# Can an Immersive Virtual Reality Simulation Increase Students' Interest and Career Aspirations in Science?

Accepted in the British Journal of Educational Technology in the special issue Immersive Virtual Reality in Education on April 22. 2020

Guido Makransky[1], Gustav B. Petersen[1], Sara Klingenberg[1]

[1] Department of Psychology, University of Copenhagen, Copenhagen, Denmark

Corresponding Author: Guido Makransky, University of Copenhagen

Address: Øster Farimagsgade 2A, 1353 København K, gm@psy.ku.dk

- Guido Makransky (gm@psy.ku.dk) is an Associate Professor at the Department of Psychology at the University of Copenhagen and the leader of the Virtual Learning Lab (virtuallearninglab.net).
- Gustav B. Petersen (gbp@psy.ku.dk) is a PhD candidate at the Department of Psychology at the University of Copenhagen.
- Sara Klingenberg (sara.klingenberg@sund.ku.dk) is a research assistant at the department of Psychology at the University of Copenhagen.

This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. The final article will be available, upon publication, via its DOI: 10.1111/bjet.12954

## Abstract

Science-related competencies are demanded in many fields, but attracting more students to scientific educations remains a challenge. This paper uses two studies to investigate the value of using Immersive Virtual Reality (IVR) laboratory simulations in science education. In Study 1, 99 (52 male, 47 female) 7$^{th}$ (49) and 8$^{th}$ (50) grade students between 13 and 16 years of age used an IVR laboratory safety simulation with a pre- to post-test design. Results indicated an overall increase in interest in science and self-efficacy, but only females reported an increase in science career aspirations. Study 2 was conducted with 131 (47 male, 84 female) second (77) and third (54) year high school students aged 17 to 20 and used an experimental design to compare the value of using an IVR simulation or a video of the simulation on the topic of DNA-analysis. The IVR group reported significantly higher gains from pre- to post-test on interest, and social outcome expectations than the video group. Furthermore, both groups had significant gains in self-efficacy and physical outcome expectations, but the increase in career aspirations and self-outcome expectations did not reach statistical significance. Thus, results from the two studies suggest that appropriately developed and implemented IVR simulations can address some of the challenges currently facing science education.


**Keywords:** Immersive virtual reality, gender, academic choice, science interest, science career aspirations

**Practitioner notes:**

*What is already known about this topic:*

- Science-related skills are becoming increasingly important as these are in high demand, not only in traditional science occupations, but also in other fields of work and in our daily lives. Thus, it is desirable to inspire students to pursue careers within science.
- According to the social cognitive career theory (SCCT), students' educational choice goals (i.e. career aspirations) are shaped by their interests, self-efficacy, and outcome expectations.
- Students report low levels of interest in science and several studies find that positive attitudes toward science decline with age, from primary through the secondary school years.
- Unfavorable attitudes towards science could be attributed to science education failing to engage students at a satisfactory level.
- Immersive Virtual reality (IVR) is touted for its potential to offer inspiring learning experiences that increase interest and self-efficacy.

*What this paper adds:*

- A systematic investigation of how IVR laboratory simulations can increase science interest and career aspirations in middle school (aged 13 to 16) and high school (aged 17 to 20) students.
- Evidence that IVR-based learning experiences can significantly increase students' interest in science topics.
- An indication that an IVR-based simulation led to a significant pre- to post-test increase in science aspirations among 13- to 16-year-old female students.

*Implications for practice and/or policy:*

- IVR-based simulations are specifically relevant when the goal of an educational intervention is to increase students' situational interest and social outcome expectations in a science topic.
- Provided the right instructional design, IVR might help bridge the gender difference within science education in middle school (that is, students between ages of 13 and 16).
- Although IVR-based simulations can increase situational interest, longitudinal interventions are needed to create lasting effects on career aspirations in science.

## Introduction

Students from developed countries generally report a low level of interest in science, and few express a desire to become a scientist (Sjøberg & Schreiner, 2010). Several studies report that positive attitudes toward mathematics and science tend to decline with age, from primary through the secondary school years (Marginson, Tytler, Freeman, & Roberts, 2013). These unfavorable attitudes could be attributed to the fact that science education in secondary and high school fails to engage students at a satisfactory level (Barmby, Kind, & Jones, 2008; Brotman & Moore, 2008). This may result from financial, safety, and practical limitations restricting the possibility for implementing educational experiences that truly inspire students. This is particularly relevant in science education, where scientific phenomena can be difficult to experience (e.g., the inner workings of a cell, or the consequences of ocean acidification for marine ecosystems are abstract or "hidden"; Markowitz et al., 2018). Consequently, there is often a disconnect between certain science learning activities and real-world consequences (Amgen Foundation, 2016; Lyons, 2006). Research therefore suggests that utilizing more context-based instruction, in which real-world consequences and applications of science are emphasized can increase positive attitudes towards science (Bennett, Lubben, & Hogarth, 2007). Thus, it seems that in order to inspire young people to pursue a career within the fields of science, the future of science education lies in creating engaging learning experiences that are interesting and relevant to their daily lives.

One potential cost-effective way of providing engaging learning experiences is through the use of novel technologies such as immersive virtual reality (IVR). IVR constitutes a way of digitally simulating or replicating an environment that is qualitatively very different from what other media typically deliver, because the experience of being in IVR "feels real"

(Bailenson, 2018). The visual, auditory, and sometimes sensory output devices allow learners to experience the virtual world in first person, as if they were part of it. The virtual environment can either be a model of the real world; or it can be an abstract world such as a chemical molecule or a cell. This allows students to be exposed to a different world or reality that might be too dangerous (e.g. working with nuclear power or dangerous chemicals), expensive (e.g. space traveling), or impossible in the real world (e.g. experiencing the inner workings of a cell; Bailenson, 2018; Meyer et al., 2019), making abstract phenomena easier to relate to. Thus, one of the most distinctive features of VR as an educative tool is that it provides the potential to develop a connection between learning material and students' realities (Fitzgerald & Riva, 2001). In science education, students are thereby provided an opportunity to see, explore, and experiment with scientific phenomena in a realistic 3D space that enables them to experience the marvels of science in first person.

From a societal perspective, basic scientific competencies are now in demand in many fields (e.g. to solve issues related to global warming, health and medical care; Carnevale, Smith, & Melton, 2011). Therefore, there is a strong societal incentive to attract more talented students to pursuing a science education, and a first step in that direction is to spark their interest in science at a young age (Thisgaard & Makransky, 2017).

Another important issue facing science education is related to gender differences (e.g., Sjøberg & Schreiner, 2010). According to a report by the American Association of University Women, men outnumber women in almost every science and engineering field by graduation from college, with women earning as low as 20 percent of the bachelor's degrees in some fields, such as physics, engineering, and computer science (Hill, Corbett, & St Rose, 2010). Similar statistics have been reported regarding Denmark, where the current study took place.

In Denmark, 72% of students enrolling in STEM educations in 2018 were male (Statistics Denmark, 2019). Research examining the lack of science interest among young people suggests that interest specifically drops during secondary school (Barmby et al., 2008; George, 2006; Reid & Skryabina, 2002), with females' interest decreasing more than males' (Barmby et al., 2008; Reid & Skryabina, 2002; Smith, Pasero, & McKenna, 2014).

**Main objectives of this paper:**

In this paper, we conduct two studies with the purpose of investigating the value of using IVR laboratory simulations to increase science interest and career aspirations in students aged 13 to 16 (middle school) and students aged 17 to 20 (high school). In Study 1, students between 13 and 16 years of age use an IVR laboratory safety simulation with a pre- to post-test design at a science camp. The goals of this study are to investigate if the simulation can increase interest, self-efficacy, and science aspirations; and if there are differences based on gender. In Study 2, 17 to 20-year-old high school students at a science camp use a simulation based around a crime scene investigation (CSI), where the main topic is DNA analysis. In this study we use an experimental design to compare the results of using either an IVR simulation or a video of an optimal run through the simulation. Specifically, we investigate measures of interest and science aspirations, as well as self-efficacy, and outcome expectations, and explore gender differences on these variables.

In the following, we provide an overview of relevant theory and research, and present the hypotheses investigated in Study 1. This includes an introduction to the theoretical framework for this paper, which is based on the social cognitive career theory (SCCT; Lent, Brown, & Hackett, 2002; Lent et al., 2018).

