Breaking Bad Behaviors: A New Tool for Learning Classroom Management Using Virtual Reality

Jean-Luc Lugrin[1]*, Marc Erich Latoschik[1], Michael Habel[1], Daniel Roth[1,2], Christian Seufert[3] and Silke Grafe[3]

[1] HCI Group, University of Würzburg, Würzburg, Germany, [2] Media and Communication Psychology Group, University of Cologne, Cologne, Germany, [3] University of Würzburg, Würzburg, Germany

This article presents an immersive virtual reality (VR) system for training classroom management skills, with a specific focus on learning to manage disruptive student behavior in face-to-face, one-to-many teaching scenarios. The core of the system is a real-time 3D virtual simulation of a classroom populated by twenty-four semi-autonomous virtual students. The system has been designed as a companion tool for classroom management seminars in a syllabus for primary and secondary school teachers. This will allow lecturers to link theory with practice using the medium of VR. The system is therefore designed for two users: a trainee teacher and an instructor supervising the training session. The teacher is immersed in a real-time 3D simulation of a classroom by means of a head-mounted display and headphone. The instructor operates a graphical desktop console, which renders a view of the class and the teacher whose avatar movements are captured by a marker less tracking system. This console includes a 2D graphics menu with convenient behavior and feedback control mechanisms to provide human-guided training sessions. The system is built using low-cost consumer hardware and software. Its architecture and technical design are described in detail. A first evaluation confirms its conformance to critical usability requirements (i.e., safety and comfort, believability, simplicity, acceptability, extensibility, affordability, and mobility). Our initial results are promising and constitute the necessary first step toward a possible investigation of the efficiency and effectiveness of such a system in terms of learning outcomes and experience.

Keywords: virtual reality training, immersive classroom management, immersive classroom, virtual agent interaction, student simulation

1. INTRODUCTION

In a classroom, disruptive student behavior can have far-reaching detrimental effects on the experience and emotional state of both teachers and students, hindering the achievement of teaching goals and diminishing the overall efficacy of learning for one or all in the classroom (Brouwers and Tomic, 1999; Emmer and Stough, 2001). As such, preempting, controlling, and mitigating disruptive behavior are vital skills for anyone hoping to effectively teach in face-to-face and one-to-many teaching situations. Competence in establishing and maintaining order, engaging students and eliciting their trust, respect, and cooperation are essential aspects of classroom management (CRM) (Emmer and

Stough, 2001), which in turn is an important topic in educational research (Evertson and Weinstein, 2013) and a fundamental module during teacher training (Kunter et al., 2015).

Effective training depends on three major elements:

1. Exposure to realistic training scenarios and stimuli. In the case of CRM, the training stimuli are a classroom full of students displaying a wide variety of realistic normal and disruptive student behavior. Generally speaking, realistic training stimuli can be attained by either training *in vivo*, that is, in a real classroom with real students, or through simulation.
2. Fine control over training stimuli and scenarios. This includes the capacity to finely adjust the difficulty of training to match the current competence of the trainee because it is important that the training scenario be neither too far beyond or below their current capabilities. It also includes the capacity to expose trainees to identical training stimuli multiple times.
3. Fine performance feedback. Providing trainees with a fine-grained, unambiguous, timely measure of their current performance enables them to adjust their behavior to achieve better results.

In terms of the first element, the *gold standard* of CRM training is a real classroom with real students. There might, however, be times when it is not possible to provide trainee teachers with sufficient time in real classrooms. Further, it should be apparent that, in certain scenarios, a compromise exists between the first element (realism) and the second and third elements (control of stimuli and feedback). This is certainly the case with CRM. The unpredictability of a real classroom greatly diminishes any control over the exact nature and difficulty of training stimuli, and feedback is largely restricted to deferred reviews in which feedback is decoupled from the actual situational context, impeding trainees' capacity to adjust their performance in response. Ideally, control of stimuli and provision of feedback would occur in a closed real-time loop between trainer and trainee, allowing the trainer to provide stimuli that finely match and gradually extend the trainees' capabilities and skills. For example, suitable reactions to disruptive behavior have various communicative and interpersonal aspects (e.g., choice or wording, tone of voice, loudness, non-verbal signals by body posture, gestures, movement, eye contact). All of these aspects are important and have to be mastered for successful class management. A failing reaction to a disruption does not necessarily mean that all of the aspects of the counter action have been wrong; hence, targeted feedback is necessary. Providing such feedback in a real classroom is difficult to achieve without inadvertently influencing the state of the classroom or diminishing the realism of the situation.

An alternative to real-world *in vivo* training is *simulation*. In the context of CRM, virtual training environments (VTEs) have been successfully used in training and education domains for many years (Tichon, 2007; Gupta et al., 2008; Dieker et al., 2013). VTEs often provide alternatives to various teaching setups concentrating on the knowledge transfer of the subjects taught (Schutte, 1997; Keppell, 1998; Mahon et al., 2010). Stress exposure training delivered via a VTE has been used across many domains, including military, aviation, and health care (Schuemie, 2003; Baker et al., 2005). We followed this approach and developed an immersive virtual reality (VR) environment for CRM training that generates appropriately stressful situations as expected in front of classes. Stress exposure training rests on the simulation's ability to elicit emotional responses from the teachers (Tichon, 2007). The ability of the system to realistically elicit stress similar to a real classroom atmosphere is therefore paramount.

The simulation medium must therefore be capable of invoking realistic responses to stressful stimuli, which is hard to grasp and master with only video analysis and/or role-play games. One of the main technical challenges is then the simulation and control of a high number of virtual students, which is essential not just for realism but also for provoking realistic levels of stress. In terms of feedback, VTEs offer a rich variety of possibilities, ranging from continuous real-time feedback to fully deferred (Hale et al., 2014). Real-time feedback helps users to identify their weaknesses during their performance (Lopez et al., 2012) and to continuously adapt their behavior to efficiently reach training goals. Previous research has demonstrated that effective feedback systems should reinforce the gamification aspects of the training, which is based on the gradual increase of challenges, perceptual support, and finely tuned scoring systems (Charles et al., 2011; Honey and Hilton, 2011). However, how best to provide effective feedback within an immersive CRM training system remains an open question. In this research, we investigated, developed, and evaluated a VR system combining the three elements of effective training: realism, fine control of stimuli, and real-time, fine-grained feedback.

1.1. Context and Requirements

This article presents a VR training system as an apparatus for the training of CRM skills: *breaking bad behavior* (henceforth 3B) employs a one-to-one teacher/instructor paradigm, with the trainee teacher's entering a visually and aurally immersive virtual environment, while the instructor controls training tasks, monitors the teacher's performance, and provides feedback to the teacher using a non-immersive graphics console.

The 3B system is intended as a complement to traditional CRM teaching methods. It was designed as a companion tool for existing CRM seminars at the University of Würzberg, Germany. (Specifically, they were used in two seminars: *Classroom Management and Videobased Reflection of Education*, both part of the initial teacher syllabus for primary and secondary school teachers.) The system allows lecturers to link theory with practice, using the medium of VR to concretely illustrate the theory, techniques, and examples discussed in lectures and seminars.

The 3B platform aims to better prepare trainee teachers for future *in vivo* training by letting them experience and practice their coping strategies in a safe environment. It is designed to be used by practitioners in the field of educational training without expert knowledge in computer programing, virtual reality, or other technical domains. Not only can the non-technical expert run training sessions, but they can also create new training scenarios

without requiring deep technical knowledge. The system has been conceived in close collaboration with experts in pedagogy and CRM training from the University of Würzberg School of Pedagogy. These experts in the field of CRM training designed the training scenarios and virtual student behaviors and were actively involved in the design of the instructor interface and its interaction techniques. The system underwent an iterative one-year development process, with a strong focus on *user-centered* design. Before being validated by CRM training experts, a team of thirteen HCI bachelor students and two supervisors developed and evaluated four main prototypes. They also received formal and informal feedback from our pedagogic partners every two to four weeks.

This article describes the final iteration of the 3B system, its features, its internal mechanisms, and its first formal evaluation with the students of the CRM seminar. The system provides the following main characteristics for training tasks and teacher educations.

1. Control of the behaviors of individuals and of an entire class while creating dialog phases to allow a more realistic and responsive classroom.
2. Provision of adequate synchronous and asynchronous feedback to teachers by an instructor monitoring their performance.
3. Representation of the teacher's avatar as a means to
   - increase believability and hence immersion and emotional response for the teacher; and
   - provide a visualization of the teacher's body language for the instructor.
4. Usable in a classroom. It can be installed in any room without special infrastructure using low-cost hardware and free software. The system relies on consumer marker less tracking technology, head-mounted displays (HMDs), and a current game engine.
5. Come as you are. The physical self is not artificially augmented with sensors and devices, and preparation and rigging times are largely reduced. The teacher should wear the minimum equipment to be immersed in the environment. The system does not impede natural body movement. It allows the teachers to naturally express themselves as in their everyday life, with no physical constraints or additional fatigue.

Our main objective was to identify any usability issues, especially in terms of ease-of-use and potential VR-side effects (e.g., cybersickness), as well as measuring system reception and acceptability. Our second objective was to evaluate the effectiveness of feedback and evaluation criteria in terms of guiding and motivating trainees while in the VR simulation. We therefore evaluated the system with respect to three main aspects, taking into account both user groups that could considerably affect our system's integration into current teacher education curricula:

E1. Simulation effects by the immersive teacher interface
   i. Believability
   ii. Cybersickness and potential side effects
   iii. Effect of feedback cues

E2. Usability of the instructor interface
   i. Task and cognitive load
   ii. Intuitive usage of the interface
E3. Technology acceptance by teachers and instructors

1.2. Organization

After the initial motivation and brief summary of the contribution of the work presented here, we will continue with a reflection of related work. This will be twofold: section 2.1 reflects on current aspects of CRM and disruptive behavior as known from didactics and teacher education to get an understanding of the use-case scenario. Section 2.2 reflects on recent computer-supported stress exposure and social skills training systems. This is followed by a detailed description of the developed system and its main architectural and design aspects in section 3. Section 4 illustrates the evaluation method, the design of the user study, the procedure, and the measures used. The evaluation results are presented in section 5, followed by a reflection on current limitations as well as on future work in section 6.

2. RELATED WORK

2.1. Classroom Management and Disruptive Behavior

Research on CRM is a well-established topic in educational research (Evertson and Weinstein, 2013). Current models of teacher competencies have integrated CRM competencies as one important aspect (Kunter et al., 2015). As a consequence, it is an important task to integrate CRM into initial teacher education at universities and give students opportunities to develop CRM strategies during internships in schools. Jones (2006) reports that in US initial teacher US education programs, classroom management as a topic has not yet been implemented systematically. As a consequence, novice teachers do not feel sufficiently prepared with regard to CRM (ibid.). The number of publications about classroom management has increased considerably in Germany in the last years, although there is no systematic implementation of classroom management courses in current curricula, to our knowledge. Even though practicing in school in reality is most preferable for most facets of teacher education, classroom management might be considered an exception. Classroom disturbances in the real classroom are unpredictable, making it a challenge to train CRM competencies in a systematic way. Hence, training currently relies on learning the theory of CRM, often accompanied with video analysis and/or role-playing games.

One of the most relevant aspects of CRM is the prevention and management of students' misbehavior. Acting-out or aggressive children disrupt the flow of a lesson or make it impossible to teach the lesson. Effective teachers anticipate classroom disturbances to deflect them or, if that is not possible, react in an appropriate way with the right coping strategies (Borich, 2011). There are various studies about the effectiveness and efficiency of classroom management with regard to behavioral and ecological perspectives. For example, Canter and Canter (1992) developed the assertive discipline program that emphasized specifying

clear rules for student behavior, tied to a system of rewards and punishments. They have not conducted systematic research on their program, however (Brophy, 2006). From an ecological point of view, Kounin (1970) discovered different classroom activities preventing students from becoming disruptive, such as with-it-ness, overlapping, signal continuity, momentum reduction, group alerting, and accountability during lessons. His findings have been supported and enriched by recent process-outcome (teacher effects) studies, which identified and proved the assumption that reducing disruptive behavior has – in alignment with many other influencing variables – the strongest effects on the learning outcomes and their development of children and adolescents (e.g., Hattie, 2003).

With regard to modeling classroom misbehavior in a realistic way, there are various studies on different types of classroom misbehavior (Thomas et al., 1968; Levin and Nolan, 1991, 2013; Mayr et al., 1991; Seitz, 1991, 2004; Canter and Canter, 1992; Walker, 1995; Laslett and Smith, 2002; Borich, 2011; Canter, 2011). The well-known typology of Borich (2011) provides a single dimension with simple discretization level (mild, moderate, and severe) as well as numerous examples of typical observed behavior falling into these categories. Therefore, we adopted and extended Borich's classification. It provides a convenient way to systematically control a large set of virtual students to easily raise or lower the level of stress perceived by a teacher. **Table 1** summarizes the different levels of bad behaviors as inspired by Borich (2011) as well as their associated animations of the virtual students.

There is a significant body of research on the effects of teachers' non verbal communication in real classrooms (Alibali and Nathan, 2007; Kelly et al., 2008, 2009; Mahon et al., 2010; Wang and Loewen, 2015) and also in virtual ones (Barmaki and Hughes, 2015a). Teachers tend to use a variety of non verbal behaviors to communicate knowledge, including hand gestures (iconics, metaphorics, deictics, and beats), head movements, affect displays, kinetographs, and emblems (Wang and Loewen, 2015). More interesting, CRM strategies appear to significantly rely on non verbal cues like eye contact, prolonged gaze, and proximity (Laslett and Smith, 2002). It is therefore important for the instructor/evaluator to perceive the teacher's body, as well as its current point of focus. We thus decided to include avatar embodiment techniques (Spanlang et al., 2014) and a system to highlight the teacher's focus, as illustrated in **Figure 1**. The avatar embodiment is also supported for the teacher because it is an important factor of presence (Lok et al., 2003; Kilteni et al., 2012).

