Do You Need to Travel? Mapping Face-to-Face Communication Objectives to Technology Affordances

Rachel E. Dianiska[1], Charles J. Peasley[1], Nicholas Wilson[1], Neil Barnett[1], Leilani Hammel[1], Ben Purdy[1], Peggy Wu[2], Elizabeth Shirtcliff[1], James H. Oliver[1], Stephen B. Gilbert[1]
[1]Iowa State University
[2]Raytheon Technologies Research Center

Computer-mediated communications (CMC) can be used as a substitute for face-to-face (FtF) meetings but their effectiveness is highly context dependent. This paper describes a theoretical framework and initial experimental design for characterizing a travel replacement threshold. This effort begins with a use case of remote engineering maintenance training, conducted in three conditions: side-by-side (physically proximate), teleconference (using off-the-shelf software), and a custom VR/AR system designed to provide the apprentice with a virtual view of both the instructor's larger scale lab and smaller scale workbench. The research hypotheses, experimental protocol, and dependent measures are described. The task involves an instructor demonstrating a circuit board troubleshooting task to a remote apprentice. The apprentice then completes the trained task independently, and performance and subject preferences are compared across conditions. The details of this paper, the result of extensive literature review and winnowing of variables, may assist researchers exploring CMC, training, or social communication.

INTRODUCTION

The transportation sector is the largest source of US emissions related to climate change, with air travel being a major contributor. Despite recent advancements in telecommunication technologies, travel-related energy expenditures have been steadily increasing (Mokhtarian, 2009). World energy demand for aircraft is estimated to increase by a total of 10 quadrillion BTU from 2012 to 2040 (U.S. Energy Information Administration, 2016). The aviation industry has recognized this environmental impact and is addressing it in a multi-prong approach, both investing in research for increasing fuel efficiency as well as communications efforts for increasing public awareness (e.g., KLM, 2019). The transportation of passengers accounts for approximately 70% or 568 GW (17 quads/yr) of transportation energy consumption (Transportation Energy Data Book).

This research focuses on developing evidence-based recommendations for when to adopt Computer-Mediated Communications (CMC) as a reasonable alternative to physical travel–a *travel replacement threshold.* The decision to travel or replace it with telecommunications is multi-faceted and highly subjective. Assumptions about CMC are often based on intuition and opinion, and this research seeks to ground those assumptions, if appropriate, with empirical evidence. E.g., one assumption is that one must be co-located in order to build trust, rapport, and shared mental models with customers and colleagues. Such assumptions may have face validity, but they have not yet been systematically characterized or tested.

Given the prevalence of remote training across industries, this project first explores the use case of a remote expert training a field technician or apprentice. This paper describes the experimental design and dependent measures (physical, physiological, and software-based) that will be used to evaluate a new communications system designed to promote travel replacement. While the study has yet to be conducted, it is anticipated that the definitions of dependent measures below may be beneficial to others research communications technology and virtual collaboration.

PREVIOUS WORK

Extensive previous research has explored the impact of communications technologies on distributed team performance (see Ens et al., 2019, for review). Researchers including Jeremy Bailenson have led thorough explorations of human-human communications using virtual media within lab settings (e.g., Oh, et al., 2018). The NASA ANSIBLE project (Mars simulation) explored the use of VR and avatars for communication over multiple months in a more isolated context (Wu et al., 2016).

This project builds on that work by attempting to develop evidence-based benchmarks for communication and task performance goals when using telecommunications. After an extensive literature review, we have documented the psychosocial constructs that are relevant in face-to-face and computer-mediated communication in the Communications Objective Model (COM; Dianiska et al., under review). The COM characterizes the psychosocial constructs underlying effective communications, how these constructs manifest in behavior, as well as appropriate metrics for capturing these constructs. In this particular paper, the researchers describe the process of using a subset of the COM to design a study that will evaluate the affordances of specific communications system (a novel extended reality (XR) system) in a specific context (*On The Job Training*).

APPROACH

This line of work will explore potential interactions of individuals in virtual and augmented reality in fully crossed fashion across four scenarios. In the current paper, we will examine the first scenario, in which a local instructor in augmented reality (AR) demonstrates a training task with physical equipment to a remote apprentice, who is observing the task in virtual reality (VR).

