Using Augmented Virtuality to Understand the Situational Awareness Model

Siddharth Bhandari[1]; Matthew R. Hallowell[2]; Leaf van Boven[3]; Mani Golparvar-Fard[4]; June Gruber[5]; and Keith M. Welker[6]

[1]Ph.D. Candidate, Dept. of Civil, Environmental, and Architectural Engineering, Univ. of Colorado Boulder, 1111 Engineering Dr., Boulder, CO 80309-0428. E-mail: siddharth.bhandari@colorado.edu
[2]Beavers Endowed Associate Professor of Construction Engineering, Dept. of Civil, Environmental, and Architectural Engineering, Univ. of Colorado at Boulder, 1111 Engineering Dr., Boulder, CO 80309-0428. E-mail: matthew.hallowell@colorado.edu
[3]Professor, Dept. of Psychology and Neuroscience, Univ. of Colorado at Boulder, 1905 Colorado Ave., 345 UCB Muenzinger D347C, CO 80309-0428. E-mail: leaf.vanboven@colorado.edu
[4]Associate Professor, Dept. of Civil and Environmental Engineering and Dept. of Computer Science, Univ. of Illinois at Urbana-Champaign, Urbana, IL. E-mail: mgolpar@illinois.edu
[5]Assistant Professor, Dept. of Psychology and Neuroscience, Univ. of Colorado at Boulder, 1905 Colorado Ave., 345 UCB Muenzinger D321C, CO 80309-0428. E-mail: june.gruber@colorado.edu
[6]Assistant Professor, Dept. of Psychology, Univ. of Massachusetts Boston, 100 Morrissey Blvd. Boston, MA 02125-3393. E-mail: Keith.Welker@umb.edu

Abstract

Situational awareness concept has been long used in aviation and medical fields to understand and improve an individual's ability to interact and comprehend complex dynamic environments. This study attempts to empirically test situational awareness model proposed by Endsley within construction domain. Specifically, we will test how different levels of situational awareness interact with each other and how they interact with decisions. To achieve this goal, a controlled experiment was designed and conducted where the investigators (1) induced positive, negative, or neutral emotions in 66 subjects; (2) exposed participants to construction hazards within a high fidelity virtual environment; and (3) measured participant's hazard recognition skills, their understanding of those hazards, severity assessment, and subsequent decisions. The results revealed there is moderate positive correlation among each level of situational awareness. Linear mixed-effects model analysis revealed that the three levels of SA positively predicted each other. Also, the analysis revealed that there were differences in the degree of influence from each level of situational awareness on the final decision of an individual.

INTRODUCTION

A new report by the Midwest Economic Policy Institute (MEPI) found that construction industry suffers an average of 868 annual fatalities (16 fatalities per week) that cost nearly $5 billion per year in losses (MEPI, 2017). Moreover, fatality rates since 2011 have been on the rise (U.S. BLS 2016), which is a tragic trend facing both academics and practitioners. One aspect of human factors research as it relates to safety that has begun to receive attention in the past decade is situational awareness (SA).

SA is popular concept of human factors engineering that is used in various fields (e.g., nursing, aviation, and military) that involve complex and dynamic environments. Perhaps the most pervasive definition of SA is, "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*" (Endsley 1995, pg. 65). In other words, SA can be understood as a motivation for active and continuous extraction of information and assessment of that information that guides our decisions and subsequent actions.

SA is considered particularly critical in the aviation industry where poor SA was the critical causal factor identified in over 200 aircraft accidents (Hartel et al. 1991). Furthermore, individuals struggling to acquire and maintain SA might not detect and respond to problems with the system they are controlling or present in (Endsley and Kiris 1995).

Despite the fact that construction work is dynamic and complex role of SA with safety and decision-making has not been empirically explored in the industry. This is problematic because of the vulnerability of construction worker's well-being to even small errors. In this study, we will aim to test the predictive nature of Endsley's (1995) model. Specifically, we are aiming to test the relationships between each individual level of SA as it applies to construction safety and how each factor relates to an ultimate decision.

