

Trusting and Learning From Virtual Humans that Correct Common Misconceptions

Journal of Educational Computing Research

2023, Vol. 0(0) 1–27

© The Author(s) 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/07356331221139859

journals.sagepub.com/home/jec



Noah L. Schroeder¹ , Erin K. Chiou², Robert F. Siegle², and Scotty D. Craig² 

Abstract

Virtual humans are on-screen characters that are often embedded in learning technologies to deliver educational content. Little research has investigated how virtual humans can be used to correct common misconceptions. In this study, we explored how different types of narrative structures, refutation text and expository text, influence perceptions of trust, credibility, and learning outcomes. In addition, we conducted exploratory analyses examining how different measures of trust and credibility are related to each other and how these measures may mediate learning outcomes. Results showed that the type of narrative used did not influence any measure. However, the trust and credibility measures, while related to one another, were measurably distinct. In addition, only perceptions of message trust were significantly related to learning. Perceptions of message trust did not mediate learning outcomes, but were significantly predictive of learning at nearly the same effect as prior knowledge.

Keywords

trust, pedagogical agent, virtual human, explainability, refutation, credibility, learning

¹Department of Leadership Studies in Education and Organizations, Wright State University, Dayton, OH, USA

²Human Systems Engineering, Ira A. Fulton Schools of Engineering, Arizona State University, Mesa, AZ, USA

Corresponding Author:

Scotty D. Craig, Arizona State University, Human Systems Engineering, Ira A. Fulton Schools of Engineering, 7271 E. Sonoran Arroyo Mall, Mesa, AZ 85212, USA.

Email: scotty.craig@asu.edu

Introduction

In today's world we are faced with countless sources of information. Some of this information is scientifically accurate, yet some of it is not. It is therefore critical that we can evaluate what information is trustworthy and credible. Researchers have examined how beliefs in misinformation are formed and why they resist correction (Ecker et al., 2022; Lewandowsky et al., 2012), and they have also examined how constructs such as source credibility can influence how we learn (Van Boekel et al., 2017). Yet, these types of questions are not commonly explored in relation to learning technologies.

Educators and learning technologists alike strive to provide learners with trustworthy and credible sources of scientifically accurate information. One educational technology that is increasing in popularity are virtual humans. Schroeder and Craig (2021) defined virtual humans as "embodied agents with a human-like appearance" (p. 1) and noted how they are often used in learning technologies. While they highlighted the more than 20-year history of integrating virtual characters in learning technologies, gaps remain in our understanding of how to design virtual humans to be effective learning tools.

Many aspects of virtual human design and implementation have been well-investigated, but there is little known about what influences a learner's trust in, and perceived credibility of, the virtual human. In addition, it is not known how different types of verbal messaging may influence the learner's trust in, and perceived credibility of, the information presented by the virtual human. Finally, there is little work around how the concepts of trust and credibility influence learning outcomes. As such, this study explores how the design of the verbally delivered narrative influences learners' perceptions of trust, credibility, and learning outcomes. We also explore the relationships between these variables.

Literature Review

Designing Effective Virtual Humans to Aid Learning

Research has widely shown that virtual humans can be effective learning aids in some situations and for some populations (Craig & Schroeder, 2018; Schroeder et al., 2013; Castro-Alonso et al., 2021), although this is not always the case (Heidig & Clarebout, 2011). Meta-analyses have shown that their effects can vary significantly depending on the specific use case or design (Schroeder et al., 2013; Castro-Alonso et al., 2021). As such, researchers have been examining different aspects of virtual human design to determine which components of design are essential, and which are not as important. Much of the work in the area has focused on what Heidig and Clarebout's (2011) framework termed as the medium level, focusing largely on technical decisions. For example, there has been an increasing amount of research around the voice used by the virtual human (Craig & Schroeder, 2017; Baylor et al., 2003; Beege et al., 2020;

Davis et al., 2019), and recent work has begun investigating virtual human emotional expression (Lawson et al., 2021a; 2021b; 2021c).

While these areas of research often focus on learning outcomes and are relevant to virtual human design, there is less work around examining features of the virtual human that influence constructs such as trust and credibility. There is a long history and recent theoretical developments in trust scholarship for understanding what factors of information technology can influence others in their decision-making, or in achieving particular human-technology performance outcomes (Chiou & Lee, 2021; Muir, 1987). There is also a long history of credibility scholarship when it comes to understanding how to influence others, with a clearer distinction between how sources compared to their messages influence others (McCroskey & Young, 1981). Yet, this type of work could be critical for understanding and designing virtual humans because if the virtual human is not perceived as trustworthy and credible, the learner could disregard or ignore the information being taught. In the next section, we explore the concepts of trust and credibility, with an emphasis on research around the use of artificial intelligence and virtual humans.

Trust in Virtual Humans

Trust has long been a topic of interest among technology scholars, who have proposed various definitions based on our understanding of trust in society and honing it for the unique contexts of trusting technology in more industrial settings (Sheridan, 1975; Lee & See, 2004). One robust definition is that trust is the attitude that an entity will help achieve an individual's goals under conditions of uncertainty and vulnerability (Lee & See, 2004). Defining trust as an *attitude* distinguishes it from a belief, intent, or behavior, and thus avoids confusing its effect with the effects of other factors that can also affect behaviors such as values, self-confidence, or workload. As of this writing, a search on Google Scholar for "trust in technology" returns roughly 3.78M results, suggesting sustained scholarly interest in this topic. A common purpose underlying some of this work is to use trust as a lens for understanding what affects people's decisions or performance with technology including its appropriate use or adoption, and reasons for disuse or misuse. Improved understanding not only lends deeper insights into the human condition and our relationships to technology, but also informs the design of technology in our everyday lives.

Importantly, trust is distinguished from trustworthiness in this line of research, with the former referring to an attitude (which is usually assessed based on a combination of human perception and behavior) and the latter referring to qualities of the technology that can be observed or tested (e.g., reliability, responsiveness), and to qualities of the system in which that technology is embedded (e.g., the existence of consumer protection agencies, third-party verification), that would merit a person's trust in that technology (Chiou & Lee, 2021; Shneiderman, 2020). This distinction between trust and trustworthiness is important to avoid confusing a person's trust in technology with the separate but overlapping judgment that technology should be trusted.

More recently, researchers have used trust to understand the effects of virtual humans on learning outcomes (Chiou et al., 2020). In doing so, different streams of research on technology perceptions converged – streams that span application domains (from work settings to education settings) and disciplines (from engineering psychology to education technology). The insight from these converging streams is that trust and credibility may be related insofar as trustworthiness determines trust, and insofar as trustworthiness and credibility are equivalent factors when it comes to perceptions of learning technologies and their effects on learning outcomes (Schroeder et al., 2021).

Credibility and Trust

Whereas much of the trust in technology scholarship geared toward improving system design was developed in and for industrialized settings (Lee & Moray, 1992; Muir, 1987; Sheridan, 1975), credibility has a longer history in the communication literature and more recently extending into human-computer interaction (Fogg & Tseng, 1999; McCroskey & Young, 1981; Carter & Greenberg, 1965). Historically, communication and human-computer interaction have focused more on the study of information content, delivery, and source, and their effects on related outcomes such as persuasion, learning, and adoption. Similar to how trust in decision support systems has been shown to mediate performance outcomes in work settings (Madhavan & Wiegmann, 2007), credibility has been shown to mediate learning outcomes in educational or instructional settings (Finn et al., 2009).

Much of the research using credibility to explain performance outcomes has taken place in educational settings. For example, in a study comparing a teacher operating a telepresence robot with an autonomous social robot teacher, researchers found that differences in source credibility ratings mediated self-reported behavioral and affective learning outcomes (Edwards et al., 2016). Source credibility was also used to investigate how the type of subject taught and physical characteristics of a virtual teacher affected learning outcomes, finding that credibility of the virtual teacher seemed important for gaining students' interest in a subject (Matsui & Yamada, 2019). From this research on credibility in instructional settings, "trustworthiness" is more often described as a sub-dimension of credibility, stemming from Aristotle's concept of *ethos* in persuasive speech, or *source credibility*, which is said to be shaped by a speaker's intelligence, character, and good will (McCroskey & Young, 1981).

