iafor

journal of education technology in education

Volume 9 – Issue 4 – 2021

Editor: Michael P. Menchaca

Editor-in-Chief: Joseph Haldane

ISSN: 2187-0594

iafor

IAFOR Journal of Education: Technology in Education
Volume 9 – Issue 4 – 2021

IAFOR Publications

The International Academic Forum

IAFOR Journal of Education: Technology in Education

Published by The International Academic Forum (IAFOR), Japan
IAFOR Publications, Sakae 1-16-26-201, Naka-ward, Aichi, Japan 460-0008

IAFOR Journal of Education: Technology in Education
Volume 9 – Issue 4 – 2021

Published August, 2021

ISSN: 2187-0594

ije.iafor.org

Cover Image: Christina @ wocintechchat.com Unsplash

IAFOR Journal of Education: Technology in Education

Volume 9 – Issue 4 – 2021

Edited by Michael P. Menchaca

Associate Editor: Daniel Hoffman

Foreword

The global coronavirus pandemic has highlighted the crucial role of technology in education as educational systems around the world have been forced to respond to the needs and demands of students across the world.

However, the responses possible have differed depending on the local and national contexts within and between different countries as existing and inherent educational inequities and inequalities have been exacerbated by many of the modern technological solutions and tools. Where infrastructure has been poor, sometimes certain technologies once written off as obsolete have shown their enduring role and resilience.

The difference between online teaching and emergency response teaching has been documented over this period as many academics, me included, grappled with teaching online for the first time, and were forced to adapt to the "new normal" with great speed. Delivery, accountability, assessment, assurance, and countless other areas of responsibility were suddenly questioned and subject to new rules as educational institutions grappled with overnight change.

This ongoing global pandemic provides the background context to this issue, and indeed its very publication on the IAFOR platform is testament to power of technology in education. The range of quality of articles in this issue is a testament to the work of the many contributors, reviewers and collaborators, and I would like to acknowledge and thank them all here. I would especially like to thank my good friend Dr Michael Menchaca, of the University of Hawaii at Manoa, and his associate editor, Dr Daniel Hoffman, of the same institution, for their dedication to the journal.

I hope you enjoy reading the articles presented in this issue.

Joseph Haldane
Editor-in-Chief
IAFOR Journal of Education

Editorial Advice

Preparing a submission to the *IAFOR Journal of Education* is more than writing about your research study: it involves paying careful attention to our submission requirements. Different journals have different requirements in terms of format, structure and referencing style, among other things. There are also some common expectations between all journals such as the use of good academic language and lack of plagiarism. To assist you in reaching the review stage for this or any other peer-reviewed journal, we provide the following advice which you should check carefully and ensure that you adhere to.

1. Avoiding Plagiarism

Plagiarism is a practice that is not acceptable in any journal. Avoiding plagiarism is the cardinal rule of academic integrity because plagiarism, whether intentional or unintentional, is presenting someone else's work as your own. The *IAFOR Journal of Education* immediately rejects any submission with evidence of plagiarism.

There are three common forms of plagiarism, none of which are acceptable:

1. ***Plagiarism with no referencing***. This is copying the words from another source (article, book, website, etc.) without any form of referencing.
2. ***Plagiarism with incorrect referencing***. This involves using the words from another source and only putting the name of the author and/or date as a reference. Whilst not as grave as the plagiarism just mentioned, it is still not acceptable academic practice. Direct quoting requires quotation marks and a page number in the reference. This is best avoided by paraphrasing rather than copying.
3. ***Self-plagiarism***. It is not acceptable academic practice to use material that you have already had published (which includes in conference proceedings) in a new submission. You should not use your previously published words and you should not submit about the same data unless it is used in a completely new way.

2. Meeting the Journal Aims and Scope

Different journals have different aims and scope, and papers submitted should fit the specific journal. A "scattergun" approach (where you submit anywhere in the hope of being published) is not sound practice. Like in darts, your article needs to hit the journal's "bullseye", it needs to fit within the journal's interest area. For example, a submission that is about building bridges, will not be acceptable in a journal dedicated to education. Ensure that your paper is clearly about education.

3. Follow the Author Guidelines

Most journals will supply a template to be followed for formatting your paper. Often, there will also be a list of style requirements on the website (font, word length, title length, page layout, and referencing style, among other things). There may also be suggestions about the preferred structure of the paper. For the *IAFOR Journal of Education* these can all be found here:
https://iafor.org/journal/iafor-journal-of-education/author-guidelines/

4. Use Academic Language

The *IAFOR Journal of Education* only accepts papers written in correct and fluent English at a high academic standard. Any use of another language (whether in the paper or the reference list) requires the inclusion of an English translation.

The style of expression must serve to articulate the complex ideas and concepts being presented, conveying explicit, coherent, unambiguous meaning to scholarly readers. Moreover, manuscripts must have a formal tone and quality, employing third-person rather than first-person standpoint (when feasible), placing emphasis on the research and not on unsubstantiated subjective impressions.

Contributors whose command of English is not at the level outlined above are responsible for having their manuscript corrected by a native-level, English-speaking academic prior to submitting their paper for publication.

5. Literature Reviews

Any paper should have reference to the corpus of **scholarly** literature on the topic. A review of the literature should:

- Predominantly be about contemporary literature (the last 5 years) unless you are discussing a seminal piece of work.
- Make explicit international connections for relevant ideas.
- Analyse published papers in the related field rather than describe them.
- Outline the gaps in the literature.
- Highlight your contribution to the field.

Referencing

Referencing is the main way to avoid allegations of plagiarism. The *IAFOR Journal of Education* uses the APA referencing style for both in-text citations and the reference list. If you are unsure of the correct use of APA please use the Purdue Online Writing Lab (Purdue OWL), – https://owl.english.purdue.edu/owl/resource/560/01/ – which has excellent examples of all forms of APA referencing. Please note APA is used for referencing not for the general format of the paper. Your reference list should be alphabetical by author surname and include DOIs whenever possible.

This short guide to getting published should assist you to move beyond the first editorial review. Failure to follow the guidelines will result in your paper being immediately rejected.

Good luck in your publishing endeavours,

Dr Yvonne Masters
Executive Editor, *IAFOR Journal of Education*

Table of Contents

From the Editors

Thank you for perusing this latest *IAFOR Journal of Education: Technology in Education* issue. This is the *Journal*'s second offering in the specialized area of educational technology. The *Journal* overall remains committed to inter-disciplinary, international, diverse articles and this issue is no exception. The seven articles included represent: (a) research in both higher education and secondary settings, (b) a mixture of design-based, qualitative, quantitative, and mixed-methods approaches, and (c) five different countries. The articles cover diverse areas such as: artificial intelligence, teacher education, documentary film-making, language learning, mobile learning, virtual manipulatives, and learning task selection.

When perusing this issue, in particular, please note the authors' incorporation of theoretic or conceptual foundations to help guide their questions and conclusions. As editors, we receive many quality manuscript submissions whose main shortcoming is a lack of foundational theory. Consider framing your research with theory when submitting manuscripts.

Here is a summary of each article in this issue.

In the first article, "Teaching during a pandemic: Elementary candidates' experiences with engagement in distance education," Monica Gonzalez Smith and Nicole Schlaack conducted an interpretive phenomenological study of the lived experience of teacher candidates promoting student engagement while delivering emergency remote teaching in the Pacific Southwest during the pandemic. Conclusions provided by the study included that teacher candidates need more support with questioning, assessment, and technology tools to keep students engaged and that teacher preparation programs need to provide teacher candidates opportunities to practice engagement strategies in distance settings.

Onur Oymak and Feral Ogan-Bekiroglu, in their article, "Comparison of Students' Learning and Attitudes in Physical versus Virtual Manipulatives using Inquiry-Based Instruction," employ a pre-post comparison design in a Turkish secondary setting to study the effects of virtual versus physical manipulatives on student learning. An interesting conclusion was that students had fewer measurement errors when relying on virtual manipulatives.

Finland's Stina Westman, Janne Kauttonen, Aarne Klemetti, Niilo Korhonen, Milja Manninen, Asko Mononen, Salla Niittymaki, and Henry Paananen use a convergent parallel mixed-methods design in their research on using artificial intelligence to support career guidance in higher education. Their study looked at the viewpoints of students and career guidance staff while they experienced the inclusion of AI for career guidance. The study indicated AI can help support the agency of students to prioritize, choose, and consider important decisions for their professional lives.

In the fourth article, "Five tips from filmmakers: An online instructional module for documentary film research," Patsy Iwasaki, an instructor and filmmaker from Hawaii, utilized convergent mixed-methods to study how an instructional module on documentary filmmaking informed by experts provided an educational foundation for appropriate subject research and data collection among higher education students. The study revealed that providing information and data literacy helped scaffold creativity in the arts while adhering to high standards for research.

Seohyun Choi, Jaewon Jung, and Dongsik Kim's article, "The effects of task selection approaches to emphasis manipulation on cognitive load and knowledge transfer," used an experimental methodology to study student task selection in a secondary school in South Korea. The study suggested an outside agent (such as an instructor or computer) can provide better task selection by emphasizing only areas of weakness compared to students having to choose for themselves among a list of tasks that might provide redundant knowledge.

In the sixth article, "Machine translation in the language classroom: Turkish EFL learners' and instructors' perceptions and use," Murat Ata and Emre Debreli compared student versus faculty perceptions of the use of Online Machine Translation (e.g., Google Translate) for learning English as a Foreign Language in higher education. Their findings suggested institutions should establish policies regarding the use of OMT tools, specifically to guide students in language learning as well as to better understand the ethics of using such tools.

Finally, Eunice Ofori and Barbara Lockee's article, "Next generation mobile learning: Leveraging message design considerations for learning and accessibility," implemented a design and development methodology to establish guidelines for mobile learning (mLearning) in higher education in the U.S. Their guidelines were informed from prior research on message design and Universal Design for Learning and could assist in creating optimal mLearning instruction.

Overall, these articles provide appropriate direction for incorporating complex technology in educational settings. We trust you will find the articles as useful as they are enjoyable to read.

Michael P. Menchaca, Editor, and **Daniel L. Hoffman**, Associate Editor
IAFOR Journal of Education: Technology in Education
tech.editor.joe@iafor.org

Notes on Contributors

Article 1: Teaching during a Pandemic: Elementary Candidates' Experiences with Engagement in Distance Education

Dr Monica Gonzalez Smith is an Assistant Professor at the University of Hawai'i at Mānoa. She prepares teacher candidates who are earning initial teacher licensure for working with multilingual learners. Her teaching and research center on sociocultural theory, school-university partnerships, culturally responsive teaching, and reflective practice.
Email: monicags@hawaii.edu

Dr Nicole Schlaack has a PhD in Educational Psychology from the University of Hawai'i at Mānoa. She has an interest in evaluating instructional strategies and educational programs for student development within the context of a diverse society. Her research takes a cultural-historical approach to human development with a focus on school-university partnerships and organizational change.

Article 2 Comparison of Students' Learning and Attitudes in Physical versus Virtual Manipulatives using Inquiry-Based Instruction

Onur Oymak received his BSc from Dokuz Eylul University, Department of Physics Education and obtained his M.Sc. from Marmara University. He is currently working as physics teacher at Deniz Military High School.
Email: onuroymak99@gmail.com

Dr Feral Ogan-Bekiroglu obtained her BS from Ankara University, Department of Physics Engineering and MSc from the same department. She received her PhD in Science Education from Florida State University, Department of Curriculum and Instruction. She worked as expert engineer in Criminal Laboratory for three years. She is currently working as professor in the Physics Education division of Science and Mathematics Education Department of Marmara University. She participated in development and revision of curriculum studies for high school physics and undergraduate physics education programs. She has been carrying out various duties related to evaluation and accreditation of teacher education programs and is continuing her work as head of Education Committee. Her research focuses on teacher professional development, STEM, argumentation, modelling, and epistemic cognition.
Email: feralogan@yahoo.com

Article 3: Artificial Intelligence for Career Guidance – Current Requirements and Prospects for the Future

Dr Stina Westman (DSc (Tech.)) has an extensive research and development background in user-centered research and service development in the areas of media, information systems, data management, open science as well as education and learning. Currently she works as Development manager in South-Eastern Finland University of Applied Sciences.
Email: stina.westman@xamk.fi

Dr Janne Kauttonen (PhD Physics), is a researcher and data-scientist with a broad experience on data-analysis and ML/AI. His research background is inter-disciplinary, including aspects from statistical physics, computational sciences, neuroscience and behavioural sciences. Currently he works as Researcher in Haaga-Helia University of Applied Sciences.

Aarne Klemetti (MSc (Tech.)) has wide experience and background in RDI and entrepreneurship. His target areas are in the field of ICT and media engineering and especially AI/ML (+Ops), data science, edge computing, scalability, and high-performance computing. For the last two decades he has been working for Metropolia University of Applied Sciences and its predecessor as Researching lecturer in the fields of data science and related systems, printed intelligence, and AIoT.

Niilo Korhonen (MA, Pedagogy) works as a teacher and expert in digital and online learning at Häme University of Applied Sciences. He is also responsible for the design and maintenance of learning environments (LMS) at the university.

Milja Manninen (MEd) has wide experience and background in guidance and education development, especially in the areas of career and workplace guidance. Her target areas are relationship between education and work life, developing education and guidance, including AI in guidance. Currently she works as Research manager in South-Eastern Finland University of Applied Sciences.

Asko Mononen, (Master of International Business and Management) has 25 years of experience in capability development in organisations. His areas of expertise are in digital service development and modern learning methodologies. Currently he works in Digital Living Lab at Laurea University of Applied Sciences in Finland.

Salla Niittymäki (MBA) has a wide experience from the universities of applied sciences, especially in the areas of guidance of students as well as the relations between higher education and work life. Currently she works as a Guidance counselor in the degree program of Mechanical Engineering in Häme University of Applied Sciences.

Henry Paananen (MSc) works as a development manager (digital solutions) in Häme University of Applied Sciences. His expertise includes educational technology, eLearning, management, leadership and strategic planning.

