British Journal of Educational Technology Vol 52 No 2 2021 824–841
doi:10.1111/bjet.13058

Situating presence within extended reality for teacher training: Validation of the extended Reality Presence Scale (XRPS) in preservice teacher use of immersive 360 video

Enrico Gandolfi , Karl W. Kosko and Richard E. Ferdig

Enrico Gandolfi, Ph.D., is an assistant professor in Educational Technology at Kent State University with the Research Center for Educational Technology. Previously he worked as an associate researcher at Luiss "Guido Carli" University of Rome, Italy. His research interests span games and simulations for learning, augmented and virtual reality, accessibility and critical design in education, digital interfaces for special populations and instructional streaming. He is author of several journal articles (eg, for Journal of Gaming & Virtual Worlds, Convergence, Games and Culture, Simulation and Gaming, E-learning and Digital Media, Information Visualization, Feminist Media Studies and International Journal of Information and Learning Technology) and book chapters (for publishers like Routledge, MIT Press, ETC Press) about these topics and of the monographs Piloti di Console (Edizioni Paoline, 2011), Nerd Generation (Mimesis, 2014) and Independent Videogames (Unicopli, 2015). He was co-P.I. and key personnel for grants awarded by National Science Foundation, GAR Foundation and National Endowment for Humanities. He teaches Instructional Design, Educational Technology, Games and Simulations, Virtual and Augmented Reality and Researching Issues in Educational Technology. Karl W. Kosko earned his undergraduate degree in Elementary Education, with a minor in mathematics, at Winthrop University in Rock Hill, SC. While teaching in the Rock Hill School District, he also received his M.Ed. in Middle Level Education with an emphasis in mathematics. Later, he attended Virginia Tech and earned his doctorate in Curriculum and Instruction with an emphasis in Mathematics Education. He then completed a post-doctoral fellowship at the University of Michigan with the GRIP project before taking a position at Kent State University. His research interest focuses on how mathematical meaning is conveyed, including studies of mathematical argument & discourse, multiplicative reasoning and using representations of practice in elementary mathematics teacher education. Richard E. Ferdig is the Summit Professor of Learning Technologies and Professor of Educational Technology at Kent State University. He works within the Research Center for Educational Technology and also the School of Teaching, Learning and Curriculum Studies. He earned his Ph.D. in Educational Psychology from Michigan State University. He has served as researcher and instructor at Michigan State University, the University of Florida, the Wyzsza Szkola Pedagogiczna (Krakow, Poland) and the Università degli studi di Modena e Reggio Emilia (Italy). At Kent State University, his research, teaching and service focus on combining cutting-edge technologies with current pedagogic theory to create innovative learning environments. His research interests include online education, educational games and simulations, the role of faith in technology and what he labels a deeper psychology of technology. In addition to publishing and presenting nationally and internationally, Ferdig has also been funded to study the impact of emerging technologies such as K-12 Virtual Schools. Rick was the founding Editor-in-Chief of the International Journal of Gaming and Computer Mediated Simulations, is the current Editor-in-Chief of the Journal of Technology and Teacher Education and also serves as a Consulting Editor for the Development Editorial Board of Educational Technology Research and Development. Address for correspondence: Enrico Gandolfi, Kent State University, 800 Hilltop Dr, Kent, OH 44242, USA. Email: eganndol1@kent.edu

Abstract
The use of video is commonplace for professional preparation in education and other fields. Research has provided evidence that the use of video in these contexts can lead to increased noticing and reflection. However, educators now have access to evolving forms of video such as 360 video. The purpose of this study was to adapt and validate an instrument for assessing immersive 360 video use in an undergraduate preservice teacher university training program. Data provided evidence of the validity of the *Extended Reality Presence Scale* (XRPS) for 360 video research in preservice teacher professional development. Moreover, evidence from the study suggests that those with higher feelings of presence are less likely to jump around (or twitch) while watching 360 videos. The main implications are that: a) the XRPS is a validated and reliable instrument and b) more research is needed to examine the presence and practices for in-service and preservice teachers while watching 360 video.

Key words

360 video, presence, scale validation, teacher professional development

Introduction

The use of instructional videos is a well-established practice in preservice teacher training (Gaudin & Chaliès, 2015). Videos are used for several reasons including the beneficial outcomes of improved noticing and increased reflective attitudes (eg, Fadde & Sullivan, 2013). Videos have also been widely used for introducing current and future educators to pedagogical strategies, improving self-observation and increasing attitudes towards and knowledge of teaching techniques like evaluation.

Recently, technological innovations have led to novel techniques and features for experiencing audio-visual content like 360 videos. The use of 360 videos for professional preparation adds a layer of complexity and promise. For instance, watching a 360 video of a lesson allows users to observe more actions due to its expanded scope; moreover, if the 360 video is watched with a headset or head-mounted display (eg, immersive 360 videos), factors like sensory engagement and embodied interaction are introduced (Ferdig & Kosko, 2020). Unfortunately, research on the use of 360 video for professional preparation of teachers is nascent.

This paper aims to address this gap with an emphasis on immersive 360 videos in the context of undergraduate preservice teachers training at the university level. This focus is motivated by two specific problems in current literature.

First, although the research on 360 video that does exist is promising (eg, Theelen, van den Beemt, & den Brok, 2019), the field lacks instruments for understanding its impact on immersive technology concepts like presence. Presence or the sense of being there (Lee, 2004), plays a relevant role in informing understanding and engagement within mediated environments.

Second, although video is a significant part of teacher training, there are few studies that address this immersive tool within preservice teacher education. As such, this study addresses both gaps by (a) presenting a new instrument called the extended Reality Presence Scale (XRPS), adapted from an existing presence scale and then, tested with 44 undergraduate preservice teachers who

Practitioners notes

What is already known about this topic?

- Instructional videos are widely used in preservice teacher training.
- 360 videos show promise for improving preservice teacher professional development in terms of immersion and presence.

What this paper adds?

- An instrument for assessing 360 video teacher presence is presented (XRPS), targeting a current gap in the literature.
- Data provided evidence of the validity of the tool for future 360 video research and integration.

Implications for practice and/or policy

- Practitioners can use XRPS for assessing preservice teachers' experiences in immersive environments and evaluating 360 videos.
- Higher feelings of presence are associated with more focused viewpoints. Therefore, practitioners should support and facilitate this watching behavior.
- Higher scores of presence are associated with a perceived sense of agency and emotional attachment. Therefore, 360 videos should include design elements promoting these feelings.

watched 360 videos of classroom practice, (b) exploring the concept of presence itself in this context of professional development and beyond through XRPS validation, and c) investigating if presence is correlated to specific watching patterns in 360 videos in preservice teacher training. In accomplishing these goals, this study will also shed light on the concept of presence in immersive environments for professional training (an undertheorized construct) and the related impact on learners' behaviors within these mediated contexts.