Source credibility has been treated as a central factor in persuasion because no message is received independently from its presumed source (McCroskey & Teven, 1999). However, later research has shown that although there is a strong relationship between source credibility and message credibility (Pornpitakpan, 2004), there may be a need to distinguish between source credibility and message credibility in measurement (McCroskey et al., 2004; Appelman & Sundar, 2016). Furthermore, there may be fundamental differences between in-person instruction and instruction enacted

by a virtual human when it comes to their relational and rhetorical qualities that affect learning (Ramlatchan & Watson, 2020; Hollan & Stornetta, 1992).

Similar to how trust in technology has been studied, credibility has been an outcome of interest following the implementation of a particular design intervention (e.g., Ramlatchan & Watson, 2020), but few experimental studies directly address the overlap or gaps between trust and credibility. From an information processing perspective, trust and credibility seem to stem from the same understanding of informational factors that influence people's perceptions and behaviors. Similar to trust, credibility has been defined as an attitude toward another entity that is shaped by multiple dimensions of information, including perceptions of the entity's performance (e.g., intelligence, expertness; ability), process (e.g., character, trustworthiness; integrity), and purpose (e.g., good will, intention; benevolence). Performance, process, and purpose are concepts from the trust in technology literature (Lee & See, 2004) that loosely parallel the intelligence, character, and goodwill concepts of credibility from the communication literature (McCroskey & Young, 1981); as well as the ability, integrity, and benevolence concepts of interpersonal trust from the organizational management literature (Mayer et al., 1995). From this brief review, it seems that when it comes to studying the types of information that inform attitudes toward another entity, credibility and trust are overlapping concepts.

However, when comparing how the two constructs have been used in the literature, trust seems to encompass interactions more broadly, because it has also been described as an evolving relationship (Lewicki et al., 2006; Mayer, et al., 1995; Rempel & Holmes, 1985) whereas credibility has more often focused on the features of an entity, its messaging (content), and its means of messaging (media). In other words, trust and its associated outcomes are not just determined by perceived information, but trust can also be shaped by task-related signals, behaviors, and common goals (Chiou & Lee, 2021). Whereas source credibility and message credibility have been studied separately, drawing from *either* relational factors *or* rhetorical factors (McCroskey et al., 2004) or with one as a subset of the other (Pornpitakpan, 2004), integrative models of trust account for *both* rhetorical and relational factors in the same model (Chiou & Lee, 2021; Mayer et al., 1995). Although many open questions remain about the relationship between trust and credibility, it seems that agentic educational technology such as a pedagogical virtual human could serve as an initial bridge between the two streams of research. A goal is to clarify how trust and credibility in technology may be related, and to what extent either construct would be useful for predicting human-technology performance outcomes. For this study, that would be learning performance.

Exploring the Design of Virtual Humans to Facilitate Learning based on Trust and Credibility

Based on the aforementioned research, critical questions arise, such as how do trust and credibility with respect to the virtual human influence learning? Accordingly, we will first ask how does the design of virtual humans influence trust and credibility?

Though there are myriad ways to present a virtual human, one understudied area is how the character speaks to the learner, or what we will refer to as the *narrative structure* of the instructional content being delivered. Outside of intelligent tutoring systems or conversational agents that allow for interactive dialog, it is quite common for virtual human-based learning environments to teach learners through an expository narrative, such as in an instructional video or slideshow. For example, [Yung and Paas \(2015\)](#) created an instructional animation that taught students about the cardiovascular system, [Park \(2015\)](#) created a multimedia lesson around intellectual property, and [Davis and Vincent \(2019\)](#) created a video about lightning formation.

[Schroeder et al. \(2022\)](#) recently asked if structures that have been successful for promoting learning in text format retain their effectiveness when transitioned into a narrative format. Specifically, they examined a text structure known as refutation text (RT). Refutation texts are a specific text structure that state a commonly held misconception, refute it, and then provide the scientifically-accepted understanding of the phenomenon ([Hynd, 2001](#); [Mason et al., 2019](#); [Schroeder & Kucera, 2022](#); [Tippett, 2010](#)). A recent meta-analysis found that RT presented in text format provided superior learning gains compared to every other text structure it was compared to ([Schroeder & Kucera, 2022](#)). Not only was RT more effective than other text structures, but it was also broadly effective across various test types, test timings, and subject areas.

Why the Refutation Text Structure as Narrative Holds Promise. While empirical evidence shows that RT can be effective, it is not clearly agreed upon as to why it is effective. There are a number of theoretical reasons why RT may aid learning. Many of these theories are focused on the notion of conceptual change, and there are a wide variety of views on conceptual change (see [Duit & Treagust, 2003](#); [Treagust & Duit, 2008](#)). For the purposes of this study, we focus on the Knowledge Revision Components (KReC) Framework ([Kendeou et al., 2014; 2019; Will et al., 2019](#)). KReC focuses on how we revise the knowledge we hold in the long-term memory, a process which we will refer to as learning. Based on the fundamental assumptions that information held in the long-term memory can never be erased and that reading can activate this information ([Kendeou et al., 2014; 2019; Will et al., 2019](#)), KReC suggests that for learning to occur, the inaccurate prior knowledge must be activated at the same time as the correct conception, the correct conception must replace the incorrect information in the learners' existing knowledge, and this newly formed, correct understanding be activated in the future rather than the pre-existing, incorrect understanding ([Kendeou et al., 2014; 2019; Kim & Kendeou, 2021](#)). [Kim and Kendeou \(2021\)](#) suggest that, "During reading of refutation texts, this is accomplished by providing readers with causal explanations of the correct idea. This causal information, in conjunction with the correct idea, draws activation to itself and away from the previously acquired incorrect knowledge." (p. 2). Importantly, while the KReC framework is cognition-oriented, it can accommodate how attitudes, emotions, perceptions, and other affective factors may influence the learning process ([Schroeder & Kucera, 2022; Kim & Kendeou, 2021; Trevors et al., 2017](#)).

A key question then, is does the RT structure facilitate learning or learners' perceptions when it is used in verbal narrative rather than in written text? Research in this area is limited and has shown inconsistent results. Muller et al.'s (2008) work showed that a RT narration facilitated learning better than an expository narration. In contrast, Schroeder et al. (2022) compared a RT narration to three other text structure verbal narratives and found largely no significant differences in learning outcomes, emotions, attitudes, or learners' perceptions of the learning experience between groups. Given the limited scope of work in this area and the conflicting findings, more research is needed to better understand the extent to which RT narration influences learners' learning and perceptions.

The Present Study

While a growing body of literature exists around the topics of virtual humans, trust in technologies, and perceived credibility, there is a very limited pool of studies that brings these concepts together to investigate how they may influence learning outcomes. As such, the purpose of this study is to build on our existing understanding of how to design effective instructional messaging delivered by virtual humans.

In this study, we have two parts. Part one examines how the structure of a verbal narrative influences learners' perceptions of trust, perceptions of credibility, perceptions of the virtual human, and learning outcomes. Part two is more exploratory in nature, and examines how perceptions of credibility and trust may relate to one another and mediate learning outcomes.

More specifically, part one of our study examines the differences between learning from an expository narrative and learning from a RT narrative. The research questions for this part of the study include:

To what extent does the use of a narrative with a refutation text structure influence:

RQ1 – the learners' perceptions of trust in the virtual human?

RQ2 – the learners' perceptions of trust in the message?

RQ3 – the learners' perceptions of credibility of the virtual human?

RQ4 – the learners' perceptions of credibility of the message?

RQ5 – learning compared to standard expository narrative?

In addition, we investigate the question:

RQ6 – To what extent does the use of refutation text influence participants' perceptions of the virtual human?