On The Job (OTJ) Training tends to use physically in-place tools, machines, and equipment in a master-apprentice learning model outside of a traditional classroom or computer-based training. An experienced instructor provides instructions and demonstrations of task-relevant knowledge. To mimic OTJ in an operational equipment, the research study avoids pre-created digital content and instead relies on physical equipment. The virtual and augmented reality tools are used only for telecommunications purposes and do not in themselves provide task knowledge or task aides.

The context of interest, a remote training scenario, can be specified based upon three primary functional spaces with varying levels of agency within each of three spaces (Tang et al., 2010). First, a *person space* allows for participants to assess one another's attention, facial, expression, and gestures. Second, a *shared space* provides a reference space for gaining awareness of the presence of others. Finally, a *workspace* provides an area for performing work independently.

In this specific research task, each of these spaces can be examined for the extent to which participants can modify the objects within it or modify their own perspective within it. "High" denotes complete control; "medium" represents limited control, in which the participant has limited ability to modify some objects or attributes within the digital environments; and "low" represents very limited abilities in which the participant can only direct their own attention from a constrained perspective.

This experimental context examines synchronous interactions when individuals are remote, rather than co-located. The instructor will reside and be represented as a physical embodiment within the workspace (while using augmented reality to demonstrate the task performance), whereas the apprentice will be completely digital and observe the instructor's workspace in virtual reality. Individuals recruited as apprentices will be unacquainted with the instructor, leading to a neutral *a priori* social distance and power differential present between the two individuals. The training medium conditions will allow for simultaneous manipulations of *agency* (the extent to which the participant can make modifications to his or her view of the world), *embodiment* (the visual and auditory characteristics of the human operator) and *media richness* (the extent to which the work environment is represented in terms of accuracy, video resolution and audio quality; see Ens et al., 2019, for review). See Table 1 for a classification of the experimental conditions according to agency, embodiment, and media richness.

In addition to the training platforms, a primary goal was identifying an ecologically-valid experimental scenario that parallels the task requirements that are necessary in the real world. Ideally, the task would be modifiable to equate cognitive workload. Also, the ideal task required little prior background knowledge for the participant pool (university undergraduate students) and could accommodate the limited ambulatory range of the wired physiological measurement technology.

Table 1. Agency, embodiment, and media richness classification for experimental conditions

	Low	**Medium**	**High**
Agency	Telecon	XR	Side by Side
Embodiment	XR	Telecon	Side by Side
Media Richness	Telecon	XR	Side by Side

In light of these criteria, the experimental task ultimately selected involves performing maintenance on various breadboard circuit systems, and thus is called the Circuit Study. After viewing an initial five-minute training video on how to insert wires and components into a breadboard, each participant will receive training on a set of circuit systems including a rotating motor, a soft potentiometer color changer, and a dynamic LED display (see Figure 1). In each trial, the participant first completes a brief training session with an instructor (a confederate research assistant) who will describe the particular system and demonstrates the relevant troubleshooting procedures. There are always two circuit faults, one of which requires finding a replacement part located in a specific storage area within the workspace. A tablet-based mobile app Diagnostic Tool (Figure 2) has been developed for the instructor and apprentice to obtain diagnostic fault information during the task. After the training, the participant apprentice then completes a knowledge quiz over the content that was learned and then performs the maintenance task without the supervision of, or assistance from, the instructor. Finally, before beginning the next trial, participants will complete a brief survey describing their qualitative experience with the instructor and the task.

Importantly, planned analyses examine objective task performance and subjective ratings of the learning experience as a function of their training medium. Each participant will complete the circuit training in three counterbalanced conditions: i) in-person, side-by-side with the instructor at his or her workbench; in a remote physical space from the instructor that is mediated by ii) a teleconferencing platform or iii) a custom VR/AR system.

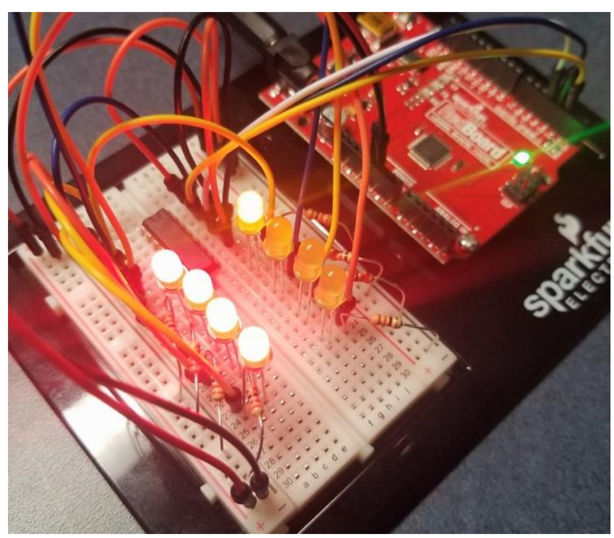

Figure 1: One of circuit boards used in the Circuit Study. Participants are trained in typical ways this circuit needs maintenance and then given a broken circuit to fix.