**Theoretical background:**

The SCCT provides a framework for understanding the processes through which people "(a) develop basic academic and career interests, (b) make and revise their educational and vocational plans, and (c) achieve performances of varying quality in their academic and career pursuits" (Lent, Lopez Jr, Lopez, & Sheu, 2008, p. 53). According to the SCCT, students' academic and career interests develop from learning experiences that enhance their self-efficacy and outcome expectations. Interest refers to the focused attention and affective reaction that is activated in the moment by certain environmental stimuli (Hidi & Renninger, 2006). Self-efficacy refers to beliefs about one's ability to successfully perform particular behaviors or courses of action (Lent et al., 2008), and outcome expectations refer to beliefs about the consequences of given actions (Lent et al., 2008). Thus, when students perceive themselves as being competent in performing an activity and they expect positive outcomes from this activity, they are likely to develop interest in a related field. In this way, the SCCT explains how students' career-related interests, along with self-efficacy and outcome expectations, shape their educational and occupational choice goals (in this case science aspirations), which in turn influence their choice actions (i.e., seeking entry into a certain university degree or job; Lent et al., 2018).

**Immersive Virtual Reality:**

Virtual reality (VR) is defined here as "a computer-mediated simulation that is three-dimensional, multisensory, and interactive, so that the user's experience is 'as if' inhabiting and acting within an external environment" (Burbules, 2006, p. 37). VR simulations may vary in their level of immersion, which is defined as the objective level of sensory fidelity a VR system provides (Bowman & McMahan, 2007). In this sense, immersion is highly dependent on the particular system. Thus, a VR lesson accessed through a desktop computer (often referred to as desktop VR) can be classified as being low-immersive, whereas VR accessed through head-mounted displays (HMD) offer a high level of immersion (Cummings & Bailenson, 2016; Lee & Wong, 2008). The latter, we will refer to as IVR throughout this article.

**Theory of change based on IVR research:**

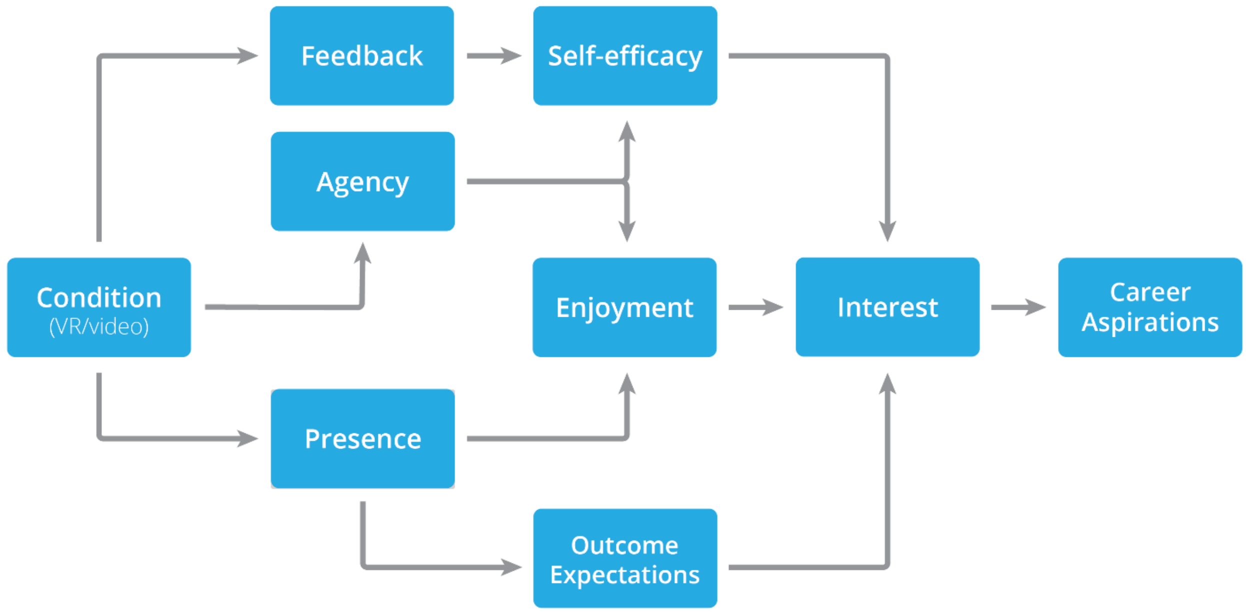



**Figure 1:** ***Illustration of the theory of change.***

Figure 1 illustrates the theory of change based on existing empirical and theoretical research that we propose in this study. Specifically, the Labster IVR interventions used for Study 1 and 2 (see Table S2 for descriptions) are assumed to increase the outcome variables of self-efficacy, interest, outcome expectations, and career aspirations through feedback, agency, presence, and enjoyment.

The two central psychological affordances of using IVR as a learning platform are a high sense of psychological presence and agency (Johnson-Glenberg, 2019; Makransky et al., 2020). The experience of being in IVR elicits high levels of psychological presence, understood as the subjective experience of being in a place or environment even when one is physically situated in another (Makransky, Lilleholt & Åaby, 2017; Witmer & Singer, 1998). Two key factors in determining the sense of presence in IVR are immersion and interaction (Mantovani, 2003). There is abundant research suggesting that IVR simulations lead to higher presence than a video which is illustrated in Figure 1 as the path between the condition and presence (e.g., Makransky et al., 2020; Meyer et al., 2019; Parong & Mayer, 2018). Presence is also known to increase enjoyment of learning activities (Makransky & Lilleholt, 2018; Makransky & Petersen, 2019) and IVR simulations are consistently reported as being more enjoyable than more conventional media (e.g., Buttussi & Chittaro, 2018; Makransky et al., 2020; Makransky & Lilleholt, 2018; Makransky et al., 2019c; Meyer et al., 2019; Parong & Mayer, 2018), which supports the path between presence and enjoyment in Figure 1. Moreover, high levels of presence in an inspiring narrative may promote positive outcome expectations of working with STEM: therefore, we propose a relation between presence and outcome expectations.

Agency is the degree of freedom and control that a learner is given to perform meaningful actions in a learning environment (Wardrip-Fruin et al., 2009). Meta-analyses have found that learners favor navigating through simulations themselves (Vogel et al., 2006), and high levels of user control result in higher estimates of self-efficacy and transfer in simulation based training (Gegenfurtner et al., 2014). Meta-analyses have found that computer-based simulation games increase self-efficacy by up to 20% (Sitzmann, 2011); and that higher levels of interaction and user control also result in higher estimates of self-efficacy (Gegenfurtner, Quesada-Pallarès, & Knogler, 2014). Furthermore, VR based lessons have been found to increase self-efficacy when VR simulations provided guided discovery learning opportunities with appropriate feedback (e.g., Buttussi & Chittaro, 2018; Makransky et al., 2016; Mayer & Mayer, 2005; Parong & Mayer, 2018; Tompson & Dass, 2000). Figure 1 depicts the relation between the learning condition and agency, since a high level of agency is expected as an outcome of the ability to interact with and perform meaningful actions in the IVR simulation (e.g., to clean a laboratory work station or perform DNA analysis and get immediate feedback from a pedagogical agent; see Table S2). Agency is related to control, and thereby to affording mastery experiences, which explains its effect on learners' self-efficacy. Furthermore, the Control-Value Theory of Achievement Emotions (CVTAE) states that achievement activities (e.g., learning activities) with high subjective value and controllability lead to enjoyment (Pekrun, 2006). Therefore, agency is suggested to also affect enjoyment.

Another essential part of how IVR simulations can increase self-efficacy is through immediate high fidelity feedback on one's actions and choices. Feedback is important for shaping self-efficacy through mastery experiences (Bandura, 1997). There is evidence that allowing students to learn by performing activities in a high-fidelity environment and gaining relevant real-time feedback can increase self-efficacy compared to more traditional learning methods (Buttussi & Chittaro, 2018; Meyer et al., 2029; Thisgaard & Makransky, 2017). There is also evidence of the importance of feedback in learning through the Labster simulations used in this study (Makransky et al., 2019b; 2018). This feedback can be accomplished by providing retrieval practice activities through multiple-choice questions (Makransky et al., 2019b) and by providing feedback using a pedagogical agent in IVR (Makransky, Wismer & Mayer, 2018).

Additionally, enjoyment is also suggested to affect interest. This is based on a study that analyzed sources of situational interest via path analyses (Chen et al., 2001), as well as the CVTAE (Pekrun, 2006). The rationale is that using IVR for presenting lessons will make learning a fun experience. Finally, the SCCT's suggested connections between self-efficacy, outcome expectations, interest, and career aspirations are depicted in the model in Figure 1.

