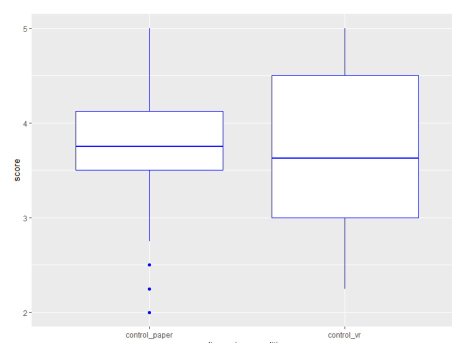

Figure 37
Box-plot of Sense of Control

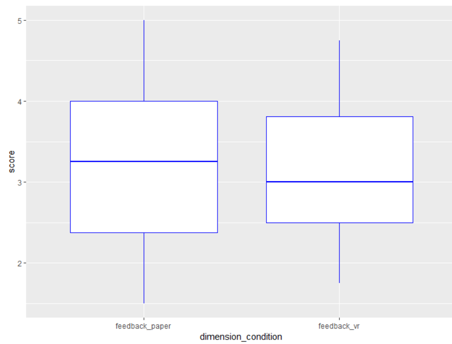

Figure 38
Box-plot of Unambiguous Feedback

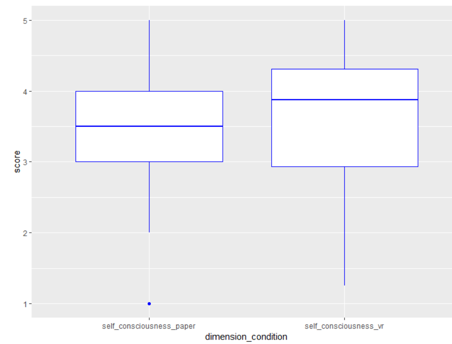

Figure 39
Box-plot of Loss of Self-Consciousness

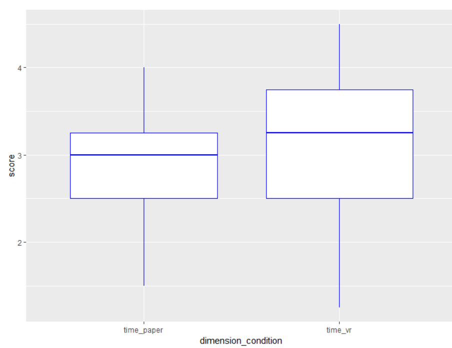

Figure 40
Box-plot of Transformation of Time

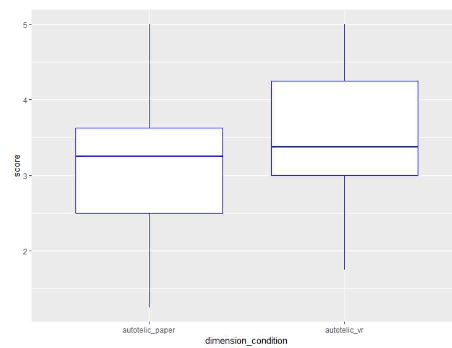

Figure 41
Box-plot of Autotelic Experience

E. PCA Understanding Test

1. What is the function of an eigenvalue?
2. What is the most widely used application of Principal Component Analysis? Explain the application.
3. What does a Principal Component of a data-set represent?
4. See Figure 1. Study the data, the red line is the first principal component. Based on what you learned, which line would logically be the second principal component? The green, yellow or blue line? Why?

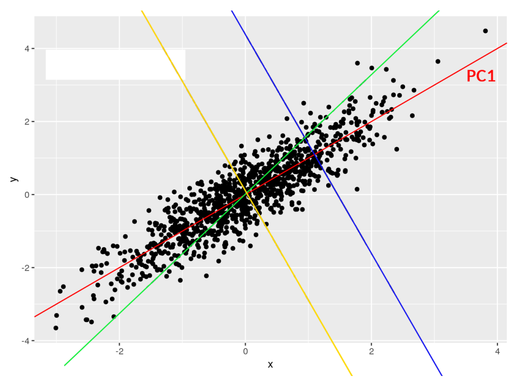

Figure 42
Question 4

5. How do we find the first principal component in a data-set?
6. What is the variance of an eigenvector?
7. See figure 2. Study the plot. How many Principal components would we have to find for this data-set? Why?

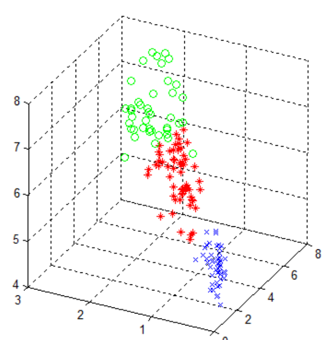

Figure 43
Question 7

8. See figure 3. Based on what you learned about principle components, which of the three principal components in the figure below would you expect to have the highest eigenvalue and why?

