Analysis of YouTube Comments to Inform the Design of Virtual Reality Training Simulations to Target Emotional Arousal

Karan R. Patil[1]; Siddharth Bhandari, A.M.ASCE[2]; Ameeta Agrawal[3]; Steven K. Ayer, A.M.ASCE[4]; Logan A. Perry, A.M.ASCE[5]; and Matthew R. Hallowell, A.M.ASCE[6]

Abstract: Workplace safety remains a concern in the construction industry as fatality rates continue to rise. While hazard recognition training programs have been implemented using multimedia-based modules, their effects have not been broadly reflected on construction sites. In an effort to provide realistic and engaging training, a recent focus on virtual reality (VR) for an immersive learning experience has been shown to offer benefits to improve traditional lecture-based training. Such virtual environments can be especially useful for simulating hazard recognition tasks that are inaccessible in real-life settings due to the potential dangers they pose for trainees. However, due to the focus on applications for performance assessment and procedural training, strategic elicitation of emotional arousal, which has been shown to be a precursor to desired learning outcomes in hazard recognition training, has not been explored for construction-specific VR applications. To guide the development of such VR environments that target emotional arousal for learning, this study used opinion mining to catalogue the features that yield or inhibit an emotional reaction in similar video simulations posted on a public video sharing platform (YouTube). Design insights such as the need to provide agency in the simulations, introducing nonplayer characters in the scene, and the like, are presented. Here the authors discuss specific implementation strategies derived from the study findings that developers can use to elicit emotional arousal in a construction-specific virtual environment. **DOI: 10.1061/JCEMD4.COENG-13245.** *© 2023 American Society of Civil Engineers.*

Practical Applications: Virtual reality (VR) training programs are increasingly being used in the construction industry to improve hazard recognition and reduce fatalities. VR provides an immersive learning experience that can simulate tasks that would be dangerous for trainees to perform in real life. However, previous VR training programs did not focus on leveraging emotional arousal, which has been shown to be beneficial for learning outcomes. This study used natural language processing to analyze user comments and identify key features that lead to an emotional reaction in virtual simulations on YouTube, and provided insights and strategies for developers to create construction-specific VR environments that elicit emotional arousal.

Introduction

Construction fatality rates remain the highest among US industries, accounting for almost one in five work-related deaths (BLS 2018). Studies have found workers on job sites to be unprepared to identify (Perlman et al. 2014; Albert et al. 2014b) and manage (Jeelani et al. 2017) risks appropriately. While safety practitioners and academics have focused on improving hazard recognition skills of workers through training modules, their expected effects have not been broadly reflected on the construction site (Mostafa et al. 2016). Indeed, most training modules validated in the literature either provide only inferential evidence on improving hazard recognition skills or lack the requisite external and ecological validity (Albert et al. 2014b; Jeelani et al. 2020; Bhandari et al. 2019). Additionally, a number of these training modules neither transfer knowledge nor facilitate long-term retention of information due to poor delivery methods (Demirkesen and Arditi 2015) because they do not incorporate the settings and context in which adults engage with learning (Mostafa et al. 2016). Finally, traditional safety training programs that do not generate sustained interest can leave workers with a negative attitude towards safety (Haslam et al. 2005).

To combat the aforementioned limitations, recent efforts have sought to explore the potential of virtual reality (VR) to provide workers with an immersive and effective learning experience (Jeelani et al. 2020; Eiris et al. 2018; Pham et al. 2018; Zhang et al. 2020). VR has shown the potential to offer benefits to improve upon traditional lecture-based training (Aggarwal et al. 2006; Farra et al. 2013; Mastli and Zhang 2017), providing a more realistic and engaging training experience by inducing a sense of presence, which produces the sense of being in a coherent virtual space with the capability to perceive and react to sensory stimuli (Slater 2003). This is especially useful in situations where placing workers in real-life settings would be highly dangerous or logistically impossible (Kaplan et al. 2020). VR has also been shown to increase emotional

[1]Ph.D. Student, School of Sustainable Engineering and the Built Environment, Arizona State Univ., Tempe, AZ 85281 (corresponding author). ORCID: https://orcid.org/0000-0002-8553-8621. Email: krpatil@asu.edu

[2]Associate Director of Construction Safety Research Alliance, Dept. of Civil, Environmental, and Architectural Engineering, Univ. of Colorado, Boulder, CO 80309. Email: siddharth.bhandari@colorado.edu

[3]Assistant Professor, Dept. of Computer Science, Portland State Univ., Portland, OR 97201. Email: ameeta@pdx.edu

[4]Associate Professor, School of Sustainable Engineering and the Built Environment, Arizona State Univ., Tempe, AZ 85281.

[5]Assistant Professor of Engineering Education, Dept. of Civil and Environmental Engineering, Univ. of Nebraska–Lincoln, Lincoln, NE 68508. ORCID: https://orcid.org/0000-0003-1558-2579

[6]Beavers Professor of Construction Engineering and Executive Director of Construction Safety Research Alliance, Dept. of Civil, Environmental, and Architectural Engineering, Univ. of Colorado, Boulder, CO 80309.

Note. This manuscript was submitted on October 13, 2022; approved on March 20, 2023; published online on June 23, 2023. Discussion period open until November 23, 2023; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Construction Engineering and Management***, © ASCE, ISSN 0733-9364.

arousal (Balban et al. 2021), which can in turn facilitate learning through neuroplasticity (Duman 2004; Green and Bavelier 2008). VR is highly effective in generating and sustaining targeted emotional arousal in a number of contexts, such as gaming, art, entertainment, and training (Bates 1992; Cheok et al. 2009; Hawkins 1995; Lateef 2010; Riva et al. 2007; Steadman et al. 2006). This is relevant because emotional engagement in a learning environment is a critical antecedent to improved learning outcomes (Deppermann et al. 2014; Sagayadevan and Jeyaraj 2012; Simpson and Marshall 2010). From the construction safety training context, maintaining contextually appropriate emotional arousal is beneficial as it can not only facilitate self-driven or intrinsic motivation to learn (Bradley et al. 1992; Cahill and McGaugh 1998; Goetz et al. 2007) but also condition risk-taking behavior (Bhandari and Molenaar 2020; Loewenstein et al. 2001).

Despite the potential of VR to leverage emotional arousal, the literature on VR's potential for safety training has not explored a systematic strategy to elicit a targeted emotional reaction in the simulation. In other words, these training modules do not address why adults learn or how knowledge is retained. While there has been user-experience research in VR generally, no studies have explored the generation of integral emotions in the context of hazardous and potentially fatal conditions that are the focus of construction safety research. Therefore, continued innovation to accurately replicate construction work environments and their associated risks in a realistic and emotionally engaging manner in VR can unlock new, and potentially better, modes of safety education that translate into reduced safety-related construction incidents.

To guide the development of emotion-eliciting VR-based hazard recognition training experiences, this study uses an opinion-mining approach to catalog the features that yield or inhibit an emotional reaction in other, nonconstruction virtual experiences that simulate similar high energy hazards and their consequences. YouTube was chosen as the source for these videos as it allows the public to watch and comment on various videos, including those targeted in this work. The comments section offers a wealth of information about the elements in those videos that elicit various reactions from viewers (Yasmina et al. 2016). The findings of this study can be used by safety practitioners and researchers to design, develop, and test construction-specific VR simulations that target emotional arousal to facilitate learning outcomes.

Background

Safety Training in the Construction Industry

Existing construction training programs have been criticized for their lecture-based, impersonal, and unengaging delivery methods (Mostafa et al. 2016; Demirkesen and Arditi 2015). This approach incorrectly assumes that learning is a passive activity and that all learners in an environment are alike in strengths, weaknesses, and personal expectations (Jeelani et al. 2020). Unsurprisingly, most adult learners need a clearly established purpose in learning. Such standardized training sessions lead to irrelevant information, poor instruction, lack of motivation, and emotional frustration (Hannum 2009). Over the past decade, several studies have confirmed the limitations of traditional training programs, which have resulted in construction workers being unable to identify, on average, half the hazards on a job site (Albert et al. 2014a).

VR has emerged as a leading platform for a new generation of safety training programs. It has the potential to deliver a safe learning option for workers in an ecologically valid environment, as it enables immersion in a realistic work setting, thereby allowing workers to interact with hazards and potential consequences without being exposed to any actual risk. This philosophy is rooted in the knowledge that human beings are more cautious and risk-averse after experiencing some direct or indirect loss (Shum and Xin 2020). Furthermore, training in environments that closely replicate the actual work environment has been shown to result in better results in practice (Cockrell 1979; Kowalski-Trakofler and Barrett 2003). This indicates the potential for VR to circumvent the limitations posed by traditional safety training seminars by making hazards, likelihood of accidents, and potential severity of negative outcomes less abstract and more experiential. Additionally, VR-based approaches have been shown to support training by simulating an ecologically valid scenario in other domains (Lopez et al. 2016; Paljic 2017). This has inspired safety academics and practitioners to implement VR-based technologies to leverage their potential in improving the safety skills of the workforce in the construction industry (Perlman et al. 2014; Hadikusumo and Rowlinson 2002; Park and Kim 2013; Zhao and Lucas 2015). However, the efficacy of VR for safety training purposes in occupational settings can be further enhanced by incorporating emotional engagement strategies that support improved learning outcomes and long-term learning (Cahill and McGaugh 1998; Pekrun et al. 2002). Thus, the literature highlights the potential for VR to provide meaningful advancements in current safety training programs.

Emotional Arousal to Facilitate Training in VR

The importance of emotion elicitation in learning contexts has been deeply explored in the literature (Chung et al. 2015; Immordino-Yang 2015). What has been consistently demonstrated is that learners who experience emotional engagement during a learning experience not only achieve learning objectives more rapidly but also recall knowledge more readily (Chung et al. 2015; Mayer and Estrella 2014; Plass et al. 2014; Um et al. 2012). For example, a study by (Kim and Kines 2018) found that the use of emotionally arousing video scenarios in safety training is effective in increasing the safety knowledge and intentions of construction workers. Similarly, Bhandari et al. (2019) found that construction workers who receive safety training that includes an emotional component (e.g., personal stories from workers who have experienced accidents) have better safety knowledge and attitudes compared with those who receive training that is purely cognitive (i.e., knowledge-based) (Salame and Göransson 1989). Furthermore, emotional arousal has been shown to generate motivation and promote detail-oriented problem solving Pekrun et al. (2002) and Pittenger and Duman (2008). There are many schools of thought on why these relationships are observed, from flashbulb memory theory (Finkenauer et al. 1998) to theories that consider neurobiological effects of emotional arousal (McGaugh 2000; Kuo et al. 2007) on learning and memory. In occupational environments, Wang and Liao (2021) found emotional arousal to significantly mediate and affect attention. This highlights the necessity of developing training environments that yield targeted emotional arousal for improved learning outcomes (Mauri et al. 2015; Lang et al. 1995).