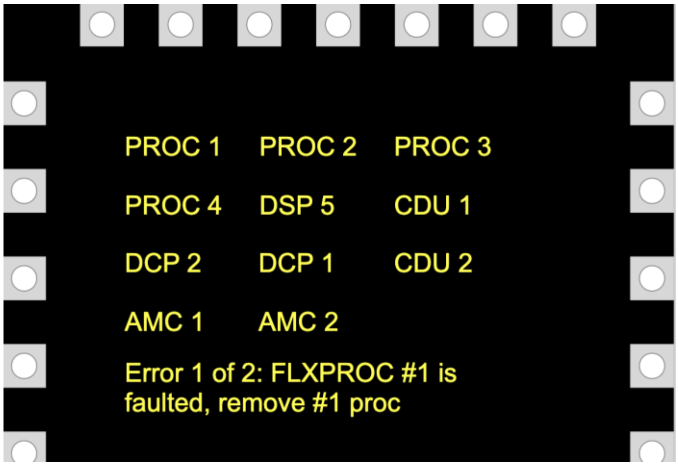

Figure 2: A screenshot of the tablet-based Diagnostic Tool modelled after an aerospace industry tool used by maintenance technicians. This tool is used for identifying errors and faults in the circuit.

Side by Side (Condition 1)

For one training trial, participant apprentices will be trained in the same room without the instructor's use of any communication medium. This is the only condition in which the instructor and participant apprentice will be co-located for training; all other conditions involve remote communication. Agency, embodiment, and media richness are assumed to be highest in this condition, given the training will be completed in-person.

Telecon Software (Condition 2)

The participant apprentice will receive training via commercial off-the-shelf teleconferencing software on a laptop for one trial. The apprentice will receive instruction through an audio stream and a webcam perspective of the physical workspace from the instructor's forehead. This condition has low media richness (and the apprentice has low agency) because of the inability to manipulate the perspective or move around in the instructor's work space. Both the instructor and the apprentice are partially represented by the teleconferencing medium on the laptop. Therefore this condition affords medium embodiment realism.

Extended Reality – XR (Condition 3)

The final training condition involves a custom VR/AR system (Figure 3). The remote apprentice will don an HTC Vive and the remote instructor will don a Microsoft HoloLens. The instructor's environment will be photographed with a 360° camera placed in the apprentice's virtual environment. The apprentice will have viewpoint control–they can choose between a workspace view that contains an avatar representing the instructor or a live video feed from the instructor's forehead camera. The instructor will also see an avatar representing the apprentice in the shared space. Eye gaze and gestural information will be conveyed by these avatars, and audio streams will be delivered through in-house speakers and microphones. Therefore, the custom VR/AR system allows for medium agency, medium to high embodiment, and high media richness.

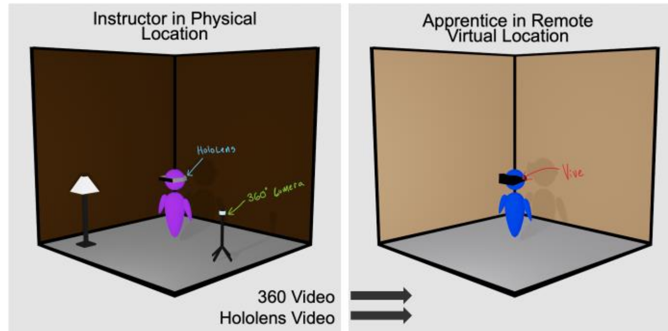

Figure 3: The instructor works in a real workshop captured by a 360° camera while wearing a Microsoft Hololens. The apprentice wears an HTC Vive and can choose to view video of the room with an Instructor avatar or the Instructor's headcam (the workspace). Audio flows in both directions.