Article 4: Five Tips from Filmmakers: An Online Instructional Module for Documentary Film Research

Dr Patsy Y. Iwasaki has a PhD in Learning Design and Technology from the University of Hawaii at Manoa. She also holds an MEd in Education and is an English Department faculty member at the University of Hawaii at Hilo. Her research interests include instructional design, online course and program development, documentary film research and practice, English studies, comics/manga and migration stories, diversity, place-based, culturally relevant resources in education, and cross-cultural exchange and collaboration. Dr Iwasaki has conducted extensive research activities, published articles and given presentations in the United States, Japan and Spain in these areas. She has published a graphic novel, and is currently creating and producing a documentary film, about a historical figure in the Japan-Hawaii immigration narrative with timely relevance today.
Email: piwasaki@hawaii.edu

Article 5: The Effects of Task Selection Approaches to Emphasis Manipulation on Cognitive Load and Knowledge Transfer

Seohyun Choi received her MA in the Department of Educational Technology from Hanyang University and now she is a PhD student in the Department of Educational Technology at Hanyang University. Her research interests are in the field of instructional design, computer-supported collaborative learning. Her current research interests are concerned with productive failure (PF) in collaborative learning.
Email: sseohyunchoi@gmail.com

Dongsik Kim is a professor in the Department of Educational Technology at Hanyang University. He earned a PhD in Educational Technology from Florida State University. His research areas are in the field of instructional design, computer-supported collaborative learning, and e-learning. He has researched regarding complex learning, computer-supported collaborated learning (CSCL) and currently a member of the Educational Technology Research and Development editorial board.
Email: kimdsik@hanyang.ac.kr

Jaewon Jung is an associate research fellow at Korean Educational Development Institute. She holds a PhD in Educational Technology from Hanyang University, South Korea. She has researched educational issues such as multimedia learning, computer-supported collaborative learning (CSCL), and instructional design. Her current research interests include exploring productive failure in CSCL to build effective knowledge construction and optimize the collaboration load.
Email: jungj5@gmail.com

Article 6: Machine Translation in the Language Classroom: Turkish EFL Learners' and Instructors' Perceptions and Use

Murat Ata is a senior instructor at Middle East Technical University Northern Cyprus Campus School of Foreign Languages. He holds an MA in English language teaching and his research interests include technology and ICT in language education. Address: Middle East Technical University Northern Cyprus Campus, School of Foreign Languages, Güzelyurt, Northern Cyprus TR-10 Mersin, Turkey.
Email: muata@metu.edu.tr

Dr Emre Emre Debreli holds a PhD from the University of Reading, United Kingdom, and is an associate professor at the European University of Lefke, Northern Cyprus, where he teaches academic literacy and teacher education courses. His experience includes EFL teacher education, teacher/student cognition, literacy and social research methods. Address: European University of Lefke, Faculty of Education, Department of ELT, Lefke, Northern Cyprus TR-10 Mersin, Turkey. **Email**: edebreli@eul.edu.tr

Article 7: Next Generation Mobile Learning: Leveraging Message Design Considerations for Learning and Accessibility

Dr Eunice Ofori holds a PhD in Curriculum and instruction with an emphasis on Instructional Design and Technology from Virginia Polytechnic Institute and State University (Virginia Tech). Currently, Eunice serves as Senior Instructional Designer in the Center for Engaged Learning and Teaching (CELT) at Tulane University in New Orleans, Louisiana. She also

serves as Adjunct Faculty in the Teacher Preparation and Certification program in the School of Professional Advancement there. Her career has centered around facilitating the use of instructional technology and sound pedagogical strategies in teaching and learning. She also specializes in the design of effective instruction through varied modalities, including online and blended education. Her career goal is to provide equal access to education for ALL, supporting academic communities using evidence-based pedagogical strategies and innovative instructional technologies.
Email: eofori@tulane.edu

Dr Barbara Lockee is Professor of Instructional Design and Technology in the School of Education and Faculty Fellow in the Office of the Provost at Virginia Tech. For more than 25 years, her research and teaching has explored the intersection of instructional design and distance learning, with the goal of advancing knowledge regarding the creation of effective distance and online education based on the principles of human learning. Dr Lockee is Past President of the Association for Educational Communications and Technology and is a Vice-President of IAFOR, serving as Chair of the Education & Language Learning division of the International Academic Advisory Board. She earned her PhD at Virginia Tech in Curriculum and Instruction (Instructional Technology) in 1996.
Email: lockeebb@vt.edu

Teaching during a Pandemic: Elementary Candidates' Experiences with Engagement in Distance Education

Monica Gonzalez Smith
University of Hawaiʻi at Mānoa
USA

Nicole Schlaack
University of Hawaiʻi at Mānoa
USA

Abstract

The following research reports on a collaborative effort between two university field supervisors for an elementary teacher preparation program in the Pacific southwest. Utilizing practitioner inquiry and situated learning as conceptual frameworks, the authors qualitatively examine the experiences ten elementary education teacher candidates have with promoting student engagement during emergency response teaching because of the COVID-19 pandemic. An interpretive phenomenological analysis of 20 lesson reflections and supervisor observation notes reveals teacher candidates (TCs) need more support with questioning, feedback and formative assessment, and technology tools to keep students engaged when teaching at a distance. Recommendations suggest a need for teacher preparation programs to provide TCs with opportunities to practice engagement strategies in distance education settings.

Keywords: emergency remote teaching, student engagement, student interaction, teacher candidate, teacher preparation, pandemic

In March 2020, the COVID-19 pandemic led to social distancing and nationwide school closure. As a result, emergency remote teaching (ERT) ensued, marking a temporary shift of instruction to an alternate mode of delivery due to the crisis circumstances (Hodges et al., 2020). While widespread use of ERT appears to be a consequence of the COVID-19 pandemic, distance education (DE), an organized instructional program in which teachers and learners are physically separated (Keegan, 1980), is a pre-pandemic phenomenon. For example, the states of Michigan, Alabama, New Mexico, and Idaho passed legislation nearly a decade ago requiring that all K-12 students complete distance learning experiences before graduating from high school (Kennedy & Archambault, 2012). In addition, most states in the United States (US) offer free virtual schooling alternatives for students who cannot attend in person (Littlefield, 2020). However, only 1.3 % of teacher preparation programs (TPPs) in the US address DE in teacher preparation coursework (Barbour et al., 2014).

While ERT is often used synonymously with the terms DE or distance learning (students learning at a distance), they describe different settings. The primary objective of ERT is to provide students with *temporary* access to instruction, while DE aims to provide students with educational content and interaction for a designated period (Hodges et al., 2020). Additionally, DE requires that instructors prepare, plan, and design instruction well in advance of enactment, while ERT marks a quick, unplanned transition of instruction to an alternative mode of delivery. During the pandemic, sudden shifts to ERT led to stigmas about DE as lower quality and efficacy than face-to-face instruction. However, studies have shown no significant difference in learning outcomes between face-to-face and DE (Zhao et al., 2005).

As the nation continues to grapple with the COVID-19 pandemic, higher education institutions are shifting courses entirely online or partially online (hybrid) for the unforeseeable future. While the spring 2020 semester utilized ERT because instructors transitioned to online instruction haphazardly mid-semester, the fall 2020 semester, instructors had time to prepare for online course content to deliver DE. Nevertheless, the transition from ERT to intentional, well-planned DE did not trickle down to Hawaii's public-school sector (Lee, 2020; Hawai‘i Department of Education, 2021). Hawai‘i public schools continue to use ERT due to increases in the number of positive COVID-19 cases and concerns about the recent Delta variant.

In the fall 2020 semester, we (two university field supervisors) transitioned a face-to-face field experience course to DE. Ten undergraduate, third-semester elementary education teacher candidates (TCs) worked with mentor teachers to provide online synchronous ERT to elementary (K-5) students. Even though TCs had two previous semesters of face-to-face K-5 classroom instruction, distance learning was a new instructional format. Throughout the fall 2020 semester, TCs met with us to complete course assignments to reflect on their ERT experiences. At this time, we noticed that TCs overwhelmingly perceived difficulties with promoting online student engagement. To support TCs' pedagogical needs, we merged tenets of practitioner inquiry with situated learning to guide formal research design.

This research seeks to fill a gap in the literature on TCs' experiences with ERT to inform TPPs on the skills TCs need to prepare for teaching online during a global crisis. This phenomenology seeks to investigate the following research question: What are undergraduate TCs' experiences promoting elementary student engagement during ERT?

Conceptual Framework

Practitioner inquiry is the systematic study of an educational problem or experience where the practitioner is the researcher. The professional context is the research site, and the practice itself is the focus of the study (Dana & Yendol-Hoppey, 2019). To inquire about teaching and learning, practitioners analyze instructional experiences to identify and explicitly work on questions that matter most to students. Practitioner inquiry aims to develop alternative ways to understand, assess, and improve teaching and learning so students benefit. Through collaboration, practitioners pose questions and gather data to become students of teaching. Two or more teachers work together to examine personal assumptions and collect and analyze data to develop local knowledge.

Situated learning theory (Lave & Wenger, 1991) proposes that learning is inseparable from real-world activity. In a situated learning context, one engages with or experiences an authentic activity. Authentic activities are real-life situations that involve cognitive apprenticeship, where an expert models a concept or skill to a learner, then slowly releases learner support so the learner may demonstrate acquisition of knowledge. A cognitive apprenticeship occurs when a TC works with a mentor teacher. Situated learning occurs because there is a transfer of knowledge from mentor to mentee (Catalano, 2015). The TC changes due to experiences within the classroom and can act to affect and modify the classroom. Thus, the complexities of learning lie not solely within the TC but in the complex dynamics of the TC-classroom interaction.

Literature Review

At the onset of the pandemic, survey research investigated school ERT readiness and the challenges perceived in education, exposing that teachers felt overwhelmed with online learning resources and the number of online tools available (Alea et al., 2020; Huber & Helm, 2020). Teachers also reported challenges with students' poor internet connections and students' overall lack of preparation for ERT (Trust & Whalen, 2020). However, before the pandemic, scholars (Foulger et al., 2017; Lai, 2017) shed light on students' and educators' lack of preparation and competency for DE, calling for technology-infused programs to address teaching with technology throughout the curriculum. However, programs such as TPPs and school districts must make several considerations before implementing quality DE.

First, educators need time to adapt to new technologies before being expected to use them effectively. For example, Teachers' Technological Pedagogical Content Knowledge (TPACK) emphasizes that it is essential for teachers to use (a) technological knowledge (TK), (b) pedagogical knowledge (PK), and (c) content knowledge (CK) to teach students online effectively (Koehler & Mishra 2009). All three components of TPACK work together to make distance learning successful. For example, a teacher may have a keen understanding of content (CK) and can maintain online student behavior and engagement (PK) but does not understand how to properly use the online software to disseminate a lesson to students (TK). TPACK emphasizes the binding relationship between technology, content, and pedagogy as the key to an online teacher's success. Teachers may use TPACK to reflect on the three components of their lesson (CK, PK, and TK) and analyze where improvement is needed. In addition, schools may use TPACK to guide teacher professional development where teachers join a professional development group according to their needs in CK, PK, or TK.

Second, educators need to be provided with unstructured professional developments (such as mentoring or online forums) to explore learner-centered activities that will help them teach *with* technology (Trust & Whalen, 2020; Zweig & Strafford, 2016). Teachers who feel comfortable and competent in the technology tools include more technology for student use in their instructional practices and are more prone to adopt technological advancements in their instruction (Ertmer, 1999). Teachers need tech support available to them when they need it to resolve technology issues immediately. Additionally, teachers need to engage with colleagues in a community of practice. DE educators often feel lonely and isolated from a lack of interaction with teaching peers and the absence of real-time feedback (Zhang, 2020). An online or face-to-face community space mitigates teacher's feelings of isolation, so they may collaborate, plan instruction, and discuss challenges they are facing (Kear et al., 2012).

Third, educators need fluid communication skills. Even though popular opinion proposes that teachers can switch instruction mediums and just "jump right in", this is not the case (Davis & Roblyer, 2014). Teacher involvement is the most defining difference between DE and face-to-face education. In the face-to-face setting, the teacher delivers content live to students and can interact with students inside and outside the classroom. But, in DE, the level of teacher involvement varies greatly depending on the program used to teach, the size of the online classroom, and if the teacher has the time to interact with students independently once content is delivered (Zhao et al., 2005). DE educators need to know when they need to interact with a small group of students versus giving whole group lectures and need to have the communication skills required for both learning contexts.

Lastly, teachers need to know how to engage students with technology so students stay interested while learning about a given skill or concept. As noted by Lai (2017), a "supportive online learning environment entails teachers using effective pedagogical practices to support their students and develop a positive teacher-student relationship to foster learner motivation and engagement" (pp. 322–323). Teachers need to design DE activities that (a) provide students with flexibility and autonomy, (b) allow for students to exhibit skill, concept and subject competence, and (c) relate to students' interests and learning needs (Stroet et al., 2013). Students achieve autonomy and flexibility in DE environments when teachers give students agency (choice in deciding on the technological tools they will use to complete course assignments). Student competence occurs when there is a straightforward course structure where instructors provide constructive, encouraging, and guiding feedback rather than evaluative feedback.

Student Engagement and Interaction

Student attitude, classroom climate, motivation, and self-regulated learning conceptualize student engagement (Fredricks, 2011). In a face-to-face instructional setting, engagement includes behavioral, emotional, and cognitive factors (Fredricks, 2011; Wang & Eccles, 2013). However, DE requires a different conceptualization of engagement since the learner has fewer opportunities to engage with the instructor (Martin & Bolliger, 2018). DE student engagement provides incentives to motivate student participation, interaction with teachers, peers, and course material (Al-Freih, 2021).

In DE literature, student engagement is students' interaction with the online classroom environment (Anderson, 2003). Others (Wagner, 1994) provide a more technical description, describing DE student interaction as a "reciprocal event that requires at least two objects and two actions…[that] mutually influence one another" (p.8). Other literature identifies three DE interaction types: student-content (SC), student-teacher (ST), and student-student (SS)

(Bernard et al., 2009; Moore, 1989). SS interaction refers to interaction among individuals or students working in small groups and is the more desirable interaction type for offering students cognitive and motivational support (Anderson et al., 2000). ST interaction focuses on the classroom-based dialogue between student and teacher where the teacher motivates the student to learn or seeks to stimulate student interest (Moore, 1989). Finally, SC interactions develop students' mental and physical skills by providing opportunities for students to connect with the subject matter under study. Some examples of SC interaction include reading informational texts, watching videos, or using simulations.

Researchers may use Moore's (1989) three interaction types to identify DE student interaction types. However, Anderson (2003) suggests that as long as a distance educator utilizes one of Moore's interaction types "at a high level" (p.4), the other interaction types are not essential. Nevertheless, Anderson does not describe how to determine if student interaction occurs with high-level fidelity. Others (Martin & Bollinger, 2018) underscore the importance of the instructor's presence in promoting students' DE interactions, arguing that it is not the type of interaction that matters but how the instructor provides the interaction type.

Gaps in the Literature

Educational technology research reports on the benefits that student engagement and interaction have on student outcomes (Al-Freih, 2021; Lear et al., 2010), describes DE interaction types (Anderson, 2003; Moore, 1989) and the skills teachers need to be effective DE educators (Zhao et al., 2005). However, most literature reports on higher education contexts where DE educators enact instruction for adult learners. In addition, educational technology literature on DE in K-12 teaching and learning contexts is limited. This research has the potential to inform TPPs on the skills TCs need to be effective DE educators. Therefore, this study seeks to fill a gap in the literature and investigate undergraduate TCs' experiences promoting elementary DE student engagement.

Methods

Phenomenology investigates the commonality of a lived experience within a particular group. The purpose is to describe the universal essence of individuals' experiences with a phenomenon (Moustakas, 1994). Classic phenomenology focuses on first-person experiences and intentionality (direction of experience towards things in the world) and aims to understand how established ways of seeing are brought into being. Understanding the researchers' presuppositions of the phenomenon is a central feature of phenomenological research. *Epoché* or bracketing (Moustakas, 1994) is done with memos or reflexivity (Rodham et al., 2015). To ensure trustworthiness in phenomenological research, the researchers need to develop a curious stance towards the data. Bracketing requires that researchers engage in reflexivity to become mindful of their biases and personal experiences on the research to self-monitor the impact of their role in creating knowledge.