Videos for preservice teaching training

The use of videos for instruction have been widely adopted in preservice teacher training. This is due, in part, to the research-based evidence of the growth of self-reflection and teaching efficacy in a number of content areas and disciplines (Blomberg, Stürmer, & Seidel, 2011; Ottenbreit-Leftwich, Glazewski, Brush, Aslan, & Zachmeier, 2018; Weber, Gold, Prilop, & Kleinknecht., 2018). Video clips have been used to prepare novice as well expert teachers in improving their self-reflection skills and ability to notice relevant learning events, from language to math (Eröz-Tuğa, 2013; Fadde & Sullivan, 2013). Han, Eom, and Shin (2013) found that video clips can improve predispositions to educational technologies. Wiens, Hessberg, LoCasale-Crouch, and DeCoster (2013) successfully used videos for assessing preservice teachers' ability to understand meaningful understanding learning environments, finding that academic factors (eg, year of college, type of classes taken, etc.) might have an impact on students' performance.

Videos have been effective in comparing and measuring ways that experienced and novice teachers attend to particular moments in recorded classrooms (Dessus, Cosnefroy, & Luengo, 2016; Cortina, Miller, McKenzie, & Epstein, 2015; van den Bogert, Bruggen, Kostons, & Jochems, 2014). For example, Dessus *et al*. (2016) used eye-tracking technology in comparing less and more experienced teachers' viewing of a classroom video. They found that less experienced teachers attempted to focus on multiple students, whereas more experienced teachers were able to perceive all students but focus on select sets of students.

Seidel, Blomberg, and Renkl (2013) explored different approaches to video adoption in teacher professional development and found that illustrative clips worked better than example-based clips in terms of reproducing factual knowledge and evaluating videotaped classroom situations, and vice versa or lesson planning. Finally, Gaudin and Chaliès (2015) provided a meta review of the literature on video viewing in teacher education and professional development. Their results pointed to three main fronts to develop: the ability to transfer what was watched to real-life experiences, how to personalize instructional videos and how to make videos a common practice in preservice training. However, immersive reality lacks such a corpus of evidence.

Immersive reality refers to a set of digital technologies that aim to situate the user within their virtual boundaries (Ferdig, Gandolfi, & Immel, 2018); in other words, immersive environments immerge the user in an alternative reality for providing experiences that are challenging to access in real life. This can be done via either artificial settings (eg, a computer-generated setting) or recorded videos (eg, 360 videos). A 360 video is one that can be viewed in any direction at the same time (360 degrees); they can be watched on a flat screen or using a headset-mounted display. The latter option has the potential to make them immersive, allowing the user to turn his/her head around and observe the surroundings in any direction (Rupp *et al.*, 2019). This technology can have a potential impact on preservice teacher noticing and reflection (Ferdig & Kosko, 2020; Kosko, Ferdig, & Zolfaghari, in press), making training videos more engaging and multi-faceted due to the broader viewpoint and the feeling to *be there* (Lee, 2004).

There is preliminary evidence of the potential of 360 video for professional preparation. For instance, Theelen *et al.* (2019) used 360 videos about classroom events with 141 first year preservice teachers. They found a significant increase in terms of noticing and use of theory-based terminology. Roche and Gal-Petitfaux (2017) explored 360 videos for future physical education teachers and found that this technology can provide more immersive and richer experiences to explore and analyze. Ferdig and Kosko (2020) found that preservice teachers perceived 360 videos as more immersive than standard videos. Finally, Walshe and Driver (2019) investigated preservice teacher self-reflection with 360 videos with an interpretive case study based on think-aloud protocol and interviews. Their study results showed promise in improving micro-teaching practice understanding and self-efficacy.

Although these studies are promising, they are exploratory. More troubling, the field lacks reliable instruments to address the implementation of the 360 technology. There is, therefore, a need of criteria to refer to while analyzing immersive 360 videos for current and future educators. There are two leading reasons for such a need. First, 360 videos are becoming increasingly used in preservice teaching practice (Ferdig & Kosko, 2020); there is strong evidence that the current COVID-19 situation will probably strengthen their use (Zolfaghari, Austin, Kosko, & Ferdig, 2020; Ferdig & Kosko, 2020). Second, this technology is more accessible and user-friendly than other solutions (eg, virtual reality) due to its relatively low cost and high usability. As both technology and implementation approaches improve, there is promise of more immersive videos addressing Gaudin and Chalies' (2015) fronts for development.

Immersive virtual reality has been often evaluated in terms of presence, which can be defined as the sense of *being there* or *naturalness* (Bianchi-Berthouze, Kim, & Patel, 2007; Lee, 2004; Mestre, 2005). The core focus of immersive technologies is indeed on capturing the user's senses in the most complete way (Freina & Ott, 2015; Lorenzo, Pomares, & Lledó, 2013), generating a feeling of presence where the mediation of technology disappears. Therefore, presence becomes a desired outcome for these technologies. This effect would help immersive virtual reality in allowing users to experience situations and learning outcomes that are challenging to access in real life (Lau & Lee, 2015; Lee & Wong, 2014; Webster, 2016).

Operationalizing presence is a challenge for at least three main reasons. First, the concept of presence itself is vague and interpreted in a variety of different ways, being conflated with terms like immersion and embodiment (Calleja, 2011; Farrow & Iacovides, 2012). Second, only a few instruments have been produced that have a focus on artificial virtual environments (Bianchi-Berthouze, Kim, & Patel, 2007; Makransky, Lilleholt, & Aaby, 2017). Third, almost no efforts have been made in understanding how presence has an impact on concrete actions and behaviors in immersive virtual environments (eg, exploring whether users act differently according to their perceived sense of presence), which can shed light on the construct itself. More important, existing presence scales do not distinguish between 360 videos experienced via a head-mounted display and 360 videos watched on a flat screen (Ferdig & Kosko, 2020), despite evidence of observable differences (Kosko *et al.*, in press).

To the best of our knowledge, no studies have contextualized presence within preservice teacher training with immersive 360 videos. These 360 videos can be considered a *mixed* or *extended reality* (Bower, Lee, & Dalgarno, 2017; Nardi, 2015) due to the realism conveyed by real-life recordings. This technology is becoming increasingly accessible and, therefore, it can represent a feasible option for teacher professional development. This paper aims to present the first attempt to address presence-related issues and challenges, validating a potential instrument about presence and exploring how immersive environments for training are experienced by preservice teachers. Moreover, an additional focus is on how 360 videos are experienced by relying on the concept of a viewer's *twitch*. A twitch could best be conceived as an irregular change of focus in the participant's perspective, thus indicating a disruption of the user's viewpoint.

Therefore, the emphasis of this paper is on preservice teacher preparation via 360 video and related presence, which becomes a key criterion to investigate but also problematize for supporting students' engagement and knowledge building. By uncovering the construct of presence itself (which is relevant but not well defined yet in the literature about immersive environments), the objective is also to set the stage for broader discussions about the impact of this measure in immersive environments for professional development in education and beyond.

