



# Motivational Effects of Virtual Reality in Music Education: A Quasi-Experimental Study

Received: February 27, 2024  
Reviewed: July 23, 2024  
Published: July 1, 2026

Pedro Iglesias\*    
Pablo Chávez\*\*  

**Article type:** Review article

## Abstract

The study examines the motivational effects observed in a pedagogical intervention in the context of a school music class, with participants aged 14 to 16 years old, in which virtual reality (VR) was used as a didactic resource. The effects of different audiovisual projection methods —immersive VR, desktop VR, and traditional projection— were compared using a quasi-experimental design with a mixed-method approach. Data from a questionnaire operationalizing motivation, learning outcomes, and a focus group were triangulated.

Results indicate that VR increased participants' motivation, with statistically significant differences and a large effect size. A positive correlation was found between some dimensions of motivation and learning outcomes. These findings diverge from the predictions of current cognitive theories, such as Multimedia Learning Theory, which suggest that VR might increase cognitive load and limit learning. It is argued that VR, by incorporating the kinesthetic perceptual channel through movements and haptic interfaces, can enhance learning in a manner not considered by this theory.

## Keywords

educational technology; virtual reality; multimedia instruction; music education; cognitive load

---

\* Instituto de Música, Universidad Alberto Hurtado, Chile. [ipedro@uahurtado.cl](mailto:ipedro@uahurtado.cl)  
\*\* Universidad Mayor, Escuela Educación, Chile. [pablo.chavez@umayor.cl](mailto:pablo.chavez@umayor.cl)

## Efeitos motivacionais da realidade virtual no ensino de música: um estudo quase-experimental

### Resumo

O estudo examina os efeitos motivacionais observados em uma intervenção pedagógica no contexto de uma aula escolar de música, com participantes de 14 a 16 anos, na qual a realidade virtual (RV) foi utilizada como recurso didático. Os efeitos de diferentes métodos de projeção audiovisual — RV imersiva, RV em desktop e projeção tradicional — foram comparados por meio de um delineamento quase-experimental com abordagem de métodos mistos. Os dados provenientes de um questionário que operacionalizou motivação, resultados de aprendizagem e um grupo focal foram triangulados.

Os resultados indicam que a RV aumentou a motivação dos participantes, com diferenças estatisticamente significativas e tamanho de efeito elevado. Foi encontrada correlação positiva entre algumas dimensões da motivação e os resultados de aprendizagem. Esses achados divergem das previsões de teorias cognitivas atuais, como a Teoria da Aprendizagem Multimídia, que sugerem que a RV poderia aumentar a carga cognitiva e limitar a aprendizagem. Argumenta-se que a RV, ao incorporar o canal perceptivo cinestésico por meio de movimentos e interfaces hápticas, pode potencializar a aprendizagem de uma forma não considerada por essa teoria.

### Palavras-chave

tecnologia educacional; realidade virtual; instrução multimídia; educação musical; carga cognitiva

## Efectos motivacionales de la realidad virtual en la educación musical: un estudio cuasi-experimental

### Resumen

El estudio examina los efectos motivacionales observados en una intervención pedagógica en el contexto de una clase escolar de música, con participantes de entre 14 y 16 años, en la que se utilizó la realidad virtual (RV) como recurso didáctico. Se compararon los efectos de distintos métodos de proyección audiovisual — RV inmersiva, RV de escritorio y proyección tradicional — mediante un diseño cuasi-experimental con enfoque de métodos mixtos. Los datos obtenidos de un cuestionario que operacionalizó la motivación, los resultados de aprendizaje y un grupo focal fueron triangulados.

Los resultados indican que la RV aumentó la motivación de los participantes, con diferencias estadísticamente significativas y un tamaño de efecto elevado. Se encontró una correlación positiva entre algunas dimensiones de la motivación y los resultados de aprendizaje. Estos hallazgos divergen de las predicciones de teorías cognitivas actuales, como la Teoría del Aprendizaje Multimedial, que sugieren que la RV podría incrementar la carga cognitiva y limitar el aprendizaje. Se argumenta que la RV, al incorporar el canal perceptivo cinestésico mediante movimientos e interfaces hápticas, puede potenciar el aprendizaje de un modo no considerado por dicha teoría.

### Palabras clave

tecnología educativa; realidad virtual; instrucción multimedial; educación musical; carga cognitiva

### To cite this article:

Iglesias, P., & Chávez, P. (2026). Motivational Effects of Virtual Reality in Music Education: A Quasi-Experimental Study, *Revista Colombiana de Educación*, (100), e21167, <https://doi.org/10.17227/rce.num100-21167>

## Introducción

Virtual reality (hereinafter VR) devices are currently being widely used. Large technology companies promote them as an icon associated with high technology (Oyelere et al., 2020), where through the possibilities of immersion they promise the possibility of exploring new worlds and attempting seemingly unimaginable feats (Meta, 2021). VR, associated with the Metaverse concept, also promises to substantially transform education through the implementation of new instructional and communication methodologies (Clegg, 2023; Contreras et al., 2022). The combination of the efforts of technological giants and the decrease in hardware prices have generated high expectations, so that in the next few years VR could massively enter educational institutions (Bower et al., 2020; Shute et al., 2017).

The promise of technology companies seems attractive, however, questions arise about the relevance of its use in the formative process of digital natives. Can these new devices really lead a revolution in education or is it simply a strategy of large technology companies to promote their products? What instructional methods are best suited to these new devices? Is there content that is more appropriate to teach through VR? What cognitive theories support the promise of VR to transform education? Reflecting on these questions is crucial to avoid technology-centric implementations (Mayer, 2009; 2022). The history of the integration of technological devices in education has been extensive and has repeatedly promised to revolutionize learning processes. Devices that at the time were novel and attractive, such as cinema, radio or television, ultimately failed to meet expectations when they entered the classroom (Cuban, 1986, 2001). Mayer (2009) proposes the implementation of technologies through a student-centered approach. This implies first understanding human cognition and then exploring how technology can enhance learning.

For all these reasons, the promises made by large technology companies about the educational potential of VR should be viewed with some suspicion and subjected to rigorous evaluation and scrutiny.