In part two of the study, we sought to conduct exploratory analyses to gain a better understanding of the empirical relationships between trust and credibility, and how these perceptions may mediate learning outcomes. Our research questions for this part of the study include:

RQ7 – To what extent are measures of credibility and trust in the message and in the virtual human related, and to what extent are they related to prior knowledge and learning outcomes?

RQ8 – To what extent do the perceptions of trust or credibility mediate the relationship between prior knowledge and learning outcomes?

Methods

Design

A randomized post-test design with a pre-test was implemented to investigate the impact narrative variations have on trust, credibility, and learning. The narration structure of the virtual human's presentation was modified to create expository and refutation condition. Trust was measured using Jian et al.'s trust scale (Jian et al., 2000). Credibility was measured using Appelman and Sundar's (2016) credibility scale. Participants knowledge was measured by the knowledge of genetically modified food measure (Trevors et al., 2017).

Participants

Ninety-two undergraduate students (including 73 male participants and 19 female participants; $M_{\text{age}} = 18.54$ years, $SD_{\text{age}} = 1.50$ years) from a large, research intensive university in the western United States volunteered to participate in this online study, completed the study, correctly answered an attention check item, and did not spend an excessive amount of time completing the study (defined as > 3600 seconds, which was twice the allotted time to complete the study). The average amount of time it took to complete the study was 19 minutes ($SD = 7$ minutes). Participants were recruited from an undergraduate subject pool populated by students taking an introductory class on human factors related topics. The study provided credit toward a class requirement for participating in experiments.

Materials

Design of the Virtual Human. We created a virtual human using Adobe Fuse (Beta) and Mixamo from Adobe Systems Inc., and embedded the character in an instructional video. The virtual human's voice was narrated by an American woman, and we used SALSA LipSync v2 suite from Crazy Minnow Studio to synchronize the virtual human's mouth to the recordings. The virtual human-based instructional video is shown in Figure 1.

Learning Materials. The instructional video (approximately 5 minutes long) featured a female virtual human (Figure 1) and focused on misconceptions about genetically modified foods (GMF). Adapted from Heddy et al. (2017) and previously used by



Figure 1. A screenshot of the virtual human-based learning environment.

[Schroeder et al. \(2022\)](#), the narrative addressed four common misconceptions: that GMF involves cloning, that GMF involves injecting hormones into organisms, that GMF only occurs through laboratory science, and that GMF is the result of new scientific research.

The videos were visually the same between conditions, with the only difference being the narrative by condition. The expository narrative contained 589 words and the RT narrative contained 622 words. Specifically, the expository narrative explained the concepts in a scientific way without directly highlighting the misconception as a misconception (e.g., “Genetically modifying foods is not the same process as cloning”). This differs from the RT condition, which highlighted the misconception, directly refuted it, and provided the scientifically-accepted explanation (e.g., “You may think that genetically modifying foods is the same process as cloning. This belief is not correct. Cloning involves...”). See [Heddy et al. \(2017\)](#) and [Schroeder et al. \(2022\)](#) for the specific texts used as the narrative.

Demographic Questionnaire. The demographic questionnaire was brief, asking participants their age and gender.

Knowledge Assessment. [Heddy et al.’s \(2017\)](#) multiple-choice knowledge assessment, validated in previous work ([Trevors et al., 2017](#)), was used as both a pre-test ($\alpha = 0.56$) and post-test ($\alpha = 0.74$) in this study. We used the sum of the participant’s correct answers as their score on these tests.

Credibility Measures. Participants' perceptions of message credibility were measured using a three-item questionnaire rated on a 7-point scale. Participants were asked to answer, *How well do the following adjectives describe the content you just heard?* (From 1 = describes very poorly to 7 = describes very well). The three items address the following qualities: accurate, authentic, and believable. This instrument was developed and validated in Appelman and Sundar's (2016) study that aimed to address the absence of an instrument exclusively measuring message credibility. The internal consistency reliability of this measure was $\alpha = 0.75$.

This same instrument was administered a second time, but slightly adapted to measure the participants' perceptions of credibility in the virtual human. Specifically, the question stem read as *How well do the following adjectives describe the virtual character that you just learned from?* (From 1 = describes very poorly to 7 = describes very well). Otherwise, the three items and 7 point scale were the same as the administered questionnaire for measuring message credibility. The internal consistency reliability of this measure was $\alpha = 0.78$.

Trust Perception Measures. Participants' trust in the virtual human's message was measured using a 12-item questionnaire. The question order was randomized for each participant, according to recommended best practices (Gutzwiler et al., 2019). These items were rated on a 7-point scale, from 1 = not at all to 7 = extremely. This 12-item instrument was developed through word elicitation and factor analysis by Jian et al. (2000) to measure people's general trust perceptions of automation. This instrument has been independently validated for measuring trust in automation (Spain et al., 2008) and adapted for understanding trust in specific types of automated technologies, including previous studies with learning technology (Chiou et al., 2020). The current study adapted the scale to measure trust in the message, i.e., the information presented by the virtual human, or the script. The questions were adapted by replacing the word original word *system* ("I am confident in the system") with the word *information* ("I am confident in the information"). The internal consistency reliability of this measure was $\alpha = 0.84$.

This same questionnaire was administered a second time but adapted to measure participants' perceptions of trust in the virtual human. The questions were adapted by replacing the word original word *system* ("I am confident in the system") with the word *instructor* ("I am confident in the instructor"). The internal consistency reliability of this measure was $\alpha = 0.87$.

Agent Persona Instrument-Revised. We measured participants' perceptions of the virtual human using the agent persona instrument-revised (API-R; Schroeder et al., 2017; 2018). This measure is a modified version of the instrument developed by Ryu and Baylor (2005), and consists of twenty-five 5-point Likert items ranging from 1 – "Strongly Disagree," to 5 – "Strongly Agree." These questions were separated into four subscales: *facilitates learning* (10 items, $\alpha = 0.92$), *credibility* (5 items, $\alpha = 0.82$),

human-like (5 items, $\alpha = 0.90$), and *engagement* (5 items, $\alpha = 0.87$). For each subscale of the API-R, we calculated an average of the participant's ratings.

In order to ensure the participant was fully reading each item on the API-R, we included an attention check measure: *To ensure you are paying attention please answer this with "Agree."* Participants who did not answer "agree" to this question were removed from the dataset.

Data Analysis

In part one, a series of ANOVAs were conducted on the learning, trust, and credibility measures. The first five items on the trust questionnaire (as initially ordered in [Jian et al., 2000](#)) were reverse-coded before averaging the responses for a final score. The three items from the credibility scale were averaged for a final score. In part two, we conduct correlational analyses and a mediation analysis.

Procedure

This was a fully online study, where participants first consented to participate, and then completed an audio check (to ensure their computer volume was appropriately set), completed the demographic questions, and then completed the pre-test. Next, participants were randomly assigned to view one of the two video conditions, after which they completed the message credibility questionnaire, the trust in the message questionnaire, the agent credibility questionnaire, and the trust in the agent questionnaire. Next, they completed the post-test and then completed the API-R. The procedure is visualized as a flowchart in [Figure 2](#).

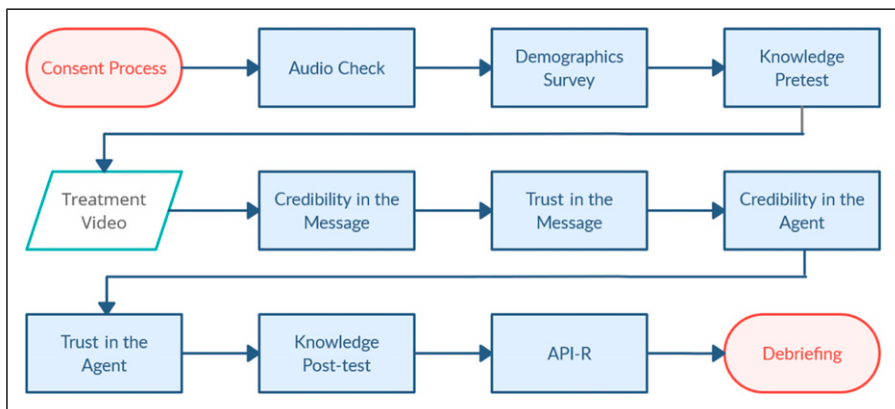


Figure 2. Flowchart of the procedure.