2.2. Stress Exposure and Social Skill Training

Stress exposure training delivered via a VTE has been used across several domains, including the military, aviation, and health care (Schuemie, 2003; Baker et al., 2005). The ability to replicate real-world scenarios in highly controlled virtual environments has proven ideal for learning and practicing decision-making skills in high-affect and dangerous situations (Tichon et al., 2003). For example, Williamon et al. (2014) proposed a system to prepare musicians to manage performance stress during auditions, recitals, or live concerts. The system simulates a virtual audience (twenty-four members) and judges, recreating level of stress comparable to real auditions. The virtual audiences and judges are interactive video-footage displayed on a semi-immersive large screen and manipulated using preset control commands from a computer located in the backstage area.
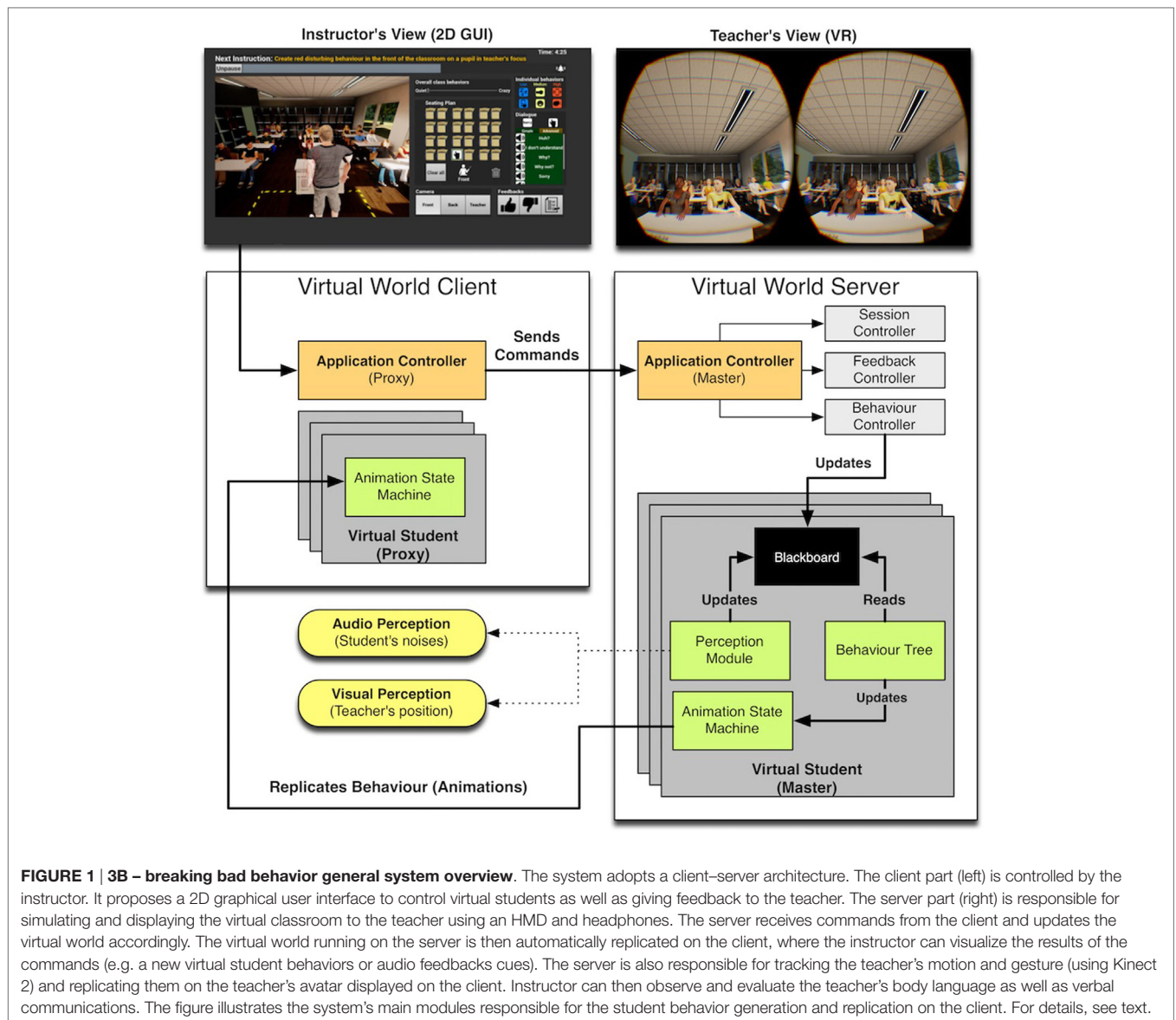
The benefits of virtual environment training in non- or semi-immersive VR systems, such as the TLE TeachLivE™ Lab (Hayes et al., 2013b), are well known and have been well studied over the last few years (Hayes et al., 2013b; Straub et al., 2014; Barmaki, 2015). TLE TeachLivE™ is built on the framework AMITIES™ [avatar-mediated interactive training and individualized experience system (Nagendran et al., 2013)]. It permits users to interactively control avatars in remote environments. It connects people controlling avatars (inhabiters), various manifestations of these avatars (surrogates), and people interacting with these avatars (participants). This unified *human surrogates* framework has also been used in multiple projects to control humanoid robots (see Nagendran et al., 2015, for an overview). TeachLivE™ has been integrated and adopted by 55 universities and was used with over 12,000 teacher candidates during the 2014/15 academic year (Barmaki and Hughes, 2015b). Its ability to improve teacher education has been demonstrated by many case studies (Hayes et al., 2013b; Straub et al., 2014).

Kenny et al. (2007) designed a virtual environment to train novice therapists to perform interviews with patients having important conduct disorders such as aggressive, destructive, or deceitful behaviors. A virtual patient was displayed on a low-immersive screen monitor and responded to users' utterances using natural language parsing with statistical text classification (Leuski et al., 2009), which itself was driving a procedural animation system. The evaluations confirmed the system's ability to replicate real-life experience. Real-time 3D embodied conversational agents have been used to prepare for other stressful social situations such as job interviews (Jones et al., 2014).

TABLE 1 | Student bad behavior classification in terms of intensity of disruption perceived − *null*: correspond to normal behaviors (e.g., quiet and listening student) − in *italic* animations were not directly controllable by the instructor during the evaluation.

Bad behavior levels				
Null	Low	Mid	High	Extreme
Listening-and-quiet	Sleeping-on-table	Laughing	Making-funny-noises	Dancing
Raising-arm-for-question	Typing-on-mobile	Pocking-neighbors	Receiving-phone-call	Screaming
Head-Scratching	*Whispering*	*Playing-with-pen*	*Throwing-object*	*Fighting*
Leg-over	*Looking-around*	*Talking-neighbors*	*Refusing-to-work*	*Making-phone-call*
Writing	*Leaning-back*	*Playing-with-pen*	*Leaving-classroom*	*Breaking-object*

FIGURE 1 | 3B – breaking bad behavior general system overview. The system adopts a client–server architecture. The client part (left) is controlled by the instructor. It proposes a 2D graphical user interface to control virtual students as well as giving feedback to the teacher. The server part (right) is responsible for simulating and displaying the virtual classroom to the teacher using an HMD and headphones. The server receives commands from the client and updates the virtual world accordingly. The virtual world running on the server is then automatically replicated on the client, where the instructor can visualize the results of the commands (e.g. a new virtual student behaviors or audio feedbacks cues). The server is also responsible for tracking the teacher's motion and gesture (using Kinect 2) and replicating them on the teacher's avatar displayed on the client. Instructor can then observe and evaluate the teacher's body language as well as verbal communications. The figure illustrates the system's main modules responsible for the student behavior generation and replication on the client. For details, see text.

This job interview simulator supports social skills training and coaching using a virtual recruiter and measuring applicants' conversational engagement, vocal mirroring, speech activity, and prosodic emphasis.

More recently, a new platform combines TeachLivE™ with a large mixed-reality room: the human surrogate interaction space (HuSIS) targets de-escalation training for law enforcement personnel (Hughes and Ingraham, 2016). The system provides a 4 × 4 m room equipped with projectors and surround sound, where virtual agents can appear among or next to real objects in the room (e.g., tables, shelves). This augmented virtuality platform aims at reproducing stressful scenarios based on an increased immersion involving highly agitated individuals, where human lives might be at risk. Examples include, dealing with highly stressed people who may harm themselves or others. To defuse such situations as rapidly and effectively as possible, the teacher must learn how to quickly assess a situation under

stress. The system should elicit such stress and allow teachers to practice and improve their coping strategies in a safe environment. The system's evaluation should provide interesting insights.

2.3. Discussion

CRM skills are an important aspect of face-to-face teaching situations typically found in classrooms. Unfortunately, training CRM with the available methods based on a pure theoretical understanding or role-play does not match all the aspects found in the real-world scenario, including all of the embodiment and stress aspects. Likewise, a real-world scenario fails in terms of fine-tuned online stimulus control and feedback required for successful training. VTEs are a promising alternative for the real-world scenario. As a computer-generated environment, they provide good control of the presented stimulus, and, in theory, they open up various feedback channels. At the same time, less

research has been dedicated to studying fully visually and auditory immersive training environments, despite their capacity to provide a more realistic emotional response and memorable training (Tichon, 2007; Slater, 2009); and the efficiency and feasibility of such systems for CRM have not yet been demonstrated with a fully immersive virtual environment.

With the recent advent of the VR consumer market, the low-cost products open up novel perspectives to integrate VR-based learning platforms to current school or university curricula. Consumer systems now provide a reasonable rendering quality, end-to-end latency, number of input/output channels (including tracking capabilities), and level of comfort for an acceptable price, making them usable and affordable for institutions such as schools (approximately 2.500 € for a computer and VR headset, with head and hands tracking). Our objective is then to provide a new apparatus for CRM training, enabling further research on such novel applications of VR stress exposure.

However, the consumer hardware and software are still not capable of providing a truly interactive photo-realism for the environments and the virtual humans and avatars, which should faithfully replicate a person's appearance, movement, and facial expression in an interactive real-time experience. In addition, despite recent progress (Waltemate et al., 2015), photo-realistic avatar and agent creation based on scanning or photogrammetry is still time consuming, which is important if one wants to simulate large crowds, such as multiple students/pupils. A higher degree of realism requires higher levels of detail (number of polygons and shaders), which in turn increases rendering time and might affect latency. Numerous user studies have demonstrate the negative impacts of high latencies, temporal jitter, and positional error on user performances, satisfaction, discomfort, and sense of immersion [see LaViola (2000) and Lugrin et al. (2013), for an overview]. Hence, it is important to find the right balance between realism and overall performance.

2.3.1. Requirements

Dieker et al. (2013) suggest three important factors for successful VTEs: (i) personalized learning, (ii) cyclical procedures to ensure impact, and (iii) suspension of disbelief; that is, *it suspends our belief of the real world into one that is altered*. Among these factors, the suspension of disbelief is critical, especially in VR where the feeling of immersion can easily be broken or deteriorate with technical issues, such as bad, slow, or inaccurate head tracking or cybersickness. However, for a virtual CRM training system, suspension of disbelief can also be affected by non realistic behavior of the simulated students. How to efficiently and realistically control dozens of students at the same time is not a trivial problem, especially while one has to react to, evaluate, and guide the teacher's reactions by appropriate stimuli and feedback.

Most existing CRM or social skills training systems are able to simulate only a relatively small number of virtual students (e.g., five for TeachLivE). Such numbers are not representative of teaching scenarios typically encountered in the real world (twenty-to-thirty students for a typical classroom, and up to several hundred for a university lecture hall). Simulating more realistic class sizes is essential because class size is a vital factor for eliciting stress in

the teacher. Additional limitations of existing systems are their lack of mobility and their price. (Most of them require special infrastructure, equipment, and intensive maintenance.)

In previous systems, virtual students are often controlled by experts or actors impersonating students. This not only increases the manpower required to run each training session but also makes it difficult to present different teachers with identical stimuli or to expose a single teacher to the same sequence of events multiple times.

Another important difference with 3B and existing educational training simulations, such as TeachLivE (Hayes et al., 2013b), is the high level of visual and auditory immersion provided by our system. The visualization of the classroom through a head-mounted display (HMD) and headphones increases the teachers' illusion of place and plausibility, resulting in more realistic teacher behavior (Slater, 2009) and eliciting stronger emotional response to and involvement in virtual student behaviors (Sanchez-Vives and Slater, 2005). In addition, the evaluation of the teachers' performance should be more systematic, standardized, and controlled by the system. Finally, instructors should be able to send feedback and guidance during the session and after the sessions. Consequently, the main system characteristics are translated into functional and non-functional requirements of the immersive CRM training system as follows:

R1 affordability and mobility: low-cost and easily installed in any room without requiring additional infrastructures

R2 extensibility and adaptability: possibility to add more scenarios, virtual agents, and behaviors without re-implementing large parts of the system

R3 simplicity: simple control of student behaviors to trigger different levels of stress, ability to provide simple method to give synchronous (after-action-review) and asynchronous (after-session-review) feedback

R4 believability: realistic (in a behavioral view), compelling immersive simulation of classrooms with a large number of students

R5 safety and comfort: does not induce cybersickness or discomfort during or after the session

3. SYSTEM DESCRIPTION

3.1. Overview

Our system is a collaborative virtual reality (VR) platform where both a teacher (trainee) and an instructor (trainer and operator) are interacting in a shared virtual environment. The teacher is immersed in a virtual classroom environment using a head-mounted display (HMD), while the instructor is interacting with the same environment through a 2D graphical user interface (GUI). As depicted by **Figure 1**, our system adopts a distributed architecture where a server is responsible for the VR simulation, and a client is sending commands updating the server's virtual world state. The client is also simulating its own version of the virtual world. However, it constantly receives updates from the server during a process called *replication*, which guarantees the synchronization of virtual object states on both virtual worlds.