Statement of Reflexivity

The first author is an Assistant Professor in multilingual learning, elementary education. Her beliefs about student engagement and distance learning align with sociocultural concepts of learning and come from her personal experiences growing up as a bilingual and her professional experiences working as a former public elementary school teacher and now a university professor. She believes tools mediate learning and that humans make sense of the world through their interactions with others. The first author is knowledgeable in preparing TCs to design and enact instruction for multilingual students or those who speak a language other than English as a first language. The first author also assumes it is easier to teach in person than it is to teach virtually.

The second author is supervising TCs in the field and works as a lecturer in education. A former elementary and secondary teacher, she developed an immersive language program and engaged her students in numerous art projects. Her teaching philosophy is grounded in social constructivism while providing students a sense of self-efficacy and agency. COVID-19 required a transfer to distance learning and the second author explored synchronous and asynchronous teaching formats. The second author believes that teaching in a distance learning environment requires educators to adapt to new modes of instruction, student participation, and engagement.

Research Context

The TPP uses a cohort model and is a four-semester long program that results in an undergraduate degree in elementary education with initial teacher licensure. TCs begin the TPP in their junior year of college and complete four semesters of field experiences and methods course instruction. A cohort coordinator assigns TCs to an elementary school and mentor teacher. Semesters one through three include two days of field experience and methods course instruction. Semester four includes full-time student teaching. In addition, a university field supervisor observes TCs in the field and teaches bi-monthly seminars. Seminars provide opportunities for TCs to discuss field experience events, review relevant theory, receive lesson plan support, engage in reflective practice, and practice instructional strategies.

On March 13. 2020, nationwide lockdown ensued. Face-to-face instruction in the spring 2020 semester abruptly transitioned to ERT. Prior to the pandemic, TCs worked with a mentor teacher two days per week to provide face-to-face instruction to K-5 students. When the pandemic hit the US, TCs were in the midst of the second semester field experience. Mentor teachers and TCs transitioned K-5 instruction to ERT with no preparation in DE.

At the same time, universities worked to accommodate the quick escalation of ERT. TPP instructors referred to empirical research on online field experiences. According to the literature (Hixon & So, 2009), there are three types of online field experiences. *Type I* is characterized by concrete, direct experiences where the TC works in a live classroom setting with a mentor teacher and students. In a Type I field experience, TCs interact face to face with students and mentor teachers. Technological tools, such as video conferencing, videotaping, and video analysis facilitate supervision, reflection, and communication. In *Type II* virtually enhanced field experiences, TCs observe students and teachers in classrooms remotely through video conferencing software such as *Zoom*, *Google Meets*, and *WebEx*. Examples of Type II field experiences are synchronous lesson observations or non-real-time pre-recorded videos. Simulated environments create Type III virtually-enhanced field experiences. In *Type III* “virtual practicums,” TCs learn about and practice pedagogy using artificial reality, such as publicly available video-recorded lessons or via teaching channels. The different types of online field experiences may be used in conjunction with one another. For example, TCs may remotely observe a classroom (Type II experience) for an assignment while working in direct (Type I) field experience. The TCs reported on in this research participated in Type II field experiences because they worked with mentor teachers and students to enact instruction at a distance. Likewise, all collaboration and dialogue between mentors, students, and university supervisors took place online.

Participants

The research took place when TCs were in the third semester of the TPP. At the start of the fall 2020 semester, instruction remained online. However, because of the pandemic, TCs were allowed to choose one of three pathways to complete the third-semester (spring 2021) field experience requirement: (1) work face-to-face with a mentor teacher, (2) work online with a mentor teacher, or (3) complete field simulation tasks provided by the field supervisor with no mentor teacher assignment. As a result, 10 of 18 TCs (56%) chose to work with a mentor teacher, either face-to-face or remotely, and are reported on in this research because they planned and enacted ERT for students.

The ten TCs reported on in this study completed two formal observations that included a lesson pre-conference, formal observation, and post-conference. To accommodate online instruction, field supervisors (the authors) joined TCs' live online lessons or watched a recording of TCs' lessons. TCs used the university's lesson plan template to plan online instruction and chose a Charlotte Danielson Framework (CDF) (Danielson, 2013) domain as a professional development goal for their lesson. The public school system used the CDF for in-service teacher evaluation. The TPP used the CDF to develop TCs' fluency in CDF language for professional development . For instance, TCs shared their lesson plan with their field supervisor, using CDF language to establish a goal (i.e., I want to create a culture for learning). Then the field supervisor used TC's CDF goal to provide lesson plan suggestions in pre-conference meetings. TCs were required to record their lesson (even if the field supervisor attended their live instruction online) and analyze their video using a video reflection framework (Smith, 2019). To culminate the formal observation assignment, TCs used the video reflection framework to guide post-conference reflective dialogue with the field supervisor.

Data Collection

Data was collected over a 16-week semester of online instruction in fall 2020. Data included 20 TC lesson reflections (two per TC). TCs completed written reflection prompts on their enacted DE lessons. TCs recorded ERT using the *Screen-Cast-o-Matic* (a web-based screen recording tool), then reflected on their audio-video recording using a free video annotation tool (v-note.org). Written reflection prompts asked TCs to comment on how their DE met or did not meet a personal pedagogical goal. TCs referred to the CDF (Danielson, 2013) to select a pedagogical goal. Before watching and analyzing their video, TCs decided what instructional elements would evidence them having met or not having met this goal. Secondary data sources included university supervisor notes (n=40). We (the authors and university field supervisors) took field notes when meeting with TCs in post-lesson conferences. During post-lesson conferences, TCs read excerpts from their written reflections to us, described how they analyzed their video recording, and played segments of their recording to exemplify what they noticed about their instruction and how it related to their CDF goal.

Explicitation of the Data

Phenomenology uses the term "explicitation" instead of "analysis" because when we analyze we break data into smaller discernable parts instead of keeping it whole (Groenwald, 2004). We used Interpretative Phenomenological Analysis (IPA) to explore TCs' experiences with promoting elementary student engagement during ERT. IPA involves a light form of thematic analysis where the data are kept intact through a process of phenomenological reduction or bracketing so a phenomenon may become evident (Smith & Osborn, 2015). While IPA is a popular analysis approach, it is essential to note that a researcher's values, perceptions inevitably influence the process of understanding someone else's life or experience, and biases, a process known as "double hermeneutic" (double interpretation) because the researcher is trying to make sense of the participant trying to make sense of their own experience (Rodham et al., 2015). To address IPA methodological issues, we referenced personal interpretive resources throughout the entire explicitation process. In addition, we read our statements of reflexivity before reading the data and took memos as we read the data to "bracket" prior experiences from coloring our interpretations of the data to ensure trustworthiness.

Data were analyzed in three stages: initial note-taking, transferring notes into themes, and connecting themes to generate findings. During initial note-taking, we employed a free analysis approach (Smith & Osborn, 2015) to explore TCs' experiences with promoting online student engagement. To do so, we opened TCs' lesson reflections on *Google Docs* and created a three-column table (listing initial notes, data excerpts, and themes). We independently reviewed lesson reflections applying *In Vivo* (Saldaña, 2021) codes to the left-hand column to capture participants' own words. We used the comment tool on Google Docs to take memos to document our emotions, interpretations, and thoughts about the data to bracket personal biases as we took initial notes.

In stage two, we met to discuss our In Vivo codes using an idiographic analytical approach (Grbich, 2013) to transfer notes into themes. This allowed us to focus on the unique experiences of each participant prior to moving towards general claims. Our conversation resulted in three themes: "questioning", "formative assessment", and "technology" (Figure 1). At this time, we noticed that data within each theme contained the following chronological pattern: the participant describes an engagement strategy, the participant reveals their beliefs about the strategy, the participant shares the challenges they encountered enacting the engagement strategy, the participant considers a plan of action for improvement. Seeing this pattern, we chose to add the following subthemes: "beliefs", "challenges", "plans" to our codebook.

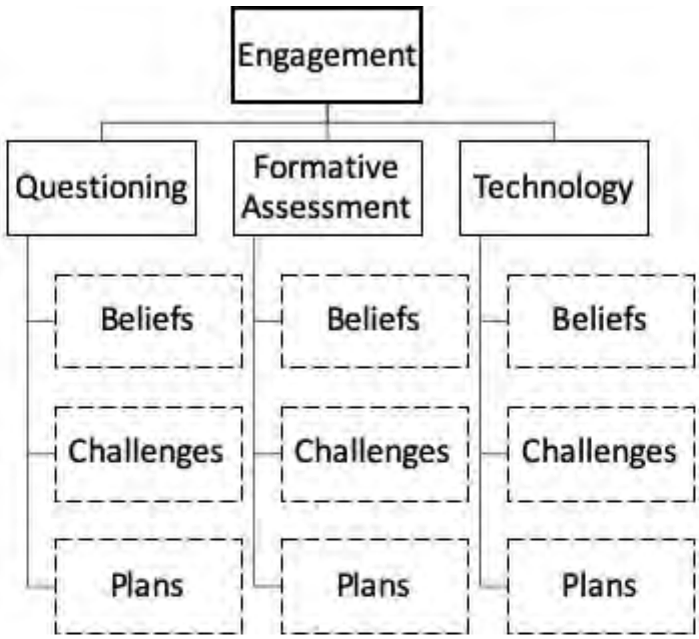


Figure 1: Transferring notes into themes.

In stage three, we looked for patterns across themes and subthemes to generate findings. First, we used the themes and subthemes to independently re-code reflections. Then we met to discuss our coding, noting similarities and differences for inter-rater reliability. When we encountered differences, we turned to our observation notes for clarification. In three instances, our observation notes were not enough to reconcile our coding discrepancies, so we phoned participants for member-checking (Lincoln & Guba, 1984). We took notes as participants clarified their experiences then connected themes and subthemes to represent findings. We continued this process until 100% inter-rater reliability was achieved and used the themes: questioning, formative assessment, and technology to report on TCs experiences with student engagement in ERT.

Findings

This study explored TCs experiences with student engagement in ERT. To understand TCs' experiences with ERT as a phenomenon, we examined the engagement strategies TCs utilized, the beliefs TCs had about the strategies they used for student engagement, challenges TCs perceived, and any plan of action for subsequent instruction.

Questioning

The questioning theme involved instances where TCs relied on question-response student-teacher (ST) interactions. TCs used questioning to promote student engagement and added rigor to online lessons as a way to invite students to think critically. For example, one TC explained, "*[Students] had the most attention when I asked them questions. They could even regain their focus if I asked them to answer a question or give me an answer to a problem*" (TC, September 2020).

Another TC shared, *"I should have worked with a small group of struggling students so that they could ask clarifying questions and get help with some of the workbook questions."* (TC, September 2020). Overall, TCs expressed the belief that questions were an essential engagement strategy that teachers should use to keep students on task as a way to help "struggling" students receive content support.

TCs combined strategies they learned about in their methods courses with questioning. TCs used think-aloud and open-ended questions such as: "*What should I do now? Can you help me*?" along with Total Participation Techniques (TPTs) such as "think-pair-share" and "thumbs-up" an attempt to keep students engaged online.

TCs noticed TPTs did not work in the online setting with the same tenacity as the face-to-face setting. This was largely due to the fact that TCs did not use breakout rooms and relied on whole group question-response. One TC explained, "*I model how to solve the problems then give them [students] time to practice on their own, but it's hard to create the "pair" time online. I know I should do breakout rooms, but I don't think I can do this on WebEx* [distance learning software used by the school]" (TC, September 2020).

TCs revealed they needed to reconceptualize the participation strategies they learned about in their methods classes to "fit" the online platform they were using. The reconceptualization involved considerations about how to support student engagement online and thinking about how the features and limitiations of the online leanring platforms (Webex; Google classroom) affected instructional intentions. TCs shared feelings of being un-prepared and unsure of their teaching competency, evidencing a lack of technological knowledge (TK) in the TPACK framework. As one TC shared, "*I know I have to get better at teaching online. I don't feel confident anymore* (TC, October 2020)."

TCs expressed strong desires to improve on the types of questions they used to promote online student engagement. One TC summed it up, "*I should have added more variety to my questions.*

In this way, I could introduce new vocabulary and expose students to different styles of questioning" (TC, December 2020). As TCs taught online, live video feed of K-5 students, provided evidence of student disengagement; this evidence prompted TCs to see a need for improvement. As one TC expressed, "*I can see them doing other things as I'm speaking, so I know I need to improve on how I ask them questions to keep them interested in the lesson*" (TC, November 2020).

An overall feeling of improving questioning techniques via online tools that could promote SS interactions was expressed by all TCs. For example, TCs shared a desire to use breakout rooms to engage students in collaborative discussions, while four TCs shared they wanted to use the chat feature to keep track of student responses.

Feedback and Formative Assessment

TCs used formative assessment in the form of feedback to promote online student engagement. Praise such as "*thank you*" or "*good job*" kept students interested in a lesson. However, TCs noted that praise was not enough to develop online student discussions for critical thinking. For instance, one TC stated: "*I'd like to take the comments my students make and have them elaborate or turn their comments into a discussion. Instead of saying a mere thank you, I'd like my feedback to become more detailed"* (TC, November 2020). TCs expressed that they used ST interaction and wanted to move towards student-to-student (SS) interaction. A noteworthy finding that may mean that TCs learn how to facilitate SS interaction online once TK and PK are acquired.

In other instances, TCs shared beliefs about the role the teacher has in facilitating SS interactions and their lack of TK to do so. For instance, TCs analyzed students' independent work samples to keep track of student progress and understanding. A TC shared, "*I use practice problems to see what students know how to do before I move on*" (TC, October 2020). However, when completing practice problems online, TCs noted that they did most of the talking and shared that they needed to learn how to facilitate students sharing their work online.

A TC explained, "*It's hard for me to create a way for students to work together in small groups online. It's always a whole group with me doing most of the talking. I know I need to change this*" (TC, November 2020).

Formative assessment became challenging for TCs to implement online: "*I created a rubric to use to check off when students complete a specific task, but I cannot see each student individually, so I observe a few students at a time; usually the ones I know are struggling*" (TC, October 2020). The distance learning environment challenged TCs in keeping track of student progress during instruction. While some TCs used online games like *Kahoot*! to formatively assess students, TCs noticed they needed to rethink formative assessment in their lessons to consider the use of technology as an instructional tool. As one TC shared, "*I can't walk around to check on the progress anymore and offer support*" (TC, October 2020). TCs' experiences revealed that while they wanted to use technology to support all students, they were only able to help students who were vocal about their learning needs.

Technology

Respective schools provided video conferencing apps such as WebEx or Google Me*et* as distance learning platforms. TCs frequently used the chat feature and camera features to keep track of student engagement in the form of verbal responses and facial gestures. Other candidates relied on the Google Suite and used Google Slides to present lesson content with animations (pictures and emojis on the slides). TCs felt confident using the Google suite for instruction. Still, they expressed issues with internet connectivity (the video freezing), visibility (not being able to see all students in a grid view), and multitasking (showing the Google Slides and teaching while also observing students). One TC expressed frustration, "*It's so hard to see all of the students when I'm teaching*," "*I can't see everything on my one screen*" (TC, October 2020). Another TC voiced, "*I could not successfully explain to the students how to take a screenshot"* (TC, December 2020).