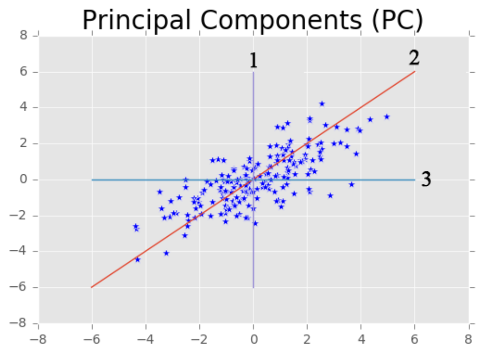

Figure 44
Question 8

F. PCA understanding Distribution Overall and Per Condition

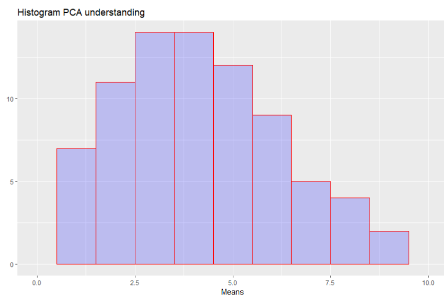

Figure 45
Histogram of PCA understanding

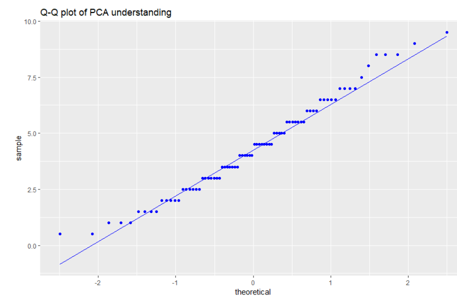

Figure 46
Q-Q plot of PCA understanding.

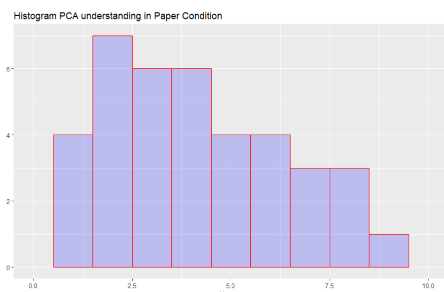

Figure 47
Histogram of PCA understanding in paper condition

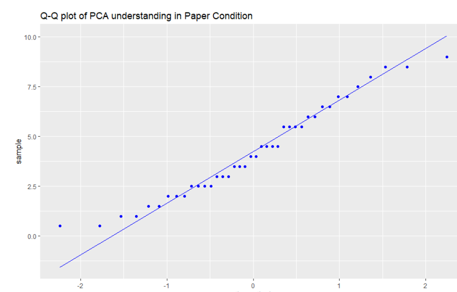

Figure 48
Q-Q plot of PCA understanding in paper condition

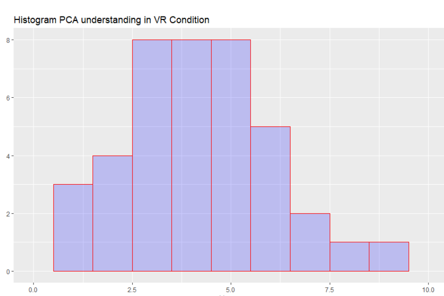

Figure 49
Histogram of PCA understanding in VR condition

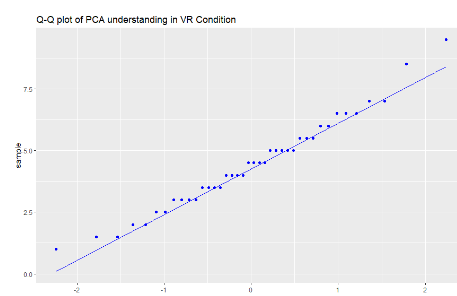

Figure 50
Q-Q plot of PCA understanding in VR condition

G. Spatial Ability Score Distribution

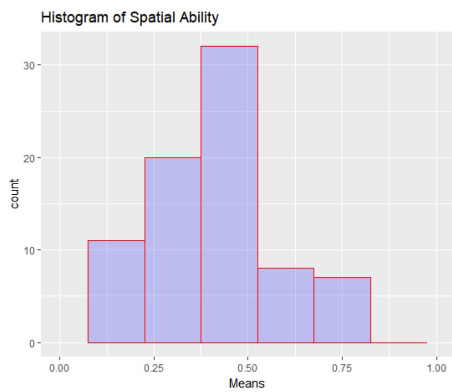

Figure 51
Histogram of Spatial Ability

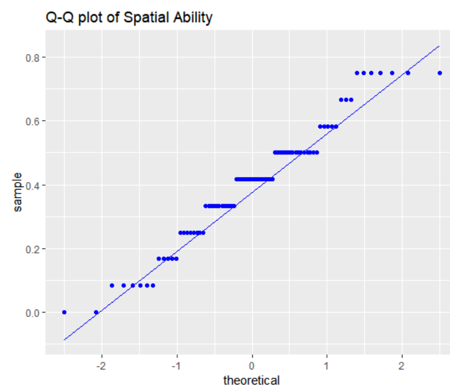

Figure 52
Q-Q plot of Spatial Ability.

H. PCA Flow Questionnaire

Question number	Component 1	Component 2
1	-0.11	0.02
2	-0.17	-0.20
3	-0.22	-0.21
4	-0.11	0.16
5	-0.23	0.09
6	-0.18	0.03
7	-0.01	0.41
8	-0.21	0.08
9	-0.21	-0.14
10	-0.20	-0.16
11	-0.20	-0.24
12	-0.19	0.23
13	-0.21	0.14
14	-0.18	0.10
15	-0.06	0.29
16	-0.17	0.11
17	-0.24	-0.12
18	-0.15	-0.19
20	-0.17	0.24
21	-0.24	0.11
22	-0.14	0.12
23	-0.10	0.34
24	-0.24	0.10
25	-0.18	-0.02
26	-0.16	-0.17
27	-0.18	-0.18
28	-0.19	0.12
29	-0.15	-0.00
30	-0.16	0.01
31	-0.02	0.16
32	-0.17	-0.01

Table 9
PCA flow questionnaire.