For instance, to request the virtual agent of a particular student to perform a disrupting behavior, such as *texting on a mobile phone*, the client will send a *command* to the server, also called a *remote event invocation*. On reception of this event, the server will interpret it and will start playing the corresponding animation on the targeted virtual agent. The name of the animation as well as its parameters will then be sent to the proxy version of this virtual agent on the client (replicating agent behavior on the server). Objects running on the server are thus referred to as *Master*, and their copies on the client(s) are named *Proxy*. Consequently, both client and server are then displaying the same animations at the same time. The instructor and teacher both observe the same behavior on different machines with different views customized for their individual tasks.

3.2. Software and Hardware

The overall system is built on the top of the Unreal Engine 4×™ (Epic Games Inc, 2015). The teacher's view is rendered to the *Oculus Rift DK2* HMD (Oculus VR LLC, 2016), and the teacher's movements are captured by the Rift for head movements and by the *Microsoft Kinect v2* (Microsoft C, 2016) for body movements. Movement data are embedded into the main system by the *Kinect4Unreal* plugin (Opaque Multimedia C, 2016). The GUI as been developed using the unreal motion graphics UI designer (UMG), a visual UI authoring tool, which can be used to create UI elements (e.g., in-game HUDs, menus, buttons, check boxes, sliders, and progress bars) (Epic Games Inc, 2016).

3.3. Teacher's Interface

The teacher's interface fully immerses the user by appropriate head-tracking and perspective projection and display to the HMD as can be seen for the two individual per-eye views on the right in **Figure 1**. The teacher appears standing in front of a crowd of virtual students in a classroom. The teacher can freely move around in a zone equivalent to 2.5×2.5 m and can interact with the students using speech and gesture. The walking zone corresponds to the maximum head and body tracking zone. The limits of this zone are represented in the virtual environment by warning-strip bands located on the floor.

3.4. Instructor's interface

The layout of the instructor interface is illustrated on the left in **Figures 1** and **2**. The GUI design has gone through multiple iterations and user evaluations during its development. The interface consists of five main parts (here called boards):

1. *Instructions board*: this board is used to display the sequence of instructions to follow, as well as time left for the session (visualized using a progress bar), together with buttons to start, stop, and pause the session.
2. *Virtual student control board*: this board has multiple functions. It permits control of the overall level of disruption of the entire class using a simple slider, from *Null* to *Extreme*. It also permits activation or deactivation of bad behavior or dialog for individuals. The instructor can simply enable/disable behaviors or dialog by drag-and-drop of respective icons from the behavior and dialog store on the right of the board

to the seating plan on the left or to the bin (see **Figure 3**). For convenience, a *Clear All* button is present to remove all current behaviors in one click. This feature was added after observation of instructor behavior during our pre-study and informative evaluations. Our system provides six different bad behavior types, and twenty different dialogs divided into a simple and an advanced category. The former contains generic questions, such as *Why?*, and *Why not?*, or *When?*, and short responses, such as *Yes, No*, and *Not sure*. The *advanced* dialogs are more complex and specific, such as *Is mobile phone usage authorized during the trip?* The system has been developed to easily insert new dialogs by importing sound files to specific directories and following a simple naming convention.

3. *Virtual environment viewport*: it displays the virtual class environment in a non-immersive 3D rendering from one of three potential points of view to monitor the overall situation, the behaviors of the students, and the reactions of the teacher. The students currently observed by the teacher are always highlighted in red in all of these views. We refer to this as the *teacher's focus*.
4. *Camera control board*: it allows users to switch the point of view of the *virtual environment viewport* camera between (i) front, (ii) back of the class, and (iii) teacher's viewpoint (see **Figure 4**).
5. *Feedback board*: it allows two types of feedback (see **Figure 3** left and center):
   (i) Synchronous feedback during the session by pressing "Thumb-up"/"Thumb-down" buttons. Audio cues with positive or negative connotation are associated with these buttons. For instance, when pressing the *Thumb-up* button to communicate positive feedback, the teacher hears a *ding* sound, representing the instructor's approval. The *Thumb-down* button, triggering a *buzz* sound, communicates negative feedback (see **Figure 3** left and center). The purpose of synchronous feedback provides a simple mechanism for guiding the teacher during the lecture without completely interrupting the session. There is an option that permits the instructor to enable or disable these features, thereby muting the feedback.
   (ii) Asynchronous feedback, which consists of a more detailed evaluation form, initially summarizing the teacher's performance for the session. This feedback uses a list of items and scales to assess the teacher's overall performance by the instructor (see **Figure 3** center for the list of items).

Both feedback types are automatically logged for each session. They are used as performance indicators to measure teacher progress from one session to another.

3.5. Virtual Environment

The virtual training environment is modeled as a typical classroom, as illustrated in parts of the **Figures 1**, **2**, and **4**. It is equivalent to a room with physical dimensions of 15 m × 20 m × 3 m, capable of accommodating up to twenty-four seated students. Our system currently provides a pool of thirty different student characters. From this pool, the system randomly generates a class
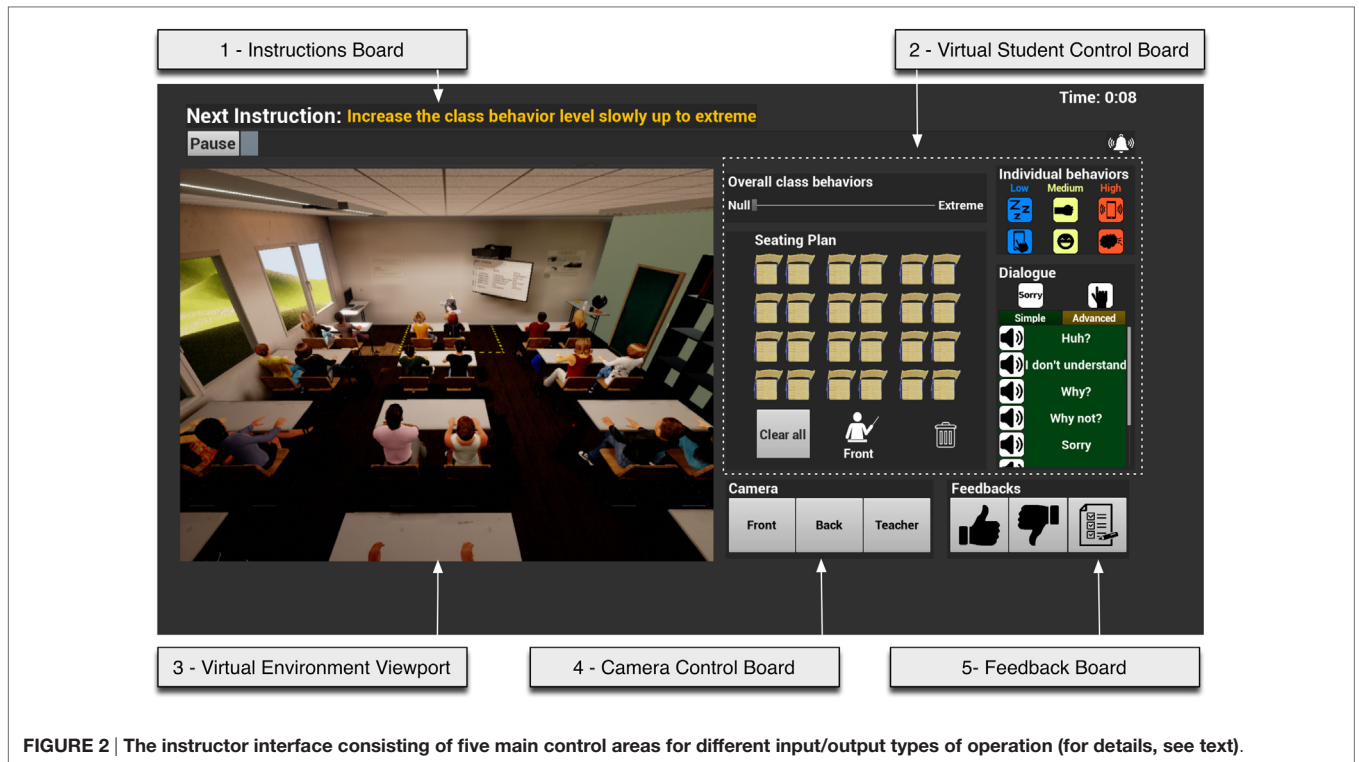
FIGURE 2 | The instructor interface consisting of five main control areas for different input/output types of operation (for details, see text).
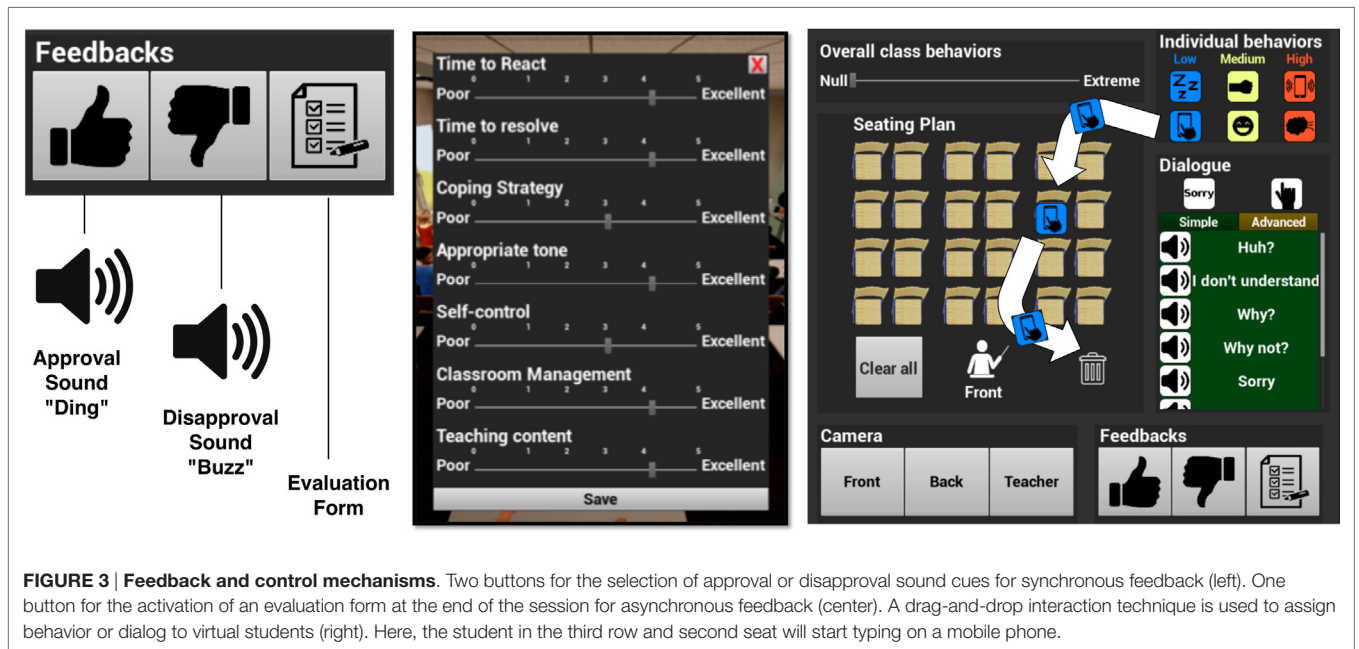

FIGURE 3 | Feedback and control mechanisms. Two buttons for the selection of approval or disapproval sound cues for synchronous feedback (left). One button for the activation of an evaluation form at the end of the session for asynchronous feedback (center). A drag-and-drop interaction technique is used to assign behavior or dialog to virtual students (right). Here, the student in the third row and second seat will start typing on a mobile phone.

of twenty-four students at the start of each session (providing over 500,000 possible configurations). The characters have been designed to give each of them a distinct individual appearance, so that the teacher can identify and name them clearly by shirt, hair, and so on. Additional variety is generated to represent different ethnicities, body proportions, and sizes as well as a small

collection of stereotypical personae (e.g., fashion-conscious, intellectual, athletic).

The software used to model these characters was Autodesks Character Generator (Autodesk, 2016a). To achieve realistic animations, we used the IpiSoft Recorder (iPiSoft I, 2016b) and Mocap Studio (iPiSoft I, 2016a). Results were then edited and
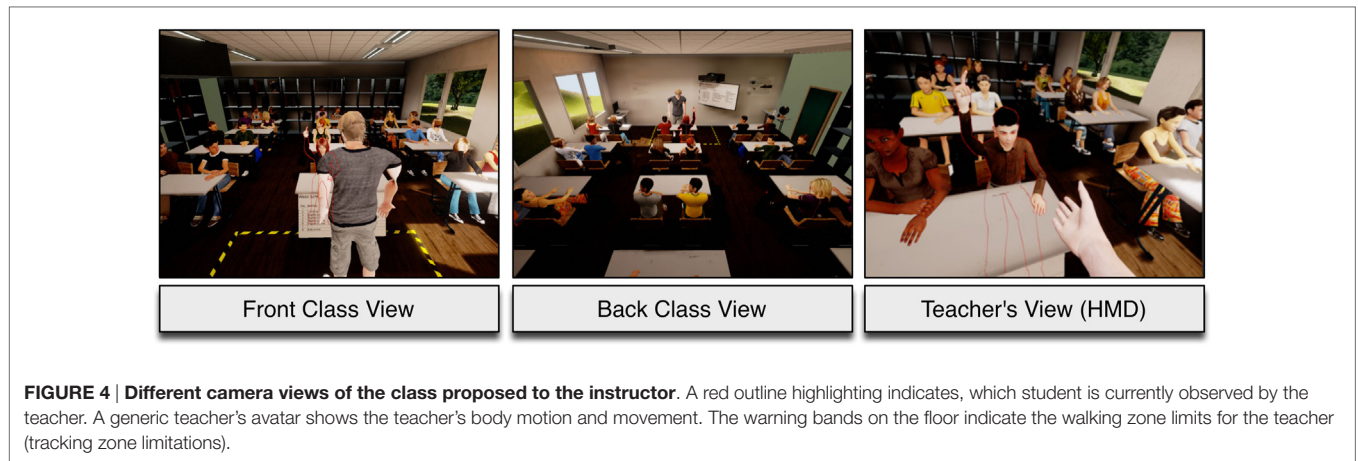
FIGURE 4 | Different camera views of the class proposed to the instructor. A red outline highlighting indicates, which student is currently observed by the teacher. A generic teacher's avatar shows the teacher's body motion and movement. The warning bands on the floor indicate the walking zone limits for the teacher (tracking zone limitations).

applied to our characters in Autodesk MotionBuilder (Autodesk, 2016c) and Maya (Autodesk, 2016b) before being finally imported in the Unreal Game engine using the *FBX* file format. The final number of triangles per character is between 7,500 and 12,500 triangles, with a total of approximately 843,000 triangles for the whole environment (including furniture, walls, and window). The environment also includes ambient sound coming from the outside countryside (e.g., wind and bird sounds).