To summarize, this study includes the following research foci:

1. Validating an instrument for measuring presence in immersive environments for preservice teacher training (eg, the extended Reality Presence Scale or XRPS).
2. Understanding presence from the XRPS validation (eg, what are the immersive presence parameters in this context?).
3. Analyzing the recorded videos for detecting viewpoints patterns in relation with presence scores (eg, is presence correlated to specific watching behaviors?).

Materials and Methods

This study relies on a Rasch analysis for validating a new instrument (the *extended Reality Presence Scale* or XRPS) and a video analysis for collecting watching patterns. Preservice teachers were recruited for viewing 360 videos and answering a related questionnaire about demographics, previous experience with immersive technologies and XRPS.

Table 1: Distribution of academic majors among participants

Academic Majors of Participants		*n*	%
Early Childhood/Elementary		7	15.9
Middle Childhood	Math & Science	2	4.5
	Language Arts & Social Studies	2	4.5
Secondary	Math	3	6.8
	Science	1	2.3
	Language Arts	7	15.9
	Social Studies	9	20.5
Multi-Grade	Art	7	15.9
	French	2	4.5
	American Sign Language	2	4.5
	School Health	1	2.3
	Special Education	1	2.3

Sample

Data were collected from 44 undergraduate students enrolled in an undergraduate educational technology course. Such a sample size is useful for Rasch modeling, used in this paper, and allows for item calibrations and/or person measures that are stable within 1 logit (Linacre, 1994). The majority of participants identified as white (90.9%) and female (63.6%), with 54.5% of the sample identifying explicitly as white female, 36.4% as white male, 6.8% as black female and one student (2.3%) as Latinx female. Participants ranged in age from 19 to 23 years of age ($M = 20.31$, $SD = 0.96$) and were at different academic ranks (*Freshman* = 2.3%, *Sophomore* = 29.5%, *Junior* = 56.8%, and *Senior* = 11.4%). The educational technology course participants were enrolled in is a requirement for various education-related majors and this is reflected in the diversity of participants' majors (see Table 1). Half of participants reported having used a virtual reality (VR) headset prior to the study (50.0%); most participants reported having viewed 360 photos (77.3%) while 59.1% had viewed 360 videos.

Measures and data

In addition to demographics, data were collected from two primary sources: an adaptation of the *Multimodal Presence Scale* (MPS) for VR environments (Makransky *et al*., 2017) and recorded video of participants' viewing experience with the 360 video. Although analysis of the adapted MPS is the primary focus in this paper, the analysis of the video recorded sessions was useful in providing additional validity evidence towards the adapted scale.

The XRPS is an adaption of the MPS initially created by Makransky *et al*. (2017). The MPS was designed to explore presence in virtual environments; it was informed by Lee's (2004) theoretical framework about presence, which suggests the construct is distinguishable through physical, social and self-presence. Makransky *et al*. (2017) used confirmatory factor analysis and item response theory to develop a validity argument for the MPS. The MPS includes 15 items (5 for each subscale) which ask for a Likert scale-type response (1 = completely disagree to 5 = strongly agree). Given Makransky *et al*. (2017) strong psychometric analysis and theoretically driven approach to their constructs, we elected to adapt the MPS to the context of watching non-responsive VR in the form of 360 video.

Adapting the MPS into the XRPS involved several changes due to the specific focus on 360 videos as opposed to interactive VR settings. A total of 15 cognitive interviews were completed in order to revise its wording and structure (Gandolfi, Ferdig, Kosko, 2020). This cognitive testing mainly helped in eliminating unclear terms (eg, *mixed reality embodiment* or *computer interface*) and removing a redundant statement ("I felt to be emotionally attached to the persons and events during the 360 video"). Both changes made the scale more understandable and concise. The final version of XRPS included 29 items mirroring the initial MPS items but also adding additional statements addressing 360 videos and their main differences from artificial settings. Participants are asked to respond to the XRPS by evaluating the frequency they experienced each scale statement (from 1 or never to 5 or always).

To supplement the data collected from the XRPS, participants' 360 video viewing sessions were recorded. Prior to watching 360 video, participants were instructed to enable the record feature on the Oculus Go. After participation, videos were downloaded and named with participant IDs to link them to survey responses. A total of 33 videos were acquired from 44 participants. The remaining clips were not viable either due to a recording error on the part of the Oculus software (n = 1) or due to user error on the part of participants (n = 10; eg, user took the headset off mid-recording, did not begin the recording session properly, etc.). Videos were analyzed to examine participants' viewpoints and attention foci, with an emphasis of the second video watched (see procedures section).

Procedures

Participants were all preservice teachers recruited via an undergraduate research pool hosted by the authors' university. After completing a brief demographic survey, participants were provided a brief overview of how to use the Oculus Go; a headset designed as an introductory device for VR. Specifically, participants were provided a reference sheet with color images illustrating the use of the controller, how to video record their session and how to access the two 360 videos they were to watch. Participants then put on the headsets, adjusted them to fit properly, placed headphones on and began screen recording and viewing of videos.

The first video lasted 2 min and 45 sec long. It served as a tutorial for how to watch 360 videos, attempting to engage viewers in turning their heads up, down, left and right to look at specific things in the tutorial. In the past, we found that many participants new to viewing 360 content with a VR headset would not look in different directions. Thus, this tutorial video was meant to teach participants that they could look around and not straight ahead (as if viewing a standard video). The second video was 6 min and 58 sec long and was content based. The recording was a video of a third-grade mathematics activity in which students engaged in informally learning about the *Commutative Property of Multiplication*. After viewing both videos, participants removed the headphones and headset and completed the XRPS.

Analysis

Analysis focused on collecting and examining evidence towards a validity argument for the XRPS. The *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) suggests that an appropriate validity argument "integrates various strands of evidence" (p. 21). To validate the XRPS, we examined evidence related to test content, response processes, internal structure and generalization. Validity evidence for *test content* considers how well the survey items used in this study measure the construct of presence. Validity evidence for *response processes* refers to "the fit between the construct and the detailed nature of the performance or response actually engaged in by test takers" (American Educational Research