**Study 1 Predictions:**

Based on our theory of change, we predict that the IVR simulation will increase middle school students' interest (Hypothesis 1), and self-efficacy (Hypothesis 2) related to laboratory work and safety. Additionally, the SCCT provides a useful framework for understanding how immersive VR can influence students' science-based academic and career aspirations through its influence on self-efficacy and interests. Therefore, we also predict that the IVR laboratory safety simulation will lead to a significant increase in science aspirations (Hypothesis 3). Previous research has documented the existence of a gender gap in science, technology, engineering and math (STEM) education (e.g., Brotman & Moore, 2008; Leaper et al., 2012; Shapiro & Williams, 2012). Therefore, it is relevant to investigate whether there are differences between genders on the above-mentioned variables when using IVR simulation (open research question).

## Materials and methods (Study 1)

**Sample and procedures:**

The sample consisted of 99 $7^{th}$ (49) and $8^{th}$ (50) grade students (52 male and 47 female) between the ages of 13 and 16. The study took place as part of two one-week long science camps, where students participated in different mandatory workshops – one of them being the IVR learning experience. Students had been selected by their teachers to take part in the science camp based on their interest in natural sciences. The sessions in the two different camps followed the same setup: in a lecture hall, all students were given the pre-test which included demographic characteristics as well as measures of interest, self-efficacy, and science aspirations. This was followed by a five-minute oral introduction on how to use the IVR headsets, and how to navigate in the IVR laboratory. Then students were asked to complete the IVR simulations individually. A total of 20 Samsung Gear VR headsets with matching phones had been prepared, and approximately 9 to 12 students entered the IVR simulation at the same time. Immediately after finishing the simulation, students were given the post-test which included the same measures as in the pre-test.

**Pre- and post-test:**

The pre-test was in Danish and contained demographic items concerning age, gender and grade; and scales to measure interest (Cronbach's alpha reliability of 0.77 and 0.78 in the pre-and post-test respectively), self-efficacy (alpha of 0.85 and 0.91), science aspirations (alpha of 0.92 and 0.93). A confirmatory factor analysis (Kline, 2011) indicated acceptable fit of the model (see Table S1 for more information and for a list of items).

**Immersive VR laboratory safety simulation:**

The IVR simulation was on the topic of laboratory safety and was built by the EdTech company Labster (2019a; see Figure 1). The IVR simulation was administered on Samsung Galaxy S7 or S8 phones, and stereoscopically displayed through a Samsung Gear VR HMD. The experience was optimized for IVR, so players could use the full potential of the virtual space which featured circular workbenches (see Figure 1, II). The likelihood of motion sickness was reduced by high frame-rates. The simulation was in English and lasted approximately between 15 and 20 minutes depending on how quickly students navigated through the tasks. Students interacted with the virtual environment by actively handling different lab equipment with a touch pad on the right side of the HMD. Please see Table S2 for a further description of the simulation.

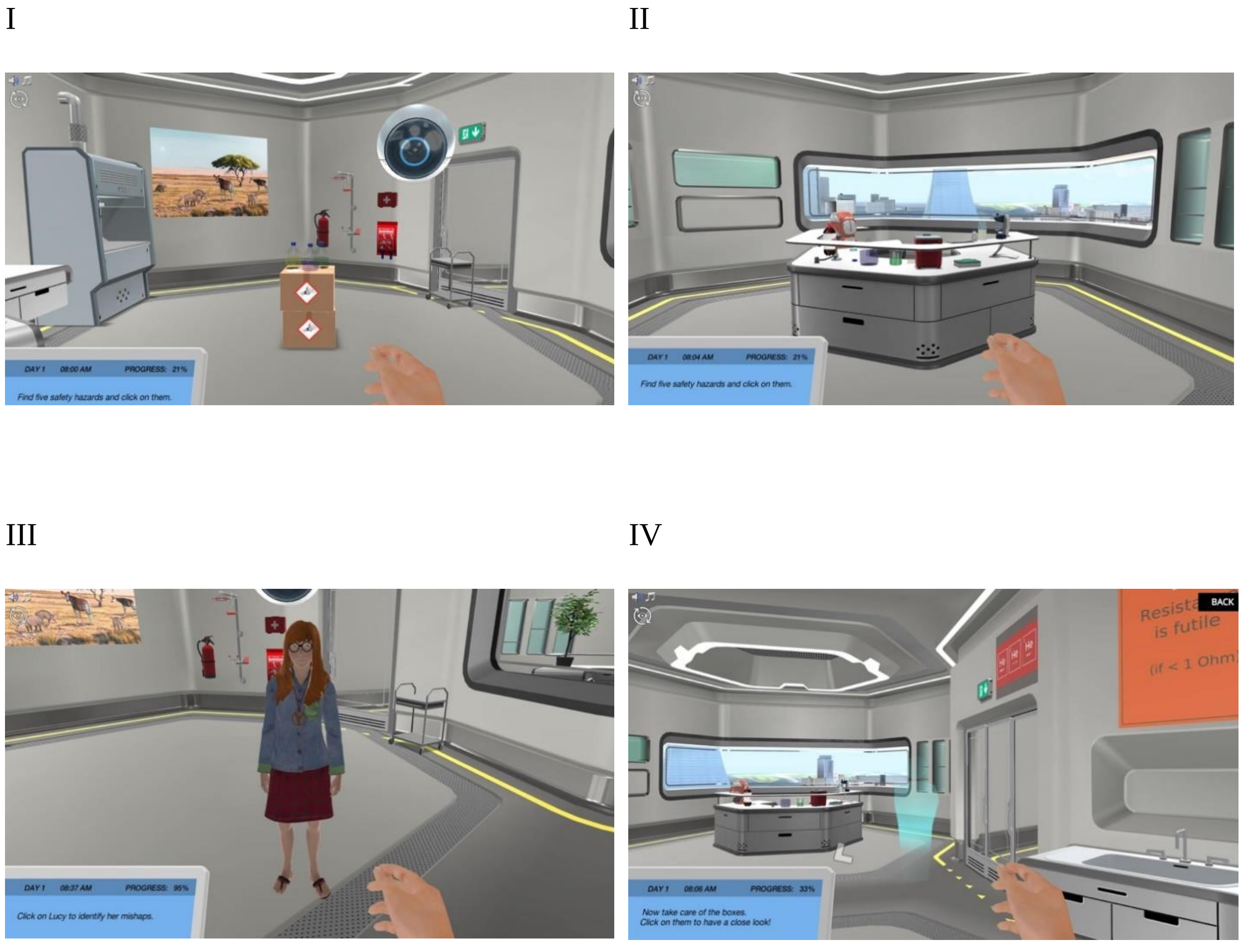



Figure 2. *Screenshots of Lab Safety Virtual Lab Simulation*. *Taken from Labster (2019a).*

## Results (Study 1)

A total of 33 students reported that they had never used IVR before, 53 reported that they had used it but for less than 2 hours, and 12 reported that they had used IVR for more than 2 hours. There were no significant differences on any of the dependent variables based on previous IVR use.

Table 1. *Mean and Standard Deviation on the pre- and post-test for the measures*

| | | Pre-test | | | Post-test | | | Significance | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Total** | **Males** | **Females** | **Total** | **Males** | **Females** | **Main effect** | **Interaction** |
| H1 | Interest | 3.99 (0.57) | 4.00 (0.59) | 3.98 (0.54) | 4.09 (0.57) | 4.06 (0.66) | 4.13 (0.45) | **p = 0.013** | p = 0.282 |
| H2 | Self-efficacy | 3.57 (0.58) | 3.67 (0.65) | 3.47 (0.47) | 3.90 (0.68) | 4.02 (0.71) | 3.77 (0.62) | **p < 0.001** | p = 0.718 |
| H3 | Science aspirations | 3.68 (0.74) | 3.82 (0.71) | 3.52 (0.74) | 3.77 (0.76) | 3.79 (0.81) | 3.73 (0.72) | p = 0.107 | **p = 0.048** |

**Hypothesis 1: The IVR simulation will lead to a significant increase in interest for laboratory work and safety**

Hypothesis 1, which was investigated with a two-group (female/male) by two time points (pre-/post-test) repeated measures ANOVA, was supported. The results which are presented in Table 1 show that there was a main effect indicating a significant increase in interest for safety from before (M = 3.99, SD = 0.57) to after using the IVR safety simulation (M = 4.09, SD = 0.57), $F_{(1,97)} = 6.354$, p = 0.013, $\eta_p^2 = 0.061$. Although females gained more interest from using the IVR safety simulation, the interaction was not significant $F_{(1,97)} = 1.171$, p = 0.282, $\eta_p^2 = 0.012$.