3.6. System Modules and Student Behavior Control

Our system uses a head-mounted display (HMD) to provide the trainee teacher with visual and auditory immersion into the virtual environment. **Figure 1** illustrates the overall control flow and modules involved in the simulation of the virtual environment and, in particular, in the student behavior generation and replication.

We adopted a client–server design, where the server renders an immersive VR version of the virtual world (i.e., a head-tracked perspective–correct stereoscopic 3D rendering) while the client provides a non-stereoscopic view of the same virtual world. The server is a *Listen* game server, which accepts connections from remote clients while locally rendering the game world for one player on the same machine. This local player is directly connected to the HMD, which provides the best immersive experience for the teacher in terms of high frame rate and low end-to-end latency [i.e., here, the delay between users' head movements and virtual cameras' update (Lugrin et al., 2013)]. The server is the *authoritative master* simulating the virtual environment, which means that it is controlling the world state update on the clients. The clients are *predictive clients* independently simulating their own version of the virtual word in between the server's replication messages. In our system, we currently use one client representing the instructor's view.

The additional synchronization overhead for the client is tolerable due to the reduced real-time requirements for the non-immersive instructor interface. It is important to note that the system could accommodate multiple instructor views at the same time, which could be locally or remotely connected machines.

The system architecture is decomposed into the following modules developed on top of the game engine (see **Figure 1**).

- Application controller master is receiving commands from the *application controller proxy* running on the client, which translates GUI interactions to commands and sends them to the server via the network (UDP). The *application controller master* acts then as a relay module to the *session*, *feedback*, and *behavior controller*. Additionally, it is responsible for logging all events and commands.
- Session controller is responsible for handling commands related to the training session such as starting, terminating, or pausing.
- Feedback controller is handling positive and negative feedback given by the instructor and whether it is audible by the teacher or not.
- Behavior controller is interacting with the *Blackboard* of each *virtual student* for categorized behaviors or individual behavior.
- Virtual student master: all simulated virtual students have two technical incarnations: the *Master* incarnations are running on the server while their copies running on the client are labeled *Proxy*. Each *Master* is composed of three main sub modules:
  - A blackboard used as a central shared storage of values to be accessed by all modules responsible for student behavior.
  - A *behavior tree* permanently watching *blackboard* values updated by other modules to react to changes. As explained in the following section, it controls behavior by logical decisions based on certain conditions.
  - A perception module that recognes and locates audible and visual events inside the virtual world, such as, the position of the teacher, the *funny noises*, or any disrupting behavior made by other students around them. This module updates the *blackboard* accordingly, possibly triggering other branches in the *behavior tree*.
  - An animation state machine that controls the animations and their transitions on the server, triggering sounds and spawning objects. The *animation state machine proxy* is automatically receiving the animation state of the *Master* entity and starting to play the animation on the client.

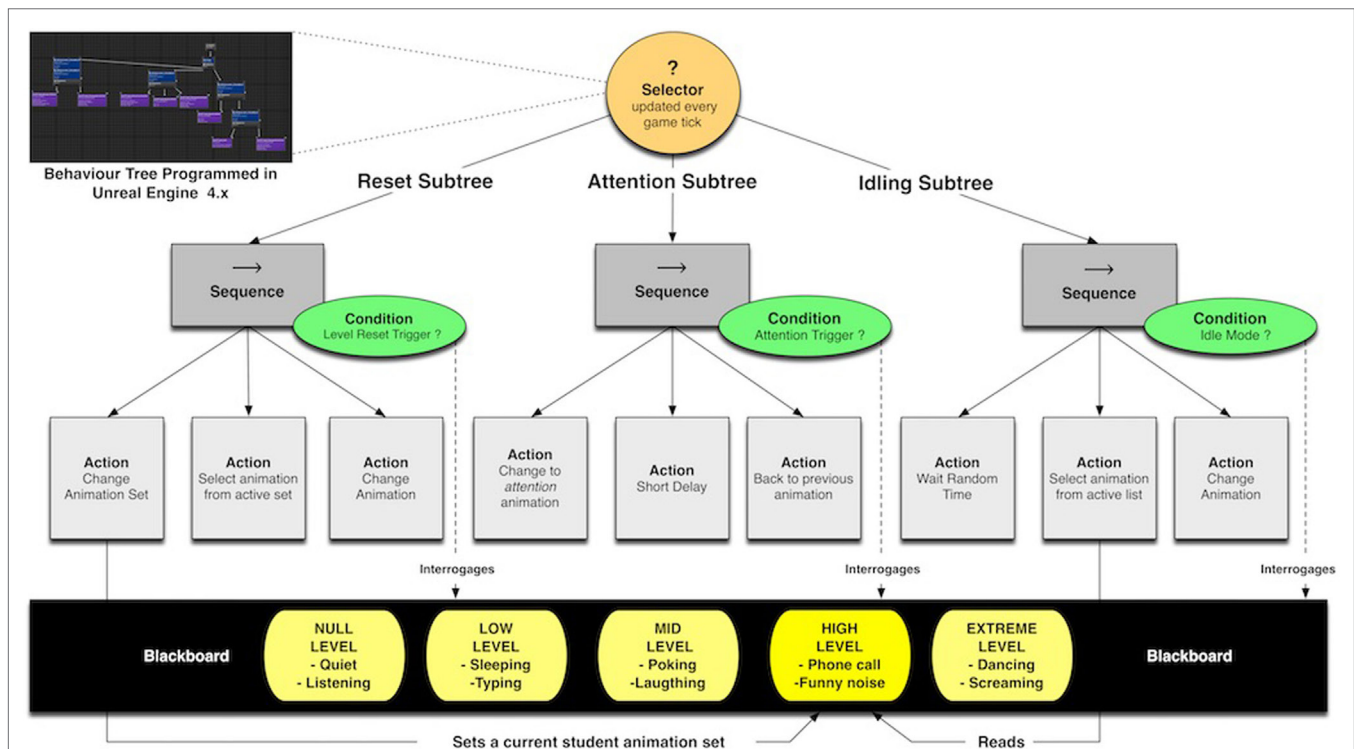### 3.6.1. Behavior Tree: Virtual Student Behavior Algorithm

Each virtual student is driven by a behavior tree connected to the virtual world through individual *perception* and *behavior controller* modules; both of them are using a *blackboard* data structure to communicate (see **Figure 1**). The behavior tree's main role is to determine the active animation to play for a virtual student's *animation state machine*. This finite state machine handles animations, their transitions, and the control of corresponding sounds.

The *perception* module helps to create a more believable environment and adds more dynamics to the scene by adding the ability for virtual students to react to each other. It basically emulates humanoid senses, in this case hearing and sight, which can be limited to realistic values. When a sound event is recognized, the *perception* module decides whether the perceiving student should react to it or not. In case of a reaction, the sound location and a trigger are stored on the *blackboard* to tell the *behavior tree* to react to the event as well as to know where it came from. Simultaneously, sight is used to locate the teacher and make the student face the teacher in suitable animation states. As seen in **Figure 1**, the *behavior tree* is scanning the stored blackboard values for changes and selects a reaction based on logical conditions. This scan is performed every five seconds. This interval is customizable to allow fine-tuning of the agents' overall responsiveness with respect to the available computing resources.

The *behavior tree* has been implemented using the unreal engine's behavior tree editor and *blueprint visual scripting*. The *behavior tree* itself can be split into three subtrees, which represent one task each, as illustrated in **Figure 5**. The target subtree to be executed is determined by checking for certain conditions on variables contained in the *blackboard* as well as the tree's hierarchical order from left to right. The tasks of the three subtrees are as follows:

The *reset subtree* (**Figure 5**, left) is executed if a new level of class disruption is selected (changing the current bad behavior category). Once the user moves the overall class behavior slider, a *level reset trigger* variable is set to true inside the *blackboard*. It actives the corresponding animation set associated with a particular level of disruption (i.e., null, low, mid, high, or extreme). It then immediately chooses an animation from this new animation set. This provides immediate behavior changes and, hence, highly reactive students when changing the class behavior by instantly stopping their current animation and starting to play an appropriate new one.

The *attention subtree* (**Figure 5**, middle) is executed when the *attention trigger* is set to true by the *perception* module. As a result, a special *attention* animation is displayed by the student. It makes the student turn his/her head toward the student who has emitted the sound event. It then plays as random animation as a reaction to it. After a randomly short amount of time, the attention stops and the student returns to it previous animation state.

FIGURE 5 | The overall behavior tree controlling the behavior of the virtual students. It is decomposed into three main parts: (i) one for immediately changing the student behavior when requested by the instructor (reset subtree, left), (ii) one for automatically making the virtual student attending a disruption source, such as a student making a noise (attention subtree, middle), and (iii) one for automatically selecting a behavior according to the current bad behavior level (idling subtree, right). The conditions to activate a subtree are checked by scanning-specific blackboard values, which reflect the overall state of the simulated environment and instructor's commands.

The *idle subtree* (**Figure 5**, right) is executed when none of the others are currently active and no individual behavior is selected by the instructor. Its task is to create a living environment by occasionally changing the animation to a randomly selected one from the active animations list.

4. EVALUATION

The *affordability* and *mobility* requirements of our system are met by the use of low-cost, off-the-shelf equipment with no need for calibration of hardware or software. Its *extensibility* and *adaptability* come from simple and fast integration of new assets (e.g., student character models, animations, sounds) into our AI system without requiring systemic modifications. For instance, once a character model or animation has been imported into an *Unreal* asset package, a simple reference to its name and associated categories of bad behavior (i.e., null, low, medium, high, and extreme) is all that is required for the model to be included in the system. For the other requirements—*simplicity*, *believability*, and *safety* and *comfort*—a user study was executed. As part of this study, we evaluated the system's *acceptability* from the perspective of the future user target group and the effect of synchronous feedback cues on trainees' experience and acceptance. Our evaluation methods and results are described in the rest of this section.

4.1. Pre-User Study: Latency Measurements

We first evaluated the system performance in terms of the average frame rate and end-to-end latency perceived by the teacher. This is the critical technical component to simulate a believable environment for the teacher and to ensure a safe and healthy environment that does not induce cyber sickness and which provides an acceptable user experience.

To guarantee that our system provides synchronous temporal visuomotoric stimulation, we performed video-based measurements of the end-to-end latency using a frame-counting method as described by He et al. (2000). This method is less accurate than the pendulum method discussed by Steed (2008) but is better adapted to immersive game measurements (Lugrin et al., 2012). The average end-to-end latency between movements of the participant's head and corresponding updates of the projected images was evaluated to approximately 73 ms ($\pm$SD 45), which is below the threshold required for real-time interactions [$\leq$150 ms (Lugrin et al., 2012)]. Measurements were realized with videos recorded at 480 Hz with the Casio EX-ZR200 camera at a resolution of 224×160. The overall system delivered an average frame rate of 75 frames per second for an average number of 400 K triangles per frame.

4.2. User Study: Experimental Design

The overall setup of the experimental design is illustrated in **Figure 6**. The idea of the user study follows the proposed method by Barmaki and Hughes (2015b). Our experiment simultaneously involved participant pairs of teacher and instructor, each one having a different task: (i) the teacher task was to perform a short presentation in front of our virtual classroom; and (ii) the instructor task was to slowly increase the class agitation, interleaving bad

FIGURE 6 | An overview of the experimental setting. The instructor station running the 3B client interface is located in the foreground. The teacher, equipped with a VR headset, can be seen in interaction space in the background.

behaviors and dialog phases during the teacher's presentation, and evaluating the teacher's performance.

Each session took seven minutes. We tested two conditions, with and without synchronous feedback to the participants as factors using an in-between design. Feedback was initiated by the instructor using the interface illustrated in **Figure 3** left and center. Instructors were repeatedly reminded to give feedback on their teacher's reaction: the instructor interface displayed an animated message with sound, "Please give feedback," thirty seconds after each new bad-behavior-activation instruction. It was then the instructor's responsibility to judge the teacher's reaction or absence of reaction as good or bad. One advantage of virtual training is that it can easily support different forms of continuous feedback to guide, encourage, or motivate the trainees during a session. Therefore, one important aspect of our experiment consisted of investigating the possible impacts of synchronous audio feedback cues on the teachers' experience. Previous work (Barmaki and Hughes, 2015b) demonstrates the positive impact of continuous feedback on teachers' body language also using simple visual indications in a mixed-reality setting. Their system informs teachers in real-time when they are adopting a closed posture (i.e., defensiveness or avoidance body posture, such as hands folded in front or hands clasped in back). The feedback took the form of a simple square's turning orange when detecting a closed-posture. This visual indication was displayed on a screen next to the virtual classroom visualization screen. Results suggested that this feedback was successful in providing participants with a better awareness of the message they were sending though their body pose.