Feelings of frustration led TCs to notice the importance of practicing DE lessons in advance to detect and solve technical problems. TCs explicitly stated that they practiced distance learning lesson procedures at home "*with stuffed animals*," "*with a sibling,*" or "*with a peer*" before they taught their lessons to students, albeit home practice did not mimic the reality of the virtual classroom. As one TC shared, "*I practice my lessons in advance, but on the day of my lesson, everything changed. I needed to keep 15-second graders engaged. This is something I can't practice at home*" (TC, December 2020). TCs expressed that they would like to improve their multitasking skills. Improvement plans included monitoring students via camera and chat through an additional screen or asking their mentor teacher what they do to multitask.

Discussion

The following research sought to investigate undergraduate TCs' experiences with promoting elementary student engagement during ERT. Findings reveal TCs focused on questioning, feedback and formative assessment, and technology to keep students engaged online.

When using questioning, TCs heavily relied on student-teacher (ST) interactions that only motivated some students to participate in the discussion. TCs noticed a need to create more student-to-student (SS) interactions. They began to consider breakout rooms or group chats but were missing the technological knowledge needed to enact these ideas. Even though scholars argued that ST interaction has the highest perceived value among students (Anderson, 2003; Martin & Bollinger, 2018), this research showed that SS interaction is critical when working at a distance. Agreeing with Lear et al. (2010), findings from this research support the notion of building interactivity and community online to promote student engagement and learning outcomes.

When it came to feedback and formative assessment, TCs relied on strategies learned in methods coursework that worked in the face-to-face setting (i.e., praise). However, when teaching synchronously at a distance, TCs noticed that they could not monitor students' work. TCs missed opportunities to provide "struggling" students with feedback because students were silent, not working, had their cameras off, or the TC did not notice the student was experiencing difficulty. TCs began to consider technological tools that could help them monitor student work (i.e., Jamboard) and use explicit feedback to promote students' cognitive engagement. Findings from this research reiterate the need for TPPs to reconceptualize teacher competency. Teacher competency should include a teacher's ability to teach at a distance. A suggestion is for TPPs to consider using the TPACK framework (Koehler & Mishra, 2009) to provide TCs with feedback when working in the field or as a TC self-assessment tool. With TPACK, TCs may learn the importance of demonstrating interactions between content, pedagogy, and technological bodies of knowledge.

The technology theme revealed TCs' lack of understanding of DE and the overall need for TK professional development. TCs shared that they needed to practice using technology more and believed that they needed to learn how to "multitask" with multiple screens to see students while delivering content. While TCs demonstrated content (CK) and pedagogical (PK) knowledge, findings suggested that TCs struggled with applying PK and CK in DE due to a lack of TK. Teachers need to be able to use technology for student instruction when graduating from a TPP. An implication for the field is for TPPs to apply Lave's and Wenger's (1991) situated learning theory to provide TCs with ample situations to practice DE; this could include virtual simulations of a DE classroom. Davis and Roblyer (2005) recommended that TCs learn about DE while in a TPP. However, current frameworks used to evaluate TCs who are working in the field do not contain language about DE (Danielson, 2013). TPPs need to address the ways TCs teach with technology (Foulger et al., 2017) and should use teaching evaluation frameworks that address TCs' DE competence.

A limitation of this study is the sample size; therefore, we want to reiterate that findings from this study are made to provide a description only and should not be used to make general claims about other TCs. Additionally, because this research used a Type II online field experience where TCs did not have direct contact with us, their mentor teachers, or students, we acknowledge that TCs' experiences with teaching at a distance may have been different if intermittent face-to-face interactions were included. More research is needed to highlight how TCs may be prepared for distance learning and should consider larger sample sizes, and face-to-face interactions or online community-building spaces for mentor teachers, TCs, and university supervisors. Additionally, more empirical research is needed examine the connections between SS, and TK and PK to describe how TCs obtain TK and PK while in a TPP. More research is also needed on K-12 synchronous DE formats. A continuation of this study would be an examination of how TCs use SS interaction in DE to work with K-12 students.

Conclusion

This phenomenological study reports on TCs' experiences with promoting elementary student engagement during ERT. Findings report on TCs' use of questioning, feedback and formative assessment, and technology engagement strategies. Recommendations are made to support a need for DE evaluation tools in TPPs, along with the addition of explicit opportunities for TCs to teach at a distance.

References

Alea, L. A., Fabrea, M. F., Roldan, R. D. A., & Farooqi, A. Z. (2020). Teachers' Covid-19 awareness, distance learning education experiences and perceptions towards institutional readiness and challenges. *International Journal of Learning, Teaching and Educational Research*, *19*(6), 127–144. https://doi.org/10.26803/ijlter.19.6.8

Al-Freih, M. (2021). The impact of faculty experience with emergency remote teaching: An interpretive phenomenological study. *IAFOR Journal of Education 9*(2), 7–23. https://doi.org/10.22492/ije.9.2.01

Anderson, T. (2003). Getting the mix right again: An updated and theoretical rationale for interaction. *The International Review of Research in Open and Distance Learning*, *4*(2), 1–14. https://doi.org/10.19173/irrodl.v4i2.149

Anderson, T., Garrison, D. R., Archer, W., & Rourke, L. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education 12,* 8–22. https://telearn.archives-ouvertes.fr/hal-00197319

Barbour, M. K., Siko, J., Gross, E., & Waddell, K. (2014). Virtually unprepared: Examining the preparation of K-12 online teachers. In Hartshorne, R., Heafner, T.L., & Petty, T.M. (Eds.), *Teacher education programs and online learning tools* (pp. 187–208). Hershey, PA: IGI Global.

Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M. A., & Bethel, E. C. (2009). A meta-analysis of three types of interaction treatments in DE. *Review of Educational Research, 79*(3), 1243–1289. https://doi.org/10.3102/0034654309333844

Catalano, A. (2015). The effect of a situated learning environment in a DE information literacy course. *The Journal of Academic Librarianship*, *41*(5), 653–659. https://doi.org/10.1016/j.acalib.2015.06.008

Dana, N. F., & Yendol-Hoppey, D. (2019). *The reflective educator's guide to classroom research: Learning to teach and teaching to learn through practitioner inquiry*. Thousand Oaks, CA: Corwin.

Danielson, C. (2013) Teaching Evaluation Tool. Accessed from: http://www.loccsd.ca/~div15/wp-content/uploads/2015/09/2013-framework-for-teaching-evaluation-instrument.pdf

Davis, N. E., & Roblyer, M. D. (2005). Preparing teachers for the "Schools that technology built" Evaluation of a program to Train teachers for virtual schooling. *Journal of Research on Technology in Education*, *37*(4), 399–409. https://doi.org/10.1080/15391523.2005.10782445

Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: Strategies fortechnology integration. *Educational Technology Research and Development, 47*(4), 47–61. https://doi.org/10.1007/BF02299597

Fredricks, J. (2011). Engagement in school and out of school context: A multidimensional view of engagement. *Theory Into Practice, 50*(4), 327–335. https://doi.org/10.1080/00405841.2011.607401

Foulger, T. S., Wetzel, K., & Buss, R. R. (2019). Moving toward a technology infusion approach: Considerations for teacher preparation programs. *Journal of Digital Learning in Teacher Education, 35*(2), 79–91. https://doi.org/10.1080/21532974.2019.1568325

Grbich, C. (2013). *Qualitative data analysis: An introduction*. Thousand, Oaks, CA: Sage.

Groenewald, T. (2004). A phenomenological research design illustrated. *International Journal of Qualitative Methods*, *3*(1), 42-55. https://doi.org/10.1177/160940690400300104

Hawaiʻi Department of Education. (2021). *Return to learn*. https://www.hawaiipublicschools.org/ConnectWithUs/MediaRoom/PressReleases/Pages/school-year-2020-21.aspx

Hixon, E., & So, H. J. (2009). Technology's role in field experiences for preservice teacher training. *Educational Technology & Society*, *12*(4), 294–304.

Hodges, C.B., Moore, S., Lockee, B. B., & Bond, M. (2020, March 27). The difference between emergency remote teaching and online learning. *Edu-CAUSE Review.* https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Huber, S. G., & Helm, C. (2020). COVID-19 and schooling: evaluation, assessment and accountability in times of crises—reacting quickly to explore key issues for policy, practice and research with the school barometer. *Educational Assessment, Evaluation and Accountability*, *32*(2), 237–270. https://doi.org/10.1007/s11092-020-09322-y

Kear, K., Chetwynd, F., Williams, J., & Donelan, H. (2012). Web conferencing for synchronous online tutorials: Perspectives of tutors using a new medium. *Computers & Education, 58,* 953–963. https://doi.org/10.1016/j.compedu.2011.10.015

Kennedy, K., & Archambault, L. (2012). Offering preservice teachers field experiences in K-12 online learning: A national survey of teacher education programs. *Journal of Teacher Education*, *63*(3), 185–200. https://doi.org/10.1177/0022487111433651

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary Issues in Technology and Teacher Education*, *9*(1), 60–70. https://doi.org/10.1007/s11528-017-0176-2

Lai, K.W. (2017). Pedagogical practices of NetNZ teachers for supporting online distance learners. *Distance Education, 38*(3), 321–335, https://doi.org/10.1080/01587919.2017.1371830

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Lear, J. L., Ansorge, C., & Steckelberg, A. (2010). Interactivity/community process model for the online education environment. *Journal of Online Learning and Teaching, 6*(1), 71–77.

Lee, S. (September, 2020). Hawaii Schools To Continue Mostly Distance Learning Through December. *Honolulu Civil Beat.* https://www.civilbeat.org/2020/09/hawaii-schools-to-continue-distance-learning-through-december/

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.

Littlefield, J. (2020, August 27). State-by-state list of free online public schools, K–12. *ThoughtCo.* https://www.thoughtco.com/free-online-public-schools-4148138

Martin, F., & Bolliger, D. U. (2018). Engagement matters: Student perceptions on the importance of engagement strategies in the online learning environment. *Online Learning*, *22*(1), 205–222. https://doi.org/10.24059/olj.v22i1.1092

Moore, M. G. (1989). Three types of interaction. *The American Journal of Distance Education, 3*(2), 1–6. https://doi.org/10.1080/08923648909526659

Moustakas, C. (1994). *Phenomenological research methods*. Thousand Oaks: CA: Sage

Rodham, K., Fox, F., & Doran, N. (2015). Exploring analytical trustworthiness and the process of reaching consensus in interpretative phenomenological analysis: Lost in transcription. *International Journal of Social Research Methodology*, *18*(1), 59–71. https://doi.org/10.1080/13645579.2013.852368

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.), Thousand Oaks, CA: Sage.

Smith, M. G. (2019). A video-mediated critical friendship reflection framework for ESL teacher education. *TESL-EJ*, *23*(1),1–18.

Smith, J. A., & Osborn, M. (2015). Interpretative phenomenological analysis as a useful methodology for research on the lived experience of pain. *British Journal of Pain*, *9*(1), 41–42. 10.1177/2049463714541642

Stroet, K., Opdenakker, M., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review, 9*, 65–87. https://doi.org/10.1016/j.edurev.2012.11.003

Trust, T. & Whalen, J. (2020). Should teachers be trained in emergency remote teaching? Lessons learned from the COVID-19 pandemic. *Journal of Technology and Teacher Education*, *28*(2), 189–199.

Wagner, E. D. (1994). In support of a functional definition of interaction. *American Journal of Distance Education*, *8*(2), 6–29. https://doi.org/10.1080/08923649409526852

Wang, M. T., & Eccles, J. S. (2013). School context, achievement motivation, and academic engagement: A longitudinal study of school engagement using a multidimensional perspective. *Learning and Instruction*, *28*(1), 12–23. https://doi.org/10.1016/j.learninstruc.2013.04.002

Zhao, Y., Lei, J., Yan, B., Lai, C., & Tan, S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record*, *107*(8), 1836–1884. doi.org/10.1111/j.1467-9620.2005.00544.x

Zhang, C. (2020). From face-to-face to screen-to-screen: CFL teachers' beliefs about digital teaching competence during the pandemic. *International Journal of Chinese Language Teaching, 1*(1), 35–52. https://doi.org/10.46451/ijclt.2020.06.03

Zweig, J. S., & Stafford, E. T. (2016). Training for online teachers to support student success: Themes from a survey administered to teachers in four online learning programs. *Journal of Online Learning Research, 2*(4), 399–418. https://files.eric.ed.gov/fulltext/EJ1148594.pdf

Corresponding author: Monica Gonzalez Smith
Email: monicags@hawaii.edu

Comparison of Students' Learning and Attitudes in Physical versus Virtual Manipulatives using Inquiry-Based Instruction

Onur Oymak
Deniz Military High School
Turkey

Feral Ogan-Bekiroglu
Marmara University
Turkey

Abstract

This research aimed to determine whether implementation of virtual technology or implementation of physical materials in a learning environment is more efficient in understanding physics concepts and developing positive attitudes at the high school level. The theory that framed this study is the model of learning as dynamic transfer. Participants were 96 ninth grade students (n = 96) distributed randomly to the virtual or physical group. Inquiry-based instruction continued during teaching of kinematics and dynamics, which lasted for eight weeks for both groups. Data from the Force and Motion Achievement Instrument (FMAI), student worksheets, the Attitude Towards Physics Scale (APCS), and anecdotal observations were collected. This study concluded that the use of physical and virtual manipulatives in inquiry-based instruction had the same effect on students' conceptual and procedural knowledge, as well as their attitudes towards physics. However, students who dealt with physical experimentation had lower learning than their peers who experienced virtual experimentation due to measurement errors made by students. Furthermore, physical investigations left students with some irrelevant knowledge. Therefore, it is reasonable to assume that implementing virtual manipulatives is more advantageous for learning in some conditions. The final conclusion is that attitude and learning may be developed in a parallel manner.

Keywords: Attitude, learning, physics, physical manipulatives, virtual manipulatives

In an ideal situation, students are expected to formulate ideas that align with scientific explanations; however, significant constraints such as limitations of the laboratory environment work against this possibility (Marshall & Young, 2006). Educators should design the learning environment in a way that provides students with more experiences and more opportunities to understand the process of doing science so that it can facilitate learning (Vosniadou, Ioannides, Dimitrakopoulou & Papademetriou, 2001). Research suggests that laboratory and hands-on activities create effective learning environment to increase achievement in science knowledge and to influence attitudes toward science in a positive way when properly designed (Adesoji & Raimi, 2004; Freedman, 1997; Gibson & Chase, 2002). Both virtual and physical materials can be used during laboratory science activities for the construction of powerful learning environment.

Manipulatives are multisensory tools that represent ideas in more than one way to promote communication among students to enhance and deepen understanding (Shaw, 2002). De Jong, Linn and Zacharia (2013) stated that “both physical and virtual manipulatives can achieve similar objectives, such as exploring the nature of science, developing team work abilities, cultivating interest in science, promoting conceptual understanding, and developing inquiry skills, yet they also have specific affordances” (p. 305). The ability to change the values of variables and modify model characteristics (Ford & McCormack, 2000; Tao & Gunstone, 1999; Windschitl, 2000; Zacharia, 2003), make the “unseen” seen (Potkonjak et al., 2016), simplify complex and messy real-world models (Hennessy, Deaney & Ruthven, 2006; Hsu, 2008; Triona & Klahr, 2003; Trundle & Bell, 2010; Zacharia & de Jong, 2014), and conduct experiments about unobservable phenomena (Jaakkola, Nurmi & Veermans, 2011; Zacharia & Constantinou, 2008) are some of the advantages of using virtual manipulatives (VM). On the other hand, enabling learners to experience the challenges many scientists face (Balamuralithara & Woods, 2009; de Jong et al., 2013; Marshall & Young, 2006; Windschilt, 2000) and allowing them to acquire complexities and a sophisticated epistemology of science by dealing with unanticipated events and measurement errors (de Jong et al., 2013; Olympiou & Zacharia, 2012; Toth, Morrow & Ludvico, 2009) are some of the benefits of using physical manipulatives (PM).