In a construction engineering and management (CEM) context, Bhandari and Hallowell (2017) found that hyper-realistic re-creation of common workplace injuries can generate targeted emotional arousal among construction workers. The elicitation of such emotions not only improves learning outcomes (Bhandari et al. 2019) but also conditions risk-taking behavior (Tixier et al. 2019). However, due to the logistical challenges associated with developing hyper-realistic re-creations, it is difficult to adopt such replications across different sectors of the industry in a cost-effective manner. Alternatively, VR presents an opportunity to replicate realistic visceral emotional experiences in a cost-effective manner without actually causing harm but generating life-long motivation for learning. These key findings provide the impetus for a deeper introspection on the capability of VR-based learning modules that effectively generate and sustain emotional engagement among learners in learning environments.

VR's ability to elicit emotional arousal in order to support specific outcomes has been well studied in other domains like medical and military training (Jones et al. 2002; Shilling et al. 2002; Ulate 2002; Altgassen et al. 2010). One such study reported that during a military training simulation, participants in the "high arousal" condition were significantly better at encoding and recalling objects presented in the virtual environment (Shilling et al. 2002). The entertainment industry has also long recognized the use of emotion to effectively immerse a user in a virtual scenario (Ravaja et al. 2006; Villani et al. 2018) and VR's role in eliciting higher levels of emotional engagement in those scenarios (Kim et al. 2018). For example, in a comparison between film viewing in VR and traditional two-dimensional (2D) conditions, Ding et al. (2018) reported a greater emotional effect when subjects watched the film in VR. While VR has been shown to elicit higher arousal when compared with videos or pictures in construction contexts as well, the development of analogous VR experiences for construction-specific applications has not been explored as a way to leverage the potential for learning that follows an emotionally arousing experience. Therefore, by using emotionally engaging experiences from other areas, researchers in the construction domain can explore design strategies to develop an emotional VR experience for hazard recognition training.

Finally, existing frameworks such as Krathwohl's affective taxonomy (Morshead 1965) can be useful in evaluating the impact of emotional experiences in virtual reality (VR). The affective domain consists of five levels of emotional development: receiving, responding, valuing, organizing, and characterizing, which are ordered by the principle of internalization or the process whereby a person's affect toward an object passes from general awareness to a point where the affect is "internalized" and consistently guides or controls the person's behavior (Seels and Glasgow 1990).

Gonzalez and Hernandez (2014) examined the use of Krawthol's taxonomy in the development of a virtual reality safety training program for construction workers and found the program to be effective in increasing participants' knowledge of safety protocols and in promoting safe behaviors on the job. While traditional and contemporary VR may have evolved certain aspects of this hierarchy, such as the receiving stage, by presenting information to be learned, emotional and interactive construction VR may provide an opportunity to scaffold construction practitioners to higher forms of learning by engaging and experiencing emotional stimuli in a realistic VR setting. Therefore, future researchers may use such accepted taxonomies to evaluate the impact of construction-specific emotional VR for safety training.

Knowledge Gap

In recent years, VR-based training models have shown how they can be used to counter the limitations of traditional training programs. However, there is an opportunity to further increase their impact by enhancing emotional experiences in the simulation through scientifically validated virtual features (e.g., learning objectives, animation, visual and auditory realism, immersion, presence).

While the literature has shown the potential for VR and the effects of emotions on learning, there is little evidence on which features in a construction-specific virtual environment can be systematically designed to trigger emotional arousal among individuals. Because this knowledge is not well understood, VR's potential has been limited to assessment, design reviews, and procedural training in the construction industry (Azimi et al. 2018; Jenkins et al. 2020; Su et al. 2013; Teizer et al. 2013). By gathering evidence on how other domains have elicited emotional responses in virtual experiences, researchers can initiate a strategic exploration into how similar features can be embedded in hazard recognition training simulations. Identifying features that can achieve emotional arousal in VR will enable future researchers to design construction-specific VR environments in accordance with these features and test their ability to elicit desired outcomes in the targeted user group of construction professionals. The following research questions are addressed:

RQ 1: What are the features in the selected videos that lead to an emotional response?

RQ 2: What are the features in the selected videos that lead to disengagement from the experience?

Methodology

Publicly available data obtained from social media platforms are often used when obtaining objective data is either infeasible or impossible (Amarasekara and Grant 2019; Miller 2015). Such data provide an opportunity to access reactions of the general population through their virtual interactions and responses to social engagements. User reactions are comparable to user feedback, which is important for making design decisions to improve human-centered interfaces where the end user's experience guides iterative design development (Johnson 1998; Kuniavsky 2003). Specifically in construction, researchers have studied the efficacy of social media platforms to provide valuable insight and the potential for industry to leverage such insight for making decisions (Kim et al. 2014; Tang et al. 2017).

YouTube has emerged as a popular social media platform for research purposes as it allows researchers to extract the large number of user comments that may offer insights into reactions to contextually appropriate video(s) (Yasmina et al. 2016). While there were initially concerns regarding the reliability of comments due to spam and redundant content, researchers have used filtering algorithms to successfully utilize data gathered from YouTube for content analysis to identify certain exploratory and confirmatory patterns of human interactions, cognition, and behavior (Kim et al. 2010; Yoo and Kim 2012). Content analysis is a method that enables researchers to make inferences by objectively and systematically identifying specified characteristics of messages (Holsti 1969). For example, Paek et al. (2010) performed content analysis on comments left on 934 antismoking videos posted on YouTube to identify how individuals reacted to various psychological and informational features in those videos.

We used this approach to first filter and then analyze comments posted on YouTube videos simulating hazardous situations to empirically identify the features that yield emotional responses from viewers. To achieve our research objectives, machine learning algorithms were used to analyze characteristics of the comments and extract semantic patterns, which have been established in other domains (Agrawal and An 2012; Sarakit et al. 2015; Savigny and Purwarianti 2017). Additionally, a qualitative analysis was performed to interpret and seek an understanding of the initial outputs from the machine learning algorithms (Chang et al. 2009; Creswell and Poth 2007). The qualitative discourse presented allowed us to critically discuss the different features identified and why they lead to engagement or disengagement among viewers. Additionally, the qualitative findings provide a rationale for the inferential evidence reported from objective analysis. These findings may provide design guidance for future VR programs that deliver construction-specific learning experiences while generating the requisite emotional engagement.

We recognize that the videos selected were not construction-specific. This was because this study serves as an initiation for efforts that will support future development and more rigorous testing of initial findings. Due to the low number of construction-specific simulations of hazardous scenarios, constraining the selection criteria would have limited the range of responses and consequently made it difficult to draw meaningful conclusions. Therefore, we elected to leverage more popular videos but chose ones that closely replicated scenarios of interest—namely, hazardous situations—that could guide the implications for developing construction-specific emotional VR simulations. Since this paper does not make claims about the direct impact of design decisions on eventual construction-specific training, the specific differences in the context of the simulations are not relevant to the validity of the findings presented. In addition to enabling exploratory research, the selection of nonconstruction simulations enabled us to adopt a positivist approach to identifying aspects that trigger emotional reactions by avoiding any biases or expectations they may have or expectations that may influence the findings. Instead, the nonconstruction videos allowed for novel themes that naturally emerged from the data. This data-driven approach provides a robust foundation for the development of construction-specific VR.

We also want to note that this paper does not make claims about the intensity of arousal based on the different modes of consumption available to viewers posting comments, such as television, mobile phones, or VR headsets. Instead, the paper inquires about the aspects of the content consumed that elicited an arousing response among its consumers. While a difference in intensity may be hypothesized based on how the content was consumed, that is not within the scope of this paper's research question.

Natural Language Processing (NLP)

While a large quantity of data offers an opportunity to conduct an in-depth analysis, it also poses some challenges. Similar to those on other social media platforms, YouTube comments are short and informal, with numerous accidental and deliberate errors and grammatical inconsistencies. A manual analysis of such data using human coders would not only be highly inefficient but also be unreliable, given the difficulty in achieving inter-rater consistency due to irregularities in the understanding and use of language (Gergen 2015). Therefore, natural language processing (NLP), a subfield of linguistics and artificial intelligence dealing with analyzing and understanding language in a form that is natural to humans, was used in the exploratory tests performed in this study. (Table 1 provides the relevant terminology.)

Table 1. Relevant terminology

Term	Definition
Document	Unit of text in a data set for NLP algorithms
Comments-data set	Data obtained from YouTube consisting of comments posted publicly in response to selected videos
Topic-words	Output from NLP algorithm of semantically similar words in a piece of text indicating a common topic
Topic-labels	Inference on the topic being discussed based on topic-words produced by the NLP algorithm
VA-space	Two-dimensional space with all comments from the comments-data set plotted based on their valence (V) and arousal (A) rating obtained in data analysis
VA-quadrants	Four quadrants of two-dimensional VA-space: "positive valence-low arousal," "positive valence-high arousal," "negative valence-low arousal," and "negative valence-high arousal"

Although a novel method, there is precedent in the literature for the use of NLP to extract insights from social media platforms such as Youtube and Twitter based on textual content analysis (Liu 2012; Severyn et al. 2014; Kolhatkar et al. 2020; Müller et al. 2023; Das et al. 2019).

Social media platforms provide access to a large and diverse set of data, including a wide range of comments from individuals with different backgrounds, cultures, and viewpoints. Such data are typically unstructured and written in natural language, which well suits them for NLP techniques that can process and analyze text data. NLP algorithms can be used to identify patterns, trends, and sentiments in social media comments, which can be helpful for understanding how people feel about a particular topic or issue.

Furthermore, similar NLP methodologies have been used in the construction domain (Tang et al. 2017; Damirchilo et al. 2021; Lee et al. 2019). While such data consist of a high amount of noise or unrelated content, with programmatic methods to filter this out, NLP provides an opportunity to streamline the analysis of big data (Alberto et al. 2015), thereby significantly enhancing the internal validity of the reported findings.

NLP Toolbox

A significant portion of the research protocol used NLP algorithms to clean, structure, and analyze the data set of YouTube comments. The next section outlines the algorithms used and how they were trained. The set of algorithms and the approach shown has been used extensively (e.g., Zhang and El-Gohary 2016; Liu 2012; Serrano et al. 2020) on similar data sets (Kim et al. 2014; Yoo and Kim 2012). For brevity, these algorithms are referred to here as they were labeled in the data analysis, which is in parentheses.