METHODS

Situational Awareness Model in Construction Context

SA is divided primarily into three levels: *perception, comprehension,* and *projection*. These levels will be the foci of this paper:

1. The Level 1 of SA, *perception of the elements in environment*, refers to the ability to grasp relevant stimuli from environment along with their attributes and nature.
2. The Level 2 of SA, *comprehension of the current situation*, refers to the ability of the individual to understand the stimuli perceived in Level 1.
3. Finally, Level 3 of SA, *projection of future status*, is the ability to take the information from Level 1 and Level 2 to make a prediction of the outcomes associated with the elements in the environment.

Endsley (1995) defines elements as the relevant stimuli in an environment that an individual must distinguish and value within a complex system. If we look at this from construction safety lens, then a hazard is an element. Thus, Level 1 refers to *identifying hazards in the environment*, Level 2 refers to *understanding the danger associated with those hazards*, and Level 3 refers to *assessing the severity associated with hazards*. There are many other key factors summarized in Endsley (1995; 2015) such as mental models, nature of information, and cognitive workload that influence the three levels. However, for brevity, the entire Endsley (1995) model cannot be discussed here.

There has been debate among researchers regarding the legitimacy and differing interpretation of the three-step model. Recently, Endsley (2015) rebutted the various alleged misconceptions and fallacies associated with the original SA model, claiming that the three levels of SA are not linear stages and, rather, represent ascending levels. Furthermore, Endsley also disagreed with the Salmon et al. (2012)'s suggestion that there needs to be in-depth exploration on the links and interactions between the levels of SA and Sorensen et al. (2010)'s assertion that each level of SA is highly separated. Endsley (2015) claims that this is an inaccurate understanding of the model because there is a natural progress to higher levels where the linkages are determined by mental models. Therefore, this study aims to explore the relationships between each level of SA while controlling for individual's emotional state and learn about the influences of each level on the subsequent decision under uncertainty.

Experimental Design

Our experiments were conducted in a controlled environment where each subject followed the trajectory shown in Figure 1. Overall experimental protocol was (1) induce three specific emotion states (positive, negative, or neutral) among young adult participants using standardized movie clips validated in previous studies, (2) allow participants to interact with construction hazards in high-fidelity augmented virutality, (3) measure the three levels of SA, and (4) record their ultimate decision regarding safety. To measure each level of SA as it relates to construction hazards, we framed questions to participants based on each of the three levels. Specifically, we asked participants to: identify hazards (Level 1), rate the danger of each hazard (Level 2), and rate the severity associated with each hazard (Level 3). Participants continue this loop until they report there are no more hazards left. After participant report, there are no more hazards in the environment, they have to make a decision on whether or not to allow work being performed. They had option of choosing a decision on a scale of 1-5 where 1 represented *let work proceed as it is being performed*, 2 represented *let work proceed with high caution*, 3 represented *stop work and make minor changes*, 4 represented *stop work and make major changes*, and 5 represented *stop work and note as emergency condition*.

The experiment was carefully designed to ensure control over confounding factors and sources of variability. All subjects were given a private room that was closed to public and received the same laptops, headphones, comfort level, workspace, lighting, and air conditioning. Before starting the experiment, each participant was given a brief demonstration of how to navigate the AV system. Subjects were then randomly selected for the control group (i.e. neutral emotion) or randomly assigned an emotion eliciting video clip. Participants were told that the movie clip viewing was for another experiment to avoid response bias such as demand characteristics. Once the participants completed the emotion induction task, they entered the AV system that was pre-loaded on their computers. The participant's avatar was randomly placed in the environment and they were free to navigate anywhere on site. As they encounter a stimulus, they were told to click on the visual cues, which would unlock a detailed real-world image of construction task being performed. The participants responded to same embedded questionnaire for each unlocked image (Figure 1). When a participant completed the questions for the particular photograph, they were returned to the virtual site to find another stimulus. Each participant was given 30 minutes in the environment. There were 16 stimuli that a participant could encounter that were randomly distributed across the virtual site.