**Hypothesis 2: The IVR simulation will lead to a significant increase in self-efficacy for laboratory work and safety**

Hypothesis 2, was supported. The results for the repeated measures ANOVA presented in the second row in Table 1 show that there was a significant increase in self-efficacy in laboratory safety from the pre-test (M = 3.57, SD = 0.58) to post-test (M = 3.90, SD = 0.68), $F_{(1,97)} = 36.321$, p < 0.001, $\eta_p^2 = 0.272$. There was not a significant interaction between gender and time $F_{(1,97)} = 0.131$, p = 0.718, $\eta_p^2 = 0.001$.

**Hypothesis 3: The IVR simulation will lead to a significant increase in science aspirations**

Hypothesis 3, was partially supported. The results for the repeated measures ANOVA presented in the third row in Table 1 show that there was a significant interaction between gender and time, $F_{(1,97)} = 4.015$, p = 0.048, $\eta_p^2 = 0.040$. Paired samples t-tests for each gender independently showed that there was a significant increase in science aspirations for females $t_{(46)} = 2.349$, p = 0.023, d = 0.30; but not for males $t_{(51)} = 0.292$, p = 0.772, d = 0.04.

## Discussion (Study 1)

Using a pre-test/post-test design, Study 1 investigated the effect of an IVR simulation on middle school students' levels of interest, self-efficacy, and science aspirations in the context of laboratory safety. The results supported the hypotheses that the IVR simulation would significantly increase students' interest in laboratory work and safety, and self-efficacy with regard to laboratory work and safety. Furthermore, a gender difference was demonstrated with regard to science aspirations, with only females reporting a significant increase following the IVR simulation.

The finding that the IVR simulation led to an increase in interest is consistent with previous research suggesting that immersive VR-based learning experiences can instigate interest (Makransky et al., 2019a; 2019b; Meyer, Omdahl, & Makransky, 2019; Parong & Mayer, 2018). The finding that the IVR simulation increased students' self-efficacy is consistent with earlier research (Buttussi & Chittaro, 2018; Parong & Mayer, 2018) which suggests that interacting in an IVR learning environment can provide high fidelity mastery experiences. The results are consistent with mastery experiences being highlighted as the most effective way of developing self-efficacy (Bandura, 1997).

While the results of using the lab safety simulation had positive effects for both genders on interest and self-efficacy, only females demonstrated a significant increase in science aspirations in this study. According to the SCCT, the significantly increased levels of self-efficacy and interest would affect choice goals, in this case leading to increased levels of science aspirations among both genders. The fact that an increase in science aspirations is only observed in females might be a result of females initially reporting lower levels of science aspirations relative to males. Following the IVR simulation, both female and male students had similar levels of science aspirations, although Table 1 shows that females still had lower science aspirations as compared to males. Nevertheless, the result demonstrates the potential for IVR simulations to contribute to balance out the gender difference in STEM fields.

## Limitations

Study 1 has a number of limitations. Firstly, it employs a pre-/post-test design with no control group. Research suggests that participating in out of school activities such as science camps can have a positive effect on student interest in STEM (Kong et al., 2014; Young et al., 2017), meaning that it is not possible to ascribe the positive results specifically to the IVR simulation with certainty. Furthermore, a large synthesis of 1200 meta-analyses concerning various influences on academic achievements in education settings found that nearly all interventions work as compared to not implementing an intervention (Hattie, 2015). This means that any intervention could have resulted in higher post-test results and, consequently, a control group is necessary in order to investigate the relative impact of different interventions. Furthermore, several variables that are important in SCCT were not included because the teachers responsible for the camps were worried about exhausting the young students with questionnaires and tests rather than giving them rich educational activities. For instance, we did not assess outcome expectations in this study because the IVR simulation was developed specifically for increasing self-efficacy, and it did not include features targeting students' outcome expectations. The SCCT hypothesizes that students' outcome expectations play an important role in interest development. Therefore, a greater impact on interest and science aspirations could be achieved if an IVR science simulation included a narrative that provided students with a more realistic picture of a particular STEM career. Factors such as the monetary outcomes associated with a science career (physical outcome expectations), the social possibilities with such a career (social-outcome expectations), and finally the self-evaluative outcomes, such as self-satisfaction or gaining expertise associated with such a career (self-outcome expectations), all play an important role in career aspirations (Lent, Brown, & Hackett, 1994). A final concern was if the positive results would also appear with a different sample using different IVR content.

## Study 2

Study 2 was intended to be an extension of Study 1, and to address some of the above-mentioned limitations. Consequently, we wanted to investigate whether an IVR simulation (experimental group) would result in stronger pre- to post-intervention changes on the following variables: interest, self-efficacy, and science aspirations (from Study 1); as well as outcome expectations (physical, social, and self) as compared to a video of an optimal run through the same content (control group). A video of an optimal simulation experience was chosen because it would allow us to specifically isolate interaction (student interaction vs. passive viewing), immersion (immersive vs. desktop interface). These two factors are related to the psychological affordances of agency, presence, and feedback as outlined in Figure 1. Furthermore, we wanted to investigate the consequences of using an IVR simulation with an older population (17 to 20 years) as well as on a different science topic (analyzing a DNA sample by conducting a PCR experiment); and, finally, to investigate whether there were differences between genders in the above-mentioned variables.

Research comparing IVR simulations to less immersive media has found that IVR simulations are superior in terms of their effect on interest (Parong & Mayer, 2018). Therefore, we predict an interaction between media (VR/video) and time (pre-/post-test), with the IVR simulation leading to significantly higher interest development than the video (Hypothesis 1). Similarly, since a key factor contributing to higher self-efficacy is gaining experience in performing certain behaviors and receiving feedback (Bandura, 1997), we predict an interaction between media (IVR/video) and time (pre-/post-test) with the IVR simulation leading to significantly higher self-efficacy than the video due to the amount of agency it provides (Hypothesis 2). Although we could not find any literature on the use of IVR vs. a video on outcome expectations, we expect that the higher level of presence in IVR provides a more realistic picture of science activities, which could increase outcome expectations as highlighted in Figure 1. Therefore, we also predict an interaction between media (IVR/video) and time (pre-/post-test), with the IVR simulation leading to significantly higher outcome expectations than the video (Hypothesis 3). Since IVR is posited to have a superior effect on self-efficacy, outcome expectations, and interest, based on the SCCT we predict an interaction between media (IVR/video) and time (pre-/post-test), with the IVR simulation leading to significantly higher science aspirations than the video (Hypothesis 4). Finally, since we observed differences between genders in Study 1, we wish to investigate if gender similarly affects any of the predictions made in Study 2.

## Materials and methods (Study 2)

**Sample and procedures**

The sample consisted of 131 second (77) and third year (54) high school students (47 males, and 84 females) between the ages of 17 and 20. The study took place during three one-week-long science camps, where students participated in different mandatory workshops including the VR-workshop. Students had been selected by their teachers to take part in the science camp based on their interest in natural sciences. In the three different camps the sessions followed an identical setup: prior to playing the simulations, all students were gathered in a lecture hall where they received randomized ID numbers and completed the pre-test which included demographic characteristics, as well as the measures of interest, self-efficacy, outcome expectations, and science aspirations. Students were then divided into two groups based the random ID number that was generated for the study: IVR CSI simulation (N = 64), CSI video condition (N = 67). Students were separated and were not aware that the sessions were different until after the experiment. The students in the CSI video condition were gathered in one classroom and watched the CSI video on a large screen. The students using the IVR simulation were gathered in a second classroom and received a five-minute oral introduction on how to use the VR headsets and navigate in the VR laboratory before entering the simulation. As in Study 1, 20 Samsung Gear VR headsets with matching phones had been set up, and students played the VR-simulations individually with 9 to 12 students using the VR-simulation at the same time. Immediately after the lesson, all students in the two groups were given the post-test separately. The post-test consisted of the same scales as the pre-test including interest, self-efficacy, outcome expectations, and science aspirations.