Filtering Out Spam and Off-Topic Comments (Filter_Clean)

Typical sets of comments from YouTube contain some unwanted ones that could be classified as either spam or off-topic. Spam comments provide malicious/bare URLs or illegible and automatically generated text. Similarly, off-topic comments are ones that have very little informative content—e.g., those containing single words. Neither category of comments helped answer the research questions posed in this study, and it was important to distinguish them from on-topic, meaningful comments. Hence, we used a filter (filter_clean) that uses machine learning algorithms trained on a training data set to automatically classify text into one or more classes such as spam or useful (Asiri 2018). Here, the SenTube corpus (Uryupina et al. 2014) was chosen as our training data. It contains user-generated comments from YouTube and has been annotated for spam filtering. We extracted a subset "SenTube_Spam" from the original data set, yielding a training data set of 8,866 comments with inter-rater agreement scores of 0.94 and 0.56 for "spam" and "off-topic" labels, respectively. These ratings were considered acceptable given the complexities and disparities in the range of comments found under YouTube videos (Uryupina et al. 2014).

For the machine learning classifier, we adopted the logistic regression (LR) model to classify text as spam or not spam. Using LR was appropriate here since the outcome variable would be binary. However, before using LR to classify our own (unseen) data, it was necessary to validate its performance. To do so, we benchmarked its accuracy on a known data set, where we trained the LR classifier on 80% of our SenTube_Spam data set and tested it on the remaining 20% to achieve an acceptable accuracy of 76%. The LR classifier developed using SenTube_Spam was finally used on our comments data set to filter out spam comments and retain the comments labeled as not spam for the next steps, which are described in subsequent sections. Additionally, all "stop words," or words that are most common in a particular language and add very little to the semantic value (e.g., *it*, *for*, *then*, *the*), were first removed from the -data set using the Natural Language Toolkit (NLTK) library (Babanejad et al. 2020; Bird et al. 2009). This was done to avoid skewing the outcome by incorporating words that do not provide context about the emotional content of a document.

Valence-Arousal Scores (get_VA)

Valence (positive versus negative) and arousal level (activating versus passive) along with dominance (degree of control) are key dimensions of emotions (Osgood et al. 1957; Russell 1980, 2003). The valence dimension shows overall emotional polarity; that is, is the person experiencing more pleasant than unpleasant feelings towards a stimulus? The arousal dimension indicates the degree to which an individual is experiencing the discrete emotions, such as happiness, sadness, anger, joy, and surprise (Russell 1980).

From the National Research Council of Canada (NRC) comes NRC-VAD, an emotion knowledge base of valence, arousal, and dominance (VAD) numerical vectors for about 20,000 English words, with scores for V, A, and D ranging from 0 to 1 (Mohammad 2018). For instance, the VAD vector of "success" is represented as (0.96, 0.88, 0.98), indicating valence, arousal, and dominance scores, respectively. Similarly, "crying" is represented as (0.15, 0.70, 0.21), indicating high arousal but a negative valence. However, for the purposes of this study, the two dimensions of valence and arousal themselves served as strong indicators of how individuals perceive, adapt, respond, accept, and apply new knowledge (Bischoff-Grethe et al. 2008; Parthasarathy and Busso 2017; Schechtman et al. 2010). Thus, the dominance scores were eliminated and only the valence and arousal scores were considered in the get_VA algorithm.

Topic-Words Algorithm (Topic_Modeling)

Topic modeling in NLP refers to the task of identifying topics that best describe a set of textual data. The output of this algorithm is a set of topic-words, or words that are semantically similar in a given context, indicating a topic. In this study, the topic-modeling task was leveraged to obtain an indication of the topic being discussed for comments reflecting a common emotional response to a particular video. However, simply extracting the most frequent words in a document is not very useful. A better approach is to identify more meaningful words that represent the underlying semantic theme of the document. Thus, the topic_modelling algorithm was used here to obtain topic-words for the comments-data set.

We employed the well-established and popular technique of topic modeling called latent Dirichlet allocation (LDA) (Blei et al. 2003), which is a generative probabilistic model in which every document is a combination of topics and the combination of words signifies the topics.

We used the LDA with the online variational Bayes algorithm, from the Scikit-learn library (Pedregosa et al. 2011), with the number of topics and the number of words in a topic set to 10 and all other parameters set to default.

Research Protocol

Fig. 1 illustrates the research protocol, which is elaborated in the following sections.

Step 1: Select Videos and Download Comments

We were interested in the features of a video that supported or deterred engagement and emotional arousal in the viewer (Table 2 is a list of selected videos). Videos from YouTube were selected based on certain features that best aligned with the research goal of identifying psychological and technological features that affect viewers' emotional engagement in virtual simulations. The following characteristics were sought in the videos selected by the authors:

- Single event simulation: This filter eliminated videos that were simulations of multiple situations (e.g., driving to the store and walking inside the store to purchase items) instead of single ones (e.g., driving to the store) to avoid studying comments with varying focus points. This feature avoided capturing confounding effects in the data set.

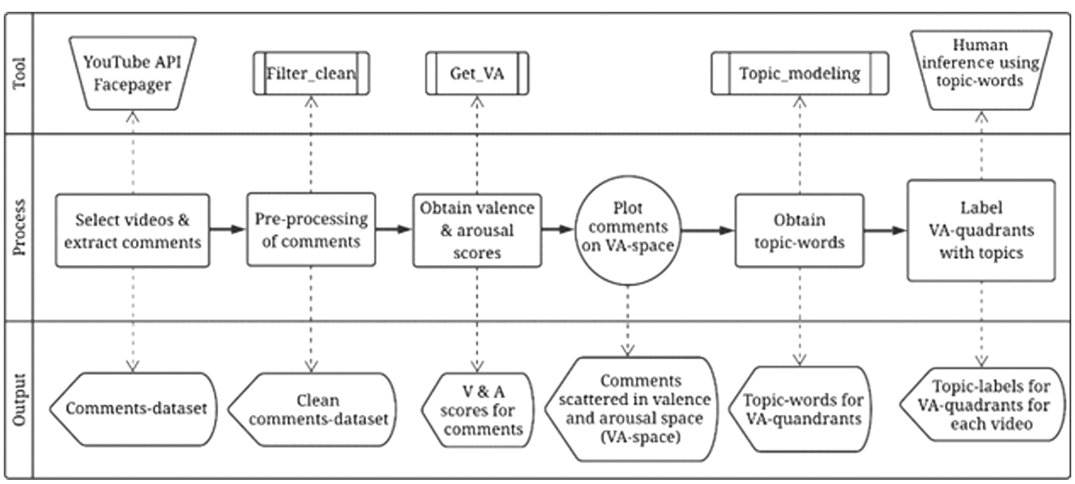

Fig. 1. Research protocol.

Table 2. Videos selected from YouTube

URL	Notation	Relevance to RQ
https://www.youtube.com/watch?v=f_pFe3Jglvw	car_flood	Simulates the emotional experience following the mistake of driving into a flooded area
https://www.youtube.com/watch?v=xjnnuubKS_s	plane_ride	Simulates the emotionally intense experience of a turbulent plane ride
https://www.youtube.com/watch?v=kOnWwE9Qq0s	plane_crash	Simulates the emotionally intense experience of surviving a turbulent plane crash from a first-person perspective
https://www.youtube.com/watch?v=pkO0jUF-e8M	weather_tornado	Simulates the experience of being caught in a tornado
https://www.youtube.com/watch?v=gOc1dAcfCfk	car_texting	Simulates an accident following a driver who continues texting while driving
https://www.youtube.com/watch?v=Vpj-JfDU-4M	escape_blocked	Simulates the panic of being caught in a high-rise with jammed elevators and doors

- First person point of view: This characteristic required the videos to be in the first-person perspective. This was important to support the development of simulations for VR-based training programs which would inevitably situate individual workers (e.g., first-person perspective) in a virtual environment with a head-mounted display (e.g., HTC Vive, Oculus Quest, Valve Index) for an immersive experience with high ecological validity.
- Experience of consequences: Videos that simulate hazardous situations (e.g., tornadoes and car accidents) with unfortunate and negative consequences (e.g., death, injury, or property damage) were selected to capture the emotional reactions and/or disengagements of users within the experience. This aligned with safety-based VR training programs that seek to generate emotional arousal to condition behavior and improve learning outcomes.
- Virtual, not real: Videos selected were all virtual simulations using virtual objects as opposed to reality-captured videos of reenactments using real people/environments. This was done because emotional reactions to real actors as opposed to virtual characters might be different and VR experiences that afford interactions are predominantly developed using virtual artifacts and characters.

After selecting the videos listed in Table 2, their posted comments were gathered using the YouTube application programming interface (API). The API was accessed using Facepager software, which enables accessing publicly available data from YouTube, Twitter, and other social media platforms (Jünger and Keyling 2020).

Step 2: Preprocess the Comments-Data Set

Before analyzing the data set, the filer_clean algorithm was used to remove any inconsistencies in formatting to enable a systematic analysis and also to remove comments that did not add value to the discussion based on the video, generally referred to as spam. As noted, this step enabled us to eliminate spam and "stop- words" from comments, which were not informative.

Step 3: Obtain Valence and Arousal Scores for the Data Set

Valence and arousal scores were obtained from the get_VA algorithm previously described. For a comment sequence consisting of n words, $d = \{\mathrm{w}1, \mathrm{w}2, \ldots, \mathrm{w}n\}$, the VAD vectors for each of its words was referred to using the NRC-VAD; thus, an overall valence, arousal, and dominance score for the comment was obtained by computing the average of the score of all the individual words. The average was calculated by dividing the sum of VAD scores for words in a comment by the number of words in that comment. This approach has been validated (e.g., Taboada et al. 2011). Once generated, the valence and arousal scores were mapped onto the two dimensions of valence and arousal rating to explore the distribution of the valency and arousal of the comments. This two-dimensional space defined by valence and arousal scores is referred to as "VA-space" in this paper.

Next, the VA-space was divided into four separate categories based on the valence and arousal scores (i.e., positive valence-low arousal, positive valence-high arousal, negative valence-low arousal, and negative valence-high arousal) similar to the four quadrants of Posner's circumplex model (Posner et al. 2005). Organizing the comments based on the combination of valency and arousal enabled the authors to explore the shared underlying antecedents that may have led to observed emotional reactions.

Before moving on to the next step, considering the negative valence of the videos and after a manual review, comments in the positive half of the valence dimension (rating >0.5) were removed from further analysis. It was decided to ignore these comments, as they generally praised the video but did not provide insight into its content. Also, the fact that these comments had a positive valence indicated that they were not specifically reactions to the content, which were naturally negative in valence.

Step 4: Obtain Topic-Words for Each Quadrant of VA-Space

After eliminating comments with a positive valence score, the LDA topic_modeling algorithm was applied to the remaining subset of comments organized according to the comments' shared emotional scores in the two quadrants of the VA-space (low valence-low arousal and low valence-high arousal).