Besides these pros, there are some cons of using virtual and physical manipulatives. Having obstacles for testing specific ideas and models in the micro-world (Roth, Woszczyna & Smith, 1996) and unfamiliar parameters (Marshall, 2002) are the constraints of virtual manipulative environment. Producing confusing and inconsistent feedback due to irrelevant information (Klahr, 2007) is the critical aspect of physical manipulative environment. Therefore, some researchers have found that VM enhance students' conceptual knowledge and attitudes of science more than PM. In contrast, other researchers have found the opposite effect. Due to the conflicting results in the literature, this research aimed to determine whether implementation of VM or PM in a learning environment is more efficient in understanding physics concepts and developing a positive attitude. In this context, the term virtual manipulative is used to refer to virtual technology such as computer-based simulations, videos, and e-books, whereas physical manipulative is used to refer to real-world concrete materials and instruments for this study. The research was interested in answering the following research question: What are the significance differences between using virtual manipulatives and physical manipulatives in terms of high school students’ learning of motion and force concepts and their attitudes towards physics? One promising method of promoting conceptual change in science learning is inquiry-based learning (Hofstein & Lunetta 2004; de Jong 2006). However, inquiry learning is only effective if students receive sufficient instructional guidance (Alfieri et al., 2011; Driver et al., 1994). Therefore, guided inquiry-based instruction was preferred for the context of the study.

Theoretical Background

The importance of learning is that it is responsible for all the skills, knowledge, attitudes, and values that are acquired by human beings (Gagne, 1977). Thus, both knowledge and attitude are important outcomes of learning. Research in physics education has revealed that students in diverse grade levels hold various conceptions of mechanics that are irreconcilable with the Newtonian understanding of motion and force (Clement, 1982; Finegold & Gorsky, 1991; Graham, Berry & Rowlands, 2013; Mildenhall & Williams, 2001; Rowlands, Graham, Berry & McWilliam, 2007). On the other hand, attitudes of students towards science are leading us towards a society with less and less scientific vocation (Aguilera & Perales-Palacios, 2020). The educational challenge is to promote positive attitudes towards science, which should be given priority in educational research (Osborne, Simon & Collins, 2003).

Transfer is the dynamic creation of associations between knowledge elements (Rebello et al., 2005). Transfer occurs when students use learning from one context in another (Reed, 1993; Singley & Anderson, 1989). According to Schwartz, Varma and Martin (2008), because learners need to go beyond their original learning to accomplish a conceptual change, dynamic transfer occurs when component competencies are coordinated through interaction with the environment to yield novel concepts or material structures. In other words, students can learn from interacting with complex, well-structured environments that may include tools, representations, other people, and so forth (Schwartz et al., 2008). Schwartz and colleagues state that a model for dynamic transfer of learning also implies that learners need to bring their attitudes which help determine whether or not they will engage the environment in productive ways. Consequently, the theory that framed this study is the model of learning as dynamic transfer (diSessa & Wagner, 2005). This model deals with (re)constructing knowledge in new context or environment (Rebello et al., 2005). This model was chosen because the study examines students' physics learning and attitudes towards physics when interacting with a virtual or physical manipulative environment. According to Rebello et al. (2005), dynamic transfer occurs when instruction attempts to change student knowledge, provides rich setting for students to express themselves, and involves groups of up to three students.

Empirical Studies Comparing Virtual and Physical Manipulatives

Researchers have compared the impact of using VM and PM on learning and attitude by taking their affordances into consideration. Plenty of research has discovered cases where using VM seemed to be as effective for student learning and attitude as using PM (Apkan, 2002; Darrah et al., 2014; Jaakkola & Nurmi, 2008; Klahr, Triona & Williams, 2007; Taghavi & Colen, 2009; Triona & Klahr, 2003; Zacharia, 2003; Zacharia & Constantinou, 2008; Zacharia & de Jong, 2014; Zacharia & Olympiou, 2011). For example, Taghavi and Colen (2009) compared and evaluated the effectiveness of computer simulated laboratory instruction versus physical laboratory instruction. Their results based on 22 college students, indicated students' attitudes were similar with regard to both the simulated and physical laboratory instruction. Similarly, Zacharia and Constantinou (2008) explored the effect of experimenting with physical or virtual manipulatives on 68 undergraduate students' conceptual understanding of heat and temperature. Their results showed that both modes of experimentation were equally effective in enhancing students' conceptual understanding. Correspondingly, there are some studies revealing that the use of VM facilitated student learning more than the use of PM (Bozkurt & Sarikoc, 2008; Finkelstein et al., 2005; Husnaini & Chen, 2019; Wang & Tseng, 2018). For instance, Husnaini and Chen (2019) investigated the effects of physical and virtual laboratories on conceptual understanding of 68 secondary school students. The participants conducted a pendulum experiment with guided inquiry-based approach. The researchers discovered that the virtual laboratory was more effective for improving difficult concepts than the physical laboratory. On the other hand, a few research studies produced opposite results, where PM created a more valuable experience than VM (Coramik, 2012; Marshall & Young, 2006). Marshall and Young (2006) studied three prospective teachers, working together in a group, as they used both Interactive Physics and physical manipulatives to explore what happens to the momentum of objects in collisions. According to their results, the participants took longer to execute cycles of exploration with the computer than with the physical manipulatives. Furthermore, they spent much more time processing feedback from the program. It is important to ask why the results of plenty of research mentioned above are not consistent with each other. The reason for the discrepancies may be that some variables such as instructional method, teacher's approach, and physical conditions were not taken under control in most of the studies. Since these factors directly affect student learning and attitude, it is hard to reach any consensus about which manipulative, physical or virtual, is better to use. More quasi-experimental studies are needed.

Reviewing the literature also points out that the majority of research was conducted with university students. Research carried out with high school students is rare. In addition, comparisons of the impact of using VM and PM on learning has been made for various physics concepts but little research has focused on motion and force concepts. This is critical because it is important to study students' conceptual understanding across several science domains (Olympiou & Zacharia, 2012). Much remains to be learned about the relative efficacy of physical and virtual materials when they are used in different science domains, with different instructional goals, approaches, outcome measures, and types of students (Klahr et al., 2007). Besides, studies comparing the effectiveness of virtual and physical experiments on different outcomes other than learning are worthy of investigation (de Jong et al., 2013). Therefore, the purpose of the present study was to compare the impact of using virtual manipulatives and physical manipulatives on students' learning of motion and force concepts and their attitudes towards physics.

Methodology

Participants

A pre–post comparison design was used for this research. Participants were randomly assigned to a virtual manipulative group or a physical manipulative group. The participants were 96 ninth graders ($n = 96$) from an all-boys military boarding school. They were already randomly distributed to four classes by the school administration. Consequently, there were 24 students in each class. The students' ages were between 15 and 16. One author was the physics teacher for the four classes and he randomly chose two classes to work with the virtual manipulatives. The other two classes worked with the physical manipulatives. In total, there were 48 students in each group.

Procedure

The research was conducted in the students' physics class. The students attended the class two hours a week. The instruction took place during a chapter on motion and force, which lasted eight weeks. This chapter included the following concepts: position, distance, displacement, speed, velocity, instant velocity, average velocity, acceleration, force, force of friction, weight, Newton's first law of motion (law of inertia), Newton' second law of motion, and Newton' third law of motion (action-reaction forces). The fact that there were various teaching resources related to one dimensional motion and force enabled the teacher to use different manipulatives.

The participants had formal education on motion and force concepts when they were students in the middle school.

Since the participants came to this boarding school from various middle schools and might have different backgrounds about inquiry, guided inquiry was employed in both groups to enable students who lacked experience to conduct research and experiments. During the guided-inquiry instruction, the problem, the background, and guidance of the procedures were given to the students but the methods of analysis, interpretation, and conclusion were for the students to generate. The same concepts were taught and same sample problems were solved in both virtual and physical manipulatives groups. The students were actively involved in exploring and constructing their own understanding and worked in groups where it was necessary to enable occurrence of dynamic transfer.

Activities in the groups started with open-ended questions to assess the students' prior knowledge and capture their attention. The students worked collaboratively in small groups and were encouraged to state their ideas in discussions held at the end of the activities. Sometimes the teachers addressed misunderstandings with the help of student explanations. Learning objectives, instructional method (inquiry), time on task, types of questions and probes from the teacher and assessment were the same for both the virtual and physical groups. The participants took their classes in the same technology-supported physics laboratory and their teacher was the same person. Moreover, since the participants were semester boarders, their learning activities after school hours were pretty much the same. Therefore, important variables that might influence learning and attitude were the same for the groups. Only the medium of presentation – virtual or physical – varied between the groups. Simulations, video recordings, interactive whiteboard, tablets and z-book were used in the virtual manipulatives group; while, experiment sets including air track, board and textbooks were used in the physical manipulatives group. For example, for students to learn position, displacement, speed, and velocity concepts, the students ran "walking man" simulation, the teacher used interactive presentations, and the teacher and the students solved some problems on the smart board by using interactive programs in the VM group. Meanwhile, the students did experiments with air tracks, the teacher made explanations by using the board, and the teacher and the students solved some problems on the board interactively to facilitate student learning of these concepts in the PM group. Likewise, the students in the VM group used a 2D freeware program and played the "maze game" online while the students in the PM group used air track sets and did hands-on activities by playing with velocity and acceleration cards to learn about acceleration.

The inventories designed to measure student learning and attitude were administered to the participants in four classes at the same time as pre- and post-tests. Other teachers in the school were observers and the researcher visited the classes during the administration.

The students were given worksheets created by the researchers that helped them explore scientific knowledge by requesting and guiding them to construct experiments and conduct various measurements. They included open-ended questions to get students' attention. Thus, six worksheets were prepared throughout the motion and force chapter based on the same performance objectives. The only difference between the worksheets used in both groups was the manipulatives in the directions and questions. The concepts and performance objectives assessed in the worksheets are presented in Table 1. The worksheets were completed by the students individually.

Table 1
Concepts and Performance Objectives Assessed in the Worksheets

Worksheet	Concepts	Performance Objectives
1	Linear motion	Collecting data by doing experiments, drawing of position-time and velocity-time graphs, interpretation of graphs, graph transformations.
2	Acceleration and two dimensional motion.	Explanation of acceleration by relating it with speeding up and slowing down. Inquiring the reasons for acceleration, collecting data by doing experiments, drawing of velocity-time and acceleration-time graphs, interpretation of graphs, graph transformations.
3	Force and friction force	Explanation of friction force, comparison of static and kinetic friction forces, exploring variables that the friction force depends on, making inferences from data, exploring advantages and disadvantages of friction force in daily life.
4	Balanced forces and Newton's first law	Calculation of the combination forces exerted on an object and explanation of motion of the object. Exploring and explaining of Newton's first law related to inertia by collecting data and doing experiments.
5	Newton's second law	Exploring and explaining of Newton's second law by collecting data and doing experiments.
6	Newton's third law	Exploring and showing action and reaction forces by using free-body diagrams.

Role of the Researcher and Teacher Intervention
The teacher of both groups was the first author. He had two roles. One role was as the participants' teacher and the other was as a researcher, who collected and analyzed the data. However, he was only a teacher throughout the instruction of motion and force concepts. He did not analyze any data until the instruction was over. Due to his teacher role, he established good communication with the students and worked to create an environment where the students felt comfortable about sharing their views. Rebello et al. (2005) argue that the researcher should be an observer and an instructor in order for dynamic transfer to occur. Therefore, he observed the students, directed them to the next step and promoted learning with the manipulatives.

Even though the teacher did not adopt the researcher role during instruction, some precautions were taken in order to prevent possible researcher bias. First, the two researchers prepared the lesson plans and worksheets together for both groups. This was an effort to make sure that the only difference between the groups was the manipulatives. Second, each lesson in both groups was videotaped and the two researchers watched and discussed the teacher's acts and performances before the next lesson to prevent any action that might affect student learning apart from the instruction. There was not any threat identified in the video recordings regarding research bias. Third, all of the data collection sources were written documents and both researchers analyzed them together. And finally, interrater reliability values were measured for the scoring of the rubrics.

Data Collection

The empirical phase of the study included the eight weeks of instruction, as well as two weeks for pre- and post-tests. In total, the study lasted ten weeks. As described below, both qualitative and quantitative methods were used to collect data.

Participant learning. Student learning was assessed formatively as well as summatively. In order to measure changes in student understanding of kinematics and dynamics concepts, the Force and Motion Achievement Instrument (FMAI) developed by Gokalp (2011) was administered as a pre-test and a post-test. This instrument was chosen among similar instruments for multiple reasons. First, it was comprehensive and specifically designed for ninth grade students. Second, the FMAI assessed both content and skill objectives by using various types of questions. Finally, the internal reliability coefficient for the FMAI was reported as .84 (Gokalp, 2011), which indicates high internal reliability. After performing both exploratory factor analysis and confirmatory factor analysis, Gokalp (2011) found that the FMAI measured students' achievements of "uniform linear motion", "fundamental forces", "Newton's laws of motion", "friction", and skill objectives as intended. Skill objectives were related with problem solving skills, information and communication technology skills, and physics-technology-society-environment skills. The instrument itself consisted of 30 questions including 16 multiple-choice, 12 open-ended, and two true-false questions. The questions on the instrument were conceptual as well as quantitative. Each question in the FMAI has an option of "I don't know / I can't do". In this way, unanswered questions can be categorized accurately. If this option was chosen, it was coded as "0". The true-false and multiple-choice questions were coded as "0" for nonscientific answers and "1" for scientific answers. There was a scoring rubric to analyze students' answers. The open-ended questions were coded as "0" for nonscientific answers, "1" for partially scientific answers, "2" for mostly scientific answers, and "3" for totally scientific answers. Therefore, possible scores ranged from 0 to 54. Students were given 50 minutes to complete the FMAI. Two open-ended questions and their scoring rubric were given. See the Appendix for examples.

Based on Rebello et al.'s (2005) suggestion for dynamic learning, student learning was also assessed during the instruction with the help of worksheets. Formative assessment integrated with instruction ideally provides a seamless process of assessment followed by instruction (Cauley & McMillan, 2010). Thus, the worksheets were used for the purpose of assessment for learning. The students completed each worksheet in one class hour. Student learning of the concepts covered during instruction was compared in detail. One two-point scoring rubric (from 0 to 2) was created for each worksheet based on the performance objectives assessed in the worksheet. As a result, six scoring rubrics were generated in total. The rubrics were the same for both groups.