Part I Results and Discussion

Part One – Examining Differences Between Conditions

RQ1 – To What Extent does the Use of a Narrative with a Refutation Text Structure Influence Participants’ Perceptions of Trust in the Virtual Human? An ANOVA was conducted on participants’ trust data to determine differences in trust of the virtual human based on use of RT narrative or an expository narrative. The participants’ trust in the virtual human was not significantly different between conditions, $F(1, 91) = 0.93, p = .34, \eta_p^2 = 0.01$. The descriptive statistics are provided in [Table 1](#).

RQ2 – To What Extent Does the Use of a Narrative with a Refutation Text Structure Influence Participants’ Perceptions of Trust in the Message? An ANOVA was conducted on participants’ trust data to determine differences in trust in the message based on the use of RT narrative or an expository narrative. The participants’ trust in the message was not significantly different between conditions, $F(1, 91) = 0.25, p = 0.62, \eta_p^2 < 0.01$. The descriptive statistics are provided in [Table 1](#).

RQ3 – To What Extent Does the Use of a Narrative with a Refutation Text Structure Influence Participants’ Perceptions of Credibility of the Virtual Human? An ANOVA was conducted on participants’ credibility data to determine differences in the credibility of the virtual human based on the use of RT narrative or an expository narrative. The participants’ perceived credibility of the virtual human was not significantly different between conditions, $F(1, 91) = 0.26, p = .61, \eta_p^2 < 0.01$. The descriptive statistics are provided in [Table 1](#).

RQ4 – To What Extent Does the Use of a Narrative with a Refutation Text Structure Influence Participants’ Perceptions of Credibility of the Message? An ANOVA was conducted on participants’ credibility data to determine differences in credibility in the message based on the use of RT narrative or an expository narrative. The participants’ perceived credibility of the message was not significantly different between conditions, $F(1, 91) = 0.80, p = .37, \eta_p^2 = 0.01$. The descriptive statistics are provided in [Table 1](#).

Table 1. Descriptive statistics for trust, credibility, and learning test by condition.

Variable	Refutation Text			Expository Text		
	Mean	SD	N	Mean	SD	N
Trust in the virtual human	4.75	0.96	47	4.56	0.96	45
Trust in the message	5.29	0.89	47	5.21	0.78	45
Credibility of the virtual human	4.65	1.42	47	4.50	1.51	45
Credibility of the message	5.93	0.90	47	5.75	1.04	45
Pre-test score	4.68	2.01	47	4.62	1.82	45
Post-test score	6.51	2.18	47	6.87	2.31	45

RQ5 – To What Extent Does the Use of a Narrative with a Refutation Text Structure Influence Learning Compared to the Standard Expository Narrative? A mixed ANOVA was conducted on participants’ learning data to determine differences in learning over time (pre-to post-testing), based on the use of RT narrative or an expository narrative. The participants’ learning over time was overall significantly different between pre-test and post-test, $F(1, 90) = 93.90, p < .01, \eta_p^2 = 0.51$. The participants’ pre-test data were not significant between conditions, $F(1, 91) = 0.01, p = .97, \eta_p^2 < 0.01$, the post-test data were also not significant between conditions, $F(1, 91) = 0.58, p = .45, \eta_p^2 = 0.01$. The descriptive statistics are provided in [Table 1](#).

RQ6 – To What Extent does the Use of a Narrative with a Refutation Text Structure Influence Participants’ Perceptions of the Virtual Human? A separate ANOVA was conducted on each of the four subscales of the API-R to see if learners’ perceptions of the virtual human differed by condition. The participants’ perception of the learning experience was not significantly different between conditions on any of the subscales: facilitates learning $F(1, 91) = 2.37, p = .13, \eta_p^2 = 0.03$; credibility $F(1, 91) = 0.00, p = 1.00, \eta_p^2 = 0.00$; human-like $F(1, 91) = 3.02, p = .09, \eta_p^2 = 0.03$; engagement $F(1, 91) = 2.99, p = .09, \eta_p^2 = 0.03$. The descriptive statistics are provided in [Table 2](#).

Discussion. Based on the results of a recent meta-analysis which showed that RT broadly facilitated learning compared to all other text formats it was compared to ([Schroeder & Kucera, 2022](#)), we questioned if structuring a verbal narrative with a RT format would influence learning outcomes and perceptions of trust, credibility, and the virtual human itself. Across all measures, we found no significant differences between groups.

The lack of significant differences in learning outcomes between groups is consistent with the work of [Schroeder et al. \(2022\)](#) that used the same instructional materials. Taken together, these findings raise an important question – does the RT structure lose its efficacy when transitioned to verbal narration? Similar to [Schroeder et al. \(2022\)](#), we posit that the lack of differences between groups could be due to two factors: the specific narrative used, and the transience of information in videos. On the specific narrative used, there are only four instances of the RT structure present. When this factor is combined with the transience of the information presented in the video, it may be that this is not enough to grab the learner’s attention and activate the learner’s

Table 2. Agent Persona Inventory – Revised Subscales.

Condition	Facilitates Learning			Credibility		Human-Like		Engagement	
	N	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Refutation text	47	3.22	0.71	3.61	0.63	2.18	0.81	2.39	0.73
Exposition text	45	2.97	0.84	3.61	0.72	1.86	0.93	2.10	0.87
Total	92	3.10	0.79	3.61	0.68	2.02	0.88	2.25	0.81

knowledge revision process more deeply than a simple expository narration. From the transient information effect, we know that information which exceeds the learner's working memory capacity should be presented as text if it cannot be segmented (Leahy & Sweller, 2011; Singh et al., 2012; Sweller et al., 2011; Wong et al., 2012). Yet, this video (being the same as two of the conditions in Schroeder et al. (2022)) was not segmented.

Recall that the KReC framework suggests that reading information can activate the learner's prior knowledge and learning occurs when the new information is activated more readily than the previous incorrect information (Kendeou et al., 2014; 2019; Kim & Kendeou, 2021). If we extrapolate the task of reading text to the case of hearing text while watching a video, our results show that knowledge revision did occur, however the RT structured narration offered no learning advantage compared to expository narration. It seems reasonable that the transience of information along with the limited number of refutations did not lead to the learner to revise their knowledge any more than those who learned from the expository narration. The key questions then are, does the RT structure lose all advantages when used to structure a verbal narrative, is the narrative we used simply not providing enough direct refutations to gain the benefits seen in RTs, or is the non-segmentation of the video playing a notable role in the findings? Further research is needed to disentangle these issues, but since this study replicated the same lack of significant differences in learning outcomes between conditions from Schroeder et al. (2022), it appears that the use of RT in this particular narrative may not retain its superiority to expository narrative when transitioned to a verbal narration, despite it being shown to work when provided in written form (Heddy et al., 2017).

We also explored the effects of different narrative structures on learners' perceptions of the credibility of the message, credibility of the virtual human, trust in the virtual human, and trust in the message. We hypothesized that due to the structure of RTs, which states a commonly held misconception, refutes it, and then provides the scientifically-accepted understanding of the phenomenon (Hynd, 2001; Mason et al., 2019; Schroeder & Kucera, 2022; Tippet, 2010), narration in this format may influence learners' perceptions of trust and credibility. However, this was not the case as we found no significant differences between the groups. In fact, examining the mean scores between groups we see remarkable similarities. As we noted about learning outcomes, we suspect this may be a result of both dosage and the transience of information, however given that there is almost no research in this area, more research is clearly needed.