The output produced was a cluster of words that occurred together in each of the two quadrants according to a certain topical pattern. These topic-words had an internal consistency—they often appear together in the documents and/or do not typically appear outside that cluster of semantic space. In this manner, topic-words were obtained for the two categories of reaction (i.e., the two quadrants of the VA-space) for all videos individually. The following is an example of the topic-words obtained for comments in a quadrant defined by low valence and low arousal scores for a car crash video: *phone*, *driving*, *looking*, *text*, *crash*, *turn*, *man*, *going*, *guy*, *scared*. As can be seen, the topic-words do not semantically or syntactically form a meaningful sentence but are indications of the common topic of the comments being analyzed. It should be noted here that the context of the video and the guiding question are important in interpreting a potential topic from the output of the NLP algorithm (Chang et al. 2009). Thus, after obtaining topic-words for each quadrant in each video's circumplex, we sought qualitative human inference to label the topics for each quadrant, relying on the indication of topic-words obtained in this step.

Step 5: Label VA-Quadrants with Topic (Topic-Labels) Based on Topic-Words

After the topic-words were obtained for each cluster of comments, human inference was required to identify the underlying common topic suggested by the topic-words (Chang et al. 2009). Understanding the context based on which topic-words have been generated is crucial to extracting relevant themes or "topic-labels." Such "themes" based on contextual understanding mimic a qualitative thematic coding process, which is inherently subjective, to identify themes that are not evident in the objective data (Maxwell 2016; Creswell and Poth 2007).

Similarly, contextual knowledge around the video scenario is important in deriving topic-labels. Thus, the video was carefully re-evaluated in conjunction with the newly derived topic-words to interpret the underlying topic, which clusters them in semantic space. Further, specific goals have a significant impact on how this step of topic labeling is carried out. Therefore, a deductive method was adopted to structure the coding based on tags that were relevant to the research question, but the codes for each tag were coded inductively based on the comments posted for each video (Maxwell 2016; Guest et al. 2011). For example, the authors first read the topic-words derived for a particular VA-quadrant posted for Video X followed by watching the entire Video X.

Next, we read the topic-words and referred to the comments that they related to for a holistic understanding of the comments in addition to the topic-words. Then, to deductively label the topic-words, we coded a group of topic-words based on three queries after reaching consensus:

- Are the comments content or video related?
- What is the key characteristic discussed in the comments?
- What is the effect expressed in the comments as it relates to the key characteristic?

The three tags together constituted the topic-label for a quadrant of comments posted for a particular video which were coded inductively from exploration of the raw data (Creswell and Poth 2007). Fig. 2 shows the process of labeling a quadrant based on the topic-words.

Responses to the first query, "Are the comments content or video related," provided information about whether the cluster of comments were discussing features of the video itself or the contents of the video. A topic labeled "video" contained comments generally discussing the quality or effectiveness of the video, whereas a topic labeled "content" contained comments discussing its specific contents. Responses to the next query, "What is the key characteristic discussed in the comments," revealed information about the fundamental feature being discussed in the cluster of comments. A full list of key features coded for all comments is provided in Table 3. Lastly, the effect tag from the third query abstracted the general reaction expressed in the cluster of comments. These ranged from intense emotion to disappointment to disengagement.

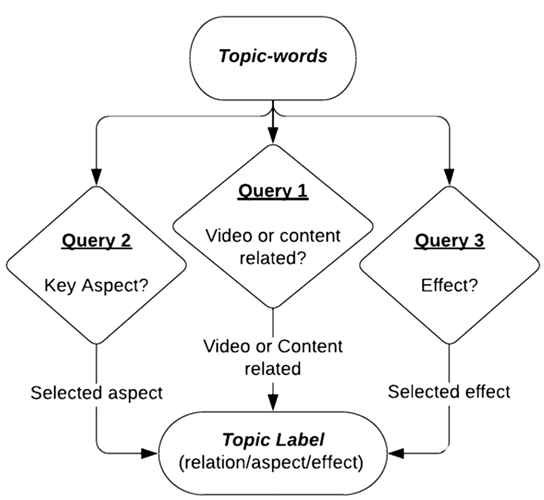

Fig. 2. Systematic approach to labeling topics based on topic-words.

Table 3. Codes for key features discussed in a set of topic-words

Key feature	Description
Audio cues	Auditory cues of abstract or remote occurrences
Bad graphics	Graphical quality of video
Missing virtual character	Presence of virtual character in experience mistake
Mistake	Mistake being made is simulated in the video
Sound effect	Sound and audio in the video
Unattempted fix	Insufficient options portrayed as solution to a problem unexplained constraint
Unexplained constraints	Certain rules of the virtual world not explained
Unreal characters	Realism of the video's characters
Unreal consequence	Realism of an action's consequences simulated
Virtual character consequence	Virtual character experiencing the consequence of an event
Virtual character uninvolved	Realism of virtual character's interaction with its environment virtual interaction
Virtual interaction	First player's ability to interact with objects in the virtual world
Visual cues	Visual cues of abstract or remote occurrences

Results

Descriptive Statistics

From the selected videos, 8,043 comments were downloaded using Facepager (Jünger and Keyling 2020). The preprocessing of the comments using filter_clean reduced the entire data set from 8,043 to 2,413. While this seems as though a majority of the data set was eliminated at this stage, it should be noted that spam comments are commonplace on social media platforms (Alberto et al. 2015; Sureka 2011) and even when they are not malicious, many can be irrelevant to the video, rendering them uninformative in terms of the research questions.

After processing and filtering the data set, the authors first determined the valence and arousal scores of the comments for each video. Next, the topic-words were obtained for each quadrant of the VA-space and finally, the topic-labels were obtained by referring to the topic-words of a quadrant and corresponding comments from the comments-data set (Table 1). The following sections detail these results in chronological order.

Valence and Arousal Scores

As noted previously, the valence and arousal scores obtained from the get_VA algorithm were plotted along the Valence and Arousal dimensions in values of 0 to 1. Each data point on the plot in Fig. 3 depicts a comment from the selected videos (Table 2 lists the video notations).

J. Constr. Eng. Manage., 2023, 149(9): 04023077

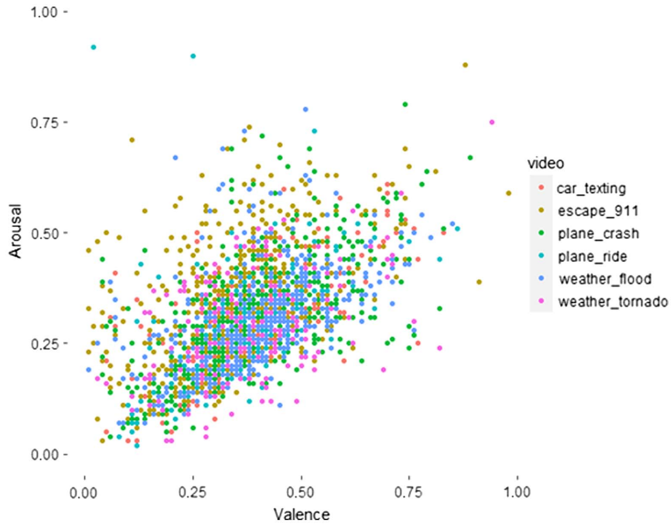

Fig. 3. Distribution of comments on Valence and Arousal dimensions (VA-space).

It is clear from the scatter plot of comments along the Valence and Arousal dimensions that a majority of the comments fell in the negative half of the valence dimension (Rating <0.5). This result was consistent with the type of video selected, portraying an undesirable situation and the consequences of events that occur in that situation. For example, during an experience in which the user gets into a car accident or loses a friend due to driving into a flooded area, negative valence responses are expected.

Topic-Words for Each Quadrant

After the comments were plotted along the Valence and Arousal dimensions (VA-space), the topic-words for each video's quadrants were obtained by applying the topic_modeling tool to cluster words in the semantic space. As described previously, comments with a positive valence score were eliminated, leaving two VA-space quadrants: "low valence-low arousal" and "low valence-high arousal.". Table 4 lists the unique topic-words obtained for all videos selected for analysis. These topic-words indicate the topic of discussions for the cluster of comments that elicited a reaction defined by the quadrants they lie in.

Underlying Topics (Topic-Labels)

The topic-words (Table 4) obtained represent the words that appear together in the comments and predicted to generally refer to the same topic for this reason. However, such pure text data are naturally complex and may pick up on idiosyncratic correlations of the words in a set of comments. Thus, conceptually spurious words can appear together not because they are related to other words in a topic semantically but because they are demographically correlated. For example, the following topic-words were obtained for a car crash video: *speeding*, *fault*, *horseman*, *guy*, *steer*, *hit*, *deploy*, and *airbag*. it is clear, after getting a contextual understanding of the video, that this topic generally refers to a car accident. However, the topic-word "horseman" seems out of place and unrelated to the underlying topic. One of the reasons for the anomaly could be a cultural reference that places the word demographically close to the topic without any real relevance to it. This limitation of topic-modeling algorithms requires some human inference to analyze the topic-words with prior knowledge of the context of the

Table 4. Results of topic-words for all selected videos

Video	Topic-words
car_texting	speeding, steer, red, fault, hit, horseman, deploy, guy, airbag, disguise
weather_flood	door, claustrophobic, hard, happen, escape, beast, crying, idiot, deep
escape_911	acting, jump, engine, voice, hit, scared, disrespectful, suicide, loud, smoke, volume
weather_tornado	realism, house, rug, love, rebuild, position, gives, destroying, anxiety
plane_ride	thing, screaming, happy, scared, hell, girl, body, die
plane_crash	scare, survive, simulation, physics, landing, explode, wreck, strip, crash, mention, accident, ironic

videos and the goals of classifying topics. Thus, topic-labels were inferred from the topic-words for each VA quadrant (van Kessel 2019; Chang et al. 2009).

To systematize the inference process for the purposes of this study, topic-labels were obtained by answering the three queries outlined in Step 5 of the research protocol (Fig. 1). The resulting labels for each comment constituted three codes corresponding to the three queries. As an example, the topic-label is provided here for comments plotted in the low valence-low arousal quadrant of the car_texting video with the following topic-words: *phone*, *driving*, *looking*, *text*, *crash*, *turn*, *man*, *going*, *guy*, *scared*. These were labeled *content*, *mistake*, and *emotional* for the three queries that constituted the topic-label. The *content* tag was selected because the comments in this cluster were generally in relation to a specific event in the video where the virtual character gets into an accident by virtue of texting while driving. The key feature of the video that these comments are reacting to is the *mistake* of texting while driving, which eventually leads to an accident.