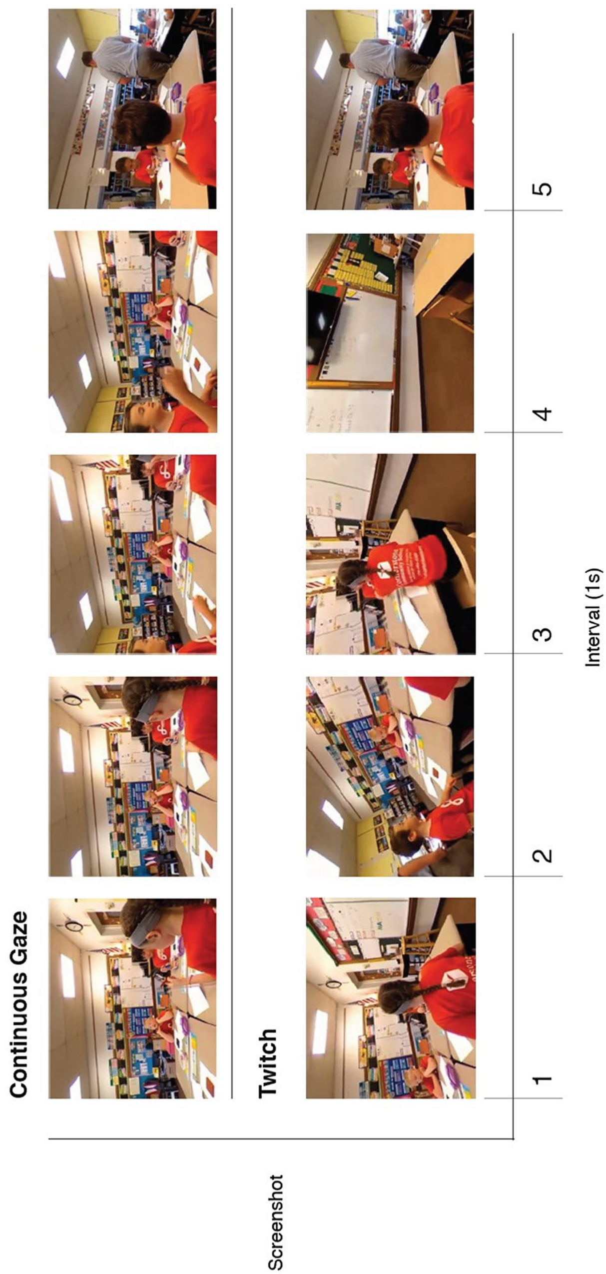

Figure 1: Image intervals from students with continual gaze versus those who jumped around the view [Colour figure can be viewed at wileyonlinelibrary.com]

Association, American Psychological Association, & National Council on Measurement in Education, 2014, p. 15).

Primary sources of evidence for both test *content* and *response processes* included cognitive interviews, analysis of participants' viewing of a 360 video and analysis of the construct key map (CKM) produced from Rasch analysis of the survey. Validity evidence *for internal structure* refers to the degree to which "relationships among test items and test components conform to the construct" (p. 16). Here, we used Rasch principal components analysis (Rasch PCA) alongside fit statistics to examine whether the presence construct assessed was uni- or multidimensional. Validity evidence towards *generalization* focuses on how a measure can generalize to new contexts or situations. In this study, we used various indicators of reliability to examine the survey's internal consistency and estimated reliability with similar samples of respondents.

Rasch modeling was used, alongside classical test theory (CTT), to examine the psychometric properties of the presence scale. The Rasch approach focuses on modeling a person's ability, an item's difficulty and examining the relationship between the two (Bond & Fox, 2015). Raw response data are logarithmically transformed to report two different logit-based statistics: a participant's ability is designated by the theta statistic, θ, while an item's difficulty is designated by the delta statistic, δ (Bond & Fox, 2015; Wang & Wilson, 2005). When applied to polytomous data, such as Likert responses used in the current study, the ordinal data are transformed such that a delta statistic is estimated for each Likert response to an item is given (*Never, Rarely, Sometimes, Often and Always*). Specifically, the data transformation converts the ordinal response data from Likert scales to continuous data. By consequence, a response of *Never* on one item may have a higher or lower score (delta statistic) than that of another item. Furthermore, while Likert scale items are often treated as having the same distance in magnitude from another (1 = *Never*, 2 = *Rarely*, 3 = *Sometimes*, 4 = *Often* and 5 = *Always*), modeling this data with a Rasch approach means that the differences between such responses will vary both within and between items. To assure that such conversions are reliable, Rasch modeling provides for various fit statistics (infit and outfit) for both the items and the individuals completing the survey. In addition to fit statistics, reliability is estimated for both items and persons, and a Wright Map is created to examine the relationship between the sample theta statistics and item delta statistics.

Following Rasch modeling of response data for the presence scale, we analyzed participants' recordings of their 360 video viewing experience. The analytic procedure adopted was based on counting the "twitches" for each video. With a twitch, the reference goes to a discontinuous change (to the left, right, up or down) in the participant's viewpoint. The assumption is that videos with less twitches would mean a more focused perspective and vice versa. Figure 1 illustrates the difference between a continuous gaze and a twitch (relying on 1-second- long time intervals) in the context of the video observed. It should be noted that twitches are often inevitable, also because of what may happen in the video itself (eg, the teacher pointing at someone or somebody out of scope); however, the hypothesis is that individuals with a more continuous viewpoint (with less twitches) in the 360 environment would feel more naturally involved and, therefore, more immersed. As such, their presence should be higher than individuals with videos characterized by many twitches. Such a focus is motivated by the concept of presence itself, defined by Mestre (2005) as a "psychological state experienced as a consequence of focusing one's energy and attention on a coherent set of stimuli" (p. 2). Therefore, presence and focus are supposed to be strongly tied. This lens has been added to better understanding how presence is related to concrete behaviors within immersive virtual environments.

Table 2: Item analysis statistics for the final model

				Infit		Outfit	
Items	*Point-biserial*	*Item difficulty*	*SE*	*Mean square*	*Z*	*Mean square*	*Z*
i5	0.30	−1.55	0.24	1.14	0.7	1.18	0.7
i7	0.73	−0.78	0.2	0.46	−3	0.47	−2.8
i8	0.54	0.24	0.18	1.08	0.5	1.06	0.3
i9	0.71	−0.28	0.19	1.02	0.2	0.98	0
i11	0.55	1.76	0.19	1.22	1.1	1.38	1.6
i12	0.76	−0.06	0.19	0.6	−2.1	0.76	0.63
i13	0.65	−1.04	0.21	1.35	1.4	1.15	0.7
i14	0.67	1.16	0.18	1.28	1.4	1.2	1
i15	0.59	−1.00	0.21	1.48	1.9	1.33	1.3
i16	0.63	0.95	0.17	1.32	1.5	1.32	1.5
i17	0.62	−1.76	0.26	0.57	−1.9	0.58	−1.5
i18	0.17	−1.77	0.25	1.33	1.3	1.31	1
i19	0.70	0.52	0.18	0.79	−1	0.84	−0.8
i21	0.73	0.23	0.18	0.47	−3.2	0.57	−2.4
i22	0.43	−1.14	0.22	1.3	1.3	1.28	1.1
i23	0.75	0.23	0.18	0.61	−2.2	0.62	−2
i24	0.81	0.86	0.17	0.44	−3.5	0.44	−3.5
i26	0.55	−0.91	0.21	1.25	1.1	1.12	0.5
i27	0.52	1.44	0.18	1.33	1.6	1.35	1.6
i28	0.74	1.16	0.18	0.77	−1.2	0.8	−1
i29	0.64	1.74	0.18	1.18	0.9	1.24	1.1

Results

Psychometric results

A Cronbach's alpha coefficient of .878 was calculated, suggesting sufficient internal reliability for the XRPS. However, Rasch provides more specific indicators of reliability by examining item and person reliability. The initial modeling of the XRPS was found to have sufficient item reliability (0.97), with a separation index of 5.38. This suggests that the XRPS can distinguish between items with lower and higher ratings. The XRPS was also found to have sufficient person reliability (0.87) with a separation index of 2.57. This suggests that the assessment is also able to distinguish between groups of people (ie, higher vs. lower sense of reported presence). Person fit statistics indicate an average mean square infit ($MNSQ = 0.98$, $Z = -0.20$), outfit ($MNSQ = 1.00$, $Z = -0.10$) close to the Rasch modeled expectations of 1 and standardized fit Z values near zero. This suggests that the variance in modeled fit scores is generally in the acceptable range.