Our intention is, then, to explore the effectiveness of simple audio feedback signals within an immersive VR system. We aim to determine if such feedback is beneficial or disruptive. Our

expectation is that it helps trainees in providing a certain aware-ness of their performance difficulties in real time, allowing them to adopt a different strategy to deal with the situation at hand. At the same time, there is a risk that such feedback leads to a diminishing of the teachers' sense of immersion, suspension of disbelief, motivation or concentration, and focus on the task at hand. Providing an answer to this question can have a critical impact on the overall system's acceptance and effectiveness, and a clear answer could not be clearly discerned from previous work alone. Consequently, we evaluate whether synchronous audio feedback is an effective feedback mechanism, or whether it is a disruptive influence on the training simulation.
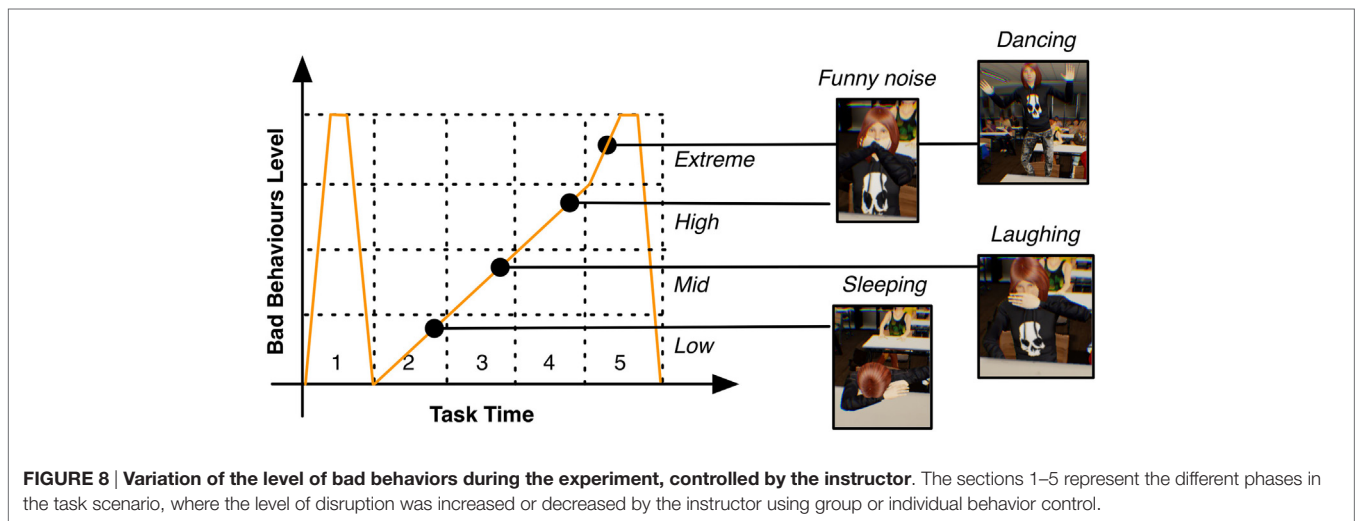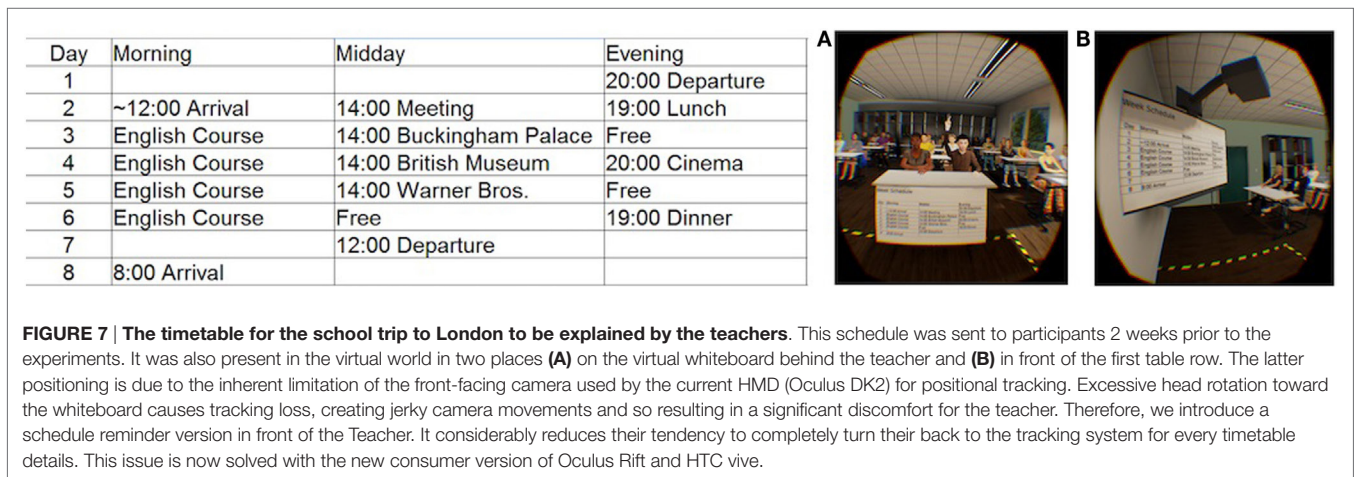
4.3. Experimental Tasks

The task of the teacher was to present the organization of a school trip to London in the context of an English course in high school for teenagers from sixteen to eighteen years of age. This task was divided into the presentation of two main types of information: the trip's weekly schedule (see **Figure 7**) and the explanation of the school's rules and policies (e.g., no smoking, no alcohol, no drugs, bed time). Teachers were also asked to reply to possible student questions. The task design incorporated interdisciplinary

applicability because it did not involve knowledge of specific sub-jects taught. As a result, it also reduced cognitive and preparation demands for the teachers. Finally, the organization of school trips or excursions is a recurring task in a teacher carrier, which they should be prepared to face. The excitation usually caused by the announcement of extra curriculum activities is usually prone to provoke extra disturbances among students.

The task of the instructor was to control and adapt the stimuli, or the (disruptive) behavior of the students, to reflect such excitation and agitation, and to evaluate the teacher's per-formance. We designed a set of instructions that should ideally reproduce a scenario in which an entire classroom could rapidly become uncontrollable. The overall objective was to simulate the typical class effervescence and fast-growing agitation among the students.

Figure 8 outlines the variations of the bad behavior level during the experiment. We defined five main phases with dif-ferent levels of disturbances. Phases 1 and 5 simulated a chaotic class at the start and end of the lecture. In between, phases 2, 3, and 4 defined a steady increase in class agitation, which was simulated by subsequently triggering sequences of two *low*, *mid*, and *high*-level bad behaviors, intertwined with *dialog* phases.

Day	Morning	Midday	Evening
1			20:00 Departure
2	~12:00 Arrival	14:00 Meeting	19:00 Lunch
3	English Course	14:00 Buckingham Palace	Free
4	English Course	14:00 British Museum	20:00 Cinema
5	English Course	14:00 Warner Bros.	Free
6	English Course	Free	19:00 Dinner
7		12:00 Departure	
8	8:00 Arrival		

FIGURE 7 | The timetable for the school trip to London to be explained by the teachers. This schedule was sent to participants 2 weeks prior to the experiments. It was also present in the virtual world in two places **(A)** on the virtual whiteboard behind the teacher and **(B)** in front of the first table row. The latter positioning is due to the inherent limitation of the front-facing camera used by the current HMD (Oculus DK2) for positional tracking. Excessive head rotation toward the whiteboard causes tracking loss, creating jerky camera movements and so resulting in a significant discomfort for the teacher. Therefore, we introduce a schedule reminder version in front of the Teacher. It considerably reduces their tendency to completely turn their back to the tracking system for every timetable details. This issue is now solved with the new consumer version of Oculus Rift and HTC vive.

FIGURE 8 | Variation of the level of bad behaviors during the experiment, controlled by the instructor. The sections 1–5 represent the different phases in the task scenario, where the level of disruption was increased or decreased by the instructor using group or individual behavior control.

The instructions were prompted at the top of the instructor interface screen (see top of **Figure 2**). The instructor had to execute the instructions and evaluate the teacher's reaction using the synchronous feedback buttons. The complete list of instructions is presented in **Table 2**. These instructions were generic (e.g., create blue disturbing behavior in the front of the classroom on a pupil outside teacher's focus). This gave a certain degree of freedom to the instructors with regard to the exact behavior and student to apply the behavior to.

4.4. Measures
4.4.1. Simulator Sickness
We measured simulator sickness for the teachers before and after the induction using the simulator sickness questionnaire (SSQ) (Kennedy et al., 1993). This questionnaire was used to evaluate

the *safety and comfort* requirements. The results were also used to sort out participants, if applicable.

4.4.2. Behavior Categories
For the evaluation of our stress and bad behavior induction scenarios, we included questions to see if the categoric levels of bad behavior matched the users' perceptions, such as, "Please indicate if these disruptions are correctly categorized as a low Level disruption."

4.4.3. Presence, Immersion, and Suspension of Disbelief
To evaluate the teachers' quality of experience, we included the teacher experience questionnaire (Hayes et al., 2013a). To allow for a quantitative comparison and analysis of potential effects

TABLE 2 | List of predefined instructions to follow, their durations, and their respective types.

Id	Instructions	Duration (s)	Type
1	Increase the class behavior level slowly up to extreme	20	Class-bad-behavior
	Wait for the teacher to react	7	Observation
	Decrease it to quiet	7	Class-bad-behavior
	Do not forget to give feedback	7	Feedback
2	Create blue disturbing behavior in the front of the classroom on a pupil outside teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
3	Create blue disturbing behavior in the front of the classroom on a pupil in teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
4	Create a dialog in front of the class on a pupil in teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
5	Create yellow disturbing behavior in the middle of the classroom on a pupil outside teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
6	Create yellow disturbing behavior in the middle of the classroom on a pupil in teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
7	Create a dialog in front of the class on a pupil in teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
8	Create red disturbing behavior in the front of the classroom on a pupil outside teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
9	Create red disturbing behavior in the front of the classroom on a pupil in teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
10	Create a dialog in front of the class on a pupil in teacher's focus	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback
11	Progressively increase the class disturbance level	20	Class-bad-behavior
	Wait for the teacher to react	7	Observation
	Decrease it to quiet	7	Class-bad-behavior
	Do not forget to give feedback	7	Feedback
12	Lesson is finished, use the feedback panel to give a feedback for the whole lesson	20	Bad-behavior
	Wait for the teacher to react	7	Observation
	Do not forget to give feedback	7	Feedback

Instructions and durations were displayed on the instructor's graphical interface and signaled on screen using both visual cue (color animation) and audio sound cue (bip).

between feedback conditions, we changed the student questions from qualitative to individually stated Likert-scale responses (1 = not at all, 5 = very much) with an additional comment of *Why?* as a qualitative text-based answer. Teachers responded to questions about their simulation experience related to suspension of disbelief, presence, fidelity, and immersion:

1. Overall, how successful do you feel your virtual rehearsal performance was?
2. How can you tell that the students are engaged or not engaged with you?
3. How did the virtual students compare to students you encounter in a classroom?
4. How did the virtual classroom compare to your experience of a physical classroom?
5. When you were teaching the virtual students, were you able to suspend disbelief?
6. When teaching the students, did you feel like you were in the same physical space as they were?

4.4.4. Intuitive Use and Task Load
For a basic usability evaluation of the instructor interface, we included the following two questionnaires:

- NASA TLX: to measure cognitive and physical workload estimates from one or more operators while they are performing a task or immediately afterward (Hart and Staveland, 1988).
- QUESI: to measure the subjective consequence of intuitive use by Hurtienne and Naumann (2010). The questionnaire consists of fourteen items dividing intuitive use in five factors (low subjective mental workload, high perceived achievement of goals, low perceived effort of learning, high familiarity, and low perceived error rate).

4.4.5. Technology Acceptance
For a user-centered evaluation of the system, we included two Likert-scale questions (1 = does not apply all, 5 = completely applies) for each requirement (mobility, price, usability, closeness to reality, study improvement, and system enhancement), such as, "The system mobile enough to be set up easily for a simulation," and "I could easily explain the usage of the system to my colleague," and both participant groups. Items were recoded where necessary.

4.4.6. Performance
We evaluated instructor and teacher's performances with different metrics:

- TeachLivE™ teacher performance (Hayes et al., 2013a): teachers responded to questions about their CRM performance. This questionnaire was also given to the instructors to evaluate the respective teacher performance.
- Teacher task performance: number of positive and negative synchronous feedbacks received as well as final score evaluation with asynchronous feedback (see **Figure 3**, center).
- Instructor task performance: number of behaviors, dialogs, and feedbacks generated, number of commands following

the instructions given, number of errors, and number of extra commands.

4.5. Apparatus
The hardware setup consisted of a Microsoft Kinect 2 sensor, one client PC (Quad core 3.7 GHz, 16 GB RAM), and one server station (Intel Core i5-6600K 3.50 GHz CPU, 16 GB of RAM, AMD Radeon R9 390 Graphics card). As depicted by **Figure 6**, *teachers* were visually immersed in a virtual environment using the Oculus Rift DK2 stereoscopic HMD, with a field of view of 100° horizontally, a resolution of 960 × 1080 pixel per eye, and a refresh rate of 75 Hz. The cost of the overall setup was approximately 2500 €.