Finally, we explored if the style of narration would have any influence on learners' perceptions of the virtual human using the API-R. Here again, we found no significant differences between groups on any measure. These results are fully consistent with Schroeder et al. (2022), and imply that at least in the case of this particular narrative, the use of a RT structure does not seem to influence learners' perceptions of the virtual human on API-R scales.

Part Two – Exploratory Analysis Results and Discussion

RQ7 – To What Extent are Measures of Credibility and Trust in the Message and in the Virtual Human Related, and to What Extent are they Related to Prior Knowledge and Learning Outcomes? A series of Pearson correlations were conducted on participants' credibility and trust responses, as well as their prior knowledge and learning outcomes, to see how the six measures were related. A correlation matrix was created on the dependent variables: perceptions of credibility in the message, perceptions of credibility in the virtual human, trust in the message, trust in the virtual human, prior knowledge, and learning outcomes (Table 3).

First, we examined the relationships between the trust and credibility measures. All four items were significantly positively correlated with each other, with moderate strength. There was a positive correlation between message credibility and virtual human credibility, $r = 0.562$, $n = 92$, $p < .01$. There was a positive correlation between the message credibility and trust in the message, $r = 0.501$, $n = 92$, $p < .01$. There was also a positive correlation between message credibility and trust in the virtual human, $r = 0.507$, $n = 92$, $p < .01$. There was a positive correlation between virtual human credibility and trust in the message, $r = 0.359$, $n = 92$, $p < .01$. There was a positive correlation between virtual human credibility and trust in the virtual human, $r = 0.604$, $n = 92$, $p < .01$. There was a positive correlation between trust in the message and trust in the virtual human, $r = 0.487$, $n = 92$, $p < .01$.

An interesting pattern emerged when we examined the prior knowledge and learning outcome scores in relation to each other and the trust and credibility measures. As expected, pretest and post-test scores were significantly related, $r = 0.549$, $n = 92$, $p < .01$. More interestingly, we found that message trust was significantly and positively related to learners' prior knowledge, $r = 0.287$, $n = 92$, $p < .01$. In addition, we found that message trust was significantly and positively correlated to learning post-test scores, $r = 0.225$, $n = 92$, $p = 0.03$. There were no other significant relationships between the prior knowledge or learning outcome measures and trust or credibility measures.

RQ8 – To What Extent do the Perceptions of Trust or Credibility Mediate the Relationship Between Prior Knowledge and Learning Outcomes? Since perceived message trust was significantly related to both prior knowledge and learning outcomes, we examined if perceived message trust mediated the relationship between prior knowledge and learning post-test scores (Figure 3). A series of linear regressions were conducted. As expected, there was a significant positive association between prior knowledge and post-test score, $\beta = 0.64$, $SE = 0.10$, $p < .01$. There was a significant positive association between pre-test and perceptions of trust in the message, $\beta = 0.13$, $SE = 0.04$, $p < .01$. Furthermore, the perceptions of trust in the message were significantly associated with post-test learning outcomes, $\beta = 0.61$, $SE = 0.28$, $p = .03$. A Sobel test was then conducted on perceptions of trust in the message's mediation effect between pre-test and post-test, which was not significant, $\beta = 1.81$, $SE = 0.04$, $p = .07$. In other words,

Table 3. Correlation Matrix of Credibility and Trust Data.

Variable	N	Message Credibility		Virtual Human Credibility		Message Trust		Virtual Human Trust		Pre-test		Post-test	
		Mean		Mean		Mean		Mean		Mean		Mean	
Message credibility	Pearson correlation Sig. (2-Tailed) N	1 92		0.562** <0.01 92		0.501** <0.01 92		0.507** <0.01 92		0.156 0.14 92		0.171 0.10 92	
Virtual human credibility	Pearson correlation Sig. (2-Tailed) N	0.562** <0.01 92		1 <0.01 92		0.359** <0.01 92		0.604** <0.01 92		-0.011 0.92 92		-0.048 0.65 92	
Message trust	Pearson correlation Sig. (2-Tailed) N	0.501** <0.01 92		0.359** <0.01 92		1 <0.01 92		0.487** <0.01 92		0.287** <0.01 92		0.225* 0.03 92	
Virtual human trust	Pearson correlation Sig. (2-Tailed) N	0.507** <0.01 92		0.604** <0.01 92		0.487** <0.01 92		1 <0.01 92		-0.039 0.71 92		-0.026 0.81 92	
Pre-test	Pearson correlation Sig. (2-Tailed) N	0.156 0.14 92		-0.011 0.92 92		0.287** <0.01 92		-0.039 0.71 92		1 <0.01 92		0.549** <0.01 92	
Post-test	Pearson correlation Sig. (2-Tailed) N	0.171 0.10 92		-0.048 0.65 92		0.225* 0.03 92		-0.026 0.81 92		0.549** <0.01 92		1 <0.01 92	

*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).

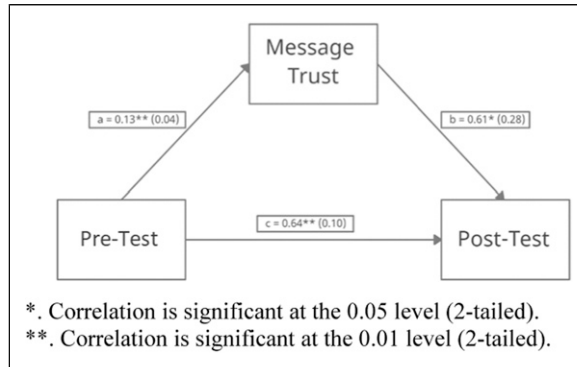


Figure 3. The mediation analysis examining if message trust mediated the relationship between pretest and post-test scores. *. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).

perceived trust in the message did not mediate the relationship between prior knowledge and learning outcomes, however perceived trust in the message did have a significant association with learning outcomes to nearly the same extent as prior knowledge.

Discussion. As noted, there is very little research investigating the relationship between trust and credibility with virtual humans in learning technologies. Broadly speaking, trust has been used to understand attitudes toward technologies in the workplace, based on the fundamental idea that trust is an evolving relationship, whereas credibility has been used to understand attitudes toward instructional and informational (including technological) features or interventions. The use of virtual humans in learning technologies thus raises the notion that the broader more comprehensive concept of trust may envelope the more specific concept of credibility, but which of the two best explains learning outcomes was an outstanding question (Schroeder et al., 2021). We thus conducted exploratory analyses to examine how the concepts of trust and credibility relate to one another and if they mediate learning outcomes.

There are two notable findings from our results. First, while learners' perceptions of trust in the virtual human, trust in the message, credibility of the virtual human, and credibility of the message are strongly related to one another, they are also measurably distinct concepts. This in itself is an important contribution to the field, as we are unaware of any studies in learning technologies that empirically demonstrate these relationships. These findings hold important implications for future research as the field begins to explore how trust and credibility may influence learning and learning-related outcomes.

Second, we found that perceived trust in the message was nearly as predictive of learning outcomes as prior knowledge. Meanwhile, we did not find statistically

significant relationships between trust in the virtual human, perceived credibility of the virtual human, or perceived credibility of the virtual human's messaging and learning. This holds important implications for the design of learning technologies broadly. If replicable, this finding highlights the importance of designing for learners' trust in instructional messaging, distinct from trustworthiness or credibility, which is a largely unexplored area of research in learning technologies.

General Discussion

Above, we have discussed the findings of Part 1 and Part 2 of the study in relation to the existing literature. In this section, we focus on the implications of our findings for theory and practice, discuss the limitations of this work, and elaborate on potential future research.