And finally, the effect query was tagged as *emotional* because most comments in this cluster are expressing anger or some form of emotion in reaction to the mistake that the character makes in the simulation. More examples of topic-labels for a group of topic-words from other videos are presented in Table 5.

Discussion

The purpose of this study was to inform the design of virtual simulations specific to the construction industry by leveraging reactions posted by individuals in the YouTube comments section.

While development of VR environments has been explored in other domains, the findings often inform VR development in general but not for specific scenarios like the hazardous situations that were of interest in this study (Hudson et al. 2019). For example, it has been found that social interactions, virtual avatars, and eye contact elicit arousal; however, these findings are reported to improve storytelling in VR (Monteiro et al. 2018; Ho and Ng 2022), not to develop hazardous scenarios. While prior findings provide value, they do not provide similar aspects to be included in a hazardous training simulation, which has limited the development of emotionally arousing VR simulations for construction training. This is a key distinction because there is a fundamental difference in what humans perceive and pay attention to in a hazardous environment compared with a nonthreatening environment. Aspects

Table 5. Examples of topic-labels

Topic-words	Topic-label		
	Content/video	Key feature	Effect
guy, water, door, dude, watching, best, omg, crying, really	Content	Virtual character consequence	Aroused emotion
scared, guy, happen, sup, bad, crying, saw, died, freaked, deep	Content	Virtual character consequence	Aroused emotion
window, break, open, windows, door, seat, roof, water, use, situation	Content	Unattempted fix	Disappointed

that may not be salient in a nonthreatening scenario may suddenly become salient in a hazardous situation and vice versa. This has been thoroughly explored in research on relevance realization, which investigates how the context of our environment dynamically affects our perception and attention systems (Vervaeke et al. 2012; Vervaeke and Ferraro 2013). It is not obvious that aspects reported in the literature are salient and consequently worth developing for a hazardous environment. It would be unwise and inefficient to allocate resources to these aspects before investigating what is appropriate and salient to users in a hazardous scenario. The findings of this study, on the other hand, focus on aspects of the simulation during hazardous scenarios to extract higher-order patterns that can be implemented in construction-specific VR development to elicit realistic reactions.

Therefore, comments posted about hazardous simulations that have high external validity were used because they are not typically restricted to a target population. Thus, finding relevant and scientifically validated patterns can be highly applicable to the general population (Savigny and Purwarianti 2017; Uryupina et al. 2014).

Considering the differences among selected videos and the topic-labels obtained for them, certain patterns were identified that provided exploratory evidence of a particular video's ability to emotionally engage viewers and also features that led to disengagement. These patterns, if confirmed by future researchers, could inform the design of construction-specific virtual simulations, especially where engagement, attention, and emotional arousal in the user are desired. This discussion is structured around these major patterns and their implications for the design of construction-specific virtual simulations.

Pattern 1: Role of Nonplayer Characters in Generating Engagement

Virtual characters in a simulation that are not embodied or controlled by the user are called nonplayer characters (NPCs) (Warpefelt 2016). Our findings revealed that there was a difference across the videos in how viewers responded to NPCs placed in the virtual environments. Videos that required NPCs to interact with viewers generated higher levels of emotional engagement than videos in which NPCs had no interaction or communication with viewers. For example, when examining the topic-labels generated from the car_flood video (Table 5), which simulated NPCs interacting with viewers through a casual conversation, there were sympathetic and empathetic responses from viewers when the NPC encountered a negative event. Topic-labels such as "virtual character consequence" were observed for this video experience that indicated words such as *friend*, *guy*, *help*, *crying*, and *scared*. Comparatively, the plane_crash video, which did not have NPCs interacting with viewers, generated more humorous, sarcastic, or even aggressive comments. Topic-labels such as "disengagement" and "unrealism" reflected jokes being made about the suffering virtual character, which highlighted the difference between the emotional engagement and experience of the viewers of the two videos showing similar simulations.

In the literature, the function of NPCs in VR applications to target emotional arousal of construction workers has not been validated or discussed in detail. Considering the difference virtual avatars can make in emotional engagement, which is a primary antecedent to learning, the authors propose the following (a summary is provided in Table 6):

- Integration of virtual construction workers and their involvement in a safety incident to trigger an emotional reaction.
- Establishing a connection to achieve social immersion with virtual construction workers in order to maintain the user's engagement. This could be done through a scripted personal conversation with the virtual character as in the escape_blocked video. Failure to establish a connection with the virtual character could lead to a missed opportunity for emotional arousal in a construction simulation when, for example, a safety incident negatively affects the NPC.
- Integration of artificial intelligence to introduce believable agents who behave in a believable manner in the audience's perception (Mateas 1999).

Pattern 2: Relevant Realism to Avoid Disengagement

Another observed pattern related to the effect of realism on the viewer is relevant realism to avoid disengagement This pattern relates to the production of the simulation itself rather than the content of the video, as indicated by the topic-labels. The quality of the physics engine of the simulation was indicated to be important. A bad physics simulator was seen to disengage viewers from the scene. Similarly, related to production of the content, graphic or visual realism emerged as an important consideration in maintaining engagement. Disengagement from videos with poor graphic quality was observed because it became the topic of the viewers' attention as reflected in their comments. However, some studies have proposed that physical realism is not important for presence in virtual environments (Loomis 1992; Rosenberg 1994). This disagreement may be due to variations in the content of the simulations. The value of graphical realism needs to be further tested for its relevance in engagement as it relates to triggering emotional arousal. This is an important design decision given that physical realism is a potentially significant cost in terms of computing capability, which is a limited resource. Therefore, VR designers

Table 6. Summary of design implications for construction-specific VR from Pattern 1

Design insight	Example
Integrate virtual characters	Presence of virtual construction workers conducting various tasks on the virtual construction site
Establish connection with virtual characters	Trigger casual scripted conversation by virtual character when in the vicinity of the user
Introduce BA using artificial intelligence	Enable BA capabilities for some virtual construction workers

may benefit from looking at specific use cases of the experience rather than adopting umbrella design strategies.

In addition to the quality of the simulation, realism in terms of the user's ability to manipulate the environment was also observed as a potentially important influencing factor in the viewer's presence in the virtual world. Videos that simulated the player's interaction with virtual objects in the environment were marked with labels that indicated higher engagement of the viewer as opposed to videos where the user was merely a point of view with no ability to manipulate the virtual environment. This is supported by evidence in the literature on virtual presence that proposes interaction in a virtual environment to be instrumental in establishing presence (Loomis 1992; Rosenberg 1994).

Following are some implications for designing construction-specific VR experiences based on findings about the value of realism in a virtual environment (a summary is provided in Table 7):

- A good physics engine must be used in the development of the VR experience. The physical rules are especially important for simulating a hazard recognition task, since it has been shown to improve performance when workers use Haddon's energy classification to identify hazards (Haddon 1973; Albert et al. 2014b).
- the development effort must focus time and resources on the development of realistic physical features of the environment.
- Internal discomfort like excessive heat, exhaustion, and heat stroke should be simulated in the virtual experience by consulting with construction workers and aiming to simulate their perceptual experience in those situations.
- To stimulate a realistic auditory experience in addition to visual realism, three-dimensional spatial ambient sound recorded on a real construction site can be integrated into the immersive virtual construction experience.
- Users prefer to have agency to interact with virtual objects and have an impact in the virtual world. In a hazard recognition task, allowing the user to have an impact on the probability of a hazard causing an accident could be facilitated by this capability.

Pattern 3: Importance of Logical Consistency

The third pattern highlights the importance of maintaining logical consistency in a simulation. Unrealistic consequences or illogical occurrences lead to disengagement and act as a deterrent to triggering emotional arousal. An example of this was observed in the weather_tornado video, where a piece of furniture remained still during a tornado when in reality it would have been thrown by the tornado's centrifugal force. This type of illogical event seemed to disengage the viewer's attention. Logical consistency in a simulation could be especially important if the user were familiar with the situation being simulated. Similarly, sometimes limited sensory access to a virtual world leads to some characteristics of the experience seeming illogical. For example, in the weather_flood video, viewers seemed to disengage because it was not clear why someone stuck in a car during a flood would not try to open the door. However, addressing how water pressure on the door would make it difficult to open during a flood might have resolved this issue among viewers to avoid disengagement. This highlights the importance of addressing nonobvious and especially nonvisual features of the environment such as pressure. Without the simulation of such nonobvious features, the virtual environment might seem unnatural, risking disengagement.

The findings regarding the importance of addressing nonvisual or nonobvious features in the simulation highlight the importance of involving features of the construction site in the environment, which might be beyond the scope of visual and auditory sense perception afforded by typical VR systems. Many aspects of the construction environment, especially ones that have a direct effect on construction site safety are nonobvious, nonphysical, or nonvisual.

Examples of these include suboptimum weather conditions, hazards due to gravitational or stored potential energy (Albert et al. 2014b), and other working conditions that are imperceivable to vision or hearing. A potential approach to addressing these features of the environment is to introduce them in a scripted dialogue with a virtual character in the environment. However, there is an apparent limit to the number of characteristics that can be introduced through dialogue and that can be held in the user's working memory. Thus, technologies that enable haptic feedback and other sensory stimuli in VR could be helpful in conveying some of these characteristics. In the simulation of a construction environment for hazard recognition training, it seems essential to involve sensations beyond sight and hearing due to the numerous nonphysical sources of hazards as showcased in the hazard energy wheel by Haddon (1973) and specifically in construction by Albert et al. (2014b). Thus, involvement of technologies like haptic feedback are especially relevant to guiding subsequent studies that explore their effect on safety learning (a summary is provided in Table 8).

Limitations

The primary limitation of this study is the reliance on nonconstruction simulations due to the lack of construction simulations of hazardous scenarios to target emotional arousal. However, VR development is labor-intensive and VR learning requires substantial professional resources that can be hard to access. Therefore,

Table 7. Summary of design implications for construction-specific VR from Pattern 2

Design insight	Example
Good physics engine	Integration of physics engine in the game development software
Simulate perceptual discomfort for nonobvious symptoms of hazardous environment	Render models with realistic textures and include ambient 3D sound recorded on a real construction site using ambisonic microphones
Provide agency to the user in a virtual world	Ability to avoid an accident by managing a hazard on a virtual construction site

Table 8. Summary of design implications for construction-specific VR from Pattern 3

Design insight	Example
Resolve logical fallacies in the simulation	Involving a SME in the design process to verify the validity of the virtual world
Introduce nonobvious characteristics of the environment in a dialogue with a virtual character	Introduce high temperatures by scripting virtual construction workers who complain about the heat or mention discomfort due to the heat
Introduce nonvisual characteristics of the environment using additional technologies	Use haptic feedback gloves to simulate vibrations of a concrete vibrator; use heat fans to simulate high temperatures

even though the design decisions in this paper may not make construction-specific claims about effects on construction learners, they can provide relevant guidance based on existing evidence to help future researchers use their VR development resources wisely, especially in the development of industrial use cases of VR that aim to elicit emotional arousal.