## Cognitive Theories Supporting the Use of VR as a Learning Resource

VR provides the user a sensory immersion, presenting visual and auditory stimuli controlled by haptic and kinesthetic interfaces. This creates the illusion that the user is immersed in a real world (Dede et al., 2017). These immersive properties have led to VR being commonly used for simulations. Several authors have

highlighted the potential benefits that using VR could have in these learning situations. First, VR would allow the facilitation of experiences that would be impossible or impractical to try in the classroom (Bhattacharjee et al., 2018; Bower et al., 2020; Dede et al., 2017; Di Natale et al., 2020; Hamilton et al., 2020; Oyelere et al., 2020; Wu et al., 2020). Second, haptic control of the device, dynamic spatial representation, and user interaction and manipulation with different objects in real time could provide opportunities for instructional approaches linked to constructivism (Chen, 2009; Dalgarno & Lee, 2010).

Simulation as a learning tool can be supported by theories that are related to constructivism. This cognitive paradigm proposes that knowledge is constructed in the individual's mind from interaction with the environment through processes of adaptation and accommodation (Flavell et al., 2002; Piaget, 2000). This assumption gives rise to the construction of several learning theories, such as *discovery learning*, which proposes that learning occurs through stimuli with which individuals construct meanings that they integrate into their cognitive schemes (Bruner, 1999).

In this framework, the *experiential learning theory* (Kolb, 2015) is proposed. This theory seeks to integrate discovery learning theories in formal learning contexts, differentiating it from everyday learning situations. Kolb (2015) proposes an instructional model that goes from the concrete to the abstract, with phases of 1) concrete experience; 2) observation and reflection; 3) construction of abstract ideas and generalizations; and 4) testing the implications of concepts in new situations. There would be evidence that experiential learning is an effective framework in different pedagogical situations, besides positively influencing the classroom climate, fostering participation and motivation (Kolb, 2015). In that sense, the immersive properties of VR could be used to enhance phases 1 and 2 of the model proposed by Kolb, located in the concrete setting. Subsequently, phases 3 and 4 of the model could be developed, where students would build abstract knowledge and generalizations from the immersive audiovisual experience with the guidance of the teacher.

In this way, the use of VR as a device that simulates reality could be supported from the perspective of cognitive psychology. However, there are two points that are important to consider. First, that learning situations based on direct instruction, such as verbal exposure to content, could not be explained with this theory. Second, the constructivist pedagogical approach has been criticized for being inefficient and for not considering the architecture of human cognition in its formulation (Kirschner et al., 2006). Although there is empirical evidence that could support constructivist pedagogical approaches, these would be effective to the extent that individuals have partially constructed cognitive schemas of the intended learning. Prior to this construction, constructivist learning methodologies that promote indirect instruction would be ineffective (Tuovinen & Sweller, 1999).

To understand how the human cognitive architecture processes information during learning, the Cognitive Load Theory (CLT) is proposed. This theory assumes that human cognitive architecture is constituted by working memory and long-term memory (Sweller et al., 2011). Working memory, also known as short-term memory, has a limited capacity in terms of time and amount of information it can process. In contrast, long-term memory can retain large amounts of information for extended periods of time. In this framework, learning is defined as an alteration of long-term memory (Paas & Swellwer, 2022).

CLT distinguishes different types of cognitive load, establishing categories in terms of their relationship with learning objectives and the resources that are necessary for an individual to build new mental schemas (Sweller et al., 2011). Intrinsic cognitive load refers to the processing of the specific content to be taught. This load is fixed, and its level is determined by the complexity of the learning content. The germane cognitive load includes the didactic methods necessary for the learning of these contents. Finally, the extraneous cognitive load includes elements that are not related to the learning contents. The total cognitive load arises from the combination of these three categories, which work additively. Since working memory has limited capacity, an effective instructional design should strive to keep extraneous cognitive load low to free resources for intrinsic load. The germane cognitive load is necessary for learning; however, the didactic method used could increase it unnecessarily and limit the learning of essential content (Moreno & Park, 2010).

Based on CLT, the Cognitive Theory of Multimedia Learning (CTML) is proposed (Mayer, 2009). This theory seeks, through empirical evidence, to understand how people learn through multimedia elements, such as presentations and simulations (Mayer, 2022). In that context, multimedia refers to the integration of images and sounds in instructional designs, which are sensorially perceived by the eyes and ears. These perceptions are processed by independent channels in working memory, which as noted above has a limited capacity (Mayer, 2022; Mayer & Moreno, 2010). Based on the same assumptions of CLT, principles are established for instructional designs and multimedia simulations to reduce germane and extraneous cognitive load. Some of these principles have to do with coherence between image and sound, avoiding redundancy of information and excluding everything that is not relevant to the essential content (Fiorella & Mayer, 2022; Mayer, 2009).

VR analyzed from a CTML point of view presents ambivalent perspectives. On the one hand, the complexity of an immersive system could increase the germane and extraneous cognitive load to critical levels, making it a less effective alternative to traditional audiovisual presentation. It could also violate the principle of coherence by presenting stimuli in the form of animations and sounds that would not be essential for understanding the content to be represented (Parong, 2022). These claims are supported by empirical research, where high rates of participant motivation when

using immersive VR do not translate into improved learning outcomes (Makransky et al., 2019; Parong & Mayer, 2018). This perspective may support low-immersive VR systems, such as desktop VR (Parong, 2022).

However, these results would be divergent with those of many other empirical studies reporting positive learning outcomes using VR (e.g., Barmpoutis et al., 2020; De Back et al., 2020; Innocenti et al., 2019). In this regard, VR is presented as an anomaly in CTML, with empirical results diverging from its predictions. While the causes of this anomaly could be multiple, some authors attribute it to the failure of CLT to incorporate motivation into its analysis of the cognitive process of learning (Feldon et al., 2019).

## Motivation in VR Learning Processes

Different research reports significant increases in participants' perceived motivation through VR, both in music (Fletcher et al., 2019; Innocenti et al., 2019) and in other disciplines (Ekstrand et al., 2018; Hewawalpita et al., 2018; Parong & Meyer, 2018). However, it should be considered that high levels of motivation do not necessarily imply improvements in learning (Schunk et al., 2014).