Implications for Theory

Our results have implications for the KReC framework, as well as the trust and credibility literature. In regards to KReC, our findings showed that regardless of the type of verbal narration learners experienced, they were able to revise their existing knowledge and learn the information. This implies that their prior knowledge was activated when they heard the narrative, which is an expansion of KReC, as KReC often speaks specifically to reading interventions rather than audio/video interventions. In addition, it is clear that the new information outcompeted the old information in the learner's memory when it was retrieved for completing the post-test. In short, our results, as well as those from [Schroeder et al. \(2022\)](#), show that KReC can be applied to the domain of auditory messaging, rather than only visual messaging (e.g., text).

In regards to the trust and credibility literature, our findings showed that measures of perceived trust and credibility in the source and in the message were consistently related, yet trust in the message stood out as a predictor of learning outcomes, in addition to a learner's prior knowledge. Because trust in the message and the learner's prior knowledge were also related, it is possible that learners who are more knowledgeable about a particular subject are also likely more able to assess whether or not they can trust the scientific accuracy of the information, regardless of the format. This harkens back to earlier research on expert trust in decision support systems and the many factors that are shown to affect human-technology performance (in this case, learning outcomes), including self-confidence, monitoring of the system, and task difficulty ([Lee & See, 2004](#); [Madhavan & Wiegmann, 2007](#); [Bitkina et al., 2020](#)). That trust in the message was the only factor among all the trust and credibility measures that predicted learning outcomes is promising for future work studying learners' trust in instructional content.

Implications for Practice

From an instructional design perspective, our results indicate that the virtual human's narrative structure delivered via RT or expository format may not be of major consequence for learning, perceived trust, perceived credibility, or the learners' perception of the virtual human. Our results showed no significant differences between groups on all measures for instructional audio/video content. We note however, that our intervention was a short (about 5 minutes), non-interactive instructional video. It seems plausible that with a longer video, or in a system with dialogue between the learner and the virtual human, the results may be different. The results may also differ in situations where the learner controls the pace of a segmented video, or in narratives where there are more direct refutations. In short, given the lack of work in the area, our findings should be seen as a starting point, rather than a broad generalization.

Another important contribution of this study is the finding that perceived trust in the message predicted the post-test scores nearly as well as prior knowledge. This implies that it is very important to find ways to build trust in the instructional messaging. However, work in the area is limited and therefore direly needed to help guide practitioners. It is unclear if different types of narrative structures (as seen here and in [Schroeder et al. \(2022\)](#)), different styles of voices used for narration (e.g., [Craig & Schroeder, 2017](#); [Davis et al., 2019](#)), or virtual humans with different features (e.g., appearance, gestures, role in the environment, etc.) may influence perceived trust in the message, as opposed to trust in the virtual human.

Finally, from a future research perspective, our results empirically show that learners can differentiate between trust and credibility, as well as trust and credibility in the virtual human as well as the message. These are important for research across the realm of educational technology as researchers seek ways to create more accessible learning technologies for all learners.

Limitations and Future Directions

To further clarify the relationship between perceived trust and credibility, future work may need to parse the different dimensions of each construct. To do this we would need to use an instrument that is developed to test this, for example the [Chancey et al. \(2017\)](#) trust questionnaire (or another version of it) in place of the [Jian et al. \(2000\)](#) questionnaire on trust in automation. The [Jian et al. \(2000\)](#) trust in automation questionnaire was designed to measure general trust attitudes toward automation, and it was developed using word elicitation from an English-language, U.S.-educated, small-sample population and then by conducting a factor analysis on the resulting terms, rather than developed based on a scholarly understanding of the construct (as in [McCroskey & Young, 1979](#)). The scholarly understanding of trust has consistently defined three types of information that inform a person's trust in an entity, which we have summarized previously as: purpose, process, and performance.

One question that may be raised from this study is how to build trust in the instructional message? This question implies that factual information for some reason is not persuasive at face value, and thus requires further modification to improve its persuasiveness. If there is a need for persuasion, this points to aspects of the task and the task environment that merit further investigation. The rhetorical question is, why don't people (want to) learn or revise their existing knowledge? There could be many reasons that are unrelated to the delivery of the content or the content itself and the conceptual change literature explores many of these factors. However, the three types of information mentioned (purpose, process, and performance) that affect a person's trust may help to pinpoint related areas to improve understanding and designing for trust in instructional technologies. For example, it may not matter whether information is delivered in a RT structure versus a more linear narrative when it comes to learning outcomes, if the *process* of information delivery is the same (e.g., passive), *performance* is the same (i.e., scientifically accurate), and the *purpose* of the system is the same (i.e., to deliver scientifically accurate knowledge). However, changes to the *process* of delivery may matter much more, such as comparing interactive versus passive consumption of the information. Or, perhaps the *purpose* of the system is misunderstood or misaligned with a learner's goals, leading to disengagement.

Importantly, understanding *purpose* in human-technology interactions has been severely understudied due to the common assumption and blanket oversimplification that technology is designed to align with human needs, and that human needs do not meaningfully vary given a particular purpose, like learning. In reality, it remains a perennial challenge to identify human needs in complex task environments like learning that a single technology can unilaterally address (Righi et al., 2017). Furthermore, it is not always that case that individual sub-goals are aligned with the shared goal (National Academies of Sciences, Engineering, and Medicine, 2021; Sheridan, 2019). Although the purpose of a system may be to deliver scientifically accurate information, a student in that moment of interaction may have more specific needs that are unmet by the mode of information delivery.

In addition, different measures of both trust and credibility may be needed that are more specific to the task context, to further investigate the role of trust and credibility in learning outcomes, and their interrelationships on various related dimensions. For example, there may be crucial differences between credibility of information for students in a learning task compared to credibility of information for an intelligence analyst in a reconnaissance task (e.g., Blasch et al., 2020), and perceptions of credibility may differ from actual decisions to rely on the information provided. There may be a need to inspect in finer-grained detail these more distinct dimensions of source and message credibility, as well as source and message trust, that are based on the role of the perceiver, and the type of task environment.

Conclusion

In this study, we examined the influence of two different types of verbally-delivered narrative structures on learners' perceptions of credibility, trust, perceptions of a virtual human, and learning outcomes. Our results failed to show any significant differences between the narrative structure types. However, our exploratory analyses revealed two novel contributions to the literature. First, we found that the perceived credibility of the message and virtual human, as well as trust in the message and virtual human, are correlated, yet four measurably distinct concepts. In addition, we found that only trust in the message was predictive of learning outcomes, and it did so nearly as well as prior knowledge, yet it did not mediate the relationship between prior knowledge and learning.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Parts of this material is based upon work supported by the U.S. Department of Homeland Security under Grant Award Number 17STQAC00001-05-00. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security.

ORCID iDs

Noah L. Schroeder  <https://orcid.org/0000-0002-3281-2594>

Scotty D. Craig  <https://orcid.org/0000-0002-4103-5547>

References

- Appelman, A., & Sundar, S. S. (2016). Measuring message credibility: Construction and validation of an exclusive scale. *Journalism & Mass Communication Quarterly*, 93(1), 59–79. <https://doi.org/10.1177/1077699015606057>
- Baylor, A., Ryu, J., & Shen, E. (2003). The effects of pedagogical agent voice and animation on learning, motivation and perceived persona. In *EdMedia+ innovate learning* (pp. 452–458). Association for the Advancement of Computing in Education (AACE).
- Beege, M., Schneider, S., Nebel, S., & Rey, G. D. (2020). Does the effect of enthusiasm in a pedagogical agent's voice depend on mental load in the learner's working memory? *Computers in Human Behavior*, 112. <https://doi.org/10.1016/j.chb.2020.106483>
- Bitkina, O. V., Jeong, H., Lee, B. C., Park, J., Park, J., & Kim, H. K. (2020). Perceived trust in artificial intelligence technologies: A preliminary study. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 30(4), 282–290. <https://doi.org/10.1002/hfm.20839>