Initial analysis of item fit statistics suggested four items demonstrated a large amount of variance: i3 (*infit MNSQ* = 1.45, *outfit MNSQ* = 1.56), i6 (*infit MNSQ* = 1.62, *outfit MNSQ* = 1.60), i10 (*infit MNSQ* = 1.56, *outfit MNSQ* = 2.08) and i25 (*infit MNSQ* = 1.45, *outfit MNSQ* = 1.44. Some of these items had demonstrated interpretability issues in cognitive interviewing sessions. For example, i25 currently reads: "I felt like my real body was affected by what was happening in the 360 video." However, it was originally worded as: "I felt like my viewpoint was an extension of my real body within the 360 video." Thus, in examining the wording of these items, it appeared that certain phrasings may have allowed for different interpretations than intended. Therefore, these items were removed and the model was rerun.

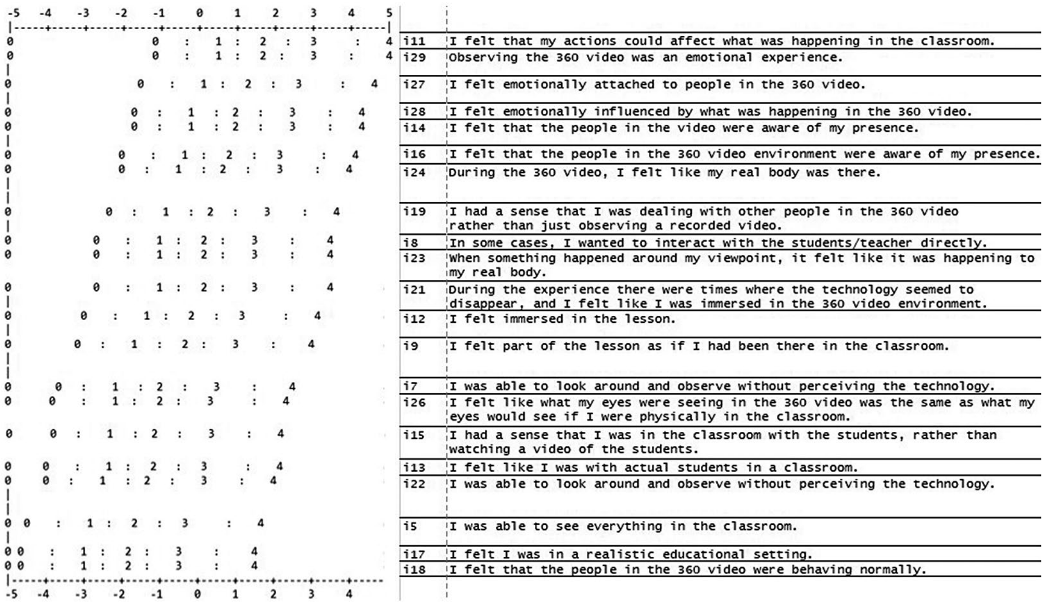

Figure 2: Construct key map (CKM) with survey items juxtaposed

Next, we examined the unidimensionality of the model through a Rasch PCA in which the factor analysis examines the standardized residuals of the Rasch modeled data. The Rasch PCA indicated that the measure explained 41.6% of the variance (36.5 of 62.5 units). The first contrast explained 6.9% of the variance with an eigenvalue of 4.29 and the second contrast explained 5.8% of the variance with an eigenvalue of 3.61. When eigenvalues are greater than 3 and the contrast explains more than 5% of the variance, further analysis of potential multidimensionality is warranted. Examination of the disattenuated correlations between the main construct and the potential construct indicated from the second contrast indicates it should be retained as part of the primary construct ($r = 0.87$). However, the disattenuated correlation between the main construct and that indicated from the first contrast was −1.00, suggesting a new and distinct construct. The items in the new construct, listed below, were all negatively worded but were reverse coded for the Rasch analysis. In essence, a higher rating of these items (without reverse-coding) correlated positively with a higher sense of presence. This suggested that the negatively worded items may, in fact, represent issues in the items' design rather than a new theoretically driven construct.

- *I thought about what was happening around me, outside of the 360 experience* (i1).
- *I was aware of my surroundings, outside of the 360 experienc*e (i2).
- *I was unable to see what every student was doing in the classroom* (i4).
- *I did not feel like I was in the classroom* (i10).

Thus, a total of eight items were removed from the scale (Table 2 contains the items and statistics of the final model; Appendix A in supporting information contains the final instrument).

Both person reliability (0.91) and item reliability (0.96) were found to be sufficient for the final Rasch model of the XRPS. Additionally, item infit ($MNSQ = 1.00$, $Z = -0.20$) and outfit ($MNSQ = 0.99$, $Z = -0.20$) and person infit ($MNSQ = 1.01$, $Z = -0.10$) and outfit ($MNSQ = 0.99$, $Z = -0.10$) were sufficient. At the item level, the majority of items were within the typically accepted range of .75–1.33. A subset of items did have infit statistics below 0.75, suggesting their fit is "too good