Secondly, the study relied on a sample of anonymous commenters on YouTube, which may not be representative of the population of interest, which is construction workers. However, future researchers can leverage these findings to develop construction simulations and gather feedback from construction workers using robust quantitative and qualitative measurements afforded in an experimental setting. These efforts will not only enable construction-specific feedback but may also encourage future use of nontraditional data like YouTube comments for the design of simulations based on intended emotional reactions.

Overall, the study was limited by its exploratory nature. Its results provide evidence of some design implications for construction-specific VR simulations, but further research is needed to confirm its findings and establish their generalizability. In particular, future research should focus on gathering feedback from construction practitioners on context-relevant virtual environments to assess the impact of design decisions on learning outcomes in a construction-specific context.

Conclusion

In an effort to address the limitations of traditional lecture-based hazard identification training by introducing immersive VR environments, this study's findings inform the design of such environments to leverage emotional arousal and maintain users' attention. The contribution of this work is a novel approach to deriving feedback from users of virtual simulations from other domains by analyzing comments about YouTube videos that resemble the desired training environment.

The patterns identified in the reactions to specific features of the videos provide evidence for certain design decisions to be made in developing a construction-specific virtual environment. The recommended design implementations relate to nonplayer characters (NPCs), realism of the virtual environment, logical validity of the environment, and integration of additional technologies to simulate sensations beyond visual and auditory perception. Based on the patterns that were observed, the authors report design implications for a construction-specific VR environment where an emotional reaction is desired.

This work makes two primary contributions. First, it contributes to the construction research body of knowledge by defining a method to leverage NLP strategies to make sense of large data sets of unstructured human-generated comments. Second, it may inform subsequent construction safety training research targeting VR to replicate the kinds of scenarios that can be hazardous.

While the findings of this work have not yet been validated by construction-specific VR safety research, they offer usable, and evidence-based, development strategies to inform the development of virtual construction simulations based on findings from other relevant domains.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

This material is based on work supported by the National Science Foundation under Grants Nos. 1917763 and 1917750.

References

Aggarwal, R., S. A. Black, J. R. Hance, A. Darzi, and N. J. W. Cheshire. 2006. "Virtual reality simulation training can improve inexperienced surgeons' endovascular skills." *Eur. J. Vasc. Endovasc. Surg.* 31 (6): 588–593.

Agrawal, A., and A. An. 2012. "Unsupervised emotion detection from text using semantic and syntactic relations." In *Proc., IEEE/WIC/ACM Int. Conf. on Web Intelligence and Intelligent Agent Technology*, 346–353. New York: IEEE.

Albert, A., M. R. Hallowell, B. Kleiner, A. Chen, and M. Golparvar-Fard. 2014a. "Enhancing construction hazard recognition with high-fidelity augmented virtuality." *J. Constr. Eng. Manage.* 140 (7): 04014024. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000860.

Albert, A., M. R. Hallowell, and B. M. Kleiner. 2014b. "Enhancing construction hazard recognition and communication with energy-based cognitive mnemonics and safety meeting maturity model: Multiple baseline study." *J. Constr. Eng. Manage.* 140 (2): 04013042. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000790.

Alberto, T. C., J. V. Lochter, and T. A. Almeida. 2015. "TubeSpam: Comment spam filtering on YouTube." In *Proc., IEEE 14th Int. Conf. on Machine Learning and Applications (ICMLA)*, 138–143. New York: IEEE.

Altgassen, M., L. H. Phillips, J. D. Henry, P. G. Rendell, and M. Kliegel. 2010. "Emotional target cues eliminate age differences in prospective memory." *Q. J. Exp. Psychol.* 63 (6): 1057–1064. https://doi.org/10.1080/17470211003770920.

Amarasekara, I., and W. J. Grant. 2019. "Exploring the YouTube science communication gender gap: A sentiment analysis." *Public Understanding Sci.* 28 (1): 68–84. https://doi.org/10.1177/0963662518786654.

Asiri, S. 2018. "MachineLearningClassifiers." Accessed June 1, 2020. https://towardsdatascience.com/machine-learning-classifiers-a5cc4e1b0623.

Azimi, E., A. Winkler, E. Tucker, L. Qian, J. Doswell, N. Navab, and P. Kazanzides. 2018. "Can mixed-reality improve the training of medical procedures?" In *Proc., Conf. IEEE Engineering in Medicine and Biology Society*, 4065–4068. New York: IEEE.

Babanejad, N., A. Agrawal, A. An, and M. Papagelis. 2020. "A comprehensive analysis of preprocessing for word representation learning in affective tasks." In *Proc., 58th Annual Meeting of the Association for Computational Linguistics*, 5799–5810. Stroudsburg, PA: Association for Computational Linguistics.

Balban, M. Y., E. Cafaro, L. Saue-Fletcher, M. J. Washington, M. Bijanzadeh, A. M. Lee, E. F. Chang, and A. D. Huberman. 2021. "Human responses to visually evoked threat." *Curr. Biol.* 31 (3): 601–612. https://doi.org/10.1016/j.cub.2020.11.035.

Bates, J. 1992. *Virtual reality, art, and entertainment*. Cambridge, MA: MIT Press Direct.

Bhandari, S., and M. R. Hallowell. 2017. "Emotional engagement in safety training: Impact of naturalistic injury simulations on the emotional state of construction workers." *J. Constr. Eng. Manage.* 143 (12): 04017090. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001405.

Bhandari, S., M. R. Hallowell, and J. Correll. 2019. "Making construction safety training interesting: A field-based quasi-experiment to test the relationship between emotional arousal and situational interest among adult learners." *Saf. Sci.* 117 (Jun): 58–70.

Bhandari, S., and K. R. Molenaar. 2020. "Using debiasing strategies to manage cognitive biases in construction risk management: Recommendations for practice and future research." *Pract. Period. Struct. Des. Constr.* 25 (4): 04020033. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000521.

Bird, S., E. Klein, and E. Loper. 2009. *Natural language processing with Python: Analyzing text with the natural language toolkit*. Sebastopol, CA: O'Reilly Media Inc.

Bischoff-Grethe, A., E. Hazeltine, L. Bergren, R. B. Ivry, and S. T. Grafton. 2008. "The influence of feedback valence in associative learning." *Neuroimage* 44 (1): 243–251.

Blei, D. M., A. Y. Ng, and M. I. Jordan. 2003. "Latent Dirichlet allocation." *J. Mach. Learn. Res.* 3 (Jun): 993–1022.

BLS (Bureau of Labor Statistics). 2018. "Employer-reported workplace injuries and illnesses (Annual)." Accessed May 15, 2020. https://www.osha.gov/pls/odi/establishment_search.html.

Bradley, M. M., M. K. Greenwald, M. C. Petry, and P. J. Lang. 1992. "Remembering pictures: Pleasure and arousal in memory." *J. Exp. Psychol.: Learn. Memory Cognit.* 18 (2): 379–390.

Cahill, L., and J. L. McGaugh. 1998. "Mechanisms of emotional arousal and lasting declarative memory." *Trends Neurosci.* 21 (7): 294–299. https://doi.org/10.1016/S0166-2236(97)01214-9.

Chang, J., S. Gerrish, C. Wang, J. L. Boyd-Graber, and D. M. Blei. 2009. *Reading tea leaves: How humans interpret topic models*, 288–296. Red Hook, NY: Curran Associates.

Cheok, A. D., M. Haller, O. N. N. Fernando, and J. P. Wijesena. 2009. "Mixed reality entertainment and art." *Int. J. Virtual Reality* 8 (2): 83–90. https://doi.org/10.20870/IJVR.2009.8.2.2729.

Chung, S., J. Cheon, and K. W. Lee. 2015. "Emotion and multimedia learning: An investigation of the effects of valence and arousal on different modalities in an instructional animation." *Instr. Sci.* 43 (5): 545–559. https://doi.org/10.1007/s11251-015-9352-y.

Cockrell, J. T. 1979. *Effective training for target identification under degraded conditions*. Alexandria, VA: Army Research Inst for the Behavioral and Social Sciences.

Creswell, J. W., and C. N. Poth. 2007. *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: SAGE.

Damirchilo, F., A. Hosseini, M. M. Parast, and E. H. Fini. 2021. "Machine learning approach to predict international roughness index using long-term pavement performance data." *J. Transp. Eng. Part B: Pavements* 147 (4): 04021058. https://doi.org/10.1061/JPEODX.0000312.

Das, S., A. Dutta, T. Lindheimer, M. Jalayer, and Z. Elgart. 2019. "YouTube as a source of information in understanding autonomous vehicle consumers: Natural language processing study." *Transp. Res. Rec.* 2673 (8): 242–253. https://doi.org/10.1177/0361198119842110.

Demirkesen, S., and D. Arditi. 2015. "Construction safety personnel's perceptions of safety training practices." *Int. J. Project Manage.* 33 (5): 1160–1169. https://doi.org/10.1016/j.ijproman.2015.01.007.

Deppermann, S., H. Storchak, A. J. Fallgatter, and A. C. Ehlis. 2014. "Stress-induced neuroplasticity:(Mal) adaptation to adverse life events in patients with PTSD—A critical overview." *Neuroscience* 283 (Jun): 166–177. https://doi.org/10.1016/j.neuroscience.2014.08.037.

Ding, N., W. Zhou, and A. Y. H. Fung. 2018. "Emotional effect of cinematic VR compared with traditional 2D film." *Telematics Inf.* 35 (6): 1572–1579. https://doi.org/10.1016/j.tele.2018.04.003.

Duman, R. S. 2004. "Neural plasticity: Consequences of stress and actions of antidepressant treatment." *Dialogues Clin. Neurosci.* 6 (2): 157–169. https://doi.org/10.31887/DCNS.2004.6.2/rduman.

Eiris, R., M. Gheisari, and B. Esmaeili. 2018. "PARS: Using augmented 360-degree panoramas of reality for construction safety training." *Int. J. Environ. Res. Public Health* 15 (11): 2452. https://doi.org/10.3390/ijerph15112452.

Farra, S., E. Miller, N. Timm, and J. Schafer. 2013. "Improved training for disasters using 3-D virtual reality simulation." *West. J. Nurs. Res.* 35 (5): 655–671. https://doi.org/10.1177/0193945912471735.

Finkenauer, C., O. Luminet, L. Gisle, A. El-Ahmadi, M. Van Der Linden, and P. Philippot. 1998. "Flashbulb memories and the underlying mechanisms of their formation: Toward an emotional-integrative model." *Memory Cognit.* 26 (3): 516–531. https://doi.org/10.3758/BF03201160.