4.6. Procedure
As illustrated in **Figure 9**, the overall experiment followed ten main stages:

1. *Pre-questionnaires*: let teacher and instructor complete a consent form and a demographic questionnaire. Let teacher fill out the pre-SSQ questionnaire.
2. *Conditions*: let teachers throw a die to determine whether or not they will receive synchronous feedback. In case of a value <4, they were informed about the feedback mechanism. We used a non-algorithmic randomization method for group assignments, and counterbalancing was applied at the end to produce a group of same size.
3. *Instructor training*: instructor was introduced to the graphical user interface (GUI), its features and controls, by the experimenter in about three minutes. Then the instructor started a two-minute training scenario, with no teacher in the virtual environment. During this time, the instructor had to follow a series of six instruction messages appearing on the screen. The experimenter assisted the instructor if questions or mistakes were made. At the end of the session, the experimenter asked if further training was necessary and went through the task flow once more (i.e., read instructions, executed them, and read the next one).
4. *Student training*: equip participants with the HMD and immerse them in the virtual classroom. Calibrate the HMD for comfort and correct stereoscopy. Ask them to walk around in the virtual room to get familiar with wearing the HMD and navigating the virtual environment. As illustrated in **Figure 4**, the walking zone was delimited by warning bands on the floor and corresponded to a zone of 2.5 × 2.5 m. Participants were also instructed to check their virtual blackboard and to report if they felt anything unnatural or uncomfortable. This session lasted from two to four minutes, depending on the participant's questions or adjustments needs. During this phase, the classroom was populated by students in a "quiet" behavior mode.
5. *Break*: ask participants if a break was necessary before continuing with the experiment.
6. *Experiment start*: once experimenters verified that both participants were ready to start, they signaled the instructor to press the *Start* button. Then a school's bell ringing sound
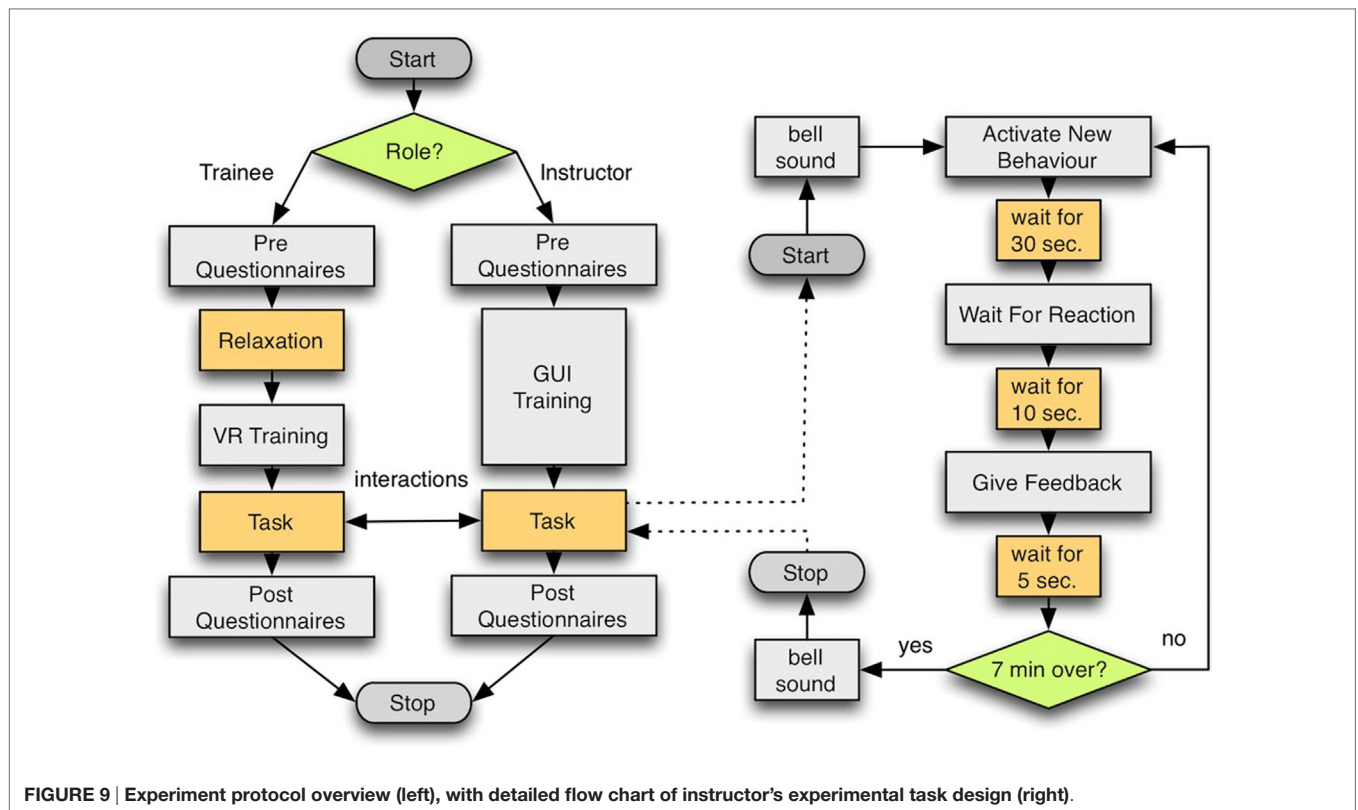
FIGURE 9 | Experiment protocol overview (left), with detailed flow chart of instructor's experimental task design (right).

informed the teacher and instructor of the start of their tasks.

7. *Experiment*: without any help from the experimenters, the instructor had to follow the instruction popping up on screen, while teachers started their presentation and were presented different students' behaviors.

8. *Experiment end*: after seven minutes, another school bell sound signaled the end of the task. Teacher were then de-equipped, taking off headphones and the HMD.

9. *Post-questionnaires*: ask participants to fill out the rest of the questionnaires:
   - For teacher: post-SSQ, TeachLivE™ teacher experience and performance, presence, bad-behavior categories, and technology acceptance questionnaires
   - For instructor: QUESI, NASA TLX, TeachLivE™ teacher performance, bad-behavior categories, and technology acceptance questionnaires

The whole experiment took approximately forty to sixty minutes, depending on the break time each participant needed.

4.7. Sample

Our total sample consisted of twenty-two (German) participants, divided into one *teacher* group and one *instructor*. All the participants were following the previously mentioned CRM seminars in the context of their bachelor of educational science and pedagogy at the University of Würzburg. The mean age of the *teacher* group was 21.45 (SD_{age} = 1.86), and six of the eleven teachers

were females. None of them had severe visual impairments, all of them were students, and none of them had previous experience with VR. In the instructor group (M_{age} = 21.73, SD_{age} = 2.87), nine subjects were females, and ten were students. None of them had previous experience with VR systems, but they used a PC regularly (M = 5.27 where 1 was not at all and 7 was very often, SD = 1.1). The participants were grouped in pairs of two: one teacher and one instructor. The eleven pairs were then divided into two groups; six subject pairs received continuous instructor feedback, and five pairs did not. The instructor's feedback consisted of short audio cues with strong positive or negative connotations. Both groups received different instructions prior to the experiment:

- Teacher: *classroom management* seminar
  - Preparation task: two weeks before the experiment, students received the task description and were required to prepare their presentation.
  - Rationale: students have spent ten weeks in a seminar discussing coping and prevention strategies for controlling bad behaviors.
- Instructor: *video-based reflection of education* seminar
  - Preparation task: two weeks before the experiment, students were required to prepare an evaluation form to quickly assess a teacher's ability to manage classroom behavior.
  - Rationale: students have spent ten weeks in a seminar analyzing teachers' behavior, techniques, and good and bad qualities based on different video sources.

However, the actual experiment procedure, objective, and system were never disclosed to the students. Students were told only that they would use these preparative works in an exercise involving a virtual classroom.

The purposes of these pre-experiment instructions were multiple. The first was to evaluate the system within its intended context of use: as part of existing seminars related to class management. Second, without preparation, the teacher's task (making a presentation in front of classroom on a particular topic) would be difficult to improvise and could create additional stress. Both would interfere with the actual system perception, its possible usefulness, and the measure of the stress induced by the system. Similarly, for the instructor's task, to quickly give correct and constructive feedback, seminar students should have already prepared certain evaluations criteria. Both student groups have acquired the theoretical knowledge, which could be applied in the 3B simulation as teacher or instructor. They are thus very representative of the future target audiences for our training simulations. In addition, they could provide us with interesting insights regarding possible improvements and the permanent integration of such a system into their current teacher education curriculum.

5. RESULTS

5.1. Simulation Effects

Concerning the teach-live questionnaire results, all dimensions were rated equal or higher than scale average (see **Figure 10**). While the highest ratings were achieved on the dimension of being in the same physical space ($M = 4.09$, $SD = 1.22$), subjects were most unsatisfied with their performance ($M = 3.00$, $SD = 1.00$) and the similarity between virtual and real students ($M = 3.09$, $SD = 1.044$). To further investigate causes, we looked at the students' ratings of the behavior categories and found that

eight of out eleven reported that the lower bad behaviors (e.g., sleeping, mobile phone usage) should be set to a higher level.

5.2. Instructor Interface Usability

The task load for the instructors to interact with the system was low [$M = 6.42$, $SD = 3.42$ compared to scale average (scale $0 =$ low task load, $20 =$ high task load, items were recoded where necessary)], indicating that users felt a low task load and the interface control and GUI were neither stressful nor complicated. Concerning the scores on the QUESI questionnaire, all dimensions as well as the total score ($M = 3.83$, $SD = 0.68$) were rated well above scale average ($1 =$ negative perception, $5 =$ positive perception), indicating that users felt comfortable interacting with the system and reached their goals easily and intuitively (see **Figure 11**). Similar score values have been reported for Apple iPod Touch (5) and iPhone (6) (Naumann and Hurtienne, 2010). The score for the perceived error rate ($M = 4.18$, $SD = 0.93$) underlines the robust functionality for the operator.

5.3. Instructors' Performances

The total number of instructor's commands during a session was on average $M = 517$ commands ($SD = 184.43$), ranging from 189 to 764 commands. On average, users issued $M = 23.54$ individual behavior commands ($SD = 10.46$, $range = 7–43$) and 9 class behavior changes ($SD = 5.12$, $range = 0–19$). Furthermore, users started an average of $M = 3.45$ dialogs per session ($SD = 2.70$, $range = 0–8$). During the simulation $M = 10.36$, sounds were played and the instructors gave $M = 5.54$ feedbacks, diversely ranging from 0 to 11 feedback messages per session ($SD = 3.32$).
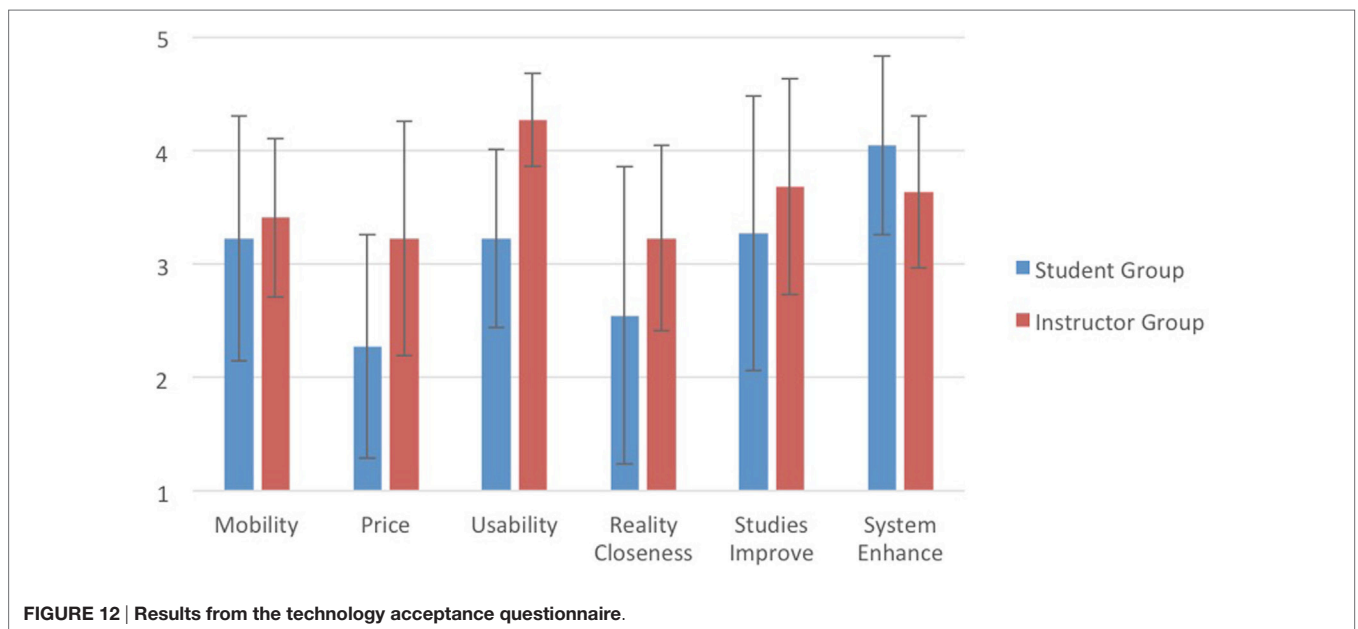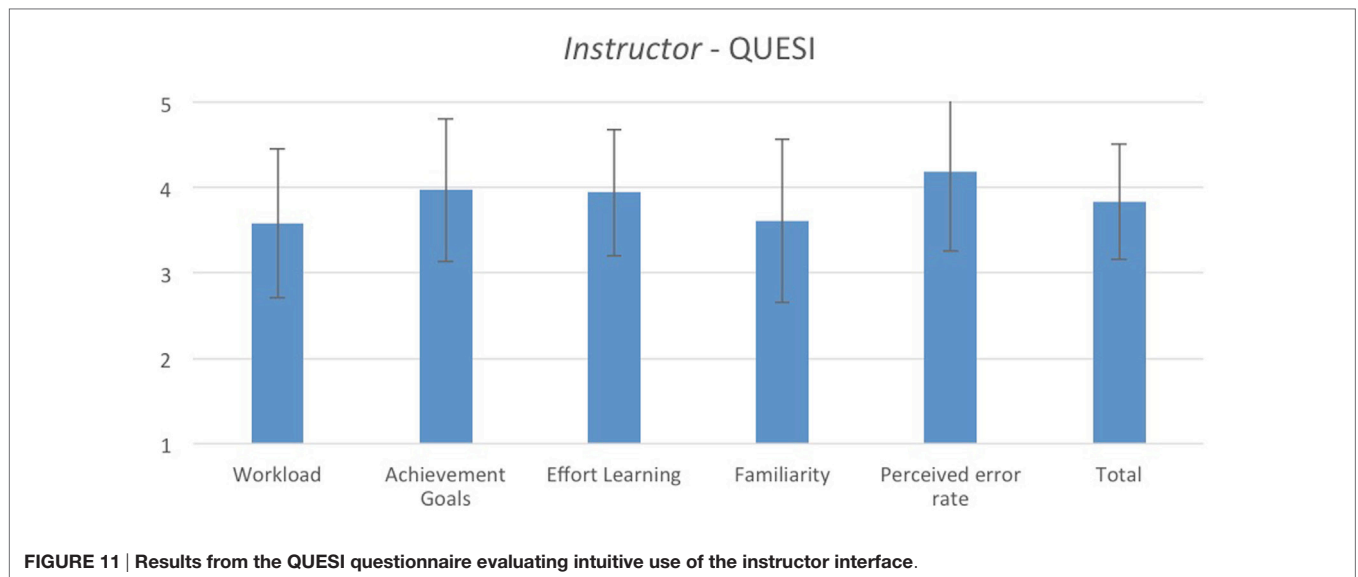
5.4. Teachers' Performances