We recruited 66 subjects by sampling from upper-level undergraduate construction engineering courses taught at author's university. This is an obvious participant bias in our sample; however, we accept the bias since the objective was to understand the relationships among levels of SA and the impact of emotional state rather than the skill of the SA process. Furthermore, students offer less experiential bias given their relatively similar professional experiences (Tixier et al. 2014). It should be noted that each participant was given the same hazard recognition training one month prior to the experiment. This training was provided because Bhandari et al. (2016) noted that when asking participants to identify hazards (i.e., the sources of energy that could injure a person), most participants reported safety violations (e.g., not wearing proper protective equipment). This is a subtle but important limitation for most studies involving hazard recognition.

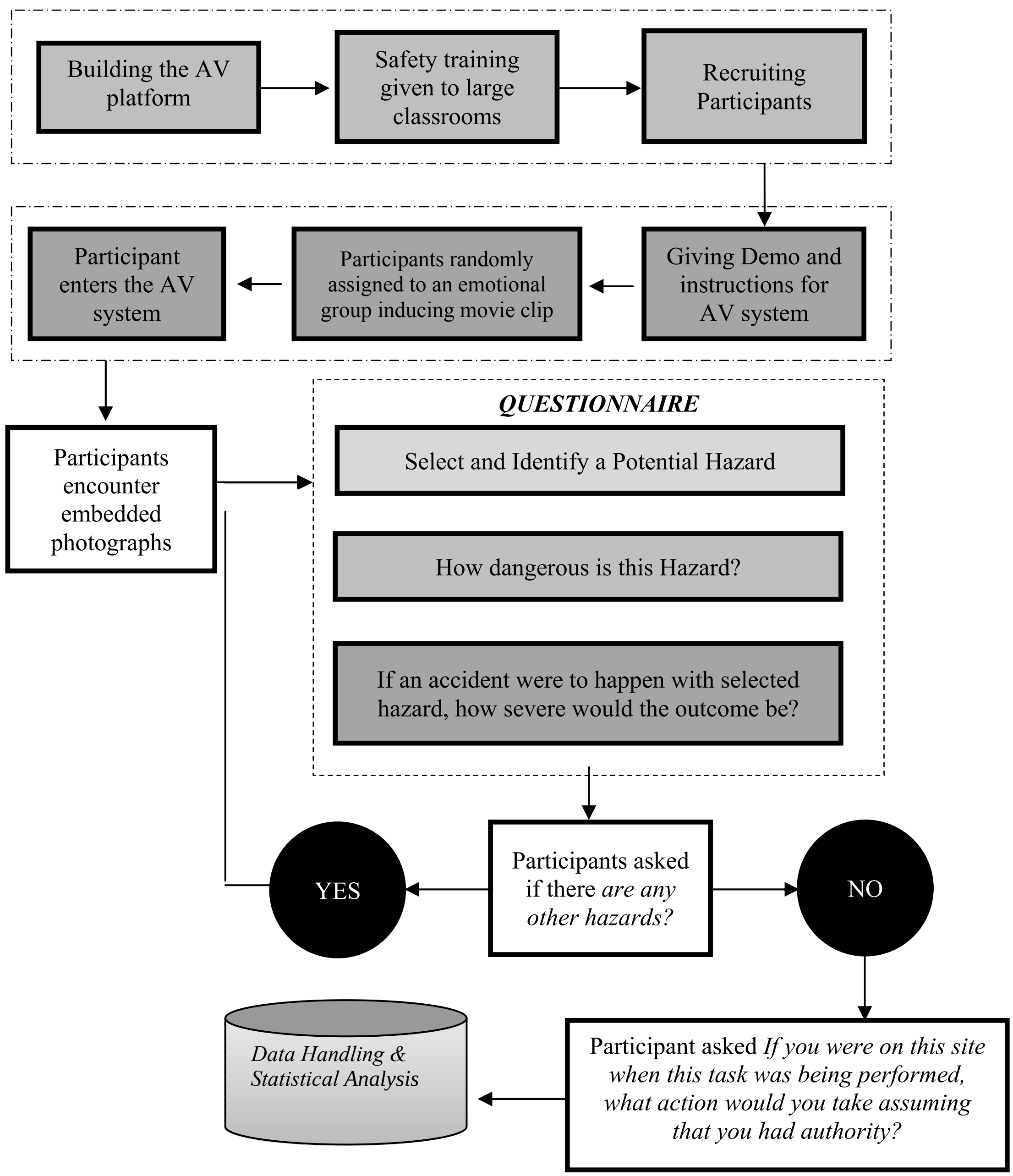

Figure 1. Overview of research process

Using High-Fidelity Virtual Construction Environment

The process of conducting experiments with safety on live construction sites is not possible when the experiment could place people in danger. Therefore, we elected to use a high-fidelity augmented virtual environment as an experimental platform (see Figure 2). Such platforms have been used to conduct successful construction safety studies in the past, serving as a suitable alternative (Albert et al. 2013; Tixier et al. 2014). Although we lose some ecological validity by using an environment, the ability to control the experimental process allows for very strong internal and construct validity.

In our experiment, the platform we have developed can be classified as *augmented virutality* (AV) since it incorporates real-world images into the virtual environment. The images selected by the research team were high-definition images that captured different types of energies that could release and could cause injuries with varying degrees of severity. We included images that showed worker(s) exposed to high and low energy to enhance the sense of realism for the participants in the environment. These images and associated questionnaires were embedded in the AV system that measured different levels of SA discussed above.