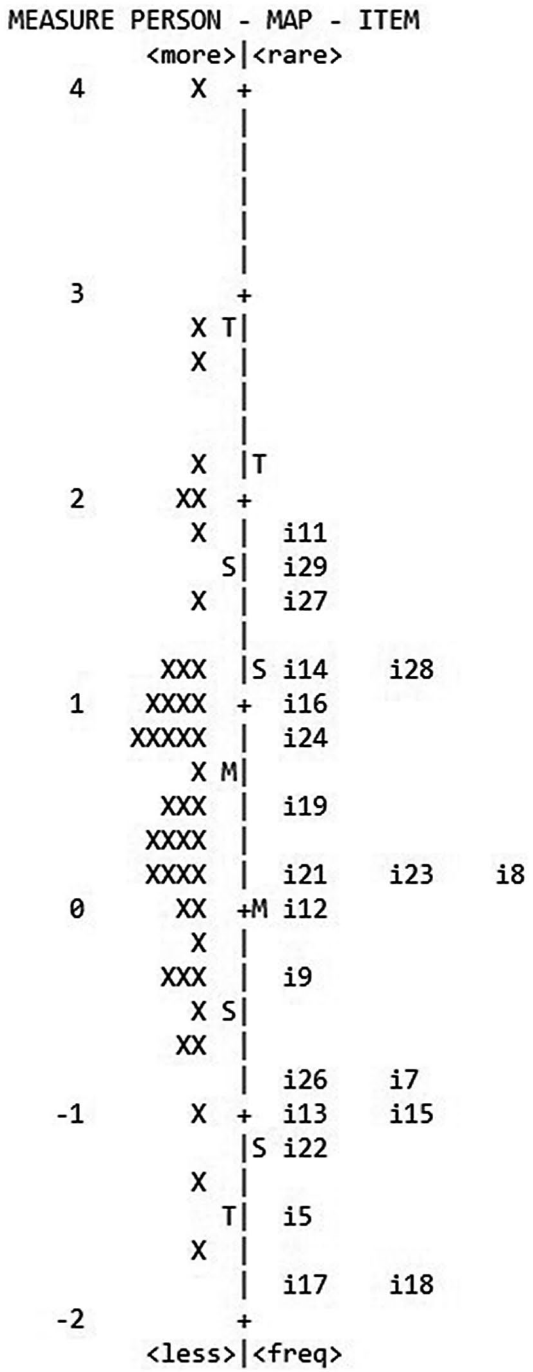

Figure 3: Wright map from data analysis

to be true" (p. 53, Bond & Fox, 2015). However, Bond and Fox (2015) urge caution in attempting to remove items in an effort to clean up one's model. "Often the flaws in items that [have under or overfit] are too small to distort the measurement in any noticeable way" (p. 43). In contrast to the

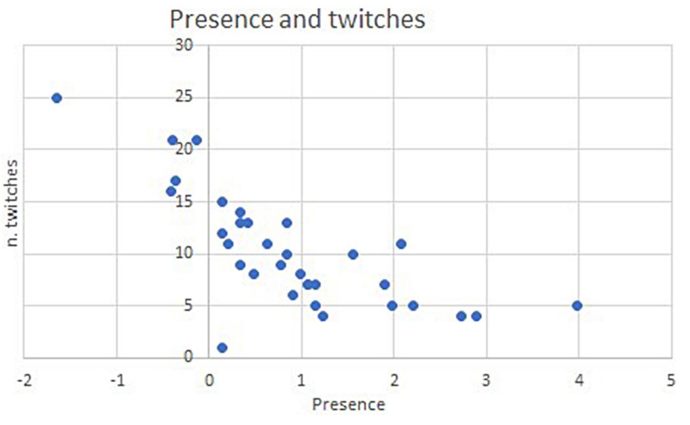

Figure 4: Scatterplot of Rasch scores for the XRPS and number of twitches [Colour figure can be viewed at wileyonlinelibrary.com]

items that were removed earlier, examination of item wording and cognitive interview data did not indicate any cause for removal. Thus, all items presented in Table 2 were retained.

Our next step was to examine the CKM, shown in Figure 2 and the Wright Map, shown in Figure 3. These representations illustrate the relative weight of different ratings for specific items (0 = Never, 1 = Rarely, 2 = Sometimes, 3 = Often and 4 = Always), as well as the relative distribution of the sample in relation to these items. In examining this interplay, we conjecture that four stages of presence may be deduced from the data. Scores approximately −1.50 logits and below probabilistically rate i5, i17 and i18 as *sometimes*, and all other times with lower frequencies. This suggests individuals with such scores demonstrate a sense of *withdrawn presence* when viewing a 360 video. Scores ranging approximately between −1.50 and 0.00 logits were considered as a similar level of presence as watching a standard video. For example, an individual with a theta score of −0.50 may respond to i12 "I felt immersed in the lesson" with *rarely*, but rate i17 "I felt I was in a realistic educational setting" as *often*. Scores approximately between 0.00 and 2.00 logits were considered as demonstrating presence of viewing something more than a video, but not fully immersed. For example, a participant with a theta score of 1.00 would likely rate i12 as *often* but would likely rate i28 "I felt emotionally influenced by what was happening in the 360 video" as *rarely*. Individuals with theta scores of 2.00 logits or higher appear to demonstrate a sense of *mesmeric presence* or a sense similar to "being there."

The Wright Map presented in Figure 3 corresponds with information presented in the CKM (Figure 2), but with two components lending to further interpretation. First, items are presented at the question level rather than for each Likert response since this latter information is already provided in the CKM. As such, the Wright Map provides a summary view of how questions, in general, compare regarding the sense of presence measured. Second, the left-hand side of the figure illustrates the distribution of participants' theta statistics. The juxtaposition of this data allows for interpretation of whether the distribution of items and persons correspond. Examination of Figure 3 suggests that there is a large degree of correspondence except at the higher end of the scale. Thus, if we seek to better assess mesmeric presence, future studies should consider writing items to target this range of the scale.

Video analysis results

For the video analysis, 33 full videos were successfully recorded by the participants. The correlation between XRPS scores and number of twitches was calculated with a Pearson correlation coefficient (r). The range of the number of twitches varied from 4 to 25 ($M = 11.45$; $SD = 5.65$). The correlation between participants' XRPS Rasch scores and number of twitches was found to be statistically significant with a strong negative coefficient ($r = -0.72$, $p < .001$). A post hoc power analysis of the correlation coefficient, using an alpha level of .05, yielded a statistical power of 0.99. The relationship is illustrated in the scatterplot provided in Figure 4. The illustrated relationship, along with the correlation coefficient, suggests that a higher sense of presence is associated with a more continuous gaze, while a participant's more frequent shifts in field of view is associated with a lower sense of presence.

Discussion

Results presented in this paper provide several pieces of evidence supporting the validity of the XRPS. First, evidence from participant responses in cognitive interviews (Gandolfi et al., 2020) combined with examination of the CKM (see Figure 2) and analysis of twitches in video viewing suggest that the XRPS items measure the construct of presence (test content) and that participants' actions align with their XRPS scores (response processes). Second, psychometric analysis using a Rasch modeling approach suggests the XRPS demonstrates evidence towards unidimensionality (internal structure) as well as person and item reliability (generalization). Finally, the distribution of item and person scores illustrated in Figure 3 provides evidence towards a continuous scale ranging from a sense of withdrawn presence (lower) to mesmeric presence (higher) when viewing 360 videos.

In reviewing the wording of items and their placement on the CKM (Figure 2), it appears that both a sense of agency and emotional attachment operate as features of increased presence. In contrast, items at the lower range of the scale tend to focus on the intermediation between the user and 360 video. Stated differently, a rating of *often* on certain statements on the lower range of the CKM (eg, "I was able to see everything in the classroom") held a similar indication of presence as a rating of rarely for statements on the higher range of the CKM (eg, "I felt that my actions could affect what was happening in the classroom"). Thus, certain items, and the characteristics of presence they embody, are better indicators of different ways participants may perceive the construct.

Referring to the literature, the transition from withdrawn to mesmeric presence is what differentiates VR and immersive VR—the feeling of being there and forgetting the mediating technology (in this case, the headset and the 360 video itself) (Ferdig *et al.*, 2018; Freina & Ott, 2015). In contrast, with a lower sense of presence, these constraints seem to be perceived. Participants saw the 360 experience as just a video outside of their presence. This outcome expands the scope of the study beyond the mere validation of XRPS, suggesting empirically based understandings of presence that can be considered and tested in additional studies and with different types of immersive technologies and environments.