HYPOTHESES & MEASURES

The ultimate goal of this research is to standardize benchmarks for communication and performance objectives when using telecommunications. In the Circuit Study, we will assess a subset of psychosocial constructs derived from the COM (Dianiska et al., under review). This study examines the effect of diverse training mediums (side-by-side, telecon, XR) on not only objective task performance, but also the communication metrics described by the COM to capture the overall user experience.

Table 2 describes the primary dependent variables in the current experiment. The COM measures of interest in the current experiment here include *trust, shared situation awareness, engagement, and rapport*. We will also examine the effect of training medium on measures of *mental workload*, *ease-of-use* ratings, and objective *task performance* metrics.

Table 2. Dependent measures categorized by acquisition stage

During Training (Synchrony)	Post-Training (Task Performance)	Post-Task Performance (Perception)
Facial EMG	Facial EMG	Trust (Mayer et al., 1999)
Autonomic nervous system (ANS) Arousal	Autonomic nervous system (ANS) Arousal	3D-SART (Selcon, et al. 1992)
	Number of errors	Rapport (Bernieri et al., 1994)
	Time on task	NASA-TLX (Hart & Staveland, 1988)
	Knowledge retention quiz	UTAUT (Venkatesh et al., 2003)

Research Design: Measurement of Physiological Arousal

For both the apprentice and the instructor, electrodes associated with a Biolab data acquisition system (Mindware, Ltd.) will be applied to the face in a validated configuration for facial electromyography (fEMG). Six small facial sensors (and one ground) will be applied to the participant's face to measure muscles associated with the expressions of anger (i.e., corrugator supercilii), happiness (i.e., zygomaticus major), and startle or surprise (i.e., orbicularis oculi). The leads for both apprentice and the instructor can continuously monitor in real-time facial muscle movements that serve as measures of emotional expression (e.g., Mahlke & Minge, 2006) as well as a proxy for mental workload (e.g., de Morree & Marcora, 2010). The facial EMG also offers a measure of engagement and rapport, with respect to emotional synchrony between the apprentice and instructor (Delaherche et al., 2012).

To complement fEMG, autonomic data is recorded to assess physiological arousal via heart rate (HR), heart rate variance (HRV) and galvanic skin conductance (GSC). These measures offer convergent evidence for how aroused or relaxed the apprentice is based on activity in the sympathetic and parasympathetic autonomic nervous system. Such autonomic arousal can also serve as a measure of mental effort (e.g., Nourbakhsh et al., 2012; Setz et al., 2009).

We define *engagement* as directing one's attention, acknowledging others, and demonstrating a readiness to interact with others (Peters et al., 2009). We will assess the extent to which there is synchrony in facial muscle movements (via fEMG) between the apprentice and the instructor, given that such synchrony is likely to be associated with shared emotions (Uzefovsky et al., 2016). When participant apprentices are afforded more agency, embodiment, and media richness, they should be better able to sustain attention and consequently show more engagement.

Research Design: Self-Report Measures

Trust. We leverage the definition of trust as a willingness to accept vulnerability by allocating time or other resources to the trustee to achieve an objective. After completing the trained task, participants will complete the self-report trust scale developed by Mayer and colleagues (1999). This scale describes the participant apprentice's perception of the instructor's ability, benevolence, and integrity. For each item, participants rate the extent to which they *strongly disagree* (1) to *strongly agree* (5) with each statement. We are expecting that as agency and realism of the avatar increases, perceptions of trustworthiness should also increase. First, the apprentice is able to control the viewing perspective and should therefore find it less likely that the instructor will cause distrust (see Marsh & Dibben, 2005). Second, in line with prior work, the realism of the avatar may increase trust by serving to increase social presence in avatar-mediated communication (e.g., Liew, et al., 2017; Pan & Steed, 2017).

Shared Situation Awareness. Shared situation awareness refers to participants' shared understanding of the environment (Endsley, 1988). Here we will assess the instructor and apprentice's shared situation awareness using the condensed 3D-SART (Situation Awareness Rating Technique; Selcon et al., 1992). This scale involves a global assessment of the demands on attentional resources, the supply of attentional resources, and an understanding of the situation from *low* (0) to *high* (100). We expect that shared situation awareness will increase with increased agency, given that the apprentice is given viewpoint control and can therefore use their time more efficiently to synchronize his or her understanding with the instructor. Further, we expect that shared situation awareness will increase to the extent that the instructor embodiment and media richness of the environment increases. Greater instructor embodiment should provide more cues to convey where one's attention is allocated, whereas higher media richness should reduce the need for verbalization of explicit descriptions.