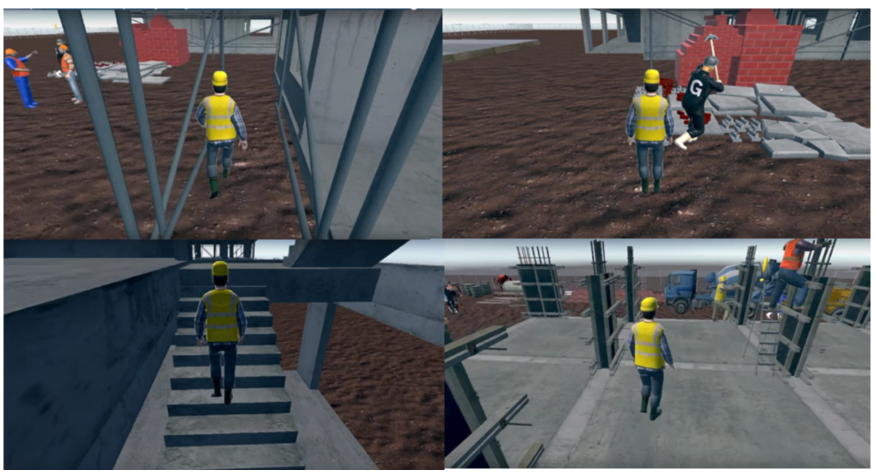

Figure 2. Screenshots of AV platform

The AV system was designed to maximize ecological validity through realism. For example, the system was designed with a non-repeating soundtrack capturing the typical noise of a site to give participants both visual and audible stimuli. Additionally, the system included embodied agents, which were virtual workers controlled by an independent algorithm (Bailenson and Blascovich 2004). It was critical for us to have realistic virtual construction workers because studies have found that subjects in these environments interact with virtual humans like they would with real people (Donath 2007). By incorporating large equipment and material, moving virtual construction workers, and sounds of a construction site, we can support a true immersion in subjects that should engage their SA. This realism in the AV system also allows the subjects to look for hazards and dangerous simulations without putting them in actual line of fire.

Inducing Target Emotions with Movie Clips

Emotions influence our behavioral, physiological, and experiential aspects of daily life by helping us interact and adapt with our surroundings (Oatley et al. 2006). Emotions consciously and sub-consciously influence our decision-making abilities and our decisions under uncertainty are made by valuing both the risk of a particular outcome and the emotional response to the outcome (LeDoux 1998). Various studies have shown that emotions influence our perception of risk and tolerance of risk (Clore et al. 1994; Tixier et al. 2014). These established findings elucidate the need to understand how emotional states influence both a person's progression through the three steps of SA and the final decision.