From the standpoint of the learning psychology, the construct motivation is addressed by multiple theories, such as self-determination theory (Ryan & Deci, 2000), social cognitive theory (Bandura, 1991), self-regulation theory (Zimmerman et al., 1996), personal and situational interest (Schiefele, 2009), to name a few. Although these theories present varied approaches to the concept of motivation, they all converge in that it is an important aspect of the cognitive process.

In the present study, motivation is defined as the “the extent to which one intends to engage in an activity” (Jones, 2020, p. 10). Motivation is a process that cannot be directly observed but is inferred from actions and behavior of the individual. The presence of motivation implies that the individual would be more likely to participate in certain types of activities related to learning, and that these activities could be maintained over time if the motivation is sustained. Later, in the methodological section, the dimensions of the motivation construct are operationalized in concrete observable variables.

## Research Objectives

Given the high potential promised using VR in learning contexts, a pilot project called “Virtual Reality in the Music Classroom” began in 2019 at the *Universidad Mayor*. The objectives were: 1) to design and produce original educational audiovisual material for the music classroom in VR format; 2) to evaluate and compare the effectiveness of the material produced in the school context; 3) to

collect the perceptions of the participants in the experience; and 4) to understand from a cognitive point of view how, through VR, learning processes in music education can be enhanced. The results of objectives 1 and 2 were reported in another publication (Iglesias & Chávez, 2022), which measures and compares the effectiveness of VR with respect to other audiovisual projection methods, a variable that was operationalized through the percentage of achievement in an evaluation with questions about the contents of the class. This article reports on objectives 3 and 4.

## Methodological Design

The study had an exploratory scope and a mixed approach. Multiple measurements were made over a period of one month. For the quantitative data, a quasi-experimental counterbalanced design with post-test was used, in which all the experimental treatments were applied to the participants (Campbell & Stanley, 1995).

A complete eighth grade class was used as a sample. Fifteen students participated ( $n=15$ ), with an age range between 14 and 16 years, ten boys and five girls. The course was randomly divided into three groups of five students. Three class sessions were conducted by the same teacher. All three classes contemplated working from a music video, which each group observed using different audiovisual projection formats: Immersive VR; desktop VR; and traditional two-dimensional (2D) projection. All groups participated in the same class. In this manner, the audiovisual projection format becomes the experimental treatment or independent variable.

Typically, VR is used with specialized hardware devices known as head-mounted displays (HMD), in which a screen covers the user's field of view, and headphones provide sound. These devices have an internal oscilloscope that allows images and sounds to move interactively, simulating the perception of reality. In this article, the term immersive VR is used for this category. Although HMDs are the most used devices for VR, there are other possibilities for representations in this format, such as desktop VR, in which the simulation is displayed on a computer and controlled by the mouse (Wu et al., 2020). Thus, there is a wide range of possibilities for VR implementation, ranging from low-immersion systems running on ordinary computers to fully immersive HMD. Table 1 details the rotation of groups across sessions with respect to the counterbalanced experimental treatment, constructed from these two categories of VR and traditional audiovisual projection.

**Table 1.**  
 Counterbalanced quasi-experimental design

|                | <b>Session 1</b>              | <b>Session 2</b>              | <b>Session 3</b>                           |
|----------------|-------------------------------|-------------------------------|--|
| <i>Group 1</i> | X <sub>1</sub> <sup>a</sup> O | X <sub>2</sub> <sup>b</sup> O | X <sub>3</sub> <sup>c</sup> O <sup>d</sup> |
| <i>Group 2</i> | X <sub>2</sub> O              | X <sub>3</sub> O              | X <sub>1</sub> O                           |
| <i>Group 3</i> | X <sub>3</sub> O              | X <sub>1</sub> O              | X <sub>2</sub> O                           |

Source: Adapted from Campbell and Stanley (1995). Own elaboration.

a X1: Immersive Virtual Reality (with HMD device).

b X2: Desktop or non-immersive Virtual Reality (with PC and mouse).

c X3: Traditional 2D projection (from PC).

d O: Observation (closed-ended questionnaire).

At the end of each of the sessions, a closed-ended questionnaire was administered. The purpose of this questionnaire was for participants to evaluate the motivation perceived during the audiovisual experience. To operationalize the concept of motivation, it was broken down into dimensions to which specific indicators were associated. The aim was to arrive at a simple instrument that the participants, schoolchildren between 14 and 16 years of age, could understand and thus evaluate the experience with consistency throughout the sessions.

Immersion in an audiovisual projection was defined as the effectiveness with which representation replaces perception of the real world (Di Natale et al., 2020). Attention as the ability to direct perception and action to perform a task and achieve a goal (Gibson & Rader, 2011; Shell & Flowerday, 2019). Interest as the psychological state of perceptual readiness to specific stimuli (Pintrich, 2003; Renninger & Hidi, 2019). The operationalization of the dependent variable motivation was constructed from these concepts and is shown in Table 2.

**Table 2.**  
 Operationalization of the motivation variable

| <b>Dimensions</b> | <b>Indicators</b>  |   |  |  |
|-------------------|--|---|--|--|
|                   | <b>Total</b>   | <b>Almost total</b>   | <b>Partial</b>   | <b>Low</b>   |
| <b>Immersion</b>  | I managed to abstract myself from all external distractions during the video playback. | I managed to abstract myself from almost all external distractions during the video playback. | I managed to partially abstract myself from external distractions during playback. | I was constantly distracted by external distractions during playback.    |
| <b>Attention</b>  | I was able to keep the attention throughout the video playback                         | I was able to keep the Attention in more than half of the reproduction                        | I was able to keep the Attention in the middle of the video playback               | I was able to keep the Attention in less than half of the video playback |

|                 |   |   |   |  |
|-----------------|---|---|---|--|
| <b>Interest</b> | I remained interested throughout the video. | I remained interested throughout most of the video. | I stayed interested in half of the playback time. | I stayed interested in less than half the length of the video. |
|-----------------|---|---|---|--|

Source: Own elaboration

For data analysis, scores were assigned to each of the indicators, transforming them into ordinal variables (total=4, almost total=3, partial=2, low=1). The data were analyzed using descriptive and inferential statistics using R-Studio software<sup>2</sup> (R Core Team, 2023). Percentage, mean and standard deviation were calculated for each of the three dimensions. In addition, the effect size was calculated from Cohen's d coefficient (Cohen, 1988). To assess the internal consistency of the measurement scales, Cronbach's alpha coefficient was calculated (Russell, 2018).