Gergen, K. J. 2015. *An invitation to social construction*. New York: SAGE.

Goetz, T., A. C. Frenzel, R. Pekrun, N. C. Hall, and O. Lüdtke. 2007. "Between-and within- domain relations of students' academic emotions." *J. Educ. Psychol.* 99 (4): 715–733. https://doi.org/10.1037/0022-0663.99.4.715.

Gonzalez, M., and J. Hernandez. 2014. "Using Krawthol's affective taxonomy to enhance safety training in the construction industry." *Saf. Sci.* 72 (Jun): 112–119.

Green, C. S., and D. Bavelier. 2008. "Exercising your brain: A review of human brain plasticity and training-induced learning." *Psychol. Aging* 23 (4): 692–701. https://doi.org/10.1037/a0014345.

Guest, G., K. M. MacQueen, and E. E. Namey. 2011. *Applied thematic analysis*. New York: SAGE.

Haddon, W. 1973. "Energy damage and the ten countermeasure strategies (IRCOBI)." In *Proc., Int. Research Council on the Biomechanics of Injury Conf.*, 489–505. Los Angeles, CA: SAGE.

Hadikusumo, B., and S. Rowlinson. 2002. "Integration of virtually real construction model and design-for-safety-process database." *Autom. Constr.* 11 (5): 501–509. https://doi.org/10.1016/S0926-5805(01)00061-9.

Hannum, W. 2009. "Training myths: False beliefs that limit the efficiency and effectiveness of training solutions, part 1." *Perform. Improv.* 48 (2): 26–30. https://doi.org/10.1002/pfi.20051.

Haslam, R. A., S. A. Hide, A. G. Gibb, D. E. Gyi, T. Pavitt, S. Atkinson, and A. R. Duff. 2005. "Contributing factors in construction accidents." *Appl. Ergon.* 36 (4): 401–415. https://doi.org/10.1016/j.apergo.2004.12.002.

Hawkins, D. G. 1995. "Virtual reality and passive simulators: The future of fun." In *Communication in the age of virtual reality*, 159–189. London: Taylor & Francis.

Ho, J. C., and R. Ng. 2022. "Perspective-taking of non-player characters in prosocial virtual reality games: Effects on closeness, empathy, and game immersion." *Behav. Inf. Technol.* 41 (6): 1185–1198. https://doi.org/10.1080/0144929X.2020.1864018.

Holsti, O. R. 1969. *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.

Hudson, S., S. Matson-Barkat, N. Pallamin, and G. Jegou. 2019. "With or without you? Interaction and immersion in a virtual reality experience." *J. Bus. Res.* 100 (Jun): 459–468. https://doi.org/10.1016/j.jbusres.2018.10.062.

Immordino-Yang, M. H. 2015. *Emotions, learning, and the brain: Exploring the educational implications of affective neuroscience (the Norton series on the social neuroscience of education)*. New York: WW Norton & Company.

Jeelani, I., A. Albert, and J. A. Gambatese. 2017. "Why do construction hazards remain unrecognized at the work interface?" *J. Constr. Eng. Manage.* 143 (5): 04016128. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001274.

Jeelani, I., K. Han, and A. Albert. 2020. *Development of virtual reality and stereo-panoramic environments for construction safety training*. Bingley, UK: Emerald.

Jenkins, J. L., G. Takahashi, W. J. Huang, J. McGraw, A. Luginbuhl, and L. Theademan. 2020. *OSHA funded fall safety VR training simulator*. Amsterdam, Netherlands: Elsevier.

Johnson, R. R. 1998. *User-centered technology: A rhetorical theory for computers and other mundane artifacts*. New York: SUNY Press.

Jones, L., C. A. Bowers, D. Washburn, A. Cortes, and R. V. Satya. 2002. *The effect of olfaction on immersion into virtual environments*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Jünger, J., and T. F. Keyling. 2020. "Facepager. An application for automated data retrieval on the web." Accessed June 20, 2020. https://github.com/strohne/Facepager/.

Kaplan, A. D., J. Cruit, M. Endsley, S. M. Beers, B. D. Sawyer, and P. A. Hancock. 2020. "The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis." *Hum. Factors* 63 (4): 706–726.

Kim, A., M. Chang, Y. Choi, S. Jeon, and K. Lee. 2018. "The effect of immersion on emotional responses to film viewing in a virtual environment." In *Proc., 2018 {IEEE} Conf. on Virtual Reality and 3D User Interfaces ({VR})*, 601–602. New York: IEEE.

Kim, D., J.-H. Kim, and Y. Nam. 2014. "How does industry use social networking sites? An analysis of corporate dialogic uses of Facebook, Twitter, YouTube, and LinkedIn by industry type." *Qual. & Qual.* 48 (5): 2605–2614.

Kim, J., and P. Kines. 2018. "The effect of emotional arousal on safety intentions in construction." *Saf. Sci.* 105 (Aug): 153–160.

Kim, K., H. J. Paek, and J. Lynn. 2010. "A content analysis of smoking fetish videos on YouTube: Regulatory implications for tobacco control."

Health Commun. 25 (2): 97–106. https://doi.org/10.1080/10410230903544415.

Kolhatkar, V., H. Wu, L. Cavasso, E. Francis, K. Shukla, and M. Taboada. 2020. "The SFU opinion and comments corpus: A corpus for the analysis of online news comments." *Corpus Pragmatics* 4 (2): 155–190. https://doi.org/10.1007/s41701-019-00065-w.

Kowalski-Trakofler, K. M., and E. A. Barrett. 2003. "The concept of degraded images applied to hazard recognition training in mining for reduction of lost-time injuries." *J. Saf. Res.* 34 (5): 515–525.

Kuniavsky, M. 2003. "Observing the user experience: A practitioner's guide to user research." In *IEEE transactions on professional communication*. New York: IEEE.

Kuo, M.-F., J. Grosch, F. Fregni, W. Paulus, and M. A. Nitsche. 2007. "Focusing effect of acetylcholine on neuroplasticity in the human motor cortex." *J. Neurosci. Soc. Neurosci.* 27 (52): 14442–14447. https://doi.org/10.1523/JNEUROSCI.4104-07.2007.

Lang, A., K. Dhillon, and Q. Dong. 1995. "The effects of emotional arousal and valence on tele- vision viewers' cognitive capacity and memory." *J. Broadcast. Electron. Media* 39 (3): 313–327. https://doi.org/10.1080/08838159509364309.

Lateef, F. 2010. "Simulation-based learning: Just like the real thing." *J. Emerging Trauma Shock* 3 (4): 348. https://doi.org/10.4103/0974-2700.70743.

Lee, J., J.-S. Yi, and J. Son. 2019. "Development of automatic-extraction model of poisonous clauses in international construction contracts using rule-based NLP." *J. Comput. Civ. Eng.* 33 (3): 04019003. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000807.

Liu, B. 2012. "Sentiment analysis and opinion mining." In Vol. 5 of *Synthesis lectures on human language technologies*, 1–167. Berlin: Springer.

Loewenstein, G. F., E. U. Weber, C. K. Hsee, and N. Welch. 2001. "Risk as feelings." *Psychol. Bulletin* 127 (2): 267–286. https://doi.org/10.1037/0033-2909.127.2.267.

Loomis, J. M. 1992. "Distal attribution and presence." *Presence: Teleoperators Virtual Environ.* 1 (1): 113–119. https://doi.org/10.1162/pres.1992.1.1.113.

Lopez, M. C., G. Deliens, and A. Cleeremans. 2016. "Ecological assessment of divided attention: What about the current tools and the relevancy of virtual reality." *Rev. Neurol.* 172 (4): 270–280. https://doi.org/10.1016/j.neurol.2016.01.399.

Mastli, M., and J. Zhang. 2017. "Interactive highway construction simulation using game engine and virtual reality for education and training purpose." *Comput. Civ. Eng.* 2017 (1): 399–406.

Mateas, M. 1999. "An Oz-centric review of interactive drama and believable agents." In *Artificial intelligence today: Recent trends and developments*, 297–328. Berlin: Springer.

Mauri, P., C. Miniussi, M. Balconi, and D. Brignani. 2015. "Bursts of transcranial electrical stimulation increase arousal in a continuous performance test." *Neuropsychologia* 74 (Aug): 127–136.

Maxwell, J. A. 2016. "Expanding the history and range of mixed methods research." *J. Mixed Methods Res.* 10 (1): 12–27.

Mayer, R. E., and G. Estrella. 2014. "Benefits of emotional design in multimedia instruction." *Learn. Instr.* 33 (Feb): 12–18. https://doi.org/10.1016/j.learninstruc.2014.02.004.

McGaugh, J. L. 2000. "Research shows the role of emotions and stress in forming long-lasting memories." *Brain Front.* 2000 (1): 2–3.

Miller, E. D. 2015. "Content analysis of select YouTube postings: Comparisons of reactions to the Sandy Hook and Aurora shootings and Hurricane Sandy." *Behav. Soc. Networking* 18 (11): 635–640. https://doi.org/10.1089/cyber.2015.0045.

Mohammad, S. 2018. *Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words*, 174–184. Melbourne, Australia: Association for Computational Linguistics.

Monteiro, D., H.-N. Liang, J. Wang, L. Wang, X. Wang, and Y. Yue. 2018. "Evaluating the effects of a cartoon-like character with emotions on users' behaviour within virtual reality environments." In *Proc., 2018 IEEE Int. Conf. on Artificial Intelligence and Virtual Reality (AIVR)*, 229–236. Los Alamitos, CA: IEEE Computer Society.

Morshead, R. W. 1965. *Taxonomy of educational objectives handbook II: Affective domain*. New York: David McKay Publications.

Mostafa, N., A. Albert, and Z. M. Carlos. 2016. "Improving hazard-recognition performance and safety training outcomes: Integrating strategies for training transfer." *J. Constr. Eng. Manage.* 142 (10): 04016048. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001160.

Müller, M., M. Salathé, and P. E. Kummervold. 2023. "COVID-Twitter-BERT: A natural language processing model to analyse COVID-19 content on twitter." *Front. Artif. Intell.* 6: 1023281. https://doi.org/10.3389/frai.2023.1023281.

Osgood, C. E., G. J. Suci, and P. H. Tannenbaum. 1957. *The measurement of meaning*. Champaign, IL: Univ. of Illinois Press.

Paek, H.-J., et al. 2010. "Content analysis of antismoking videos on YouTube: Message sensation value, message appeals, and their relationships with viewer responses." *Health Educ. Res.* 25 (6): 1085–1099. https://doi.org/10.1093/her/cyq063.