Participant attitude. Changes in student attitude towards physics was assessed by applying the Attitude Towards Physics Scale (APCS) developed by Geban et al. (1994). This instrument was administered before and after the instruction. The instrument consisted of 15 items and a 5-point Likert scale (1 = "strongly disagree" to 5 = "strongly agree"). Possible scores on the APCS ranged from 15 to 75. Four items were related to enjoying physics, seven items were about interest in physics, and four items were related to necessity of physics. The internal reliability coefficient for the APCS was 0.83. This scale was chosen to because of its high internal reliability and shortness.

In addition to the FMAI, the APCS, and the worksheets, anecdotal observations were recorded while the students were working with manipulatives.

Data Analysis

Normality analyses were done separately for the learning and attitudinal data. Shapiro-Wilk tests were performed to determine if the pre- and post-tests data gathered from the APCS and the FMAI were normal. The significance values for pre-FMAI, post-FMAI, pre-APCS, and post-APCS were greater than 0.05 (p = 0.11, p = 0.50, p = 0.22, p = 0.55 respectively); therefore, all data followed normal distributions within a 95% confidence interval. Skewness and kurtosis were also calculated. Skewness values were between -1.0 and -0.5. Values of kurtosis fell between 0.5 and 1.0. Therefore, they supported normality. Independent *t*-tests were performed to analyze the data and compare the groups statistically. Dependent *t*-tests were used to analyze the data within groups. Effect sizes were calculated for the changes in the groups (Cohen, 1988).

The reliability analyses of the FMAT and the APCS were performed for this study. The worksheets were evaluated by one of the researchers based on the rubrics. In order to assess the reliability of scoring, the other researcher randomly selected 32 students (30%) from the PM and VM groups and scored their worksheets independently. Then, the two researchers compared their scoring and calculated the agreement for each group of worksheets separately. The researchers were able to reach 91% agreement for the first worksheet. The reliability measured by Cohen's κ was 0.71. Agreement percentages for the remaining five worksheets were 96%, 94%, 91%, 92%, and 93%. The following Cohen's κ values were reached regarding these agreement values: 0.86, 0.85, 0.70, 0.80, and 0.80. Fleiss (1981) characterizes Kappa values over 0.75 as excellent, values between 0.40 to 0.75 as fair to good, and below 0.40 as poor. Consequently, the scoring of students' knowledge reflected in the worksheets had adequate reliability. The authors re-scored the items on the rubrics that did not have initial agreement and the final scoring scheme was constructed by reaching consensus.

Results

Results of Student Learning

The internal reliability coefficient for the pre-FMAI was 0.40 indicating low reliability and 0.67 for the post-FMAI indicating medium reliability. Some students might have forgotten some parts of the force and motion domain after several years.

The results of independent samples *t*-tests showed a small difference between the two groups' performance on the FMAI (see Table 2). Before instruction, the physical manipulative group scored slightly higher (M = 12.16, SD = 3.30) than the virtual manipulative group (M = 11.77, SD = 3.74), a difference that was not statistically different. After instruction, the physical manipulative group scored lower on the FMAI (M = 27.20, SD = 4.81) compared to the virtual manipulative group (M = 27.56, SD = 5.81). Again, this difference was not statistically significant.

Table 2
Comparison of the FMAI Scores Between Groups

	Group	*M* (*SD*)	t	df	p
Pretest	PM	12.16 (3.30)	-0.51	85	.608
	VM	11.77 (3.74)			
Posttest	PM	27.20 (4.81)	0.31	87	.757
	VM	27.56 (5.81)			

However, the results of paired samples *t*-tests specified that both the PM and VM groups' post-instruction FMAI scores were significantly higher than their pre-instruction FMAI scores (see Table 3). There was an increase in the VM group's FMAI scores from pre-instruction to post-instruction ($M_{pre\text{-}post} = -15.79$) and the pre-to-post difference was statically significant for the VM group: $t_{(75)} = -15.28$, $p < .001$. The PM groups' FMAI scores also increased from pre-instruction to post-instruction ($M_{pre\text{-}post} = -15.04$), a difference that was statically significant: $t_{(75)} = -17.05$, $p < .001$. The increase in performances from pre-instruction to post-instruction was little higher for the VM group than for the PM group. Effect sizes between the pre- and the post-instruction FMAI scores were found to be 0.88 for the PM group and 0.85 for the VM group, which exceeded Cohen's (1988) convention for a large effect ($d = 0.80$).

Table 3
Comparison of the FMAI Scores Within groups

Group	Measurement	Mean Difference	*t*	*df*	*p*
PM	Pretest-posttest	-15.04	-17.05*	76	.000
VM	Pretest-posttest	-15.79	-15.28*	75	.000

The worksheets used as for formative assessment enabled the researchers to compare student understanding while they were working on the experiments and utilizing the manipulatives. Table 4 presents the results of independent samples *t*-tests for the groups' learning as assessed by the worksheets.

Table 4
Comparison of Worksheet Scores Between the Groups

Subject	Group	*M* (SD)	*t*	*df*	*p*
Linear Motion	PM	1.08 (0.11)	13.90	79	.000**
	VM	1.53 (0.17)			
Acceleration and two-dimensional motion	PM	1.56 (0.29)	2.58	51	.013*
	VM	1.69 (0.13)			
Force and friction forces	PM	1.42 (0.15)	0.94	69	.349
	VM	1.46 (0.21)			
Balanced forces and Newton's first law	PM	1.70 (0.23)	-1.71	75	.092
	VM	1.59 (0.35)			
Newton's second law	PM	1.61 (0.26)	2.14	40	.038*
	VM	1.71 (0.11)			
Newton's third law	PM	0.96 (0.52)	2.46	73	.016*
	VM	1.24 (0.47)			

* p < .05, ** p < .001

As shown in Table 4, the mean linear motion score on the rubric earned by the VM group ($M_{VM} = 1.53$, $SD = 0.17$) was higher than the PM group ($M_{PM} = 1.08$, $SD = 0.11$). This difference was statistically significant, $t_{(79)} = 13.90$, $p < .001$. The same situation occurred for acceleration and two-dimensional motion. That is, the mean acceleration and two-dimensional motion score on the rubric earned by the VM group ($M_{VM} = 1.69$, $SD = 0.13$) was higher than the PM group ($M_{PM} = 1.56$, $SD = 0.29$). This difference was statistically significant, $t_{(51)} = 2.58$, $p < .05$. Similarly, the mean Newton's second law score on the rubric earned by the VM group ($M_{VM} =$

1.71, *SD* = 0.11) was higher than the PM group (M_{PM} = 1.61, *SD* = 0.26). This difference was statistically significant, $t_{(40)} = 2.14$, $p < .05$. Finally, the mean Newton's third law score on the rubric earned by the VM group (M_{VM} = 1.24, *SD* = 0.47) was higher than the PM group (M_{PM} = 0.96, *SD* = 0.52). This difference was statistically significant, $t_{(73)} = 2.46$, $p < .05$.

Students in both groups studied linear motion by drawing position-time and velocity-time graphs, which was covered during instruction. Whereas the VM group could draw graphs on the simulations, the PM group collected data from the air track set and drew graphs on graph papers. However, according to anecdotal observations, some students in the PM group wrote time values on the vertical displacement axis instead of horizontal axis; hence, they did not draw proper constant velocity-time graphs. As a result, they might not understand the meaning of linear motion. While studying non-uniform motion, some students in the PM group could not calculate the slope of the velocity-time graph correctly. In addition, they could not draw the acceleration-time graph. This might be one of the reasons they could not conceptualize what happened if the velocity of an object was not constant and the object was speeding up or slowing down. As seen in Table 4, the mean rubric scores for the concept of acceleration were higher than the mean scores for linear motion in both groups. After working on the graphs during the application of the first worksheet, the students' graph skills increased by the second worksheet, resulting in more scientific graphs related to motion. Simulations that were used by the students in the VM group provided visualizations for the concepts of acceleration and force. This situation might have enabled students to acquire more knowledge of Newton's Second Law.

While students in the PM group were doing experiments and filling in the last worksheet about Newton's Third Law, two students connected two dynamometers to the glider improperly (opposite directions). Then, one student held one dynamometer steady whereas another student pulled the other dynamometer. They read the values on the dynamometers and recorded them, repeating this experiment three times. One student measured forces of 1.5N, 1.0N, and 2.0N, and the other student measured forces as 1.3N, 0.9N, and 1.8N respectively. Since the forces were in opposite directions and the force values were different, some students got confused and thought the glider had to move. Moreover, some students claimed that there was extra force due to the object's weight. Some students did not reset the dynamometer before starting to take new measurement and some of them could not hold it properly. At the end, these students in the PM group had difficulty making the inference that action and reaction forces are equal and opposite forces that act on different objects. That is the PM might have generated a little confusion and was not helpful for students to understand Newton's third law. Additionally, the students in the VM group did not encounter any measurement error while doing the experiments because the simulations showed numerical values of the parameters they used. Simulations provided students immediate feedback about the effect of the changes they made (Olympiou & Zacharia, 2012). This might have allowed them to investigate cause-and-effect relationships and answer questions in the worksheets more scientifically.

There was no statistical difference between the groups' learning of force and friction during instruction (M_{PM} = 1.42, *SD* = 0.15; M_{VM} = 1.46, *SD* = 0.21). Students in both groups were familiar with these concepts from the middle school science curriculum. As a result, implementing these concepts by using virtual or physical manipulatives might not have made a difference in their learning. From the dynamic transfer perspective, Schwartz et al. (2008) explained this learning situation as conceivable extension, which does not have to constitute a conceptual change alone. Although there was not a significant difference, the mean value of the PM group was higher than the mean value of the VM group for the fourth worksheet, whose performance objective was related to Newton's first law (M_{PM} = 1.70, *SD* = 0.23; M_{VM} = 1.59, *SD* = 0.35). Experiencing inertia physically and concretely during the lab activities might make it more plausible for students.

Informal observations revealed that the students in the PM group asked more questions to their teacher in order to do the experiments. At times they could not grasp exactly what to do. Hatano and Inagaki (1986) argued that if the risk attached to the performance of a procedure is minimal, people are more inclined to experiment and adapt new ways of doing things. The students in the PM group might have felt some distress while taking measurements and dealing with errors. This situation might have prevented them from acquiring new knowledge easily. Nevertheless, the students in the VM group could reach their goals after a few attempts within the simulations. They seemed more curious and involved with the lessons. Since these students are Generation Z learners and are more equipped with technology (Cilliers, 2017; Turner, 2015), they might be more open to learning with virtual manipulatives. These reasons might help explain the differences in the students' performances as assessed by the worksheets.

Results of Students' Attitude

Cronbach's alpha values for the pre- and post-instruction APCS were 0.90 and 0.93 respectively, indicating high reliability. As presented in Table 5, the results of independent samples *t*-tests showed there was no significant difference between the PM group's attitude (*M* = 49.09, *SD* = 9.70) and the VM group's attitude towards physics (*M* = 47.57, *SD* = 9.89) before the instruction. Furthermore, there was no significant difference between the groups' attitude towards physics after the instruction (M_{PM} = 56.45, *SD* = 7.80 vs. M_{VM} = 54.72, *SD* = 8.57).

On the other hand, the results of paired samples *t*-tests for the groups' attitude towards physics (see Table 6) revealed that both PM group ($M_{PMpre\text{-}post}$ = -7.36), $t_{(88)} = -3.98$, $p < .001$) and VM group ($M_{VMpre\text{-}post}$ = -7.15), $t_{(85)} = -3.60$, $p < .001$) significantly developed more positive attitudes after they received instruction with manipulatives. Effect sizes between the pre- and the post-instruction APCS scores were 0.38 for the PM group and 0.36 for the VM group. Both values were above Cohen's (1988) convention for a small effect ($d = 0.20$).

Table 5

Comparison of APCS Scores Between Groups

	Group	*M* (SD)	*t*	*df*	*p*
Pretest	PM	49.09 (9.70)	-0.73*	85	.470
	VM	47.57 (9.89)			
Posttest	PM	56.45 (7.80)	-1.00*	88	.320
	VM	54.72 (8.57)			

Table 6

Comparison of APCS Scores within Group

Group	Measurement	Mean Difference	*t*	*df*	*p*
PM	Pretest-posttest	-7.36	-3.98**	88	.000
VM	Pretest-posttest	-7.15	-3.60**	85	.000

**p < 0.01

Discussion

In this study, the aim was to compare the effects of using PM and VM within a guided inquiry approach on student outcomes. Student outcomes were analyzed on two levels: knowledge acquisition and attitude. The students worked with their peers during the experiments. Every tool utilized in the VM group was virtual including books and the board. Conditions of dynamic transfer of learning were tried to accomplish during the instruction by providing a learning environment where the students expressed themselves and worked as groups and the instructor observed them and used real-time assessment.

The comparisons made between the PM and VM groups as well as within each group in term of the FMAI scores before and after instruction revealed student learning was the same regardless of whether they were instructed with PM or VM. In other words, the students' selection of scientific choices and their scientific explanations for the content and skill questions on the FMAI were similar. These findings point out that the students' learning of force and motion concepts was elevated and they learned almost equally with either manipulation. This result was in line with findings from other scholars (Darrah et al., 2014; Jaakkola & Nurmi, 2008; Klahr et al., 2007; Triona & Klahr, 2003; Zacharia & Constantinou, 2008; Zacharia & Olympiou, 2011). However, the present study was the only study whose participants were high school students learning about the subject of motion and force.

The students who were taught dynamics concepts by using virtual manipulatives understood more concepts than the students who were taught dynamics concepts by using physical manipulatives during the instruction. This result was similar with previous research (Finkelstein et al., 2005; Husnaini & Chen, 2019) whose participants ranged from secondary school students to undergraduate students and involved an inquiry-based context. However, the findings divulged by Coramik (2012) and Marshall and Young (2006) contrast the results of this study. In Coramik (2012)'s research, the participants did not do their experiments by implementing inquiry-based approach. Therefore, his context was different from this study's context. The participants of Marshall and Young (2006), on the other hand, were not the students, they were teachers. Therefore, different context and different group of participants might result the inconsistency between this research and those.

Attitudes are tenacious over time (Hill, Atwater, & Wiggins, 1995; Koballa, 1988). Since the participants were ninth graders and took physics for the first time, the eight-week duration was enough for students in both groups to change their attitudes. Neither instruction supported with VM or PM displayed superiority. They had the same influence on the students' attitude toward physics. This result was consistent with findings that have emerged from research done by Taghavi and Colen (2009) who revealed that college students' attitudes were similar with regard to both simulated and physical laboratory instruction. The students' attitudes towards science increased no matter which manipulatives were used during the instruction.

Research implies that hands-on activities, cooperative learning, and student involvement in learning had strong influences on attitude toward science (Zacharia, 2003). In addition, Lee et al. (2020) stated that students who viewed experimental learning as achieving in-depth understanding and who perceived that experiments were guided by clear rules were prone to express a stronger sense of academic self-efficacy. The students in both groups worked in groups and followed the instructions in the worksheets as a part of the inquiry approach while they were learning. These might be the reasons that the students in the PM group as well as the VM group developed more positive attitudes towards physics.

Conclusions and Suggestions

Dynamic transfer can create conceptual change (diSessa & Wagner, 2005). Dynamic transfer depends on the environment to support coordination because it is the product of a sequence of interactions with a well-structured environment that may include tools, representations, other people, and so forth (Schwartz et al., 2008). Guided participation is based on the belief that students are active learners and the learning environment is integral to the learning process (Rogoff, 2003).