The Kruskal-Wallis test and then post hoc tests (Wicoxon pairwise tests) were performed, as these were nonparametric ordinal variables measured independently, to verify whether there were statistically significant differences from the experimental treatment (Russell, 2018).

It was also sought to find possible correlations between motivation and the learning outcomes previously reported in Iglesias & Chávez (2022). For this purpose, the Pearson correlation coefficient was calculated between the learning outcomes, defined as the percentage of achievement obtained by the participants in an evaluation of the contents covered in each session, and the perception of immersion, attention and interest.

For qualitative data collection, a focus group (Morgan & Krueger, 1998) was conducted at the end of the three sessions, to which six randomly selected participants were invited (n=6). The objective of the focus group was to deepen the participants' perception during the audiovisual experience. Triangulation was attempted based on the results obtained.

## Procedures

The sessions were conducted in the context of a school music class. A learning unit was structured throughout the three sessions, aligned with the Music Study Programs (Mineduc, 2015), in the axis of *learning to listen and appreciate*: "OA2 - Compare music with different characteristics, based on elements of musical language and compositional procedures, as well as their relationship with their

<sup>2</sup> The R Studio code and data can be found in Iglesias (2024).

expressive purpose” (Ministerio de Educación, 2016, p. 62). The learning unit aimed to develop skills in recognizing in works of different styles the instruments involved in each performance, as well as the rhythm, character and musical form.

The class was structured in opening, development and closing activities. During the development of each class, students watched and listened to an audiovisual projection in which they had to identify instruments, rhythms, character and structure. In each session they worked with a different repertoire. In this way, the audiovisual projection was framed in a pedagogical approach of experiential learning, relating and building knowledge by relating the observed with previous experiences and the contents that had been developed in the class (Kolb, 2015).

In each session the same audiovisual projection was observed by each of the three groups in a different format (immersive VR, desktop VR, 2D). Three sessions were planned to produce experimental treatment rotations so that each group could experience all three audiovisual projection formats. During the audiovisual projection, participants remained seated in rotating seats regardless of the experimental treatment applied. At the end of each class, they were asked to answer the perceived motivation questionnaire. The repertoire listened to in each session is detailed in Table 3.

**Table 3.**

*Musical repertoire listened to during audiovisual projection sessions*

|                  | <b>Repertoire listened</b>              | <b>Style</b>                        |
|------------------|---|-------------------------------------|
| <i>Session 1</i> | “Todo le canta” by Patricio Quintanilla | Traditional Chilean Christmas carol |
| <i>Session 2</i> | “Adiós Nonino” by Astor Piazzolla       | New Tango                           |
| <i>Session 3</i> | “Por una cabeza” by Carlos Gardel       | Tango                               |

Source: Own elaboration

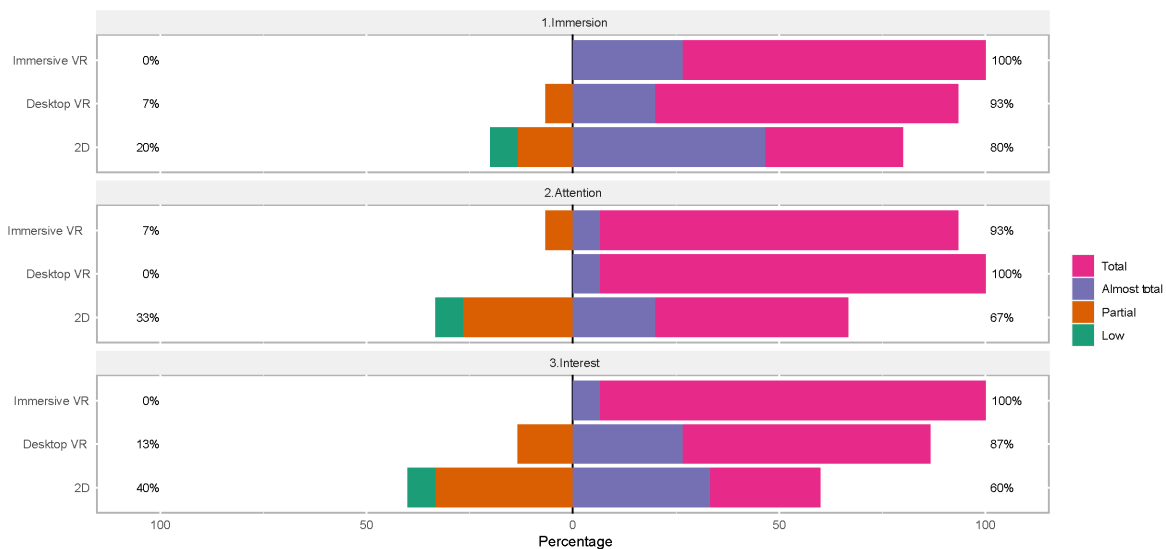
In this way, each participant evaluated the three experimental treatments, obtaining a total of 45 evaluations. This gave reliability to the measurements, an aspect that was reinforced by the teacher during each session, pointing out to participants that they should evaluate the experience by applying the same criteria. In this way, the internal validity of the quasi-experiment was controlled (Campbell & Stanley, 1995), minimizing the threats of instrumentation and administration of the measurements. There were no cases of experimental mortality, nor were any events identified in the history of the quasi-experiment that could alter the results significantly. Internal validity was also enhanced by the counterbalanced quasi-experimental design used in the investigation. According to Campbell and Stanley (1995), this design increases the precision of the measurements by applying the totality of the experimental treatments to all participants. The same authors point out

that the random assignment of groups within the sample, a procedure that as mentioned above was used in the present study, reinforces the internal validity of the quasi-experiment.

From an ethical point of view, the intervention was authorized by the school management. Since the participants were underage, their parents were informed and signed a letter of informed consent.

## Results

Figure 1 shows the results of the perception of motivation, grouped according to the categories analyzed and the audiovisual projection method.



**Figure 1.**  
*Results of the motivation questionnaire.*

Source: Own elaboration.