**Pre- and post-test**

The pre-test questionnaire was in Danish and consisted of demographic items concerning gender and grade, and scales to measure interest (Cronbach's alpha reliability was 0.65 in the pre-test and 0.74 in the post test respectively), self-efficacy (alpha = 0.80, 0.85), outcome expectations (physical; alpha = 0.71, 0.74; social; alpha = 0.63, 0.64; self; alpha = 0.77, 0.83), and science aspirations (alpha = 0.82, 0.83). A confirmatory factor analysis indicated acceptable fit of the model (see Table S1 for more information and for a list of items).

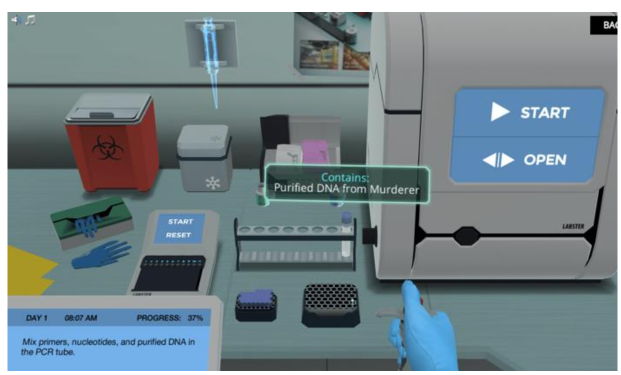
**IVR and video crime scene investigation simulations**

The IVR simulation was built by the EdTech company Labster using a CSI as frame for teaching students about DNA analysis (Labster, 2019b; see Figure 2). The virtual environment instructional approach and interaction were similar to the laboratory safety simulation in Study 1, although the topic and narrative of the simulation were different. The CSI simulation lasted approximately 12 minutes. See Table S2 for a further description of the simulation.

The video condition consisted of a recorded version of the same simulation where students observed an optimized run through the simulation lasting approximately 12 minutes. The video was carefully constructed ensuring that students had time to see all relevant information, although the video format prevented them from controlling actions themselves.

I


II



III IV


Figure 3. *Screenshots of Polymerase Chain Reaction Virtual Lab Simulation*. *Taken from Labster (2019b).*

**Results (Study 2)**

Before investigating the research questions, we investigated whether the two groups differed on basic characteristics. For the IVR group, a total of 38 students reported that they had never used IVR before, 23 reported that they had used it but for less than 2 hours, and three reported that they had used IVR for more than 2 hours. There were no significant differences on any of the dependent variables based on previous IVR use.

A chi-square test indicated that the groups did not differ significantly in the proportion of men and women, $X^2$ (N = 131) = 1.165, $p$ = .280.

The influence of gender on all of the dependent variables was investigated with two group (female/male) by two time points (pre-/post-test) repeated measures ANOVAs for the IVR and video groups separately, and as two (female/male) by two (VR/video) by two (pre-/post-test) repeated measures ANOVAs. None of the tests were significant, indicating that there were no interactions between gender and any of the dependent variables in the study. Therefore, we investigate the hypotheses independently of gender.

**Hypotheses 1-4: The interventions will have a significant positive impact on students' interest, self-efficacy, outcome expectations, and science aspirations with the greatest effects occurring for the IVR group.**

Hypotheses 1 through 4 were investigated with two group (VR/video) by two time points (pre-/post-test) repeated measures ANOVAs, and the results are presented in Table 2.

Hypothesis 1 was supported. The results presented in the first row in Table 2 show that there was a significant interaction between media and time, $F_{(1,129)} = 4.100$, $p = 0.045$, $\eta_p^2 = 0.031$ for the dependent variable of interest. Paired samples t-tests for each instructional method independently showed that there was a significant increase in interest for the IVR group $t_{(63)} = 4.644$, $p < 0.001$, $d = 0.35$; but not for the video group $t_{(66)} = 1.798$, p = 0.077, $d = 0.13$.

Hypothesis 2 was partially supported. The results indicate that there was a main effect with significant increase in self-efficacy across groups $F_{(1,129)} = 27.266$, $p < 0.001$, $\eta_p^2 = 0.174$. There was not a significant interaction between instructional method and time $F_{(1129)} = 1.669$, $p = 0.199$, $\eta_p^2 = 0.013$. Paired samples t-tests for each instructional method independently showed that there was a significant increase in self-efficacy for the IVR group $t_{(63)} = 3.819$, $p < 0.001$, d = 0.30; as well as the video group $t_{(66)} = 3.637$, $p = 0.001$, $d = 0.20$.

Hypothesis 3 was partially supported. There was a significant main effect for the outcome variable of physical outcome expectations, $F_{(1,129)} = 12.886$, $p < 0.001$, $\eta_p^2 = 0.091$ but not a significant interaction $F_{(1,129)} = 0.472$, $p = 0.493$, $\eta_p^2 = 0.004$. Paired samples t-tests for each instructional method independently showed that there was a significant increase in physical outcome expectations for the IVR group $t_{(63)} = 2.815$, $p = 0.001$, $d = 0.25$; as well as the video group $t_{(66)} = 2.217$, p = 0.030, $d = 0.15$. There was a significant interaction for the social-outcome expectations $F_{(1,129)} = 4.312$, $p = 0.040$, $\eta_p^2 = 0.032$. Paired samples t-tests for each instructional method independently showed that there was a significant increase in social-outcome expectations for the IVR group $t_{(63)} = 4.167$, $p < 0.001$, $d = 0.31$; but not for the video group $t_{(66)} = 0.980$, p = 0.331, $d = 0.09$. Finally, the main effect $F_{(1,129)} = 3.803$, $p = 0.053$, $\eta_p^2 = 0.029$ or interaction $F_{(1,129)} = 0.691$, $p = 0.407$, $\eta_p^2 = 0.005$ did not reach statistical significance for the self-outcome expectations scale. Paired samples t-tests for each instructional method independently showed that the increase in self-outcome expectations for the IVR group was on the borderline of significance $t_{(63)} = 1.994$, $p = 0.050$, $d = 0.11$; and non-significant for the video group $t_{(66)} = 0.789$, p = 0.437, $d = 0.07$.

*Table 2. Mean and Standard Deviation on the pre- and post-test for the measures used in the study.*

| | | **Pre-test** | | **Post-test** | | **Significance** | |
|---|---|---|---|---|---|---|---|
| | | **VR** | **Video** | **VR** | **Video** | **Main effect** | **Interaction** |
| H1 | Interest | 3.83 (0.51) | 3.87 (0.50) | 4.01 (0.52) | 3.94 (0.57) | ***p* < 0.001** | ***p* = 0.045** |
| H2 | Self-efficacy | 3.93 (0.55) | 4.02 (0.49) | 4.10 (0.58) | 4.12 (0.49) | ***p* < 0.001** | *p* = 0.199 |
| H3 | OE physical | 3.90 (0.47) | 3.99 (0.51) | 4.02 (0.48) | 4.07 (0.53) | ***p* < 0.001** | *p* = 0.493 |
| | OE social | 4.03 (0.48) | 4.22 (0.47) | 4.18 (0.49) | 4.26 (0.45) | ***p* = 0.001** | ***p* = 0.040** |
| | OE self | 4.38 (0.45) | 4.40 (0.43) | 4.43 (0.43) | 4.43 (0.49) | *p* = 0.053 | *p* = 0.407 |
| H4 | Science aspirations | 4.02 (0.80) | 4.14 (0.75) | 4.05 (0.81) | 4.18 (0.74) | *p* = 0.229 | *p* = 0.742 |

Hypothesis 4 was not supported. The results show that there was not a main effect $F_{(1,129)} = 1.458$, $p = 0.229$, $\eta_p^2 = 0.011$ or interaction $F_{(1,129)} = .109$, $p = 0.742$, $\eta_p^2 = 0.001$ for science aspirations. We conclude that the positive pre- to post-test changes on science aspirations did not reach significance for the IVR or video groups, and there was no interaction.

## Discussion (Study 2)

The results of Study 2 supported the hypotheses that the IVR simulation would lead to significant increases in students' interest; and that these effects would be larger than those of the video. The hypotheses regarding self-efficacy and outcome expectations were partially supported. Both the VR simulation and the video led to significant increases in self-efficacy and physical outcome expectations. Only the VR simulation, however, had a significant effect on social-outcome expectations; and neither had an effect on self-outcome expectations. No significant main effects or interactions were found for science aspirations. There were also no significant interactions between gender and any of the dependent variables in the study.