Results presented in this paper suggest participants' perceived emotional relatedness and sense of agency were indicators of higher levels of presence rather than lower. Such a finding is supported by the literature about immersion and virtual environments. Emotions can play a relevant role in immersing users and improve media experiences because they supposedly imply a full involvement of the user (Allcoat & von Mühlenen, 2018; Baños *et al.*, 2004; Marín-Morales *et al.*, 2018). Agency is an additional component that refers to the perceived ability to act meaningfully in a given environment and/or simulation (Nardi, 2015). It has been often associated with engagement in VR settings (Calleja, 2011; Guadagno, Blascovich, Bailenson, & McCall, 2007) and especially with immersive technologies (Freude, Reßing, Müller, Niehaves, & Knop, 2020;

Kong, He, & Wei, 2017). It is worth noting that agency is relevant even in experiencing 360 videos, which are often described as less interactive than artificial settings (ie, responsive vs. non-responsive VR).

The video analysis provided important findings regarding how a sense of presence interacts with physical behaviors. In this study, there was a negative correlation between the numbers of twitches and higher presence scores (as determined by XRPS scores). One potential explanation for this finding is that continuous gaze movements imply a higher degree of focus, or attention, which corresponds with a sense of being there. Recent eye-tracking research provides empirical support for this conjecture. When viewing video of classroom practice, novice teachers' eye movements are more haphazard while attempting to focus on multiple features. In contrast, experienced teachers tend to focus on select students (van den Bogert *et al*., 2014; Dessus *et al*., 2016). Results presented in this paper may extend research findings from eye-tracking studies addressing focus, attentiveness and presence.

It is possible that this twitching was an attempt to cover all 360 degrees at the same time (or at least as much as possible). So, moreover, engaged presence becomes a matter of *media transparency* (ie, forgetting the interface, therefore, forgetting that it is just a video and focusing on what is happening, etc.). In contrast, a lower sense of presence may entail a distance and need for control over the technology itself. In other words, presence could be related to the concept of *flow* (Nakamura & Csikszentmihalyi, 2009), while withdrawn presence implies a disconnection between the user and lesson because of the medium used (Aydin, Woge, & Verbeek, 2019). Looking at the videos and related XRPS scores, it seems that a higher sense of presence corresponds with more focus and attention, while a lower sense of presence suggests a *supervision anxiety* that weakens the ability to focus on what matters. This outcome sets the stage for additional inquiries focusing on how presence deals with actions in immersive environments, which is an overlooked aspect in the current literature.

Limitations

The study presents four main limitations. First, the concept of presence is multi-faceted and can be influenced by additional variables not considered in the present analysis. Those variables could include participants' attitudes towards 360 or their previous experience with the technology. It might also be related to the design of videos watched. Second, our focus was on 360 videos planned for training future educators. Therefore, findings need to be contextualized within this area, from participants' experience (preservice teachers enrolled in an US university) to content observed (a third-grade mathematics activity). Third, immersive technologies keep changing and evolving. The concepts of immersion and presence themselves are dynamic and subject to change along with their perception by users. Therefore, these findings need to be tested and stressed recursively with new devices and media experiences related to 360 videos for preservice teachers and beyond. Fourth, participants' performance (eg, in terms of noticing) during the observation was not considered in the present study, which sought to validate the XRPS. Future research could build upon these limitations, from focusing on new technologies and training contexts to considering variables important to teacher training like reflection and noticing.

Conclusion

The purpose of this study was to validate an instrument aimed at exploring presence in 360 videos. The XRPS was adapted from the MPS for VR environments (Makransky *et al*., 2017). This study provided evidence of the XRPS as a valid and reliable instrument for measuring presence when participants engage with 360 videos. As such, it provides a baseline for future studies focusing on user-experience in this particular context (eg, teacher training) but also targeting additional areas (eg, professional development). Moreover, the XRPS can act as a validated instrument for anyone interested in the use and study of 360 videos. In addition, the association between higher presence scores with a perceived sense of agency and emotional attachment suggests facilitating these feelings while developing 360 videos.

Results demonstrated a negative correlation between twitches and presence, suggesting that those who felt more presence were less likely to move suddenly, or twitch, in their video watching. Conversely, those with lower presence were more likely to attempt to consume everything happening in the 360 space. Although additional investigations are needed, this study has provided statistical evidence of the relationship and may inform further and wider reflections on the role of presence in immersive environments for teacher education and beyond.

Statement on open data, ethics, conflict of interest

The authors declare no conflict of interest related to the present study.

The study has been approved and monitored by the authors' university IRB committee.

The de-identified dataset along with some sample 360 videos will be made available on the authors' project website after publication.

References

Allcoat, D., & von Mühlenen, A. (2018). Learning in virtual reality: Effects on performance, emotion, and engagement. *Research in Learning Technology*, *26*, 1–13.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Aydin, C., Woge, M. G., & Verbeek, P. P. (2019). Technological environmentality: Conceptualizing technology as a mediating milieu. *Philosophy & Technology*, *32*(2), 321–338.

Baños, R. M., Botella, C., Alcañiz, M., Liaño, V., Guerrero, B., & Rey, B. (2004). Immersion and emotion: Their impact on the sense of presence. *Cyberpsychology & Behavior*, *7*(6), 734–741.

Bianchi-Berthouze, N., Kim, W. W., & Patel, D. (2007). Does body movement engage you more in digital game play? and why?. In A. C. R. Paiva, R. Prada, & R. W. Picard (Eds.), *International conference on affective computing and intelligent interaction*, (pp. 102–113). Berlin: Springer.

Blomberg, G., Stürmer, K., & Seidel, T. (2011). How preservice teachers observe teaching on video: Effects of viewers' teaching subjects and the subject of the video. *Teaching and Teacher Education*, *27*(7), 1131–1140.

Bond, T., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*. New York, NY: Routledge.

Bower, M., Lee, M. J. W., & Dalgarno, B. (2017). Collaborative learning across physical and virtual worlds: Factors supporting and constraining learners in a blended reality environment. *British Journal of Educational Technology*, *48*(2), 407–430.

Calleja, G. (2011). *In-game: From immersion to incorporation*. Cambridge. MA: MIT Press.

Cortina, K. S., Miller, K. F., McKenzie, R., & Epstein, A. (2015). Where low and high inference data converge: Validation of CLASS assessment of mathematics instruction using mobile eye tracking with expert and novice teachers. *International Journal of Science and Mathematics Education*, *13*(2), 389–403.

Dessus, P., Cosnefroy, O., & Luengo, V. (2016). Keep your eyes on 'em all!: A mobile eye tracking analysis of teachers' sensitivity to students. In: K. Verbert, M. Sharples, & T. Klobučar (Eds.), *European conference on technology enhanced Learning* (pp. 72–84). New York, NY: Springer.