We averaged the instructors' asynchronous feedback (five criteria: time to resolve, time to react, coping strategy, appropriate
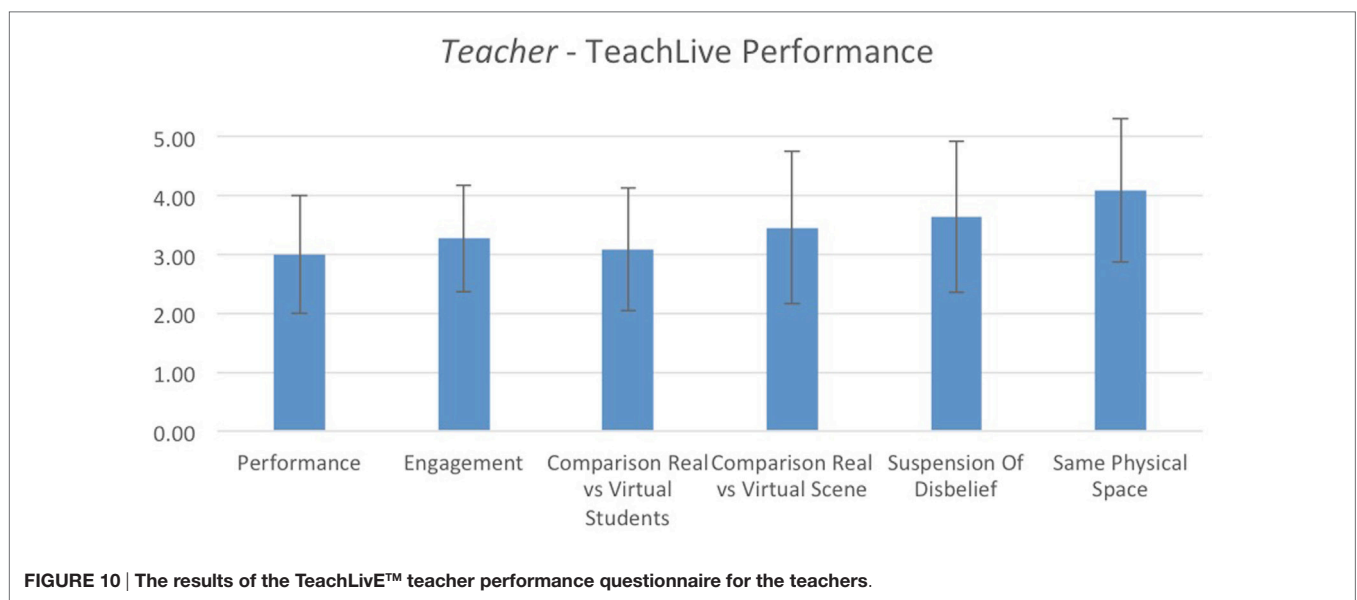
FIGURE 10 | The results of the TeachLivE™ teacher performance questionnaire for the teachers.

FIGURE 11 | Results from the QUESI questionnaire evaluating intuitive use of the instructor interface.

FIGURE 12 | Results from the technology acceptance questionnaire.

tone, and self-control). We then analyzed its correlation with the teachers' subjective ratings of their own performance, measured from TeachLivE™ teacher performance (Hayes et al., 2013a). Although not significant ($p = 294$), the analysis shows a moderate correlation between the teachers' self-evaluations and the evaluations from the instructors ($r = 348$), indicating that both teachers and instructors had similar impressions.

5.5. Technology Acceptance

We individually assessed both the student group and the instructor group for a user-centered system evaluation. In general, the instructor group had a higher acceptance for the technology than the teacher group, except for the question of whether this system could enhance their education, which students rated slightly higher. Scores were above scale average except for the price and reality closeness, which students rated lower than expected (see **Figure 12**).

The usability score for the instructor underlines the previous results. Additionally, according to the scores, both groups felt that the system could improve their studies ($M_{student} = 3.27$, $SD_{student} = 1.21$, $M_{instructor} = 3.68$, $SD_{instructor} = 0.96$) and could be enhanced with additional scenarios ($M_{student} = 4.04$, $SD_{student} = 0.79$, $M_{instructor} = 3.64$, $SD_{instructor} = 0.67$). The teachers rated an acceptable price of 3.587 € for the system, while seven of eleven students felt that the system should be introduced into their education. In contrast, the instructor group rated an acceptable price of 2.370 €,

even though nine of eleven instructors felt that the system would be beneficial for their education.

6. DISCUSSION

With respect to the three main aspects of the applicability of the proposed system from Section 1.1, the results match the evaluation criteria as follows:

E1. Simulation effects by the immersive teacher interface.

  i. *Believability*: the behavior of the virtual students was convincing and believable, as reported by teachers, and was capable of eliciting stress reactions. This was a critical requirement for a our training simulation. The above-average scores for the items suspension of disbelief and same physical space affirm a compelling reproduction of a real environment, which now provides a controlled CRM training method.

  ii. *Cybersickness and potential side effects*: we observed that no simulation sickness was induced by our system. This is a critical feature, as cybersickness could dramatically interfere with the objectives of our system (LaViola, 2000). During the session, it would prevent teachers from learning and practicing their skills. The after effects would also prevent teachers from performing any other activities such as assisting other lectures, seminars, or exams, and even to safely drive back home.

  iii. *Effect of feedback cues*: it appears that simple synchronous audio feedback cues did not affect the overall experience (i.e., no difference of immersion, believability, performance, and acceptance) to a strong degree. This result might indicate that real-time sound feedback could be used to guide the trainee without interfering with the experience, especially with its believability, which is critical for stress exposure therapy. However, our results can only be seen as a first exploratory analysis and therefore should not be generalized.

E2. Usability of the instructor interface:

  i. *Task and cognitive load and intuitive usage of the interface*: the instructor's GUI was evaluated as good in usability and did not have a high task load. The GUI allows them to simply control the overall class and individual behavior while following general instruction and giving teachers feedback. It was also a critical requirement to achieve, as the instructor's job is demanding because it requires frequent shifts of attention and fast evaluation. Our GUI design and its associated virtual agent control algorithm appear to efficiently support these requirements.

E3. Technology acceptance by teachers and instructors:

  The system's price should be low enough, and people felt the system is beneficial. Overall, our system was well received, and more than 80% of participants said our system will help them in their education and future career. Therefore, it appears that our system convinced its possible future users, and so its acceptability is fulfilled.

7. LIMITATIONS AND FUTURE WORK

Despite a limited sample size, our first usability evaluation gave us promising results and appears to demonstrate satisfaction with the critical system requirements (i.e., low cybersickness, easy to use/learn, believable, and being accepted). However, our evaluation also revealed certain limitations and improvements to consider. One important improvement of the system will consist of a integrating a larger set of possible behaviors to create richer scenarios and alternatives. It will require the exploration and evaluation of alternative interaction techniques for the instructors to allow for their fast selection and activation. Another aspect regards the integration of facial animations synchronized with text-to-speech and bad behavior animation. The possibility of recording live sessions and later replicating the exact same scenarios is an important feature under development. We are exploring the possibilities to provide an online exchange platform to support the publication and sharing of scenarios among instructors to enable widespread use of the proposed system in a teacher education curriculum.

We also would like to expand the autonomy of the virtual students and provide the instructor with an option to be automatically driven by physiological inputs such as electrodermal activity and heart rate. Students will then automatically adjust their behaviors according to the personality and the perceived stress level of the teacher. To a certain extent, the students, agitation will reflect the teachers', forcing teachers to control their stress to control their class. We are exploring a solution to couple a pleasure-arousal-dominance model (Mehrabian, 1996) with a multimodal discourse analysis (MDA) module, automatically interpreting teachers' body language and speech, as by Barmaki and Hughes (2015b) and Barmaki (2015). Finally, the question of a possible *uncanny valley* effect (Mori, 1970) should be explored, especially with more photo-realistic virtual humans. This effect could dramatically deteriorate the teacher's experience and cancel out any benefits of the immersive VR approach. In our future research, we are planning to test this hypothesis by replicating our experiment with photo-realistic virtual students.

Future prototypes will also evaluate the performances and affordances of eye-tracking HMD such as the FOVE VR headset (FOVE, Inc., 2015) or SensorMotoric eye-tracking package for Oculus DK2 (SensorMotoric, 2016). One advantage of eye tracking is that it permits users to dynamically increase the rendering quality based on the user's center focus and therefore should help to approach photo-realistic quality. It will also permit a more accurate visualization of the teacher's point of gaze. Second, the offline analysis of eye movement and focus points will provide valuable input for behavioral and cognitive studies. It could lead to the identification of visual attention patterns in stressful situations, possibly acting as indicator of the teacher's stress level. Last, it will support novel modes of hands-free interactions with the virtual students or environments (e.g., display the student's name, activate or deactivate devices inside the classroom, a menu to control the system, navigate inside the class).

Other important future work consists in realizing a field study in schools with real teachers who already possess practical

experience in class management. Now that the usability and efficiency of the system has offered promising results, we are planning collaborations with local schools on the evaluation of an extended version of our system. This includes studies comparing our new proposed method to classic theory-based or role-play-based methods.

Considering the aspect of the ever-growing mixed-cultural composition of current classrooms, we believe that our system also has great potential to help teachers enhance their current teaching practices to new challenges posed by mixed-cultural and mixed-level classrooms. Therefore, in our future work, models of different culture-related behaviors need to be modeled to drive the different goals, believes, and behaviors of potential students having different cultural backgrounds. With these additional features, the system could not only be used to prepare teachers for arising changes in classrooms but also serve as showcases to pupils who are sometimes faced with fixed rule-sets that need to be understood to integrate into existing scholar systems.

8. CONCLUSION

This article presented a novel immersive virtual reality CRM training system. The training of interpersonal skills necessary for successful CRM is critical to improve the experiences of teachers and students. Therefore, 3B has been designed as a companion tool for classroom management seminars in a syllabus for primary and secondary school teachers.

Compared to previous systems, our approach solves both important functional and non-functional requirements necessary to the successful integration of a virtual training in teacher education. It is capable of conveying controllable and replicable situations of believable classroom scenarios, including disruptive bad behaviors by virtual students, which increases stress for the teacher and hinders easy coping with these situations. The system also provides interactive control, monitoring, and evaluation of the current scenario and the teacher's performance by an instructor.

The system allows easy and everyday use in seminar rooms and teaching environments while still providing a rich and compelling experience. It uses affordable, off-the-shelf hardware and software, and it allows users to come as they are without the need for any extensive hardware and sensor rigging, which is typical of many VR systems. The 3B project is designed to help beginner, intermediate, and expert teachers to improve their techniques of recognition, anticipation, and quality of reaction to disruptive bad behavior to prevent out-of-control situations in classrooms.

Our first evaluation revealed promising results and demonstrated the successful fulfillment of the critical usability requirements (i.e., safety and comfort, believability, simplicity, acceptability, extensibility, affordability, and mobility). This constitutes the necessary first step toward a possible investigation of its efficiency, and effectiveness in terms of learning outcomes and experience with short- and long-term studies.

AUTHOR CONTRIBUTIONS

Substantial contributions to the conception or design of the work were performed by J-LL, ML, MH, CS, and SG. Initial idea by ML further pushed on and elaborated by J-LL. Technical development by J-LL and MH. Educational aspects contributed by CS and SG. Evaluation designed by J-LL, ML, MH, CS, DR, and SG. Implementation of the evaluation by MH. Analysis performed by J-LL and DR, interpretation done by J-LL, DR, MH, and ML. Drafting the work or revising it critically for important intellectual content, approval of the version to be published, and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved by J-LL, ML, MH, CS, DR, and SG.

ACKNOWLEDGMENTS

REFERENCES

Alibali, M. W., and Nathan, M. J. (2007). "Teachers' gestures as a means of scaffolding students' understanding: evidence from an early algebra lesson," in *Video Research in the Learning Sciences*, eds R. Goldman, R. Pea, B. Barron, and S. J. Derry (Mahwah, NJ: Erlbaum), 349–365.