A major finding in Study 2 was that there was a significant interaction between the instructional method and time, with the IVR group having a significantly higher increase in interest than the video group. This supports the findings in Study 1, and suggests that IVR simulations can spark interest in middle school students and are more effective than a video in creating interest in science for high school students.

The IVR simulation and the video led to significant increases in self-efficacy in Study 2. This supports the finding from Study 1, as well as other studies documenting that IVR simulations can increase self-efficacy (e.g., Buttussi & Chittaro, 2018). However, the results do not support the prediction that the increased interaction and immersion afforded by IVR would result in higher self-efficacy in high school students than the video.

The results in Study 2 indicated that IVR and video had mixed effects on the different dimensions of students' outcome expectations. Both the IVR group and the video group showed comparable significant increases in physical outcome expectations, meaning that both interventions increased the students' beliefs about the positive physical consequences, such as material goods or quality of workplace, of choosing an education within science. Only the IVR group, however, had a significant increase in social-outcome expectations (i.e. beliefs about the positive, social consequences associated with a biology education, such as helping other people or gaining high social status). Finally, none of the groups showed any increase in self-outcome expectations of choosing a biology education (such as achieving satisfaction with oneself or being able to challenge oneself). One possible explanation as to this lack of effect could be that the CSI simulation was not designed with this particular educational outcome in mind.

According to the SCCT, increasing students' outcome expectations, self-efficacy, and interest in working with genetics in the laboratory would increase their intentions to pursue a science-related career path. However, a significant increase in science aspirations was not found following the intervention. This could be explained with reference to the Four-Phase Model of Interest Development (Hidi & Renninger, 2006; Renninger & Hidi, 2016). The type of interest that predicts career aspirations could be classified as well-developed individual interest, which corresponds to the last phase of interest development. A single VR learning session, such as the one described in this study, is more likely to trigger situational interest, which corresponds to the first phase of interest development. Therefore, teachers using VR in practice should organize their lessons in a way that supports further interest development, e.g. by creating meaningful and interactive learning activities in order to foster well-developed individual interest. The fact that we did not see an increase in students' science aspirations in Study 2 could also be attributed to the educational choices the students had already made beforehand. Most of them had already selected a science-related study program as part of their high school education and had high science aspirations from the outset of the intervention.

Finally, the results of Study 2 show that there were no interactions between any of the variables and gender which suggests that the findings outlined above are consistent across genders.

## Limitations and future directions

In this study, the sample consisted of students participating at a science camp, picked out for participation by their teachers due to a keen interest in natural science. A clear limitation is therefore related to the ability to generalize the findings to other students and other settings, although our results suggest that a meaningful IVR intervention would have a larger effect on students who are low on initial levels of science aspirations. Future research should therefore investigate if a general population of students with no particular initial interest in natural science experiences similar benefits from IVR simulations.

No long-term investigation of any of the variables was conducted, meaning that another limitation concerns uncertainty with regard to the long-term effects. Especially when measuring interest, and situational interest, it is important to investigate whether this interest drops, and if so, when. Future research should therefore conduct follow-up assessments of these variables to investigate whether the changes persist. Previous research has suggested that simulations work best in combination with other approaches to teaching (Merchant et al., 2014; Smetana & Bell, 2012), that is, IVR interventions can work effectively to encourage students in developing science interests, self-efficacy, motivation, and career aspirations; but may not stand alone. Therefore, future research should also examine the effect of well-integrated IVR simulations as part of a larger study program.

## Conclusions

The two studies investigated the role of VR in enhancing science education, specifically focusing on measures of interest, self-efficacy, outcome expectations; as well as science aspirations among middle and high school students. They demonstrate that IVR can successfully increase students' self-efficacy and interests in two science-related topics: laboratory safety and DNA analysis. This is highly valuable for the future of science, inasmuch as developing a science-related interest is a crucial first step in pursuing an education within science or related fields. Nevertheless, only female middle school students (Study 1) reported an increase in science aspirations following the intervention. Thus, following the IVR simulation, female students had similar levels of science aspirations as males. This could indicate that provided with the right instructional design, IVR might help bridge the gender difference within science education. The fact that levels of interest were increased in both studies, yet only Study 1 demonstrated a change in science aspirations for female middle school students, accentuates the importance of when and how IVR based educational interventions are used. It further supports the idea that IVR can trigger interest, positioning IVR as a useful technology for improving the future of science education by providing engaging learning experiences that enhance interest, self-efficacy, and outcome expectations among students.

**References**

Amgen Foundation. (2016). Students on STEM. *Change the Equation*. Retrieved from http://changetheequation.org/students-stem%5Cnhttp://files/649/students-stem.html

Bailenson. J. (2018). *Experience on Demand: What Virtual Reality Is, how it Works, and what it Can Do.* New York, NY: Norton & Company.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Barmby, P., Kind, P. M., & Jones, K. (2008). Examining Changing Attitudes in Secondary School Science. *International Journal of Science Education*, *30*(8), 1075–1093. https://doi.org/10.1080/09500690701344966

Bennett, J., Lubben, F., & Hogarth, S. (2007). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education*, *91*(3), 347–370. https://doi.org/10.1002/sce.20186

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: How much immersion is enough? *Computer*, *40*(7), 36–43. https://doi.org/10.1109/MC.2007.257

Brotman, J. S., & Moore, F. M. (2008). Girls and science: A review of four themes in the science education literature. *Journal of Research in Science Teaching, 45*(9), 971–1002. https://doi.org/10.1002/tea.20241

Burbules, N. C. (2006). Rethinking the Virtual BT  - The International Handbook of Virtual Learning Environments. In J. Weiss, J. Nolan, J. Hunsinger, & P. Trifonas (Eds.) (pp. 37–58). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-3803-7_1

Buttussi, F., & Chittaro, L. (2018). Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics*, *24*(2), 1063–1076. https://doi.org/10.1109/TVCG.2017.2653117

Carnevale, A. P., Smith, N., & Melton, M. (2011). STEM: Science Technology Engineering Mathematics. *Georgetown University Center on Education and the Workforce*.

Chen, A., Darst, P. W., & Pangrazi, R. P. (2001). An examination of situational interest and its sources. *British Journal of Educational Psychology*, *71*, 383-400. https://doi.org/10.1348/000709901158578

Cummings, J. J., & Bailenson, J. N. (2016). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology*, *19*(2), 272–309. https://doi.org/10.1080/15213269.2015.1015740

Dewey, J. (1913). *Interest and effort in education. Interest and effort in education.* Boston, MA, US: Houghton Mifflin Company. https://doi.org/10.1037/14633-000

Fitzgerald, M., & Riva, G. (2001). Virtual reality. In L. Beolchi (Ed.), *Telemedicine Glossary* (pp. 327–329). European Commission-DG INFSO.

Gegenfurtner, A., Quesada-Pallarès, C., & Knogler, M. (2014). Digital simulation-based training: A meta-analysis. *British Journal of Educational Technology*, *45*(6), 1097–1114. https://doi.org/10.1111/bjet.12188

George, R. (2006). A cross-domain analysis of change in students' attitudes toward science and attitudes about the utility of science. *International Journal of Science Education*, *28*(6), 571–589. https://doi.org/10.1080/09500690500338755

Hattie, J. (2015). The applicability of Visible Learning to higher education. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 79–91. https://doi.org/10.1037/stl0000021

Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hill, C., Corbett, C., & St Rose, A. (2010). *Why so few? Women in science, technology,* *engineering, and mathematics.* Washington, DC: AAUW.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. Guilford Publications.

Labster. (2019a). Lab Safety. Retrieved November 30, 2019, from https://www.labster.com/simulations/lab-safety/

Labster. (2019b). Polymerase Chain Reaction. Retrieved November 30, 2019, from https://www.labster.com/simulations/polymerase-chain-reaction/

Leaper, C., Farkas, T., & Brown, C. S. (2012). Adolescent Girls' Experiences and Gender-Related Beliefs in Relation to Their Motivation in Math/Science and English. *Journal of Youth and Adolescence*, *41*(3), 268–282. https://doi.org/10.1007/s10964-011-9693-z

Lee, E. A.-L., & Wong, K. W. (2008). A Review of Using Virtual Reality for Learning BT - Transactions on Edutainment I. In Z. Pan, A. D. Cheok, W. Müller, & A. El Rhalibi (Eds.) (pp. 231–241). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-69744-2_18

Lent, R. W., Lopez Jr, A. M., Lopez, F. G., & Sheu, H. B. (2008). Social cognitive career theory and the prediction of interests and choice goals in the computing disciplines. *Journal of Vocational Behavior*, *73*(1), 52–62. https://doi.org/10.1016/j.jvb.2008.01.002

Lent, Robert W., Brown, D., & Hackett, G. (2002). Social Cognitive Career Theory. In *Career Choice and Development* (4th ed., pp. 255–312). San Fransisco: Jossey-Bass.