Table 1. Movie Clips for Emotion Induction

Target Emotion	Movie Clip (Duration)
Neutral (Control Group)	Denali National Park (3:40)
Negative (Treatment Group)	The Champ (3:21)
Positive (Treatment Group)	Whose Line is it Anyway? (2:13)

Inducing emotion through movie clips or videos is a well-established technique in the field of experimental psychology. This technique was selected to induce emotions because it has been successfully validated and used in various studies (Gruber et al. 2008). This integrated well with our experimental design, allowing us to keep subjects engaged on their computer screens throughout the different phases of experiment. Such engagement reduces external influences that could interfere with the induced emotion. The clips selected are shown in Table 1 above.

ANALYSIS

As summarized in Figure 1, participants had to identify all the hazards and finish giving each of those hazards a danger, severity of possible injury, and likelihood of accident rating. After completing those questions, they were asked to decide what action if any, would they take regarding the work shown. Our data were collected as continuous variables. Participants were asked to rate the danger and severity on a sliding scale of 1-100 and the hazard recognition data were collected as percentage of hazards identified (i.e., number of hazards identified / total number of hazards identified for a photograph). The average hazard recognition score was 30.3%, for our sample which is consistent with the average hazard recognition skill level of workers (Albert et al. 2013) suggesting strong external validity.

Relationship Between Individual Levels of SA

There is moderate correlation observed between the three levels as shown in Table 2 especially for Level 2 (Danger) and Level 3 (Severity Assessment) of SA.

Table 2. Correlation Matrix for 3 Levels of SA

	Hazard	**Danger**	**Severity**
Hazard	1.00	0.30	0.27
Danger		1.00	0.68
Severity			1.00

The emotion groups were orthogonally coded to test the two conditions: control vs. treatment (i.e. *neutral vs. positive and negative*) and *negative vs. positive* emotion groups. Response latencies were log-transformed and we used a generalized linear mixed model on the resulting latencies, treating participant as random effects and the photographs participants encountered were nested within the emotion groups. Satterthwaite approximations was used to generate degrees of freedom. The analysis revealed:

[1] Hazard perception was positively predicted by both danger rating given to each hazard [b_0= 0.1, F (1, 466.6) = 26.9, p< 0.01] and severity assessment [b_0= 0.07, F (1, 461.3) = 14.7, p< 0.01]. In other words, both Level 2 and Level 3 positively predict Level 1 of SA. There was no statistically significant difference between the emotional groups.

[2] Rating of danger had statistically significant relationship with hazard identification skill [b_0= 0.41, F (1, 408.1) = 20.25, p<0.01] and on severity assessment rating [b_0= 0.53, F (1, 481.5) = 206.2, p<0.01]. That is, both Level 1 and Level 3 positively predict Level 2 of SA. Again, there was no significant difference between the emotion groups.

[3] Severity assessment was positively related to the rating of danger [b_0= 0.55, F (1, 416.4) =228.3, p<0.01and to hazard identification skill [b_0= 0.42, F (1, 260.3) = 23.72, p<0.01]. There was no difference between control (neutral) vs. treatment (positive and negative) emotion group and between negative vs. positive emotion group.

Relationship between levels of SA and decisions

Linear mixed-effects model was used again to understand the relationship of each individual SA level with decision by participants regarding stopping work controlling for interaction between the three levels of SA. The analysis showed that decision to take action regarding safety of work had a positive relationship with Level 2 (marginal significance) and with Level 1 but no relationship with Level 3 of SA. In other words, decision regarding work being performed was significantly predicted by hazard identification skill [b_0= 0.33, F (1, 472.7) = 4.3, p = 0.039] and the danger rating [b_0= 0.28, F (1, 467.5) = 3.68, p= 0.055] that was assigned to each hazard by the participants. The two-way interaction between severity and danger assessment has a significant relationship [b_0= 0.18, F (1, 481.9) = 4.2, p = 0.041] and also the three-way interaction between all the individual levels of SA was significant [b_0= -0.25, F (1, 484.4) = 5.72, p = 0.017]. There was no statistically significant difference between control group (neutral emotion) v. experimental group (positive/negative emotion) nor was there any significant difference between positive or negative emotion induction groups.

DISCUSSION

Finding 1: hazard identification skill and danger rating had positive relationship whereas severity rating had no relationship with the ultimate decision to stop work to address safety concerns.