This study concludes that use of PM and VM in inquiry-based science has the same effect on students' conceptual and procedural knowledge, as well as their attitude towards physics. Interactions with the environment generate feedback and variability that can help students shake free of their initial interpretations and extend their knowledge (Schwartz et al., 2008). Learning environments including either VM or PM can facilitate conceptual change and dynamic transfer related to motion and force concepts.

The second conclusion is that due to measurement errors, learning of students who deal with physical experimentation is lower than learning of their peers who use virtual experimentation. Differences between formative and summative assessment results revealed that students in PM group may need time to internalize their understanding because they first resolved the problems created by measurement errors. Furthermore, physical investigations leave students with some irrelevant knowledge. Nonetheless, students in PM group were able to transfer their knowledge and use their learning gains after the instruction. An environment including virtual technology does not allow errors and this situation may maximize, as Greeno Moore and Smith (1993) elucidate, the possibilities of students' attunement to the affordances of tools. Experimenting with VM helps students grasp the motion-force relationship and understand of graphs quickly. Hence, VM can be implemented to provide authentic experiences (Steinberg, 2000) and encourage learning. Finally, although researchers have tried to explain whether attitudes influence learning or if learning influences attitudes (Zacharia, 2003), this study concludes that attitude and learning may be developed in parallel because both increased at the end of the instruction.

This study has several implications. The conclusions suggest that physical and virtual manipulatives can be used for one another when inquiry-based learning is emphasized. It is reasonable to assume that implementing virtual manipulatives has even more advantages on learning in some conditions. This suggestion is important for the science education community, especially regarding virtual schooling that came along with the recent global pandemic. However, students need to learn how to do error analysis in order to experience science phenomena in the real world; thus, using physical materials should not be abandoned. This study adds to current science education literature by using virtual materials in one group and demonstrating the effects of comparing virtual and physical materials on students' attitude and learning of motion and force concepts. Teachers might consider struggles and easiness that the students came across while dealing with virtual and physical manipulatives during this study and plan their teaching in a way that their students avoid the same struggles and experience the same easiness. Researchers would conduct exploratory studies that examine how students learn with the help of virtual and physical manipulatives.

References

Adesoji. F.A., & Raimi. S.M. (2004). Effects of enhanced laboratory instructional technique on senior secondary students' attitude toward chemistry in Oyo township. Oyo State. Nigeria. *Journal of Science Education and Technology*, *13*(3). 377–385. https://doi.org/10.1023/B:JOST.0000045465.81437.3b

Aguilera, D., Perales-Palacios, F. (2020). What Effects Do Didactic Interventions Have on Students' Attitudes Towards Science? A Meta-Analysis. *Research in Science Education*, 50, 573–597. https://doi.org/10.1007/s11165-018-9702-2

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1–18. https://doi.org/10.1037/a0021017

Apkan, J. P. (2002). Which comes first: Computer simulation of dissection or a traditional laboratory practical method of dissection. *Electronic Journal of Science Education*, *6*(4): http://wolfweb.unr.edu/homepage/crowther/ejse/akpan2.pdf

Bozkurt, E., & Sarikoc, A. (2008). Can virtual laboratory replace traditional laboratory in physics education? *Journal of Ahmet Kelesoglu Education Faculty*, *25*, 89–100.

Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 83*(1), 1–6. https://doi.org/10.1080/00098650903267784

Cilliers, E. J. (2017). The challenge of teaching generation Z. *International Journal of Social Sciences, 3*(1), 188–198. https://dx.doi.org/10.20319/pijss.2017.31.188198

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics*, *50*(1), 66–71. https://doi.org/10.1119/1.12989

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd ed. Hillsdale, NJ: Erlbaum.

Coramik, M. (2012). *The effect of teaching magnetism unit with computer-aided and experiment-aided activities on 11th grade students? Self-efficacy, metacognition, attitude, motivation and conceptual understanding* [Unpublished master's dissertation]. Balikesir University, Balikesir, Turkey.

Darrah, M., Humbert, R., Finstein, J., Simon, M., & Hopkins, J. (2014). Are virtual labs as effective as hands-on labs for undergraduate physics? A comparative study at two major universities. *Journal of Science Education and Technology*, *23*, 803–814. https://doi.org/ 10.1007/s10956-014-9513-9

de Jong T. (2006). Computer simulations: technological advances in inquiry learning. *Science*, *312*, 532–533. https://doi.org/10.1126/science.1127750

de Jong, T., Linn, M. C., & Zacharia, Z. C. (2013). Physical and virtual laboratories in science and engineering education. *Science*, *340* (6130), 305–308. https://doi.org/10.1126/science.1230579

diSessa, A. A., & Wagner, J. F. (2005). What coordination has to say about transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multi-disciplinary perspective* (pp. 121-154). Greenwich, CT: Information Age Publishing.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, *23*(7), 5–12. https://doi.org/10.3102/0013189X023007005

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: John Wiley.

Finegold, M., & Gorsky, P. (1991). Students' concepts of force as applied to related systems: A search for consistency. *International Journal of Science Education*, *13*(1), 97–113. https://doi.org/10.1080/0950069910130109

Finkelstein, N. D., Adams, W. K., Keller, C. J., Kohl, P. B., Perkins, K. K., Podolefsky, N. S., Reid, S., & LeMaster, R. (2005). When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physical Review Special Topics-Physics Education Research, 1*(1), 1–8. https://doi.org/10.1103/PhysRevSTPER.1.010103

Ford, D. N., & McCormack, D. E. M. (2000). Effects of time scale focus on system understanding in decision support systems. *Simulation & Gaming, 31*, 309-330. https://doi.org/10.1177/104687810003100301

Freedman. M. (1997). Relationship among laboratory instruction. Attitude toward science and achievement in science knowledge. *Journal of Research in Science Teaching*, *34*(4). 343-357. https://doi.org/10.1002/(SICI)1098-2736(199704)34:4<343::AID-TEA5>3.0.CO;2-R

Gagne, R.M. (1977). *The conditions of learning* (3rd ed.). New York: Holt, Rinehart & Winston.

Geban, O., Ertepinar, H., Yilmaz, G., Atlan, A., & Sahpaz, O. (1994, September 15-17). *Effects of computer-based instruction on students' science achievement and attitudes towards science* [Paper presentation]. National Science Education Conference, Izmir, Turkey.

Gibson, H. L., & Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students' attitudes toward science. *Science Education*, *86*, 693–705. https://doi.org/10.1002/sce.10039

Gokalp, M. S. (2011). *The effect of webquest based instruction on ninth grade students' achievement in and attitude towards force and motion* [Unpublished doctoral dissertation]. Middle East Technical University, Ankara, Turkey.

Graham, T., Berry, J., & Rowlands, S. (2013). Are 'misconceptions' or alternative frameworks of force and motion spontaneous or formed prior to instruction? *International Journal of Mathematical Education in Science and Technology*, 44(1), 84–103. https://doi.org/10.1080/0020739X.2012.703333

Greeno, J. G., Moore, J. L., & Smith, D. R. (1993). Transfer of situated learning. In D. K.

Detterman, & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition and instruction* (pp. 99–167). Norwood, NJ: Ablex.

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H.A.H. Stevenson, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262-272). New York: Freeman.

Hill G., Atwater, M., &Wiggins, J. (1995). Attitudes toward science of urban seventh grade life science students over time, and the relationship to future plans, family, teacher, curriculum, and school. *Urban Education*, *30*, 71–92. https://doi.org/10.1177/0042085995030001006

Hoftein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*, 28–54. https://doi.org/10.1002/sce.10106

Husnaini, S. J., & Chen, S. (2019). Effects of guided inquiry virtual and physical laboratories on conceptual understanding, inquiry performance, scientific inquiry self-efficacy, and enjoyment. *Physical Review Physics Education Research*, *15*(1), 010119, 1–16. https://doi.org/10.1103/PhysRevPhysEducRes.15.010119

Jaakkola, T., & Nurmi, S. (2008). Fostering elementary school students' understanding of simple electricity by combining simulation and laboratory activities. *Journal of Computer Assisted Learning*, *24*(4), 271-283. https://doi.org/10.1111/j.1365-2729.2007.00259.x

Klahr, D., Triona, L.M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching*, *44*(1), 183–203. https://doi.org/10.1002/tea.20152

Klahr, D., Triona, L.M., & Williams, C. (2007). Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children. *Journal of Research in Science Teaching*, *44*(1), 183-203. https://doi.org/10.1002/tea.20152

Koballa, T.R. Jr. (1988). Attitudes and related concepts in science education. *Science Education*, *72*, 115–126. https://doi.org/10.1002/sce.3730720202

Lee, M. H., Liang, J. C., Wu, Y. T. Chiou, G. L., Hsu, C. Y., Wang, C. Y., Lin, J. W., & Tsai, C. C. (2020). High School Students' Conceptions of Science Laboratory Learning, Perceptions of the Science Laboratory Environment, and Academic Self-Efficacy in Science Learning. *International Journal of Science and Mathematics Education*, *18*, 1–18. https://doi.org/10.1007/s10763-019-09951-w

Marshall, J.A. (2002). Pre-service teachers' use of technology to explore momentum conservation. *AAPT Announcer, 32*(4), 112

Marshall, J. A., & Young, E. S. (2006). Preservice teachers' theory development in physical and simulated environments. *Journal of Research in Science Teaching*, *43*, 907-937. https://doi.org/10.1002/tea.20124

Mildenhall, P., & Williams, J. (2001). Instability in students' use of intuitive and Newtonian models to predict motion: the critical effect of parameters involved. *International Journal of Science Education*, *23*, 643–660. https://doi.org/10.1080/09500690117839

Olympiou, G., & Zacharia, Z. C. (2012). Blending physical and virtual manipulatives: An effort to improve students' conceptual understanding through science laboratory experimentation. *Science Education*, 96, 21–47. https://doi.org/10.1002/sce.20463

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049–1079. https://doi.org/10.1080/0950069032000032199

Rebello, N. S., Zollman, D. A., Allbaugh, A. R., Englehardt, P. V., Gray, K. E., Hrepic, Z., & Itza-Ortiz, S. F. (2005). Dynamic transfer: A perspective from physics education research. In J. P. Mestre (Ed.) *Transfer of learning from a modern multidisciplinary perspective* (pp. 217–250). Greenwich, CT: Information Age Publishing.

Reed, S. K. (1993). A schema-based theory of transfer. In D. K. Detterman, & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition and instruction* (pp. 39–67). Norwood, NJ: Ablex.

Rogoff, B. (2003). *The cultural nature of human development*. New York, NY: Oxford University Press.

Roth W. M., Woszczyna, C., & Smith, G. (1996). Affordances and constraints of computers in science education. *Journal of Research in Science Teaching, 33*(9), 995–1017. https://doi.org/10.1002/(SICI)1098-2736(199611)33:9<995::AID-TEA3>3.0.CO;2-Q

Rowlands, S., Graham, T., Berry, J., & Mcwilliam, P. (2007). Conceptual change Through the lens of Newtonian mechanics. *Science Education*, *16*, 21–42 (2007). https://doi.org/10.1007/s11191-005-1339-7

Schwartz, D. L., Varma, S., & Martin, L. (2008). Dynamic transfer and innovation. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 479–506). New York: Routledge

Shaw, J. M. (2002). Manipulatives enhance the learning of Mathematics. Accessed August 11, 2015 from http://www.eduplace.com/state/author/shaw.pdf

Singley, K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Steinberg R.N. (2000). Computers in teaching science: to simulate or not to simulate? *American Journal of Physics*, *68*, 37–41. https://doi.org/10.1119/1.19517

Taghavi, S. E., & Colen, C. (2009). Computer simulation laboratory instruction vs. traditional laboratory instruction in digital electronics. *Journal of Information Technology Impact*, *9*(1), 25-36.

Tao, P., & Gunstone, R. (1999). The process of conceptual change in force and motion during computer-supported physics instruction. *Journal of Research in Science Teaching*, *36*(7), 859-882. https://doi.org/10.1002/(SICI)1098-2736(199909)36:7<859::AID-TEA7>3.0.CO;2-J

Triona, L. M., & Klahr, D. (2003). Point and Click or Grab and Heft: Comparing the Influence of Physical and Virtual Instructional Materials on Elementary School Students' Ability to Design Experiments. *Cognition and Instruction*, *21*(2), 149–173. https://doi.org/10.1207/S1532690XCI2102_02

Toth, E. E., Morrow, B. L., & Ludvico, L. R. (2009). Designing blended inquiry learning in a laboratory context: A study of incorporating hands-on and virtual laboratories. *Innovative Higher Education, 33*(5), 333–344. https://doi.org/10.1007/s10755-008-9087-7

Turner, A. (2015). Generation Z: Technology and social interest. *The Journal of Individual Psychology, 71*, (2), 103–113. https://doi.org/10.1353/jip.2015.0021

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction*, *11*, 381–419. https://doi.org/10.1016/S0959-4752(00)00038-4

Wang, T. L., & Tseng, Y. K. (2018). The comparative effectiveness of physical, virtual, and virtual-physical manipulatives on third-grade students' science achievement and conceptual understanding of evaporation and condensation. *International Journal of Science and Mathematics Education*, *16*, 203–219. https://doi.org/10.1007/s10763-016-9774-2

Windschitl, M. (2000). Supporting the development of science inquiry skills with special classes of software. *Educational Technology Research and Development*, *48*(2), 81–95. https://doi.org/10.1007/BF02313402

Zacharia, Z. C. (2003). Beliefs, attitudes, and intentions of science teachers regarding the educational use of computer simulations and inquiry-based experiments in physics. *Journal of Research in Science Teaching*, *40*(8), 792–823. https://doi.org/10.1002/tea.10112

Zacharia, Z. C., & Constantinou, C. P. (2008). Comparing the influence of physical and virtual manipulatives in the context of the Physics by Inquiry curriculum: The case of undergraduate students' conceptual understanding of heat and temperature. *American Journal of Physics*, *76*, 425–430. https://doi.org/10.1119/1.2885059

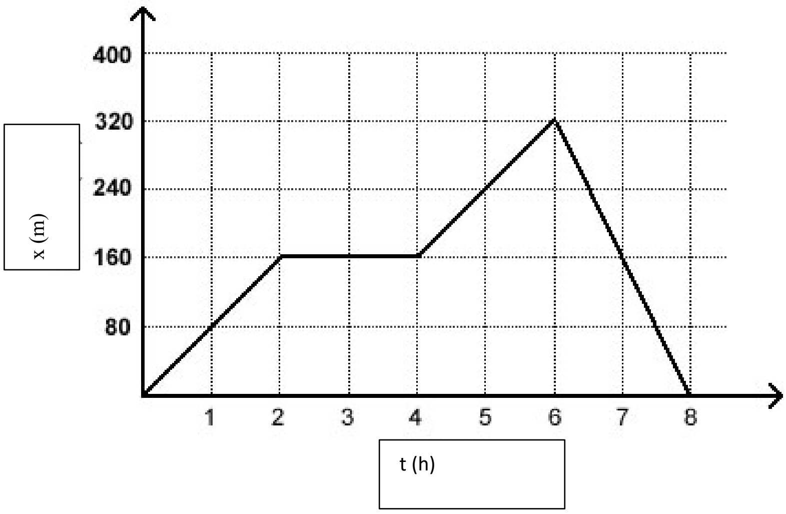
Zacharia, Z. C., & Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction*, *21*(3), 317–331. https://doi.org/10.1016/j.learninstruc.2010.03.001

Zacharia, Z. C., & de Jong, T. (2014) The Effects on Students' Conceptual Understanding of Electric Circuits of Introducing Virtual Manipulatives Within a Physical Manipulatives-Oriented Curriculum. *Cognition and Instruction*, *32*(2), 101–158. https://doi.org/10.1080/07370008.2014.887083

Corresponding author: Feral Ogan-Bekiroglu
Email: feralogan@yahoo.com

APPENDIX

Question 23



A car's position-time graph is presented above. Please describe and explain the car's motion in two hour-time interval within 8 hours by using numerical values.