The Cronbach's Alpha coefficient value was 0.76, showing high internal consistency (Russell, 2018). Kruskal-Wallis tests yielded a  $p=0.03$  value for immersion,  $p=0.005$  for attention and  $p=0.001$  for interest, thus inferring statistically significant differences between the groups in the three dimensions measured.

Wilcoxon pairwise tests allow us to determine the specific groups in which these differences occur. In the immersion dimension, differences were found between the VR and 2D projection groups, as well as between Desktop VR and 2D projection (VR/Desktop VR,  $p=0.9$ ; VR/2D,  $p=0.05$ ; Desktop VR/2D,  $p=0.05$ ). The same behavior

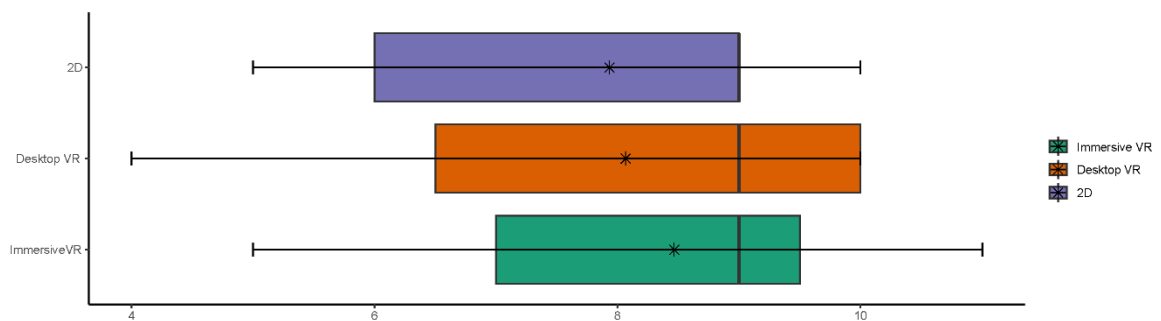
was observed in the attention measurement (RV/RV Desktop,  $p= 0.55$ ; RV/2D,  $p= 0.03$ ; RV Desktop/2D,  $p=0.02$ ). In the interest measurement, differences were found between the three groups (RV/RV Desk,  $p= 0.05$ ; RV/2D,  $p= 0.001$ ; RV Desk/2D,  $p=0.05$ ). Cohen's d coefficients and their effect size (Russell, 2018) are shown in Table 4.

**Table 4.**  
 Cohen's d coefficient and motivational effect size between groups

| Immersion                           | Attention                             | Interest                            |
|-------------------------------------|---------------------------------------|-------------------------------------|
| I-VR and D-VR: 0.12 => small effect | RI-VR and D-VR: -0.30 => small effect | I-VR and D-VR: 0.83 => large effect |
| I-VR and 2D: 0.94 => large effect   | I-VR and 2D: 0.88 => large effect     | I-VR and 2D: 1.64 => large effect   |
| D-VR and 2D: 0.78 => large effect   | D-VR and 2D: 1.15 => large effect     | D-VR and 2D: 0.78 => large effect   |

Source: Own elaboration

To compare learning outcomes with participants' perception of motivation, data reported in Iglesias & Chávez (2022) were used.



**Figure 2.**  
 Correct responses by experimental treatment, based on Iglesias & Chávez (2022).  
 The \* represent means.

The correlation between correct answers and immersion gave a moderate positive correlation index (Pearson's Coefficient= 0.37,  $p=0.001$ ). Attention obtained a weak correlation index (Pearson's Coefficient= 0.30,  $p=0.04$ ). Interest showed no correlation with learning outcomes, in addition to yielding results without statistical significance (Pearson's Coefficient= 0.18,  $p=0.2$ ) (Russell, 2018).

In focus group, students commented on different aspects of audiovisual projections, where they expressed themselves notably in favor of immersive and desktop VR. However, in some comments it is possible to infer the high extraneous

cognitive load (Sweller et al., 2011) that is produced from the use of immersive VR devices:

[The VR] was a little bit distracting, I was looking up, where I was sitting (Student 1, focus group).

With VR I was like *uh-uh*, I didn't mind watching the song (Student 3, focus group).

Despite being isolated from the outside using glasses and headphones, the device itself would produce distractions from the content presented. In this regard, from some comments it could be inferred that Desktop VR could reduce the germane and extraneous cognitive load.

Desktop VR was the one I liked the most, I was able to concentrate 100% (Student 2, focus group).

Desktop VR was the best way to retain information (Student 5, focus group).

I liked desktop VR the most, because with the glasses I can't take notes (Student 3, focus group).

Although they also highlight the benefits of immersive VR:

You can get into the world of music (Student 4, focus group).

With immersive VR you can't be distracted by anything, you didn't have your partner next to you. I couldn't see my classmate spinning around like crazy (Student 4, focus group).

You feel like you are inside (Student 5, focus group).

It is the closest experience to a concert (Student 6, focus group).

Both immersive and desktop VR have the advantage that the presentation of image and sound is controlled by the user, enhancing dynamic interaction with the content to be presented.

[The traditional 2D projection] put you on the instrument that the video wanted, instead of being able to focus on the instrument you wanted (Student 6, focus group).

This was perceived as positive for learning by the participants.

## Discussion

From the empirical point of view, the results show that the participants perceive VR, in its immersive and desktop forms, as an audiovisual projection medium that increases motivation with respect to a traditional two-dimensional projection. Statistical significance could be established between groups, with a large effect size. The participants also perceived the benefits of these new technologies for the simulation of musical situations, stating that they felt they were participating in a concert during the class. This translated into better learning results (Iglesias & Chávez, 2022), establishing a positive correlation in the dimensions of immersion and attention with respect to the percentage of achievement in an evaluation of the contents of the class.

The questionnaire used to measure motivation showed a high internal consistency (Cronbach's Alpha Coefficient = 0.76), which validates the operationalization of the construct in the dimensions of immersion, attention and interest. Although motivation could have been operationalized in much more sophisticated and complex indicators based on the different psychological theories that develop this concept, it was chosen a construct that was easy to apply and understand so that it could be answered by schoolchildren. The simplicity of the construct had another advantage, and that is that it could be applied in the same classroom context. This allowed the audiovisual experience that we sought to measure to be still latent, giving more consistency to the evaluation that each participant was able to make.