Rapport. Rapport is often conceptualized as having three distinct components: mutual attentiveness between interaction partners, positivity in one's perception of another, as well as synchrony and coordination of the interaction between the individuals (Tickle-Degnen & Rosenthal, 1990). In addition to the facial EMG measures of synchrony, we will also assess participant apprentices' self-reported perceptions of rapport with the instructor using a brief survey derived from Bernieri and colleagues (1994). This scale has been condensed to only reflect positive perceptions of the interaction (e.g., harmonious, involving, worthwhile) from *not at all* (0) to *extremely* (8). We expect to see that perceived rapport increases as embodiment increases (Smith & Neff, 2018).

Cognitive Workload. The cognitive demands associated with a task may affect an individuals' performance (de Waard, 1986) given the variability in the number of mental operations required to perform the task (see Guhe et al., 2005). In the current study, participant apprentices will provide self-report ratings based on the NASA-Task Load Index (NASA-TLX; Hart & Staveland, 1988). This instrument assesses dimensions related to the task (mental, physical, and temporal demands), to the individual's behavior (performance), and to the individual (effort and frustration). Cognitive workload is expected to decline when the training environment affords the participant greater agency, as well as when the instructor embodiment allows for faithful representation of facial expressions and gestures (Steed et al., 2016). Further, environments with greater media richness should reduce workload given that a richer environment provides sufficient cues that preclude the need for participants to perform gap-filling functions, though only if the richer media cues are helpful and not distracting (Alexander et al., 2017; cf. Badger et al., 2014).

Unified Theory of Acceptance and Use of Technology. Given our interest in exploring the utility of a VR/AR system for remote training, we will also explore participant apprentices' perceptions and acceptance of each training medium to extrapolate the likelihood of adoption. In line with the Unified Theory of Acceptance and Use of Technology model (UTAUT; Venkatesh et al., 2003), we will assess participant apprentices' perceptions of the ease of use (a user's expectations of effortful use), usefulness (a user's expectations for performance), facilitating conditions (factors that make an action easy), and social influence (a user's perception that important others believe they must adopt a system) associated with each platform. For each item, individuals will rate how *unlikely/disagree* (1) to *likely/agree* (7). We expect that training mediums with greater agency, embodiment, and media richness should increase the perceived ease of use and perceived usefulness.

SUMMARY

This experimental design offers a promising step in demonstrating the integration of research-based behavioral communication measures with a task and experimental design that can evaluate how the affordances of a remote communications system affect training performance. Future work will include data collection for this study, as well as for other scenarios, including usage of similar software with actual maintenance technicians in the field.

ACKNOWLEDGMENTS

This work is part of a project called Systematic Communication Objectives and Telecommunications Technology Investigations and Evaluations (SCOTTIE). SCOTTIE is jointly funded by the U.S. Department of Energy ARPA-E, Raytheon Technologies Research Center, and Iowa State University. We would like to thank Rachel Slaybough and Rusty Heffner for their continued guidance, as well as Kirk Thorson and Barry Cheesman from Collins Aerospace for their continued contributions to the experiment design and scenario development.

REFERENCES

Alexander, T., Ripkens, A., Westhoven, M., Kleiber, M., & Pfendler, C. (2017, July). Virtual tele-cooperation: applying AR and VR for cooperative tele-maintenance and advanced distance learning. In *International Conference on Applied Human Factors and Ergonomics* (pp. 234-244). Springer, Cham.

Badger, J. M., Kaminsky, S. E., & Behrend, T. S. (2014). Media richness and information acquisition in internet recruitment. *Journal of Managerial Psychology*.

Bernieri, F. J., Davis, J. M., Rosenthal, R., & Knee, C. R. (1994). Interactional synchrony and rapport: Measuring synchrony in displays devoid of sound and facial affect. *Personality and Social Psychology Bulletin*, *20*(3), 303-311.

de Morree, H. M., & Marcora, S. M. (2010). The face of effort: frowning muscle activity reflects effort during a physical task. *Biological Psychology*, *85*(3), 377-382.

de Waard, D., & te Groningen, R. (1996). *The measurement of drivers' mental workload*. Netherlands: Groningen University, Traffic Research Center.