Paljic, A. 2017. "Ecological validity of virtual reality: Three use cases." In *New trends in image analysis and processing*, edited by S. Battiato, G. M. Farinella, M. Leo, and G. Gallo, 301–310. Berlin: Springer.

Park, C. S., and H. J. Kim. 2013. "A framework for construction safety management and visualization system." *Autom. Constr.* 33 (Aug): 95–103. https://doi.org/10.1016/j.autcon.2012.09.012.

Parthasarathy, S., and C. Busso. 2017. *Jointly predicting arousal, valence and dominance with multi-task learning*, 1103–1107. Baixas, France: ISCA.

Pedregosa, F., et al. 2011. "Scikit-learn: Machine learning in Python." *J. Mach. Learn. Res.* 12 (85): 2825–2830.

Pekrun, R., T. Goetz, W. Titz, and R. P. Perry. 2002. "Academic emotions in students' self regulated learning and achievement: A program of qualitative and quantitative research." *Educ. Psychol.* 37 (2): 91–105. https://doi.org/10.1207/S15326985EP3702_4.

Perlman, A., R. Sacks, and R. Barak. 2014. "Hazard recognition and risk perception in construction." *Saf. Sci.* 64 (Aug): 22–31. https://doi.org/10.1016/j.ssci.2013.11.019.

Pham, H. C., N. N. Dao, A. Pedro, Q. T. Le, R. Hussain, S. Cho, and C. S. Park. 2018. "Virtual field trip for mobile construction safety education using 360-degree panoramic virtual reality." *Int. J. Eng. Educ.* 34 (4): 1174–1191.

Pittenger, C., and R. S. Duman. 2008. "Stress, depression, and neuroplasticity: A convergence of mechanisms." *Neuropsychopharmacology* 33 (1): 88–109. https://doi.org/10.1038/sj.npp.1301574.

Plass, J. L., S. Heidig, E. O. Hayward, B. D. Homer, and E. Um. 2014. "Emotional design in multimedia learning: Effects of shape and color on affect and learning." *Learn. Instr.* 29 (Jun): 128–140. https://doi.org/10.1016/j.learninstruc.2013.02.006.

Posner, J., J. A. Russell, and B. S. Peterson. 2005. "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology." *Develop. Psychopathol.* 17 (3): 715–734. https://doi.org/10.1017/S0954579405050340.

Ravaja, N., T. Saari, M. Turpeinen, J. Laarni, M. Salminen, and M. Kivikangas. 2006. "Spatial presence and emotions during video game playing: Does it matter with whom you play?" *Presence: Teleoperators Virtual Environ.* 15 (4): 381–392.

Riva, G., F. Mantovani, C. S. Capideville, A. Preziosa, F. Morganti, D. Villani, A. Gaggioli, C. Botella, and M. Alcañiz. 2007. "Affective interactions using virtual reality: The link between presence and emotions." *Cyber Psychol. Behav.* 10 (1): 45–56. https://doi.org/10.1089/cpb.2006.9993.

Rosenberg, L. B. 1994. *Virtual fixtures: Perceptual overlays enhance operator performance in telepresence tasks*. Stanford, CA: University ProQuest Dissertations Publishing.

Russell, J. A. 1980. "A circumplex model of affect." *J. Personality Soc. Psychol.* 39 (6): 1161–1178. https://doi.org/10.1037/h0077714.

Russell, J. A. 2003. "Core affect and the psychological construction of emotion." *Psychol. Rev.* 110 (1): 145–172. https://doi.org/10.1037/0033-295X.110.1.145.

Sagayadevan, V., and S. Jeyaraj. 2012. "The role of emotional engagement in lecturer-student interaction and the impact on academic outcomes of student achievement and learning." *J. Scholarship Teach. Learn.* 12 (3): 1–30.

Salame, P., and B. Göransson. 1989. "An evaluation of two types of safety training in the construction industry." *J. Saf. Res.* 20 (1): 33–40.

Sarakit, P., T. Theeramunkong, C. Haruechaiyasak, and M. Okumura. 2015. “Classifying emotion in Thai Youtube comments.” In *Proc., 6th Int. Conf. of Information and Communication Technology for Embedded Systems (IC-ICTES)*, 1–5. New York: IEEE.
Savigny, J., and A. Purwarianti. 2017. “Emotion classification on youtube comments using word embedding.” In *Proc., 2017 Int. Conf. on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*, 1–5. New York: IEEE.
Schechtman, E., O. Laufer, and R. Paz. 2010. “Negative valence widens generalization of learning.” *J. Neurosci.* 30 (31): 10460–10464.
Seels, B., and Z. Glasgow. 1990. *Exercises in instructional design*. New York: Merrill Publishing Company.
Serrano, J. C. M., O. Papakyriakopoulos, and S. Hegelich. 2020. *NLP-based feature extraction for the detection of COVID-19 misinformation videos on YouTube*. Stroudsburg, PA: Association for Computational Linguistics.
Severyn, A., O. Uryupina, B. Plank, A. Moschitti, and K. Filippova. 2014. “Opinion mining on YouTube.” In *Proc., 52nd Annual Meeting of the Association for Computational Linguistics*, 1252–1261. Stroudsburg, PA: Association for Computational Linguistics.
Shilling, R., M. Zyda, and E. C. Wardynski. 2002. “Introducing emotion into military simulation and videogame design: America’s Army: Operations and VIRTE.” In Vol. 3 of *Proc., GameOn Conf.*, 1–6. Monterey, CA: Dudley Knox Library.
Shum, M., and Y. Xin. 2020. “Time-varying risk aversion? Evidence from near-miss accidents.” *Rev. Econ. Stat.* 104 (6): 1317–1328.
Simpson, B., and N. Marshall. 2010. “The mutuality of emotions and learning in organizations.” *J. Manage. Inq.* 19 (4): 351–365. https://doi.org/10.1177/1056492610376531.
Slater, M. 2003. “A note on presence terminology.” *Presence Connect* 3 (3): 1–5.
Steadman, R. H., W. C. Coates, Y. M. Huang, R. Matevosian, B. R. Larmon, L. McCullough, and D. Ariel. 2006. “Simulation-based training is superior to problem-based learning for the acquisition of critical assessment and management skills.” *Crit. Care Med.* 34 (1): 151–157.
Su, X., P. S. Dunston, R. W. Proctor, and X. Wang. 2013. “Influence of training schedule on development of perceptual–motor control skills for construction equipment operators in a virtual training system.” *Autom. Constr.* 35 (Aug): 439–447. https://doi.org/10.1016/j.autcon.2013.05.029.
Sureka, A. 2011. “Mining user comment activity for detecting forum spammers in You tube.” Preprint, submitted March 25, 2011. https://arxiv.org/abs/1103.5044.
Taboada, M., J. Brooke, M. Tofiloski, K. Voll, and M. Stede. 2011. “Lexicon-based methods for sentiment analysis.” *Comput. Ling.* 37 (2): 267–307.
Tang, L., Y. Zhang, F. Dai, Y. Yoon, Y. Song, and R. S. Sharma. 2017. “Social media data analytics for the U.S. construction industry: Preliminary study on Twitter.” *J. Manage. Eng.* 33 (6): 04017038. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000554.
Teizer, J., T. Cheng, and Y. Fang. 2013. “Location tracking and data visualization technology to advance construction ironworkers’ education and training in safety and productivity.” *Autom. Constr.* 35 (Jul): 53–68. https://doi.org/10.1016/j.autcon.2013.03.004.
Tixier, A. J.-P., M. R. Hallowell, A. Albert, L. van Boven, and B. M. Kleiner. 2019. “Psychological antecedents of risk-taking behavior in construction.” *Saf. Sci.* 140 (11): 04014052. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000894.
Ulate, S. O. 2002. *The impact of emotional arousal on learning in virtual environments*. Monterey, CA: Naval Postgraduate School.
Um, E., J. L. Plass, E. O. Hayward, and B. D. Homer. 2012. “Emotional design in multimedia learning.” *J. Educ. Psychol.* 104 (2): 485–498. https://doi.org/10.1037/a0026609.
Uryupina, O., B. Plank, A. Severyn, A. Rotondi, and A. Moschitti. 2014. *SenTube: A corpus for sentiment analysis on You tube social media*, 4244–4249. Trento, Italy: Univ. of Trento.
van Kessel, P. 2019. “Making sense of topic models.” Accessed August 1, 2020. https://medium.com/pew-research-center-decoded/making-sense-of-topic-models-953a5e42854e.
Vervaeke, J., and L. Ferraro. 2013. “Relevance realization and the neurodynamics and neuroconnectivity of general intelligence.” In *SmartData: Privacy meets evolutionary robotics*, 57–68. Berlin: Springer.
Vervaeke, J., T. P. Lillicrap, and B. A. Richards. 2012. “Relevance realization and the emerging framework in cognitive science.” *J. Logic Comput.* 22 (1): 79–99. https://doi.org/10.1093/logcom/exp067.
Villani, D., C. Carissoli, S. Triberti, A. Marchetti, G. Gilli, and G. Riva. 2018. “Videogames for emotion regulation: A systematic review.” *Games Health J.* 7 (2): 85–99. https://doi.org/10.1089/g4h.2017.0108.
Wang, J., and P.-C. Liao. 2021. “Re-thinking the mediating role of emotional valence and arousal between personal factors and occupational safety attention levels.” *Int. J. Environ. Res. Public Health* 18 (11): 5511. https://doi.org/10.3390/ijerph18115511.
Warpefelt, H. 2016. *The non-player character exploring the believability of NPC presentation and behavior*. Stockholm, Sweden: Stockholm Univ.
Yasmina, D., M. Hajar, and A. M. Hassan. 2016. “Using YouTube comments for text-based emotion recognition.” *Procedia Comput. Sci.* 83 (Apr): 292–299. https://doi.org/10.1016/j.procs.2016.04.128.
Yoo, J. H., and J. Kim. 2012. “Obesity in the New Media: A content analysis of obesity videos on YouTube. Health communication.” *Health Commun.* 27 (Aug): 86–97. https://doi.org/10.1080/10410236.2011.569003.
Zhang, J., and N. M. El-Gohary. 2016. “Semantic NLP-based information extraction from construction regulatory documents for automated compliance checking.” *J. Comput. Civ. Eng.* 30 (2): 04015014. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000346.
Zhang, Y., H. Liu, S. C. Kang, and M. Hussein. 2020. “Virtual reality applications for the built environment: Research trends and opportunities, automation in construction.” *Autom. Constr.* 118 (Sep): 103311. https://doi.org/10.1016/j.autcon.2020.103311.
Zhao, D., and J. Lucas. 2015. “Virtual reality simulation for construction safety promotion.” *Int. J. Inj. Control Saf. Promot.* 22 (1): 57–67. https://doi.org/10.1080/17457300.2013.861853.