In the analysis of the differences between immersive and desktop VR, contradicting results are interpreted. On the one hand, while VR could generate more immersion and interest, Desktop VR could enhance attention. When triangulating the data collected through the questionnaire and focus group, convergence is seen, noting that immersive VR could distract attention from essential content. If it is considered that when using an HMD device with headphones the user is absolutely isolated from external stimuli, it can be inferred that this distraction is caused by the saturation of the germane or extraneous cognitive load (Mayer, 2022; Mayer & Moreno, 2010; Sweller et al., 2011). In this sense, these results converge with what has been raised in studies subscribed to CTML, which take a critical and skeptical stance regarding the use of immersive VR (Makransky et al., 2019; Parong and Mayer, 2018).

However, the analysis of the learning results points in the opposite direction, observing that the immersive VR group obtained better results than the Desktop VR group. Although the study did not conclude statistical significance and the effect size measured was small, this tendency is accentuated in the more difficult questions, which involved the relationship of several musical concepts and parameters (Iglesias

& Chávez, 2022). Thus, the results diverge with the predictions of the CTML, which states that the use of VR would limit learning outcomes by generating germane and extraneous cognitive overload, saturating the limited resources available to each individual to process information (Mayer & Moreno, 2010). This divergence is reiterated in several empirical studies using VR for simulations (e.g., Barmoutis et al., 2020; De Back et al., 2020; Innocenti et al., 2019), suggesting that an anomaly may exist in CTML.

The explanation of this anomaly could have significant theoretical implications. CTML seeks to understand the audiovisual learning process across a broad spectrum (Mayer, 2022). This spectrum encompasses a continuum of didactic methodologies and diverse audiovisual media. The methodologies studied range from audiovisual presentations based on direct instruction to simulations of reality. While the former seek to provide new information to the user, the latter attempt to simulate a specific experience, often without the mediation of verbal content.

Furthermore, while VR has similarities to traditional projection media, it also has notable differences, such as the ability to interact through kinesthetic and haptic channels. This represents an expansion of the auditory and visual channels found in a traditional audiovisual projection. CTML has focused its study on the latter, without considering the potential that the haptic-kinesthetic channel could offer.

In this study, the audiovisual experience was used for a simulation. This involves a method of indirect instruction, with the expectation that users would have the experience of participating in a concert. Participants were able to interact with the music presented by observing an instrumental ensemble. VR allows a performer to be visually and aurally foregrounded by the user's head movement. In other words, one can interactively control through kinesthetic channels what one wants to see and hear. In the design of the intervention, this possibility of interaction with the audiovisual projection was considered as pedagogically significant, since the learning objective of the class was the identification of the different instruments involved in the performance. In the focus group, one participant supports this idea:

[The traditional 2D projection] put you on the instrument that the video wanted, instead of being able to focus on the instrument you wanted (Student 6, focus group).

This experience was then integrated into the classroom context, framed within a constructivist experiential pedagogical approach (Kolb, 2015). This pedagogical approach is substantially different from the one that an instructional design may adopt in which the user is sought to perceive and process information through direct instruction.

There are several studies that have explored the use of haptic and kinesthetic stimuli, concluding that their use could enhance learning (Fredembach et al., 2009;

Hu et al, 2015; Kalenine et al., 2011). These results support the idea of a multisensory pedagogical approach, contradicting the premise of CLT and CTML that sensory stimuli should be limited to reduce extraneous cognitive load (Gins et al., 2020).

From the musical perspective, the idea that movement can enhance learning processes is supported by methodologies that are widely accepted in the discipline, such as the approach proposed by Orff in his *Schulwerk* or by Dalcroze in his *Eurythmy* (Abril, 2011; Álamos & Tejada, 2020). Although it could be argued that in musical learning, movement would not correspond to an extraneous cognitive load, the CTML's focus on content oriented to propositional and verbal knowledge, leaving out procedural knowledge, is evident. Together with the relationship that the development of rhythmic aspects involves with movement, it has been suggested that the execution of synchronous movements could activate attentional resources and stimulate working memory, enhancing social synchrony and other cognitive functions not directly related to music (Miendlarzewska & Trost, 2014). To date, there is no research exploring from the perspective of CLT and CTML the potential of the kinesthetic channel in audiovisual learning situations, nor guidance about its use in instructional designs. This new perspective could be particularly useful in situations in which VR is used for simulation, which seek to have the user experience different situations through immersion, to then build knowledge.

Therefore, the CTML prediction that VR can increase to critical levels the germane and extraneous cognitive load could be relevant to understand the cognitive process in situations of direct instruction, with contents that aim at propositional knowledge. However, this prediction might not be transferable to other situations, in which VR is used for simulations in a constructivist experiential learning context. In this sense, CLT is explicit and emphatic in its theoretical foundation in considering that constructivist instructional variants, such as discovery learning and other methods that promote indirect instruction, would be ineffective (Kirschner et al. 2006; Moreno & Park, 2010; Tuovinen & Sweller, 1999). As noted above, one of the great educational potentials of VR appears in simulated learning, allowing the user to experience situations that would be difficult or impossible to be actually achieved (Bower et al., 2020; Dede et al., 2017).

## Conclusions

To ensure the consistency of the research, an effort was made to control the internal validity of the quasi-experiment. Although the analysis of the validity of the data revealed no significant threats, it is important to note that the use of convenience sampling together with a small sample size could limit the external validity of the study. To mitigate this potential threat, the information obtained

was triangulated. As mentioned in the discussion, the results converge with several similar investigations.

In this sense, the results presented highlight the existing tension between different cognitive theories that allow us to understand learning through audiovisual media. However, there is also a consensus perspective: desktop VR has characteristics that make it interesting. First, it would bring most of the benefits that relate these technologies to constructivist perspectives that use simulation for knowledge construction in learning environments based on experiential learning. These involve user interaction with content through real-time manipulation of the audiovisual projection. Although desktop VR would provide less immersion, this is not necessarily a negative thing, since according to CTML immersion could limit learning by increasing the germane cognitive load. On the other hand, its implementation has advantages that make it attractive, as it requires the use of hardware that is currently already available in most educational establishments and homes: a computer and a headphone headset.