Lent, Robert W., Sheu, H. Bin, Miller, M. J., Cusick, M. E., Penn, L. T., & Truong, N. N. (2018). Predictors of science, technology, engineering, and mathematics choice options: A meta-analytic path analysis of the social-cognitive choice model by gender and race/ethnicity. *Journal of Counseling Psychology*. https://doi.org/10.1037/cou0000243

Lent, Robert W, Brown, S. D., & Hackett, G. (1994). Toward a Unifying Social Cognitive Theory of Career and Academic Interest, Choice, and Performance. *Journal of Vocational Behavior*, *45*(1), 79–122. https://doi.org/https://doi.org/10.1006/jvbe.1994.1027

Lyons, T. (2006). Different countries, same science classes: Students' experiences of school science in their own words. *International Journal of Science Education*, *28*(6), 591–613. https://doi.org/10.1080/09500690500339621

Makransky, G., Andreasen, N. K, Baceviciute S., & Mayer R. M. (2020). Immersive Virtual Reality Increases Liking but Not Learning with a Science Simulation and Generative Learning Strategies Promote Learning in Immersive Virtual Reality. *Journal of Educational Psychology*. DOI: 10.1037/edu0000473

Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019a). Motivational and Cognitive Benefits of Training in Immersive Virtual Reality Based on Multiple Assessments. *Journal of Computer Assisted Learning*. http://doi.org/10.1111/jcal.12375

Makransky, G., Mayer, R. E., Nørremølle, A., Lopez Cordoba, A., Wandall, J., Bonde, M. (2019b). Investigating the feasibility of using assessment and explanatory feedback in desktop virtual reality simulations. *Educational Technology Research and Development.* https://doi.org/10.1007/s11423-019-09690-3

Makransky, G., & Petersen, G. B. (2019). Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. *Computers & Education*, *134*, 15–30. https://doi.org/10.1016/j.compedu.2019.02.002

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019c). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Makransky, G., Thisgaard, M. W., & Gadegaard, H. (2016). Virtual simulations as preparation for lab exercises: Assessing learning of key laboratory skills in microbiology and improvement of essential non-cognitive skills. *PLoS ONE*, *11*(6), 1–11. https://doi.org/10.1371/journal.pone.0155895

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development*, *2010*, 1–24. https://doi.org/10.1007/s11423-018-9581-2

Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior*, *72*, 276-285. http://doi.org/10.1016/j.chb.2017.02.066

Makransky, G., Wismer, P. Mayer, R. E. (2018). A Gender Matching Effect in Learning with Pedagogical Agents in an Immersive Virtual Reality Science Simulation. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12335

Mantovani, F. (2003). 12 VR learning: Potential and challenges for the use of 3d environments in education and training. In G. Riva & C. Galimberti (Eds.), *Towards cyberpsychology: Mind, cognitions and society in the internet age* (pp. 208–225). Amsterdam: IOS Press.

Marginson, S., Tytler, R., Freeman, B., & Roberts, K. (2013). *STEM: country comparisons: international comparisons of science, technology, engineering and mathematics (STEM) education. Final report.* Melbourne, Vic: Australian Council of Learned Academies. Retrieved from http://dro.deakin.edu.au/view/DU:30059041

Mayer, R., & Mayer, R. E. (2005). *The Cambridge handbook of multimedia learning*. Cambridge university press.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers and Education*, *70*, 29–40. https://doi.org/10.1016/j.compedu.2013.07.033

Meyer, O. A., Omdahl, M. K., & Makransky, G. (2019). Investigating the effect of pre-training when learning through immersive virtual reality and video: A media and methods experiment. *Computers & Education*, *140*. https://doi.org/10.1016/J.COMPEDU.2019.103603

Parong, J., & Mayer, R. (2018). Learning Science in Immersive Virtual Reality Journal of Educational Psychology. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000241

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, *18*, 315-341. https://doi.org/10.1007/s10648-006-9029-9

Reid, N., & Skryabina, E. a. (2002). Attitudes towards Physics. *Research in Science & Technological Education*, *20*(1), 67–81. https://doi.org/10.1080/02635140220130939

Renninger, K. A., & Hidi, S. E. (2016). *The Power of Interest for Motivation and Engagement*. New York: Routledge. https://doi.org/10.4324/9781315771045

Shapiro, J. R., & Williams, A. M. (2012). The Role of Stereotype Threats in Undermining Girls' and Women's Performance and Interest in STEM Fields. *Sex Roles*, *66*(3–4), 175–183. https://doi.org/10.1007/s11199-011-0051-0

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology*, *64*(2), 489–528.

https://doi.org/10.1111/j.1744-6570.2011.01190.x

Sjøberg, S., & Schreiner, C. (2010). *The ROSE project: An overview and key findings*. Oslo: University of Oslo.

Smetana, L. K., & Bell, R. L. (2012). Computer Simulations to Support Science Instruction and Learning: A critical review of the literature. *International Journal of Science Education*, *34*(9), 1337–1370. https://doi.org/10.1080/09500693.2011.605182

Smith, T. J., Pasero, S. L., & McKenna, C. M. (2014). Gender Effects on Student Attitude Toward Science. *Bulletin of Science, Technology & Society*, *34*(1–2), 7–12. https://doi.org/10.1177/0270467614542806

Statistics Denmark. (2019). STEM-studerende næsten fordoblet på ti år. [STEM students almost doubled in ten years]. Accessed March 5 2020. Retrieved from https://www.dst.dk/da/Statistik/nyt/NytHtml?cid=33210.

Thisgaard, M., & Makransky, G. (2017). Virtual learning simulations in high school: Effects on cognitive and non-cognitive outcomes and implications on the development of STEM academic and career choice. *Frontiers in Psychology*, *8*, 1–13. https://doi.org/10.3389/fpsyg.2017.00805

Tompson, G. H., & Dass, P. (2000). Improving Students' Self-Efficacy in Strategic Management: The Relative Impact of Cases and Simulations. *Simulation & Gaming*, *31*(1), 22–41. https://doi.org/10.1177/104687810003100102

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. a., Muse, K., & Wright, M. (2006). Computer Gaming and Interactive Simulations for Learning: a Meta-Analysis. *Journal of Educational Computing Research*, *34*(3), 229–243. https://doi.org/10.2190/FLHV-K4WA-WPVQ-H0YM

Wardrip-Fruin, N., Mateas, M., Dow, S., & Sali, S. (2009, September). Agency reconsidered. In Breaking New Ground: Innovation in Games, Play, Practice and Theory: *Proceedings of the 2009 Digital Games Research Association Conference*,1-9. London, UK

Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *7*(3), 225–240. https://doi.org/10.1162/105474698565686

**Table S1: Items used in the pre- and post-tests in Study 1 and Study 2**

**Items used in Study 1.**

Validity: To investigate the factor structure of the measures indicated tolerable fit in the pre-test (CFI = 0.971; TLI = 0.967; RMSEA = 0.108) and acceptable fit in the post-test (CFI = 0.992; TLI = 0.991; RMSEA = 0.080). All instruments were in Danish.