It makes intuitive sense that the more hazards workers see the more inclined they would be to take action to eliminate them or minimize the damage. Similarly, assessing more danger per hazard would prompt workers to take action also. This finding confirms that to influence worker's risk-taking behavior, they should not only be trained on improving their hazard identification skills but also on assessing the appropriate danger associated with those hazards to encourage them to address safety violations or hazardous conditions on site.

However, the lack of relationship between severity assessment and subsequent decision is curious. Dread is critical aspect for training because feelings of vulnerability play primary role in motivating individuals to gain knowledge to avoid negative consequences associated with not taking action (Slovic and Peter 2006). We workers to assess and value severity associated with hazards and inform their decision regarding their work environment. Within construction context, Bhandari and Hallowell (2017) found that training focusing on dread and simulating severity associated with hazards can lead to emotional engagement that enhances risk perception. Also, it should be noted that Merkhofer (1987)'s asserted that psychometric risk dimensions such as ratings of severity are solicited from subjects over adverse situations in such a way that it misses many other factors people naturally use to make decisions. Future studies need to explore this relationship by controlling for other factors individuals rely on before taking decisions under risk.

The 2-way interaction between severity and danger suggests that relationship between decision to stop work and improve safety and danger assessment increases as the assessment of severity increases. However, 3-way interaction between percentage of hazards identified, severity and danger assessment is interesting because the analysis shows as participants reported more hazards, the interaction between severity and danger decreased. It could be possible that as number of hazards identified increase that might subconsciously or consciously prompt individuals in construction context to devalue the influence of the severity and danger assessment so that they can continue working and not lose productivity. This could be possible given the monetary and reputation factors associated with stopping work for safety hazards. Literature has found people rationalizing risk for various efficacies like emotional release and personality development (Stranger 1999). Stranger suggests aestheticizing of risks facilitates individuals to take more risks without rational justification.

These findings suggest that by simply improving hazard identification skills among workers, researchers and industry cannot expect workers to take safe decisions. Workers must be given an understanding for those hazards (i.e. ability to assess the danger associated also) which seems quite obvious but safety training sessions rarely focus on developing context required to adequately assess the characteristics of a hazard rather emphasize more on regulations that should be followed (Albert et al. 2013).

Finding 2: there is a positive relationship between each individual level of SA.

While we did not design cross-sectional study to test for bi-directional relationships between the three levels of SA, the moderate correlation levels and statistically significant relationships between each of the levels suggests that Endsley's rebuttals were indeed accurate that the levels are not linear stages. Although, the study shows that the relationship of each ascending level of SA is more nuanced with the decision in this context with working safely. This needs to be further investigated with a large study.

Finding 3: there is no statistically significant impact of induced emotion on any level of SA. This finding is contrary to existing risk and emotion studies, which have suggested that emotional state impacts risk perception (Clore et al. 2006). There are some possible explanations for why the results differ from previous studies: emotions were induced but not sustained for the duration of the experiment. Emotions are quite fickle and their typical residence time can range from minutes to hours (Verduyn et al. 2009). Also, the experience in the AV environment may have altered the induced emotional state because of the novelty of the technology (Riva et al. 2007). Finally, it is possible that we were unable to detect small or medium effects given our study was underpowered due to small sample. A future study should attempt to explore this relationship in a different experimental design to corroborate the findings.

From a theoretical standpoint, the study revealed that the interactions and linkages among the levels of SA are not straightforward and linear in nature. Although, it seems logical that individual identifies hazard, assess danger, quantifies risk and makes a decision however, such oversimplification can yield incorrect and potentially sub-optimal conclusions. The three levels of SA are ascending levels of SA as suggested by Endsley (2015) although this pilot study did not use the complete model (incorporating goals, memory structures, mental models, and attention of a person) as advocated by Endsley (2015) and future studies should scope this consideration in their experiment. From a practical standpoint, this study highlights what factors positively influence workers to make risk-averse decisions which can aid in designing better components of safety training to deliver more value to the workforce. In conclusion, this pilot study addresses the dearth of empirical research on testing a comprehensive SA model especially within construction domain and examining the relationships among SA levels and mediating impact of emotions.