Autodesk. (2016a). *Autodesk Character Generator*. Available at: https://character-generator.autodesk.com/

Autodesk. (2016b). *Autodesk Maya*. Available at: http://www.autodesk.de/store/products/maya

Autodesk. (2016c). *Autodesk Motionbuilder*. Available at: http://www.autodesk.com/products/motionbuilder/

Baker, D. P., Gustafson, S., Beaubien, J., Salas, E., and Barach, P. (2005). *Medical Teamwork and Patient Safety: The Evidence-Based Relation*. Washington, DC: AHRQ Publication.

Barmaki, R. (2015). "Multimodal assessment of teaching behavior in immersive rehearsal environment-teachlive," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (Orlando, FL: ACM), 651–655.

Barmaki, R., and Hughes, C. E. (2015a). "A case study to track teacher gestures and performance in a virtual learning environment," in *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (Orlando, FL: ACM), 420–421.

Barmaki, R., and Hughes, C. E. (2015b). "Providing real-time feedback for student teachers in a virtual rehearsal environment," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, ICMI '15* (New York, NY: ACM), 531–537.

Borich, G. D. (2011). *Effective Teaching Methods: Research-Based Practice*. Austin: University of Texas.

Brophy, J. (2006). "History of research on classroom mangement," in *Handbook of Classroom Management. Research, Practice, and Contemporary Issues*, eds C. M. Evertson, and C. S. Weinstein (New York, NY: Routledge), 17–43.

Brouwers, A., and Tomic, W. (1999). Teacher burnout, perceived self-efficacy in classroom management, and student disruptive behaviour in secondary education. *Curric. Teach.* 14, 7–26. doi:10.7459/ct/14.2.02

Canter, L. (2011). *Assertive Discipline: Positive Behavior Management for Today's Classroom*. Bloomington, IN: Solution Tree Press.

Canter, L., and Canter, M. (1992). *Assertive Discipline: Positive Behavior Management for Today's Schools*. Bloomington: Solution Tree Press.

Charles, D., Charles, T., McNeill, M., Bustard, D., and Black, M. (2011). Game-based feedback for educational multi-user virtual environments. *Br. J. Educ. Technol.* 42, 638–654. doi:10.1111/j.1467-8535.2010.01068.x

Dieker, L. A., Rodriguez, J. A., Lignugaris, B., Hynes, M. C., and Hughes, C. E. (2013). The potential of simulated environments in teacher education: current and future possibilities. *Teach. Educ. Special Educ.* doi:10.1177/0888406413512683

Emmer, E. T., and Stough, L. M. (2001). Classroom management: a critical part of educational psychology, with implications for teacher education. *Educ. Psychol.* 36, 103–112. doi:10.1207/S15326985EP3602_5

Epic Games, Inc. (2015). *Unreal Engine 4*. Available at: https://www.unrealengine.com/unreal-engine-4

Epic Games Inc. (2016). *Unreal Motion Graphics UI Designer*. Available at: https://docs.unrealengine.com/latest/INT/Engine/UMG/

Evertson, C. M., and Weinstein, C. S. (2013). *Handbook of Classroom Management: Research, Practice, and Contemporary Issues*. Routledge.

FOVE, Inc. (2015). *Fove HDM with Eye Tracking*. Available at: http://www.getfove.com/

Gupta, S. K., Anand, D. K., Brough, J., Schwartz, M., and Kavetsky, R. (2008). *Training in Virtual Environments. A Safe, Cost-Effective, and Engaging Approach to Training*. University of Maryland.

Hale, K. S., Stanney, K. M., Lindgren, R., Moshell, J. M., and Hughes, C. E. (2014). "Virtual environments as a tool for conceptual learning," in *Handbook of Virtual Environments: Design, Implementation, and Applications*, 2nd Edn (Boca Raton, FL: CRC Press), 1043–1055.

Hart, S., and Staveland, L. (1988). "Development of NASA-TLX (task load index): results of empirical and theoretical research," in *Human Mental Workload*, eds P. A. Hancock and N. Meshkati (Amsterdam: Elsevier), 139–183.

Hattie, J. (2003). *Teachers Make a Difference, What Is the Research Evidence?* Melbourne: ACER Research Conference.

Hayes, A. T., Hardin, S. E., and Hughes, C. E. (2013a). "Perceived presence's role on learning outcomes in a mixed reality classroom of simulated students," in *International Conference on Virtual, Augmented and Mixed Reality* (Berlin: Springer), 142–151.

Hayes, A. T., Straub, C. L., Dieker, L. A., Hughes, C. E., and Hynes, M. C. (2013b). Ludic learning: exploration of TLE TeachLivE™ and effective teacher training. *Int. J. Gaming Comput. Mediated Simul.* 5, 20–33. doi:10.4018/jgcms.2013040102

He, D., Fuhu, D. H., Pape, D., Dawe, G., and Sandin, D. (2000). "Video-based measurement of system latency," in *International Immersive Projection Technology Workshop* (Chicago), 111.

Honey, M. A., and Hilton, M. (2011). *Learning Science through Computer Games and Simulations*. Washington, DC: National Academies Press.

Hughes, C. E., and Ingraham, K. M. (2016). "De-escalation training in an augmented virtuality space," in *Virtual Reality (VR), 2016 IEEE* (IEEE), p. 181–182.

Hurtienne, J., and Naumann, A. (2010). *QUESI—A Questionnaire for Measuring the Subjective Consequences of Intuitive Use*. Cambridge: Interdisciplinary College, 536.

iPiSoft I. (2016a). *iPi Mocap Studio*. Available at: http://wiki.ipisoft.com/IPi_Mocap_Studio

iPiSoft I. (2016b). *iPi Recorder*. Available at: http://wiki.ipisoft.com/IPi_Recorder

Jones, H. E., Sabouret, N., Damian, I., Baur, T., André, E., Porayska-Pomsta, K., et al. (2014). Interpreting social cues to generate credible affective reactions of virtual job interviewers. *Artif. Intell.* arXiv preprint arXiv:1402.5039.

Jones, V. (2006). *How Do Teachers Learn to Be Effective Classroom Managers?* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Kelly, S. D., Manning, S. M., and Rodak, S. (2008). Gesture gives a hand to language and learning: perspectives from cognitive neuroscience, developmental psychology and education. *Lang. Linguist. Compass* 2, 569–588. doi:10.1111/j.1749-818X.2008.00067.x

Kelly, S. D., McDevitt, T., and Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Lang. Cogn. Process.* 24, 313–334. doi:10.1080/01690960802365567

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* 3, 203–220. doi:10.1207/s15327108ijap0303_3

Kenny, P., Parsons, T. D., Gratch, J., Leuski, A., and Rizzo, A. A. (2007). "Virtual patients for clinical therapist skills training," in *Intelligent Virtual Agents* (Springer), 197–210.

Keppell, C. (1998). Virtual reality: what is the state of play in education? *Australas. J. Educ. Technol.* 14, 60–74.

Kilteni, K., Groten, R., and Slater, M. (2012). The sense of embodiment in virtual reality. *Presence Teleop. Virt. Environ.* 21, 373–387. doi:10.1162/PRES_a_00124

Kounin, J. S. (1970). *Discipline and Group Management in Classrooms*. Huntington, NY: R.E. Krieger Pub. Co.

Kunter, M., Seidel, T., and Artelt, C. (2015). *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, Göttingen: Hogrefe.

Laslett, R., and Smith, C. (2002). *Effective Classroom Management: A Teacher's Guide*. Routledge.

LaViola, J. J. Jr. (2000). A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin* 32, 47–56. doi:10.1145/333329.333344

Leuski, A., Patel, R., Traum, D., and Kennedy, B. (2009). "Building effective question answering characters," in *Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue* (Stroudsburg: Association for Computational Linguistics), 18–27.

Levin, J., and Nolan, J. F. (1991). *Principles of Classroom Management: A Hierarchical Approach*. Upper Saddle River, NJ: Prentice Hall.

Levin, J., and Nolan, J. F. (2013). *Principles of Classroom Management: A Professional Decision-Making Model*. Pearson Higher Ed.

Lok, B., Naik, S., Whitton, M. C., and Brooks, F. P. (2003). Effects of handling real objects and self-avatar fidelity on cognitive task performance and sense of presence in virtual environments. *Presence* 12, 615–628. doi:10.1162/105474603322955914

Lopez, A. L., Hughes, C. E., Mapes, D. P., and Dieker, L. A. (2012). "Cross-cultural training through digital puppetry," in *Advances in Design for Cross-Cultural Activities (Part I)*, ed. D. M. Nicholson. (CRC Press), 247–256.

Lugrin, J.-L., Charles, F., Cavazza, M., Le Renard, M., Freeman, J., and Lessiter, J. (2012). "Caveudk: a VR game engine middleware," in *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology, VRST '12* (New York, NY: ACM), 137–144.

Lugrin, J.-L., Wiebusch, D., Latoschik, M. E., and Strehler, A. (2013). "Usability benchmarks for motion tracking systems," in *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology, VRST '13* (ACM), 49–58.

Mahon, J., Bryant, B., Brown, B., and Kim, M. (2010). Using second life to enhance classroom management practice in teacher education. *Educ. Media Int.* 47, 121–134. doi:10.1080/09523987.2010.492677

Mayr, J., Eder, F., and Fartacek, W. (1991). Mitarbeit und störung im unterricht: Strategien pädagogischen handelns. *Zeitschrift für pädagogische Psychologie* 5, 43–55.

Mehrabian, A. (1996). Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. *Curr. Psychol.* 14, 261–292. doi:10.1007/BF02686918

Microsoft C. (2016). *Microsoft Kinect v2*. Available at: https://developer.microsoft.com/en-us/windows/kinect/

Mori, M. (1970). Bukimi no tani [The uncanny valley]. *Energy* 7, 33–35.

Nagendran, A., Pillat, R., Kavanaugh, A., Welch, G., and Hughes, C. (2013). "Amities: avatar-mediated interactive training and individualized experience system," in *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology, VRST '13* (New York, NY: ACM), 143–152.

Nagendran, A., Welch, G., Hughes, C., and Pillat, R. (2015). "Technical report: exploring human surrogate characteristics," in *Virtual Realities* (Springer), 215–228.

Naumann, A., and Hurtienne, J. (2010). "Benchmarks for intuitive interaction with mobile devices," in *Proceedings of the 12th Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI '10* (Berlin: ACM), 401–402.

Oculus VR LLC. (2016). *Oculus Developement Kit 2*. Available at: https://www.oculus.com/en-us/dk2/

Opaque Multimedia C. (2016). *Kinect 4 Unreal Plugin*. Available at: http://www.opaque.media/kinect-4-unreal/

Sanchez-Vives, M. V., and Slater, M. (2005). From presence to consciousness through virtual reality. *Nat. Rev. Neurosci.* 6, 332–339. doi:10.1038/nrn1651

Schuemie, M. J. (2003). *Human-Computer Interaction and Presence in Virtual Reality Exposure Therapy*. Delft: TU Delft, Delft University of Technology.

Schutte, J. G. (1997). *Virtual Teaching in Higher Education*. Northridge, CA: California State University-Northridge.

Seitz, O. (1991). *Problemsituationen im Unterricht*. Cologne: Wolf.

Seitz, O. (2004). *Unterrichtsstörungen*. Nuremberg: Univ., Lehrstuhl für Schulpädag.

SensorMotoric. (2016). *Sensormotoric Instrument Eye Tracking Package for Oculus DK2*. Available at: http://www.smivision.com/

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 3549–3557. doi:10.1098/rstb.2009.0138

Spanlang, B., Normand, J.-M., Borland, D., Kilteni, K., Giannopoulos, E., Pomés, A., et al. (2014). How to build an embodiment lab: achieving body representation illusions in virtual reality. *Front. Rob. AI* 1:9. doi:10.3389/frobt.2014.00009

Steed, A. (2008). "A simple method for estimating the latency of interactive, real-time graphics simulations," in *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology, VRST '08* (New York, NY: ACM), 123–129.

Straub, C., Dieker, L., Hynes, M., and Huges, C. (2014). *Using Virtual Rehearsal in TLE TeachLivE™ Mixed Reality Classroom Simulator to Determine the Effects on the Performance of Mathematics Teachers*. 2014 TeachLivE™ National Research Project: Year 1 Findings.

Thomas, D. R., Becker, W. C., and Armstrong, M. (1968). Production and elimination of disruptive classroom behavior by systematically varying teacher's behavior. *J. Appl. Behav. Anal.* 1, 35–45. doi:10.1901/jaba.1968.1-35

Tichon, J., Hall, R., Hilgers, M. G., Leu, M., and Agarwal, S. (2003). "Education and training in virtual environments for disaster management," in *Proceedings of World Conference on Educational Multimedia, Hypermedia, & Telecommunications (EdMedia)* (Brisbane: AACE Honolulu), 1191–1194.

Tichon, J. G. (2007). Using presence to improve a virtual training environment. *Cyberpsychol. Behav.* 10, 781–788. doi:10.1089/cpb.2007.0005

Walker, H. M. (1995). *The Acting-Out Child: Coping with Classroom Disruption*. Longmont, CO: Sopris West.

Waltemate, T., Hülsmann, F., Pfeiffer, T., Kopp, S., and Botsch, M. (2015). "Realizing a low-latency virtual reality environment for motor learning," in *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology* (Bielefeld: ACM), 139–147.

Wang, W., and Loewen, S. (2015). Nonverbal behavior and corrective feedback in nine ESL university-level classrooms. *Lang. Teach. Res.* doi:10.1177/1362168815577239

Williamon, A., Aufegger, L., and Eiholzer, H. (2014). Simulating and stimulating performance: introducing distributed simulation to enhance musical learning and performance. *Front. Psychol.* 5:25. doi:10.3389/fpsyg.2014.00025