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, *3*(3), 349-365.

Dianiska, R.E., Wu, P., Oldham, E., Stonewall, J., Monaghan, Q., Peasley, C., Gilbert, S., Oliver, J., & Israelsen, B. (Under review). *A taxonomy of face-to-face communication objectives to inform telepresence technology adoption*.

Endsley, M. R. (1988, May). Situation awareness global assessment technique (SAGAT). In *Proceedings of the IEEE 1988 national aerospace and electronics conference* (pp. 789-795). IEEE.

Ens, B., Lanir, J., Tang, A., Bateman, S., Lee, G., Piumsomboon, T., & Billinghurst, M. (2019). Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*.

Guhe, M., Gray, W. D., Schoelles, M. J., Liao, W., Zhu, Z., & Ji, Q. (2005, September). Non-intrusive measurement of workload in real-time. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 49, No. 12, pp. 1157-1161). Sage CA: Los Angeles, CA: SAGE Publications.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139–183.

KLM, 2019. Retrieved from https://klmtakescare.com/en/content/co2zero-s-facts-and-figures

Liew, T. W., Tan, S. M., & Ismail, H. (2017). Exploring the effects of a non-interactive talking avatar on social presence, credibility, trust, and patronage intention in an e-commerce website. *Human-centric Computing and Information Sciences*, *7*(1), 42.

Mahlke, S., & Minge, M. (2006). Emotions and EMG measures of facial muscles in interactive contexts. *Cognition and Emotion*, *6*, 169-200.

Marsh, S., & Dibben, M. R. (2005). Trust, untrust, distrust and mistrust—An exploration of the dark(er) side. In P. Herrmann, V. Issarny, & S. Shiu (Eds.), *Trust management: Third International Conference, iTrust Proceedings* (pp. 17-33). New York: Springer.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, *20*(3), 709–734.

Mokhtarian, P. (2009). If telecommunication is such a good substitute for travel, why does congestion continue to get worse?. *Transportation Letters*, *1*(1), 1-17.

Nourbakhsh, N., Wang, Y., Chen, F., & Calvo, R. A. (2012, November). Using galvanic skin response for cognitive load measurement in arithmetic and reading tasks. In *Proceedings of the 24th Australian Computer-Human Interaction Conference* (pp. 420-423).

Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI*, *5*, 114.

Pan, Y., & Steed, A. (2017). The impact of self-avatars on trust and collaboration in shared virtual environments. *PloS one*, *12*(12): e0189078.

Peters, C., Castellano, G., & de Freitas, S. (2009, November). An exploration of user engagement in HCI. In *Proceedings of the International Workshop on Affective-Aware Virtual Agents and Social Robots* (p. 9). ACM.

Selcon, S. J., Taylor, R. M., & Shadrake, R. A. (1992, October). Multi-modal cockpit warnings: Pictures. words. or both?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 36, No. 1, pp. 57-61). Sage CA: Los Angeles, CA: SAGE Publications.

Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., & Ehlert, U. (2009). Discriminating stress from cognitive load using a wearable EDA device. *IEEE Transactions on information technology in biomedicine*, *14*(2), 410-417.

Smith, H. J., & Neff, M. (2018, April). Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).

Steed, A., Pan, Y., Zisch, F., & Steptoe, W. (2016, March). The impact of a self-avatar on cognitive load in immersive virtual reality. In *2016 IEEE Virtual Reality (VR)* (pp. 67-76). IEEE.

Tang, A., Pahud, M., Inkpen, K., Benko, H., Tang, J. C., & Buxton, B. (2010, February). Three's company: understanding communication channels in three-way distributed collaboration. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 271-280). ACM.

Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, *1*(4), 285–293.

Transportation Energy Data Book: Edition 33-2014

U.S. Energy Information Administration (EI). *Annual energy outlook 2016 with Projections to 2040* (Report No. DOE/EIA-0383). Retrieved from https://doi.org/10.2172/1329373

Uzefovsky, F., & Knafo-Noam, A. (2016). Empathy development throughout the life span. In J. Sommerville & J. Decety (Eds.) *Social Cognition* (pp. 89-115). Routledge.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 425-478.

Wu, P., Morie, J., Wall, P., Ott, T., & Binsted, K. (2016). ANSIBLE: Virtual reality for behavioral health. *Procedia engineering*, *159*, 108-111.