- Blasch, E., Sung, J., & Nguyen, T. (2020). Multisource AI Scorecard table for system evaluation. *ArXiv:2102.03985 [Cs]*. <http://arxiv.org/abs/2102.03985>
- Carter, R. F., & Greenberg, B. S. (1965). Newspapers or television: Which do you believe? *Journalism Quarterly*, 42(1), 29–34. <https://doi.org/10.1177/107769906504200104>
- Castro-Alonso, J. C., Wong, R. M., Adesope, O. O., & Paas, F. (2021). Effectiveness of multimedia pedagogical agents predicted by diverse theories: A meta-analysis. *Educational Psychology Review*, 33(3), 989–1015. <https://doi.org/10.1007/s10648-020-09587-1>
- Chancey, E. T., Bliss, J. P., Yamani, Y., & Handley, H. A. H. (2017). Trust and the compliance–reliance paradigm: The effects of risk, error bias, and reliability on trust and dependence. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 59(3), 333–345. <https://doi.org/10.1177/0018720816682648>
- Chiou, E. K., & Lee, J. D. (2021). Trusting automation: Designing for responsivity and resilience. *Human Factors*. <https://doi.org/10.1177/00187208211009995>
- Chiou, E. K., Schroeder, N. L., & Craig, S. D. (2020). How we trust, perceive, and learn from virtual humans: The influence of voice quality. *Computers & Education*, 146, 103756. <https://doi.org/10.1016/j.compedu.2019.103756>
- Craig, S. D., & Schroeder, N. L. (2017). Reconsidering the voice effect when learning from a virtual human. *Computers & Education*, 114, 193–205. <https://doi.org/10.1016/j.compedu.2017.07.003>
- Craig, S. D., & Schroeder, N. L. (2018). Design principles for virtual humans in educational technology environments. In K. Millis, J. Magliano, D. Long, & K. Wiemer (Eds.), *Deep learning: Multi-disciplinary approaches* (pp. 128–139). Routledge.
- Davis, R. O., & Vincent, J. (2019). Sometimes more is better: Agent gestures, procedural knowledge and the foreign language learner. *British Journal of Educational Technology*, 50(6), 3252–3263. <https://doi.org/10.1111/bjet.12732>
- Davis, R. O., Vincent, J., & Park, T. (2019). Reconsidering the voice principle with non-native language speakers. *Computers & Education*, 140(C), 103605. <https://doi.org/10.1016/J.COMPEDU.2019.103605>
- Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25(6), 671–688. <https://doi.org/10.1080/09500690305016>
- Ecker, U. K., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., & Brashier, N. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1(1), 13–29. <https://doi.org/10.1080/09500690305016>
- Edwards, A., Edwards, C., Spence, P. R., Harris, C., & Gambino, A. (2016). Robots in the classroom: Differences in students' perceptions of credibility and learning between “teacher as robot” and “robot as teacher. *Computers in Human Behavior*, 65(C), 627–634. <https://doi.org/10.1016/j.chb.2016.06.005>
- Finn, A. N., Schrod, P., Witt, P. L., Elledge, N., Jernberg, K. A., & Larson, L. M. (2009). A meta-analytical review of teacher credibility and its associations with teacher behaviors and student outcomes. *Communication Education*, 58(4), 516–537. <https://doi.org/10.1080/03634520903131154>

- Fogg, B. J., & Tseng, H. (1999). The elements of computer credibility. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems the CHI Is the Limit – CHI '99, Vancouver, BC, 13–18 April 1996, 80–87.
- Gutzwiler, R. S., Chiou, E. K., Craig, S. D., Lewis, C. M., Lematta, G. J., & Hsiung, C. P. (2019, November). Positive bias in the ‘Trust in Automated Systems Survey’? An examination of the Jian et al. (2000) scale. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 217–221). SAGE Publications.
- Heddy, B. C., Danielson, R. W., Sinatra, G. M., & Graham, J. (2017). Modifying knowledge, emotions, and attitudes regarding genetically modified foods. *The Journal of Experimental Education*, 85(3), 513–533. <https://doi.org/10.1080/00220973.2016.1260523>
- Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review*, 6(1), 27–54. <https://doi.org/10.1016/j.edurev.2010.07.004>
- Hollan, J., & Stornetta, S. (1992). Beyond being there. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems – CHI '92 (pp. 119–225). ACM.
- Hynd, C. R. (2001). Refutational texts and the change process. *International Journal of Educational Research*, 35(7–8), 699–714. [https://doi.org/10.1016/S0883-0355\(02\)00010-1](https://doi.org/10.1016/S0883-0355(02)00010-1)
- Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics*, 4(1), 53–71. https://doi.org/10.1207/S15327566IJCE0401_04
- Kendeou, P., Butterfuss, R., Kim, J., & Van Boekel, M. (2019). Knowledge revision through the lenses of the three-pronged approach. *Memory & Cognition*, 47(1), 33–46. <https://doi.org/10.3758/s13421-018-0848-y>
- Kendeou, P., Walsh, E. K., Smith, E. R., & O'Brien, E. J. (2014). Knowledge revision processes in refutation texts. *Discourse Processes*, 51(5–6), 374–397. <https://doi.org/10.1080/0168583X.2014.913961>
- Kim, J., & Kendeou, P. (2021). Knowledge transfer in the context of refutation texts. *Contemporary Educational Psychology*, 67(1), 102002. <https://doi.org/10.1016/j.cedpsych.2021.102002>
- Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021a). Do learners recognize and relate to the emotions displayed by virtual instructors? *International Journal of Artificial Intelligence in Education*, 31(1), 134–153. <https://doi.org/10.1007/s40593-021-00238-2>
- Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021b). The positivity principle: Do positive instructors improve learning from video lectures? *Educational Technology Research and Development*, 69(6), 3101–3129. <https://doi.org/10.1007/s11423-021-10057-w>
- Lawson, A. P., Mayer, R. E., Adamo-Villani, N., Benes, B., Lei, X., & Cheng, J. (2021c). Recognizing the emotional state of human and virtual instructors. *Computers in Human Behavior*, 114, 106554. <https://doi.org/10.1016/j.chb.2020.106554>
- Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology*, 25(6), 943–951. <https://doi.org/10.1002/acp.1787>

- Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35(10), 1243–1270. <https://doi.org/10.1080/00140139208967392>
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1). https://doi.org/10.1518/hfes.46.1.50_30392
- Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the public interest*, 13(3), 106–131. <https://doi.org/10.1177/15291006124510>
- Lewicki, R. J., Tomlinson, E. C., & Gillespie, N. (2006). Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *Journal of Management*, 32(6), 991–1022. <https://doi.org/10.1177/0149206306294405>
- Madhavan, P., & Wiegmann, D. A. (2007). Similarities and differences between human-human and human-automation trust: An integrative review. *Theoretical Issues in Ergonomics Science*, 8(4), 277–301. <https://doi.org/10.1080/14639220500337708>
- Mason, L., Zaccoletti, S., Carretti, B., Scrimin, S., & Diakidoy, I. A. N. (2019). The role of inhibition in conceptual learning from refutation and standard expository texts. *International Journal of Science and Mathematics Education*, 17(3), 483–501. <https://doi.org/10.1007/s10763-017-9874-7>
- Matsui, T., & Yamada, S. (2019). The design method of the virtual teacher. Proceedings of the 7th International Conference on Human-Agent Interaction, Kyoto, Japan, 25 September 2019, 97–101. <https://doi.org/10.1145/3349537.3351906>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, 20(3), 709–734. <https://doi.org/10.2307/258792>
- McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communications Monographs*, 66(1), 90–103. <https://doi.org/10.1080/03637759909376464>
- McCroskey, J. C., Valencic, K. M., & Richmond, V. P. (2004). Toward a general model of instructional communication. *Communication Quarterly*, 52(3), 197–210. <https://doi.org/10.1080/01463370409370192>
- McCroskey, J. C., & Young, T. J. (1979). The use and abuse of factor analysis in communication research. *Human Communication Research*, 5(4), 375–382. <https://doi.org/10.1111/j.1468-2958.1979.tb00651.x>
- McCroskey, J. C., & Young, T. J. (1981). Ethos and credibility: The construct and its measurement after three decades. *Central States Speech Journal*, 32(1), 24–34. <https://doi.org/10.1080/10510978109368075>
- Muller, D. A., Bewes, J., Sharma, M. D., & Reimann, P. (2008). Saying the wrong thing: Improving learning with multimedia by including misconceptions. *Journal of Computer Assisted Learning*, 24(2), 144–155. <https://doi.org/10.1111/j.1365-2729.2007.00248.x>
- Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies*, 27, 527–539. [https://doi.org/10.1016/S0020-7373\(87\)80013-5](https://doi.org/10.1016/S0020-7373(87)80013-5)