| | | |
|---|---|---|
| Interest 1 | I am interested in conducting experiments with chemicals in a laboratory | Source: Thisgaard & Makransky, 2017<br><br>Reliability: .77 (pre-test), .78 (post-test) |
| Interest 2 | I am interested in investigating effects and use of different chemicals | |
| Interest 3 | I am interested in evaluating, approving or analyzing laboratory results | |
| Interest 4 | I am interested in making sure safety precautions are taken in the laboratory | |
| Interest 5 | I am interested in evaluating data from the laboratory by conducting relevant mathematical or statistical calculations and analysis | |
| Interest 6 | I am interested in making sure no one is hurt or harmed during experiments in the laboratory | |
| Self-efficacy 1 | I believe that I will be very successful in laboratory work | Source: Pintrich, Smith, Garcia, & McKeachie, 1991<br><br>Reliability: .85 (pre-test), .91 (post-test) |
| Self-efficacy 2 | I am confident that I can manage the most difficult safety tasks in laboratory experiments | |
| Self-efficacy 3 | I am confident that I can understand the basic terminology concerning laboratory safety | |
| Self-efficacy 4 | I am confident that I can understand the most complex topics in laboratory safety | |
| Self-efficacy 5 | I am confident that I will be successful in assignments and at the exam in laboratory safety | |
| Self-efficacy 6 | I expect to be very successful in the current course about laboratory safety | |
| Self-efficacy 7 | I am confident that I can manage the skills I am being taught in this course about laboratory safety | |
| STEM aspirations 1 | I am planning to seek out further learning opportunities (for example classes and courses) with a scientific content (physics, mathematics, chemistry, biology, medicine, biochemistry etc.) | Source: Thisgaard & Makransky, 2017<br><br>Reliability: .92 (pre-test), .93 (post-test) |
| STEM aspirations 2 | I am planning to apply for an education with a scientific content (physics, mathematics, chemistry, biology, medicine, biochemistry etc.) | |
| STEM aspirations 3 | I am planning to investigate career opportunities within a scientific field (physics, mathematics, chemistry, biology, medicine, biochemistry etc.) | |

| | | |
|---|---|---|
| STEM aspirations 4 | I am planning to apply for a scientific study program in high school (physics, mathematics, chemistry, biology etc.) | |
| STEM aspirations 5 | I am planning to apply for a youth education with a scientific content (physics, mathematics, chemistry, biology, medicine, biochemistry etc.) | |
| STEM aspirations 6 | I am planning to apply for an internship where I have the opportunity to be engaged in science (physics, mathematics, chemistry, biology, medicine, biochemistry etc.) | |

**Items used in Study 2**

Validity: A CFA to investigate the factor structure of the measures indicated acceptable fit in the pre-test (CFI = 0.953; TLI = 0.947; RMSEA = 0.83) and post-test (CFI = 0.978; TLI = 0.976; RMSEA = 0.74). All instruments were in Danish.

| | | |
|---|---|---|
| Interest 1 | I'm interested in collecting and analyzing biological data | Source: Thisgaard & Makransky, 2017<br><br>Reliability: .65 (pre-test), .74 (post-test) |
| Interest 2 | I'm interested in evaluating, analyzing, or interpreting genetic laboratory results | |
| Interest 3 | I'm interested in examining and handling biological material. | |
| Interest 4 | I'm interested in evaluating genetic data by conducting relevant mathematical or statistical calculations and analyses | |
| Interest 5 | I'm interested in analyzing, identifying, and classifying biological material | |
| Interest 6 | I'm interested in learning about basic biological principles and theories | |
| STEM aspirations 1 | I am planning to apply for an education with biological/biochemical/biotechnological content (biology, medicine, veterinary medicine, biomedicine, genetics, zoology etc.) | Source: Thisgaard & Makransky, 2017<br><br>Reliability: .80 (pre-test), .82 (post-test) |
| STEM aspirations 2 | I am planning to pursue a career within biology/biochemistry/biotechnology (biologist, doctor, veterinarian, biomedical scientist, geneticist, zoologist etc.) | |
| STEM aspirations 3 | I am planning to pursue more learning possibilities (for instance subjects and classes) with biological/biochemical/biotechnological content | |
| STEM aspirations 4 | I am planning to explore career opportunities within the biological/biochemical/biotechnological field | |

| | | |
|---|---|---|
| Self-efficacy 1 | I believe I will get a very good grade in biology/biotechnology | Source: Pintrich, Smith, Garcia, & McKeachie, 1991<br><br>Reliability: .80 (pre-test), .85 (post-test) |
| Self-efficacy 2 | I am certain that I can understand the most difficult literature in biology/biotechnology | |
| Self-efficacy 3 | I am certain that I can understand the basic concepts that are being taught in biology/biotechnology | |
| Self-efficacy 4 | I am certain that I can understand the most complex topics I am introduced to in biology/biotechnology | |
| Self-efficacy 5 | I am certain that I will do well on assignments and at the exam in biology/biotechnology | |
| Self-efficacy 6 | I expect that I will do well in biology/biotechnology | |
| Self-efficacy 7 | I am certain that I can master the skills that are being taught in biology/biotechnology | |
| Outcome expectations (physical) 1 | I can get an attractive pay with an education within biology/biochemistry/biotechnology | Source: Thisgaard & Makransky, 2017; Bandura, 1986; Lent et al., 1994<br><br>Reliability: .71 (pre-test), .74 (post-test) |
| Outcome expectations (physical) 2 | I can have a high degree of job security with an education within biology/biochemistry/biotechnology | |
| Outcome expectations (physical) 3 | I can obtain material goods and advantages with an education within biology/biochemistry/biotechnology | |
| Outcome expectations (physical) 4 | I can have a good workplace with an education within biology/biochemistry/biotechnology | |
| Outcome expectations (physical) 5 | I can have a workplace close to where I wish to live with an education within biology/biochemistry/biotechnology | |
| Outcome expectations (social) 1 | I can help other people with an education within biology/biochemistry/biotechnology | Reliability: .63 (pre-test), .64 (post-test) |
| Outcome expectations (social) 2 | I can attain high status with an education within biology/biochemistry/biotechnology | |
| Outcome expectations (social) 3 | I can serve society and humanity with an education within biology/biochemistry/biotechnology | |
| Outcome expectations (social) 4 | I can have good colleagues with an education within biology/biochemistry/biotechnology | |
| Outcome expectations (self) 1 | I can qualify myself and attain expertise with an education within biology/biochemistry/biotechnology | Reliability: .77 (pre-test), .83 (post-test) |
| Outcome expectations (self) 2 | I can achieve satisfaction with myself with an education within biology/biochemistry/biotechnology | |
| Outcome expectations (self) 3 | With an education within biology/biochemistry/biotechnology I get the opportunity to use my head | |
| Outcome expectations (self) 4 | With an education within biology/biochemistry/biotechnology I get the opportunity to challenge myself | |
| Outcome expectations (self) 5 | With an education within biology/biochemistry/biotechnology I get the opportunity to master a discipline | |

**Table S2: Description of the Laboratory Safety simulation from Study 1 and the CSI simulation from Study 2.**

| |
|---|
| <u>Description of the tasks and narrative of the Laboratory Safety Simulation in Study 1.</u><br>This IVR simulation was on the topic of laboratory safety and was built by the EdTech company Labster (2019a). In the beginning of the IVR simulation, the learner receives information about how to interact with the virtual environment. Then after putting on safety goggles and a lab coat, the learner has to identify five safety hazards in the laboratory (chemicals stored on the floor, an emergency exit being blocked, a messy work station, a sink full of dirty glassware, and an open fume hood for an example, see Figure 1, I and II). The first task is cleaning the work station. However while cleaning, there is an acid spill, and the learner receives instructions in the appropriate procedure to handle the situation (putting on gloves, neutralizing the acid using baking powder and verifying using a pH indicator strip, cleaning it with tissue, and discarding the gloves and tissue). The next task is related to dealing with a chemical eye injury. After an accident in the lab, the learner is guided to the shower and instructed in how to use the emergency eyewash station. Afterwards, the learner is asked to identify appropriate equipment to use during a fire. In the final task, the learner has to identify reasons why a newcomer is not ready to work in the laboratory (see Labster, 2019a for a video of the simulation, or Figure Figure 1, III). A pedagogical agent guides learners through the simulation and provides real time feedback. Throughout the whole simulation, the learner receives multiple-choice questions with explanatory feedback to prime metacognition and reflection. |

Description of the tasks and narrative of the CSI simulation in Study 2.

This IVR simulation was built by the EdTech company Labster (Labster, 2019b). The main learning objective was to develop an understanding of DNA analysis, including how to perform a PCR experiment and carry out gel electrophoresis. The simulation is based around a crime scene investigation. After having witnessed the crime, students are required to collect blood samples and, lastly, they enter a high-tech virtual laboratory where they perform DNA analysis (see Labster, 2019b for a video of the simulation, or Figure 2, I and II). In the laboratory, the female virtual agent, Marie, introduces the students to a full lab bench setup and provides real time feedback throughout the simulation. They receive instructions on how to perform a PCR experiment with the purified DNA from the crime scene, and subsequently conduct gel electrophoresis and compare the results with samples from the suspect (see Figure 2, III). Finally, the learners are asked to identify the murderer. Throughout the simulation learners have to answer multiple choice questions in order to enhance reflection and metacognition (see Figure 2, IV).