In that sense, the school system would already be ready to adopt VR. However, something fundamental is missing, the development of didactic material. This is an essential point. It has been insisted that it is not the audiovisual media that teach, they are only the support of an instructional design (Clark, 1983; 1994). The potential of an audiovisual medium as a learning resource will only be achieved to the extent that there are instructional designs with a solid pedagogical foundation to support it. In this sense, a great weakness of VR is detected, since there is no massification of consistent pedagogical material that can support its entry into the classroom. It is also important to point out that although there is abundant literature on the relationship between movement and cognition (Abril, 2011; Álamos & Tejada, 2020; Miendlarzewska & Trost, 2014), a deep relationship between this knowledge and multimedia learning theories has not been established. This leaves interesting space for future research and theories; however, it is also an indicator of the lack of theoretical and pedagogical consistency in the promises of revolutionizing education through VR.

For these reasons, one should be cautious in reading the promises made by the technology giants regarding the entry of VR into the classroom, which would show a tendency towards what Mayer (2009) calls a technology-centric approach. It is also important to emphasize that while much research reports improved learning outcomes in VR pedagogical interventions, the effect size is always small.

Without consistent pedagogical foundations and adequate instructional designs, VR can only be predicted to join the list of technology implementations that failed to meet expectations upon entering the classroom.

## References

- Abril, C. R. (2011). Music, movement, and learning. In R. Collwell & P. Webster (Eds.), *MENC Handbook of Research on Music Learning, Applications* (pp.92-129). Oxford University Press.
- Álamos-Gómez, J. & Tejada, J. (2020). Interrelaciones entre acción y cognición. Aportaciones de la neurociencia a la educación rítmico-musical. *Opus*, 26(2), 1-21. doi: 10.20504/opus2020b2606
- Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational behavior and human decision processes*, 50(2), 248-287. doi:10.1016/0749-5978(91)90022-L
- Barmpoutis, A., Faris, R., Garcia, L., Gruber, L., Li, J., Peralta, F. & Zhang, M. (2020). Assessing the Role of Virtual Reality with Passive Haptics in Music Conductor Education: A Pilot Study. In J. Chen, & G. Fragomeni (Eds.), *Virtual, Augmented and Mixed Reality. Design and Interaction* (pp. 275-285). doi:10.1007/978-3-030-49695-1\_18
- Bhattacharjee, D., Paul, A., Kim, J. H., & Karthigaikumar, P. (2018). An immersive learning model using evolutionary learning. *Computers & Electrical Engineering*, 65, 236-249. doi:10.1016/j.compeleceng.2017.08.023
- Bruner, J. S. (1999). *The process of education*. Harvard University Press.
- Bower, M., DeWitt, D. & Lai, J. W. (2020). Reasons associated with preservice teachers' intention to use immersive virtual reality in education. *British Journal of Educational Technology*, 51(6), 2214-2232. doi:10.1111/bjet.13009
- Campbell D. & Stanley J. (1995). *Diseños experimentales y cuasi-experimentales en la investigación social*. Editorial Amorrortu.
- Chen, C. J. (2009). Theoretical bases for using virtual reality in education. *Themes in science and technology education*, 2(1-2), 71-90. Retrieved from <http://earthlab.uoi.gr/ojs/theste/index.php/theste/article/view/23>
- Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53(4), 445-459. Retrieved from <https://www.jstor.org/stable/1170217>
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42(2), 21-29. doi:10.1007/BF02299088
- Clegg, N. (2023). *How the Metaverse Can Transform Education*. Meta. Retrieved from <https://about.fb.com/news/2023/04/how-the-metaverse-can-transform-education/>

- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.
- Contreras, G. S., González, A. H., Fernández, M. I. S., Martínez, C. B., Cepa, J., & Escobar, Z. (2022). The importance of the application of the metaverse in education. *Modern Applied Science*, 16(3), 1-34. doi:10.5539/mas.v16n3p34
- Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*. Teachers College Press.
- Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Harvard University Press.
- Dalgarno, B. & Lee, M. J. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, 41(1), 10-32. doi:10.1111/j.1467-8535.2009.01038.x
- De Back, T., Tinga, A., Nguyen, P., & Louwse, M. (2020). Benefits of immersive collaborative learning in CAVE-based virtual reality. *International Journal of Educational Technology in Higher Education*, 17(1), 1-18. doi:10.1186/s41239-020-00228-9
- Dede, C, Jacobson, J. & Richards, J. (2017). Introduction: Virtual, augmented, and mixed realities in education. In D. Liu, C. Dede, R. Huang, & J. Richards (Eds.), *Virtual, augmented, and mixed realities in education* (pp. 1-16). Springer. doi:10.1007/978-981-10-5490-7
- Di Natale, A. F., Repetto, C., Riva, G. & Villani, D. (2020). Immersive virtual reality in K-12 and higher education: A 10-year systematic review of empirical research. *British Journal of Educational Technology*, 51(6), 2006-2033. doi:10.1111/bjet.13030
- Ekstrand, C., Jamal, A., Nguyen, R., Kudryk, A., Mann, J., & Mendez, I. (2018). Immersive and interactive virtual reality to improve learning and retention of neuroanatomy in medical students: a randomized controlled study. *Canadian Medical Association Open Access Journal*, 6(1), E103-E109. doi:10.9778/cmajo.20170110
- Feldon, D. F., Callan, G., Juth, S., & Jeong, S. (2019). Cognitive load as motivational cost. *Educational Psychology Review*, 31, 319-337. doi:10.1007/s10648-019-09464-6
- Fiorella, L. & Mayer, R. E. (2021). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (pp. 185-198). Cambridge University Press.