ACKNOWLEDGEMENTS

The material presented is based in part on work supported the National Science Foundation under Grant No. 1362263. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES:

Albert, A., Hallowell, M. R., & Kleiner, B. M. (2013). Enhancing construction hazard recognition and communication with energy-based cognitive mnemonics and safety meeting maturity model: Multiple baseline study. *Journal of Construction Engineering and Management*, *140*(2), 04013042.

Bailenson, J.N., & Blascovich, J. (2004). Avatars. In W.S. Bainbridge (Ed.), Encyclopedia of human-computer interaction (pp. 64–68). Great Barrington, MA: Berkshire.

Bhandari, S., & Hallowell, M. R. (2017). Emotional Engagement in Safety Training: Impact of Naturalistic Injury Simulations on the Emotional State of Construction Workers. Journal of Construction Engineering and Management, 143(12), 04017090.

Bhandari, S., Hallowell, M. R., Van Boven, L., Gruber, J., & Welker, K. M. (2016). Emotional states and their impact on hazard identification skills. In *Construction Research Congress 2016* (pp. 2831-2840).

Cassell, J. (2000). Embodied conversational interface agents. Communications of the ACM, 43(4), 70–78.

Clore, G. L., Schwarz, N., & Conway, M. 1994. Affective causes and consequences of social information processing. Handbook of social cognition,1, 323-417.

Construction Fatalities Cost the United States $5 Billion Per Year. 2017, May 08. (Retrieved July 15, 2017, from https://midwestepi.org/2017/05/08/construction-fatalities-cost-the-us-5-billion-per-year/)

LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster.

Donath, J. (2007). Virtually trustworthy. Science, 317, 53–54.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human factors*, *37*(1), 65-84.

Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. Journal of Cognitive Engineering and Decision Making, 9(1), 4-32.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human factors*, *37*(2), 381-394.

Gruber, J., Johnson, S. L., Oveis, C., & Keltner, D. (2008). Risk for mania and positive emotional responding: too much of a good thing?. *Emotion*, *8*(1), 23.

Härtel, C. E. J., Smith, K. A., & Prince, C. (1989). Defining aircrew coordination: Searching mishaps for meaning. In Fifth International Symposium on Aviation Psychology. Ohio State University.

Isen, A. M., Nygren, T. E., & Ashby, F. G. (1988). Influence of positive affect on the subjective utility of gains and losses: it is just not worth the risk. *Journal of personality and Social Psychology*, *55*(5), 710.

Oatley, K., Keltner, D., & Jenkins, J. M. (2006). Understanding emotions. Blackwell publishing.

Pepitone, A., & DiNubile, M. (1976). Contrast effects in judgments of crime severity and the punishment of criminal violators. Journal of Personality and Social Psychology, 33(4), 448.

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., ... & Alcañiz, M. (2007). Affective interactions using virtual reality: the link between presence and emotions. *CyberPsychology & Behavior*, *10*(1), 45-56.

Salmon, P. M., Stanton, N. A., & Young, K. L. (2012). Situation awareness on the road: review, theoretical and methodological issues, and future directions. Theoretical Issues in Ergonomics Science, 13(4), 472-492.

Sorensen, L. J., Stanton, N. A., & Banks, A. P. (2011). Back to SA school: contrasting three approaches to situation awareness in the cockpit. Theoretical Issues in Ergonomics Science, 12(6), 451-471.

Stranger, M. (1999). The aesthetics of risk: A study of surfing. International review for the sociology of sport, 34(3), 265-276.

Tixier, A. J. P., Hallowell, M. R., Albert, A., van Boven, L., & Kleiner, B. M. (2014). Psychological antecedents of risk-taking behavior in construction. Journal of Construction Engineering and Management.

U.S. Bureau of Labor Statistics, 2016. Census of Fatal Occupational Injuries (Accessed July 14, 2017 from: http://www.bls.gov/data/#injuries).

Verduyn, P., Delvaux, E., Van Coillie, H., Tuerlinckx, F., & Van Mechelen, I. (2009). Predicting the duration of emotional experience: two experience sampling studies. Emotion, 9(1), 83.