- National Academies of Sciences, Engineering, and Medicine. (2021). *Human–AI teaming: State of the art and research needs* (pp. 26355). The National Academies Press. <https://doi.org/10.17226/26355>
- Park, S. (2015). The effects of social cue principles on cognitive load, situational interest, motivation, and achievement in pedagogical agent multimedia learning. *Journal of Educational Technology & Society*, 18(4), 211–229.
- Ramlatchan, M., & Watson, G. S. (2020). Enhancing instructor credibility and immediacy in the design of distance learning systems and virtual classroom environments. *The Journal of Applied Instructional Design*, 9(2). <https://dx.doi.org/10.51869/92mrgsw>
- Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology*, 34(2), 243–281. <https://doi.org/10.1111/j.1559-1816.2004.tb02547.x>
- Rempel, J., & Holmes, J. (1985). Trust in close relationships. *Journal of Personality and Social Psychology*, 49(1), 95–112.
- Righi, V., Sayago, S., & Blat, J. (2017). When we talk about older people in HCI, who are we talking about? Towards a 'turn to community' in the design of technologies for a growing ageing population. *International Journal of Human–Computer Studies*, 108, 15–31. <https://doi.org/10.1016/j.ijhcs.2017.06.005>
- Ryu, J., & Baylor, A. L. (2005). The psychometric structure of pedagogical agent persona. *Technology Instruction Cognition and Learning*, 2(4), 291–314.
- Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research*, 49(1), 1–39. <https://doi.org/10.2190/EC.49.1.a>
- Schroeder, N. L., Chiou, E. K., & Craig, S. D. (2021). Trust influences perceptions of virtual humans, but not necessarily learning. *Computers & Education*, 160, 104039. <https://doi.org/10.1016/j.compedu.2020.104039>
- Schroeder, N. L., & Craig, S. D. (2021). Learning with virtual humans: Introduction to the special issue. *Journal of Research on Technology in Education*, 53(1), 1–7. <https://doi.org/10.1080/15391523.2020.1863114>
- Schroeder, N. L., & Kucera, A. C. (2022). Refutation text facilitates learning: A meta-analysis of between-Subjects experiments. *Educational Psychology Review*, 34, 957–987. <https://doi.org/10.1007/s10648-021-09656-z>
- Schroeder, N. L., Romine, W. L., & Craig, S. D. (2017). Measuring pedagogical agent persona and the influence of agent persona on learning. *Computers & Education*, 109, 176–186. <https://doi.org/10.1016/j.compedu.2017.02.015>
- Schroeder, N. L., Siegle, R. F., Verma, V., & Craig, S. D. (2022). When refutation and deep reasoning questions lose their edge: The case of short, narrated videos. *Journal of Educational Computing Research*, 60(5), 1197–1219. <https://doi.org/10.1177/07356331211070534>
- Schroeder, N. L., Yang, F., Banerjee, T., Romine, W. L., & Craig, S. D. (2018). The influence of learners' perceptions of virtual humans on learning transfer. *Computers & Education*, 126, 170–182. <https://doi.org/10.1016/j.compedu.2018.07.005>

- Sheridan, T. B. (1975). Considerations in modeling the human supervisory controller. *IFAC Proceedings* (pp. 223–228). Elsevier. [https://doi.org/10.1016/S1474-6670\(17\)67555-4](https://doi.org/10.1016/S1474-6670(17)67555-4)
- Sheridan, T. B. (2019). Individual differences in attributes of trust in automation: Measurement and application to system design. *Frontiers in Psychology*, 10, 1117. <https://doi.org/10.3389/fpsyg.2019.01117>
- Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Singh, A.–M., Marcus, N., & Ayres, P. (2012). The transient information effect: Investigating the impact of segmentation on spoken and written text. *Applied Cognitive Psychology*, 26(6), 848–853. <https://doi.org/10.1002/acp.2885>
- Spain, R. D., Bustamante, E. A., & Bliss, J. P. (2008, September). Towards an empirically developed scale for system trust: Take two. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 1335–1339). SAGE Publications.
- Sweller, J., Ayers, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.
- Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education*, 8(6), 951–970. <https://doi.org/10.1007/s10763-010-9203-x>
- Treagust, D. F., & Duit, R. (2008). Conceptual change: A discussion of theoretical, methodological and practical challenges for science education. *Cultural Studies of Science Education*, 3(2), 297–328. <https://doi.org/10.1007/s11422-008-9090-4>
- Trevors, G. J., Kendeou, P., & Butterfuss, R. (2017). Emotion processes in knowledge revision. *Discourse Processes*, 54(5–6), 406–426. <https://doi.org/10.1080/0163853X.2017.1312201>
- Van Boekel, M., Lasonde, K. A., O'Brien, E. J., & Kendeou, P. (2017). Source credibility and the processing of refutation texts. *Memory & Cognition*, 45(1), 168–181. <https://doi.org/10.3758/s13421-016-0649-0>
- Will, K. K., Masad, A., Vlach, H. A., & Kendeou, P. (2019). The effects of refutation texts on generating explanations. *Learning and Individual Differences*, 69, 108–115. <https://doi.org/10.1016/j.lindif.2018.12.002>
- Wong, A., Leahy, W., Marcus, N., & Sweller, J. (2012). Cognitive load theory, the transient information effect and e-learning. *Learning and Instruction*, 22(6), 449–457. <https://doi.org/10.1016/j.learninstruc.2012.05.004>
- Yung, H. I., & Paas, F. (2015). Effects of cueing by a pedagogical agent in an instructional animation: A cognitive load approach. *Journal of Educational Technology & Society*, 18(3), 153–160.

Author Biographies

Noah L. Schroeder is a Professor of Educational Technology and Instructional Design at Wright State University. His research interests are centered around the design of multimedia learning environments, with his current work largely focused on the use of pedagogical agents and the design of instructional videos. He serves as an Associate

Editor for the Journal of Engineering Education and an editorial board member for the Journal of Research on Technology in Education and Computers and Education Open.

Erin K. Chiou is an Assistant Professor of Human Systems Engineering at Arizona State University. Her research interests center around human-agent interactions in complex work systems, with current work focused on operator trust in AI-enabled decision support systems for security and intelligence applications. She serves as an Associate Editor for the Journal of Cognitive Engineering and Decision Making and is an editorial board member for the International Journal of Human-Computer Interaction.

Robert F. Siegle is a PhD student in Human Systems Engineering at Arizona State University with a background in Psychology. His current work focuses on factors that improve virtual learning environments and is currently working as a research assistant in the Arizona State University Advanced Distributed Learning Initiative Partnership Laboratory identifying best practices for eLearning environments and evaluating learning technologies. His research interests include voice and learning, virtual humans, and virtual reality.

Scotty D. Craig is an Associate Professor of Human Systems Engineering within The Polytechnic School of the Ira A. Fulton Schools of Engineering at Arizona State University and an affiliate faculty of the Mary Lou Fulton Teachers College. He is the Director of the Arizona State University Advanced Distributed Learning Initiative Partnership Laboratory. His PhD is in Experimental Psychology from The University of Memphis with a focus on cognitive learning. His research is in the learning engineering area at the intersection of psychology, learning sciences, user sciences, and technology where he has investigated the concepts of emotion and learning, discourse processing, multimedia learning, and intelligent systems in both laboratory and applied learning and training settings.