- Flavell, J., Miller, P. & Miller, S. (2002). *Cognitive development*. Prentice Hall.
- Fletcher, C., Hulusic, V. & Amelidis, P. (2019). Virtual Reality Ear Training System: A study on Spatialised Audio in Interval Recognition. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)* (pp. 1-4). IEEE. doi:10.1109/VS-Games.2019.8864592
- Fredembach, B., de Boisferon, A. H. & Gentaz, E. (2009). Learning of arbitrary association between visual and auditory novel stimuli in adults: the “bond effect” of haptic exploration. *PLoS one*, 4(3), e4844. doi:10.1371/journal.pone.0004844
- Gibson, E. & Rader, N. (2011). Attention: The Perceiver as Performer. In G. Hale & M. Lewis (Eds.), *Attention and Cognitive Development* (pp. 1-22). Springer. doi:10.1007/978-1-4613-2985-5\_1
- Hewawalpita, S., Herath, S., Perera, I. & Meedeniya, D. (2018). Effective Learning Content Offering in MOOCs with Virtual Reality-An Exploratory Study on Learner Experience. *Journal of Universal Computer Science*, 24(2), 129-148. doi:10.3217/jucs-024-02-0129
- Hamilton, D., McKechnie, J., Edgerton, E. & Wilson, C. (2020). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, 8(1), 1-32. doi:10.1007/s40692-020-00169-2
- Iglesias, P. (2024). *R Code for project Virtual Reality in the Music Classroom*. Zenodo. doi:10.5281/zenodo.13135984
- Iglesias, P., & Chávez, P. (2022). La realidad virtual en el aula de música: un estudio cuasiexperimental. *Perspectiva Educativa*, 61(2). doi:10.4151/07189729-Vol.61-Iss.2-Art.1215
- Innocenti, E., Geronazzo, M., Vescovi, D., Nordahl, R., Serafin, S., Ludovico, L. A. & Avanzini, F. (2019). Mobile virtual reality for musical genre learning in primary education. *Computers & Education*, 139, 102-117. doi:10.1016/j.compedu.2019.04.010
- Jones, B. D. (2020). Motivating and engaging students using educational technologies. In M. Bishop, E. Boling, J. Elen & V. Svihla (Eds.), *Handbook of research in educational communications and technology: Learning design* (pp. 9-35). Springer. doi:10.1007/978-3-030-36119-8
- Kalenine, S., Pinet, L. & Gentaz, E. (2011). The visual and visuo-haptic exploration of geometrical shapes increases their recognition in preschoolers. *International Journal of Behavioral Development*, 35(1), 18-26. doi:10.1177/0165025410367443

- Kirschner, P. A., Sweller, J. & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75-86. doi:10.1207/s15326985ep4102\_1
- Kolb, D. A. (2015). *Experiential learning: Experience as the source of learning and development*. Pearson Education.
- Makransky, G., Terkildsen, T. S. & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, 60, 225-236. doi:10.1016/j.learninstruc.2017.12.007
- Mayer, R. (2009). *Multimedia learning*. Cambridge University Press.
- Mayer, R. (2022). Cognitive Theory of Multimedia Learning. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (pp. 57-72). Cambridge University Press.
- Mayer, R. E. & Moreno, R. E. (2010). Techniques that reduce extraneous cognitive load and manage intrinsic cognitive load during multimedia learning. In J. Pass, R. Moreno & R. Brünken (Eds.), *Cognitive load theory* (pp. 131-152). Cambridge University Press.
- Meta. (2021). *What Is Virtual Reality All About?* Meta Quest Blog. Retrieved from <https://www.meta.com/blog/quest/what-is-virtual-reality-all-about/>
- Miendlarzewska, E. & Trost, W. (2014). How musical training affects cognitive development: rhythm, reward and other modulating variables. *Frontiers in Neuroscience*, 7. doi: 10.3389/fnins.2013.00279
- Ministerio de Educación. (2016). *Música: Programa de Estudio Primero Medio*. Ministerio de Educación, Gobierno de Chile.
- Moreno, R. E., & Park, B. (2010). Cognitive load theory: Historical development and relation to other theories. In J. Pass, R. Moreno & R. Brünken (Eds.), *Cognitive load theory* (pp. 29-47). Cambridge University Press.
- Morgan, D. L., & Krueger, R. A. (1998). *The focus group guidebook*. Sage.
- Oyelere, S. S., Bouali, N., Kaliisa, R., Obaido, G., Yunusa, A. A. & Jimoh, E. R. (2020). Exploring the trends of educational virtual reality games: a systematic review of empirical studies. *Smart Learning Environments*, 7(1), 1-22. doi:10.1186/s40561-020-00142-7
- Parong, J. (2022). Multimedia Learning in Virtual and Mixed Reality. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (pp. 498-509). Cambridge University Press.

- Parong, J. & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology, 110*(6), 785. doi:10.1037/edu0000241
- Paas, F. & Sweller, J. (2022). Implications of cognitive load theory for multimedia learning. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia learning* (pp.73-81). Cambridge University Press.
- Piaget, J. (2000). *The construction of reality in the child*. Routledge.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology, 95*(4), 667. doi:10.1037/0022-0663.95.4.667
- R Core Team. (2023) [Software de computación]. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Renninger, K & Hidi, S. (2019). Interest development and learning. *The Cambridge handbook of motivation and learning* (pp. 260-285). Cambridge University Press. doi:10.1017/9781316823279.032
- Russell, J. A. (2018). *Statistics in music education research*. Oxford University Press.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68. doi:10.1037/0003-066X.55.1.68
- Schiefele, U. (2009). Situational and individual interest. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 197-222). Routledge.
- Schunk, D., Meece, J., & Pintrich, P. (2014). *Motivation in education: Theory, research, and applications*. Pearson.
- Shell, D., & Flowerday, T. (2019). Affordances and attention. In K. Renninger & S. Hidi (Eds.), *The Cambridge handbook of motivation and learning* (pp. 759–782). Cambridge University Press. doi:10.1017/9781316823279.032
- Shute, V., Rahimi, S. & Emihovich, B. (2017). Assessment for learning in immersive environments. In Liu, Dede, Huang, & Richards (Eds.), *Virtual, augmented, and mixed realities in education* (pp. 71-87). Springer. doi:10.1007/978-981-10-5490-7
- Sweller, J., Ayres, P. & Kalyuga, S. (2011). *Cognitive Load Theory*. Springer.
- Tuovinen, J. E. & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of educational psychology, 91*(2), 334. doi:10.1037/0022-0663.91.2.334

- Wu, B., Yu, X. & Gu, X. (2020). Effectiveness of immersive virtual reality using head-mounted displays on learning performance: A meta-analysis. *British Journal of Educational Technology*, 51(6), 1991-2005. doi:10.1111/bjet.13023
- Zimmerman, B. J., Bonner, S. & Kovach, R. (1996). *Developing self-regulated learners: Beyond achievement to self-efficacy*. American Psychological Association.