Eröz-Tuğa, B. (2013). Reflective feedback sessions using video recordings. *ELT Journal*, *67*(2), 175–183.

Fadde, P., & Sullivan, P. (2013). Using interactive video to develop preservice teachers' classroom awareness. *Contemporary Issues in Technology and Teacher Education*, *13*(2), 156–174.

Farrow, R., & Iacovides, I. (2012). ‘In the game’? Embodied subjectivity in gaming environments. In: R. García Martín, & J. I. Vieco Madrid (Eds.), *6th International Conference on the philosophy of computer games: The nature of player experience*, (pp. 29–31). Madrid, Spain.

Ferdig, R. E., Gandolfi, E., & Immel, Z. (2018). Educational opportunities for immersive virtual reality. In: J. Voogt, G. Knezek, R. Christensen, & K. W. Lai (Eds.), *Second handbook of information technology in primary and secondary education* (pp. 955–966). New York, NY: Springer.

Ferdig, R. E., & Kosko, K. W. (2020). Implementing 360 video to increase immersion, perceptual capacity, and teacher noticing. *TechTrends*, *64*, 849–859. Retrieved from http://link.springer.com/article/10.1007/s11528-020-00522-3.

Freina, L., & Ott, M. (2015). A literature review on immersive virtual reality in education: State of the art and perspectives. In: I. Rocenau (Ed.), *Proceedings of eLearning and software for education (eLSE)*, (p. 133). Bucharest: National Defence University.

Freude, H., Reßing, C., Müller, M., Niehaves, B., & Knop, M. (2020). Agency and body ownership in immersive virtual reality environments: A laboratory study. In: T. Bui (Ed.), *Proceedings of the 53rd Hawaii International Conference on System Sciences*, (pp. 1530–1539). Manoa, HI.HICSS.

Gandolfi, E., Ferdig, R. E., & Kosko, K. W. (2020). The extended reality presence scale. In: D. Schmidt-Crawford (Ed.), *Proceedings of the society for information technology & teacher education*, (pp. 1011–1017). New York, NY: Association for the Advancement of Computing in Education (AACE).

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, *16*, 41–67.

Guadagno, R. E., Blascovich, J., Bailenson, J. N., & McCall, C. (2007). Virtual humans and persuasion: The effects of agency and behavioral realism. *Media Psychology*, *10*(1), 1–22.

Han, I., Eom, M., & Shin, W. S. (2013). Multimedia case-based learning to enhance preservice teachers’ knowledge integration for teaching with technologies. *Teaching and Teacher Education*, *34*, 122–129.

Kong, G., He, K., & Wei, K. (2017). Sensorimotor experience in virtual reality enhances sense of agency associated with an avatar. *Consciousness and Cognition*, *52*, 115–124.

Kosko, K. W., Ferdig, R. E., & Zolfaghari, M. (in press). Preservice teachers’ professional noticing when viewing standard and 360 video. *Journal of Teacher Education*. https://doi.org/10.1177/0022487120939544.

Lau, K. W., & Lee, P. Y. (2015). The use of virtual reality for creating unusual environmental stimulation to motivate students to explore creative ideas. *Interactive Learning Environments*, *23*, 3–18.

Lee, E. A., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education*, *79*, 49–58.

Lee, K. M. (2004). Presence, explicated. *Communication Theory*, *14*(1), 27–50.

Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, *7*(4), 328.

Lorenzo, G., Pomares, J., & Lledó, A. (2013). Inclusion of immersive virtual learning environments and visual control systems to support the learning of students with Asperger syndrome. *Computers & Education*, *62*, 88–101.

Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the multimodal presence scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior*, *72*, 276–285.

Marín-Morales, J., Higuera-Trujillo, J. L., Greco, A., Guixeres, J., Llinares, C., Scilingo, E. P., … Valenza, G. (2018). Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors. *Scientific Reports*, *8*(1), 1–15.

Mestre, D. R. (2005). *Immersion and presence*. Retrieved from: http://www.ism.univmed.fr/mestre/projects/virtual%20reality/Pres_2005.pdf.

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. In S. J. Lopez & C. R. Snyder (Eds.), *Oxford library of psychology. Oxford handbook of positive psychology*, (pp. 195–206). Oxford: Oxford University Press.

Nardi, A. B. (2015). Virtuality. *The Annual Review of Anthropology*, *44*, 15–31.

Ottenbreit-Leftwich, A. T., Glazewski, K. D., Brush, T. A., Aslan, S., & Zachmeier, A. (2018). Addressing technology integration concerns: Asynchronous video mentoring between preservice teachers and exemplary technology-using in-service teachers. *Australasian Journal of Educational Technology*, *34*(4), 1–15.

Roche, L., & Gal-Petitfaux, N. (2017, March). Using 360 video in physical education teacher education. In: P. Resta & S. Smith (Eds.), *Society for information technology & teacher education international conference*, (pp. 3420–3425). New York, NY: Association for the Advancement of Computing in Education (AACE).
Rupp, M. A., Odette, K. L., Kozachuk, J., Michaelis, J. R., Smither, J. A., & McConnell, D. S. (2019). Investigating learning outcomes and subjective experiences in 360-degree videos. *Computers & Education*, *128*, 256–268.
Seidel, T., Blomberg, G., & Renkl, A. (2013). Instructional strategies for using video in teacher education. *Teaching and Teacher Education*, *34*, 56–65.
Theelen, H., van den Beemt, A., & den Brok, P. (2019). Using 360-degree videos in teacher education to improve preservice teachers' professional interpersonal vision. *Journal of Computer Assisted Learning*, *35*(5), 582–594.
van den Bogert, N., van Bruggen, J., Kostons, D., & Jochems, W. (2014). First steps into understanding teachers' visual perception of classroom events. *Teaching and Teacher Education*, *37*, 208–216.
Walshe, N., & Driver, P. (2019). Developing reflective trainee teacher practice with 360-degree video. *Teaching and Teacher Education*, *78*, 97–105.
Wang, W. C., & Wilson, M. (2005). The Rasch testlet model. *Applied Psychological Measurement*, *29*(2), 126–149.
Weber, K. E., Gold, B., Prilop, C. N., & Kleinknecht, M. (2018). Promoting preservice teachers' professional vision of classroom management during practical school training: Effects of a structured online-and video-based self-reflection and feedback intervention. *Teaching and Teacher Education*, *76*, 39–49.
Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, *24*(6), 1319–1333.
Wiens, P. D., Hessberg, K., LoCasale-Crouch, J., & DeCoster, J. (2013). Using a standardized video-based assessment in a university teacher education program to examine preservice teachers knowledge related to effective teaching. *Teaching and Teacher Education*, *33*, 24–33.
Zolfaghari, M., Austin, C. K., Kosko, K. W., & Ferdig, R. E. (2020). Creating asynchronous virtual field experiences with 360 Video. *Journal of Technology and Teacher Education*, *28*(2), 315–320.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.