Fully compatible with scientific knowledge (3 points):

The car moved in uniform linear motion and displaced 160 km with 80 km/h velocity in the first two hours. The car stopped and did not change its position for the time interval from 2 hours to 4 hours. The car started to move in uniform linear motion again with 80 km/h velocity and displaced 160 km for the time interval from 4 hours to 6 hours. The car went back to its first position by moving in opposite direction in uniform linear motion with 160 km/h velocity and displaced 320 km.

Mostly compatible with scientific knowledge (2 points):

- Answers describing and explaining the car's motion correctly in three time intervals or
- Answers describing the car's whole motion correctly without using numbers (for example the car moved in uniform linear motion in the first two hours but did not move for the next two hours).

Partially compatible with scientific knowledge (1 point):

- Answers describing and explaining the car's motion correctly in two or less time intervals or
- Answers describing the car's some part of motion correctly without using numbers.

Nonscientific knowledge (0 point):

Answers that do not include any correct information about the car's motion.

Question 29

Design an experiment to investigate the differences between static and kinetic friction forces.

Fully compatible with scientific knowledge (3 points):

Any experiment design that enables to measure and compare an object's frictional force before and after its motion

Mostly compatible with scientific knowledge (2 points):

Although there is a complete design, some measurements cannot be taken.

Partially compatible with scientific knowledge (1 point):

Design is incomplete and some measurements cannot be taken.

Nonscientific knowledge (0 point):

All the other circumstances that do not match with the answers above.

Artificial Intelligence for Career Guidance – Current Requirements and Prospects for the Future

Stina Westman
South-Eastern Finland University of Applied Sciences
Finland

Janne Kauttonen
Haaga-Helia University of Applied Sciences
Finland

Aarne Klemetti
Metropolia University of Applied Sciences
Finland

Niilo Korhonen
Häme University of Applied Sciences
Finland

Milja Manninen
South-Eastern Finland University of Applied Sciences
Finland

Asko Mononen
Laurea University of Applied Sciences
Finland

Salla Niittymäki
Häme University of Applied Sciences
Finland

Henry Paananen
Häme University of Applied Sciences
Finland

Abstract

Career guidance in the era of life-long learning faces challenges related to building accessible services that bridge education and employment services. So far, only limited research has been conducted on using artificial intelligence to support guidance across higher education and working life. This paper reports on development on using artificial intelligence to support and further career guidance in higher education institutions. Results from focus groups, scenario work and practical trials are presented, mapping requirements and possibilities for using artificial intelligence in career guidance from the viewpoints of students, guidance staff and institutions. The findings indicate potential value and functions as well as drivers and barriers for adopting artificial intelligence in career guidance to support higher education and life-long learning. The authors conceptualize different modes of agency and maturity levels for the involvement of artificial intelligence in guidance processes based on the results. Recommended future research topics in the area of artificially enhanced guidance services include agency in guidance interaction, developing guidance data ecosystem and ethical issues.

Keywords: agency, artificial intelligence, career guidance, data, ethics, higher education

Wide-ranging and dynamic changes in working life have increased the dynamism of the labour market and transformed attitudes towards careers. There is a rising demand for learning across all education levels and age groups. This poses novel challenges for career guidance services at higher education institutions.

The focus on continuous learning highlights the need for lifelong career guidance (Toni & Vuorinen, 2020). In Finland, a national strategy for lifelong guidance positions career guidance to support individuals to be able to recognize their skills and mirror them with not only the opportunities and needs of the labour market, but also the opportunities to develop their competence (Strategy for Lifelong Guidance, 2020). This enables individuals to make meaningful plans and decisions relating to education and career paths. The expectation is that investment in guidance services can reduce dropouts, enhance completions of degrees and speed up transitions to labour market. Innovative lifelong career guidance practices can support these through upskilling and reskilling competencies and by enhancing career adaptability (Barnes et al., 2020).

Growing demands exist on the delivery and development of services for career guidance, extending the expected uses and broadening the scope of services. Various actors within the educational system, labour market as well as the social and health sector provide career guidance. As needs for guidance grow it is necessary for them to utilize digital services to save resources as well as increase value to career guidance (Toni & Vuorinen, 2020). Smart technologies can play a role in supporting both guidance practitioners and lifelong learners.

This article addresses supporting career guidance through novel technology. A multiple methods study is reported on the adoption of artificial intelligence (AI) for enhancing career guidance services in higher education. Requirements and opportunities for guidance interventions through intelligent technologies are analysed based on results from focus groups, scenario work and practical trials. Based on these, further research directions are recommended, including the effects to agency, emerging career information environment and maturity levels for leveraging AI in career guidance.

Literature Review

Guidance, Career Guidance and Lifelong Guidance

Guidance aims to support individuals building their own life paths by enhancing their ability to use their own capabilities and resources (Peavy, 2000). Guidance covers a range of individual and collective activities relating to information delivery, counselling, competence assessment, support, and teaching decision-making and career management skills (Council of the European Union, 2008).

Career guidance refers to services and activities intended to assist individuals, of any age and at any point in their lives, to make educational, training and occupational choices and to manage their careers (OECD, 2004). Within this definition, both individual and group guidance activities are included. The services range from information provision, to self-assessment and on to counselling with professional guidance staff. In recent years, the focus of career guidance has turned to needs for reskilling and upskilling within continuous education (Toni & Vuorinen 2020).

Agency in Guidance

Guidance aims to support the agency of the students. Agency is necessary and needed for students in the learning process, constructing knowledge and engaging in collaborative practices (Jääskelä et al., 2020). Agency as a concept comprises activity to prioritize, choose, and consider what is important and worth aspiring for and make decisions on one's professional identity and life (Eteläpelto et al., 2013).

Bandura (2001) describes three modes of agency: personal agency, proxy agency and collective agency. Personal agency is the direct mode of agency, exercised by the individual. Proxy agency consists of relying on others in acting and relying on other persons' resources and knowledge. Collective agency is constructed in groups through shared, collective acts.

In education these modes of agency have been modelled in pedagogical learning agents (Kim & Baylor, 2006), where such agents proved useful for modelling the social-cognitive perspectives of human and technological agents. Jääskelä et al. (2020) have also utilized the agency construct in investigating the use of learning analytics in the construction of agency.

Technology in Guidance

Digital tools can provide individuals with novel opportunities to access guidance any time or place as well as expanding the range of services offered. The potential benefits of using technology in career guidance include improved accessibility, increased access to information, assessment, and networks as well as lowered overall costs and improved cost-effectiveness (Sampson et al., 2020). The ongoing pandemic has increased the need for distance and digital services for guidance (Cedefop et al., 2020).

Guidance staff have traditionally used technology in three ways, providing: 1) learning and career information supporting career building, 2) automated interaction like career assessments, simulations or games and 3) choices of communication (Hooley et al., 2015). The development of integrated or blended guidance – guidance via digital means – requires guidance professionals and service designers to plan what technologies to use and how (Bakke et al., 2018).

The integration of new and emerging technologies into guidance services depends not only on the users' skills or technical solutions, but also on the willingness of guidance organizations and professionals to adapt (Kettunen & Sampson, 2019). The extent to which technology is integrated into guidance practices varies based on the capacity and technological orientation of staff (Kettunen et al., 2013).

AI in Education and in Guidance

In this study, artificial intelligence is defined as intelligent agents that receive percepts from the environment and take actions that affect that environment, following the definition by Russell and Norvig (2016). These agents can mimic cognitive functions such as learning, understanding, reasoning and problem solving.

The uses of AI in education have been developing for decades. Moreno-Guerrero et al. (2020) describe the trends in the study of AI in education between 1956 and 2019 based on a bibliometric analysis, concluding that while early studies centred more on technological process, more recent investigations focus on the development of AI as situated in pedagogical process.

The recent advances in AI are expected to have profound impacts on future labour markets and competence requirements, as well as enabling new ways of learning and teaching (Tuomi, 2018). According to research and review studies (e.g. Khare et al., 2018; Martiniello et al., 2020; Zawacki-Richert et al., 2019), AI can be used in education to support various functions such as student self-regulation, motivation and well-being, personalized learning support and feedback, learning process support, assessment and evaluation, profiling and prediction, usability and accessibility, resourcing, and competence management.

There exist few studies on the affordances of AI in career guidance. Khare et al. (2018) investigated the effect of artificial intelligence on the student experience, including support throughout students' studies. They concluded that AI can positively influence students and organizations, structures, processes and people that make up educational systems. While Khare et al. do not explicitly situate their study within guidance, the practical examples along the student lifecycle support the reflection on skills and learning opportunities and transitioning to working life, which comprise major functions of career guidance.

Digital services are at the core of education services for the future. However, technology does not serve only a utilitarian role in education. AI and education have a manifold relationship (Attwell et al., 2020; Roll & Wylie, 2016; European Commission, 2019). First, AI-related competences should be built up in education as they are required for future work environments where AI is utilized. Second, AI-based technology may be utilised in learning and teaching processes, integrated into existing learning environments, or by leveraging intelligent environments for educational purposes. Third, AI should also be further developed for the purposes of education.

When artificial intelligence technology is used in guidance interaction, it may also change or moderate the creation of agency. Ågerfalk (2020) posits digital agency as the capability of machines to act autonomously, but on behalf of humans, organisations and institutions. The impact of AI on digital career guidance practices could thus be further studied through agency.

Method

Framework and Research Questions

This article contributes to the body of work on digital technologies, namely artificial intelligence, in career guidance, education and lifelong learning. This study was conducted using the theoretical framework of socio-cognitive agency and its extensions to human-technology interaction.

The following research questions were posed:

- What requirements for using artificial intelligence in career guidance are identified by students and staff?
- What possibilities exist for using artificial intelligence in career guidance?

Methodological Approach

The study employed a multiple methods approach. Complementary methods were used seeking elaboration, enhancement, and clarification of the results obtained via one method with the results from other methods (Greene, 2007). Following a convergent parallel design, the different strands of the research were performed independently, with results brought together in the overall interpretation (Creswell & Plano Clark, 2011).

Research ethics practices of the participating universities were followed and necessary permits obtained from individual informants. Participants were recruited via advertisements from students and staff of participating organizations, representing potential users. For scenario workshops, public events and snowball sampling were also used. Participants gave permissions to use their personal data in trials. Privacy notices were issued according to the EU General Data Protection Regulation.

Focus groups. Seven focus group sessions on the use of AI in career guidance were run dedicated to either higher education students (total 11 persons) or guidance staff (14 persons). Discussions were facilitated on ideal guidance situations, use of technology to support guidance, and development needs in guidance services. Thematic analysis was conducted on qualitative researcher notes from focus groups.

Scenario work. Scenarios or design fictions (Cox et al., 2021) were co-created as narratives encapsulating possible futures where AI is used to support student guidance. Seven scenario workshops were organized with higher education and vocational education guidance staff (n=333). An iterative design process was employed, where the scenarios were gradually refined. First, workshop discussions were used to form initial scenario narratives. Feedback was gathered in subsequent co-design workshops based on which the scenarios were then elaborated. The scenarios serve as the output of the research as well as being used as part of a process of raising awareness (Tsekleves et al., 2017) on AI.

Trials. Two sets of practical trials were conducted utilizing AI in guidance services. AI applications were developed based on previous surveys at the institutions to support 1) course recommendations 2) job recommendations and 3) skills profile creation. Higher education students (n=179) and vocational education students (n=103) participated in the development and trials in their institutions, giving input via hackathons, design jams, workshops and qualitative user testing organized. Surveys requested feedback on the first trial and qualitative feedback was gathered for both trials. Trials utilized personal data of the students, curriculum data, national qualification data and job ads from various public and commercial web portals. The applications combined machine learning and data mining techniques with existing commercial mobile student services and AI components.

Findings

Requirements for AI in Career Guidance

In focus groups, artificial intelligence solutions were envisioned to support students in studies and career planning but also in self-management. Students welcomed the use of AI in career guidance. They emphasized the importance of accessible and timely guidance, whether delivered by AI or humans. AI was seen to have a role via detecting weak signals and potentially giving a “nudge” towards guidance interventions before either the student or staff would know to act.

Students described that any AI-enabled process should be part of everyday learning activities, not a separate application. Students brought up needs to manage their schedules and workloads, and to find suitable study methods. Students described needs for better communication and feedback with teachers and students. Students also mentioned the importance of peer support and discussed the potential role of AI in mediating this.

Students envisioned that artificial intelligence could support them in recognising their strengths and weaknesses, enabling their development. They wanted to use AI to compare their skills to the competence requirements of specific fields or positions, as well as general working life competences. They saw potential in AI applications that propose studies, thesis topics, work placements and jobs based on skills, experiences and interests. Staff envisioned a role for AI in recognition and accreditation of prior learning as well as predicting future competence needs. Staff discussed the competence or skill data used by AI. They recognized that while various data sources already exist, these are not necessarily available for students. Collating this data via AI would enable students to have a more active role in their own learning and career planning.

Guidance staff saw artificial intelligence in a supporting role to their work, balancing out the benefits and risks of incorporating technology into the guidance process. Staff hoped that AI applications could assist them in routine administrative tasks and relying information. This would free up time which they felt would be better used interacting with students in order to create relationships and build trust, and to engage in case management.

Staff also recognised that AI technology could replace human effort in some areas, changing their role and tasks. Staff cautiously welcomed this, with the expectation that the utilization of AI would enable them to allocate time to tasks where human interaction is needed. "Human touch" was considered valuable for finding opportunities, supporting decisions, detecting silent signals, interpreting affective states, motivating and encouraging, creating a safe atmosphere as well as relaying empathy and hope. Staff described these from the viewpoint of the skills needed while students described preferring human counsellors in these situations.

Finding, accessing and relaying information on for example curricula and schedules was considered difficult and time-consuming. Students and staff envisioned that AI could assist in delivering the right information at the right time. They described a proactive process, extending to information not yet needed to be known. Students and staff envisioned that AI would advise to book a guidance session when needed, supporting case management. This would enable staff to "triage" cases, taking action more quickly when needed. Staff stressed the significance of designing how to determine the urgency and importance of issues. Both students and staff hoped AI would aid in detecting and visualizing study progress based on the activity and performance data generated on online learning platforms and other digital services.

Scenarios for Future of AI for Career Guidance

Scenarios encapsulated the potential roles of AI in career guidance into narratives. The scenarios were linked to various phases along the study path, from initial application to studies, across studies and transition to employment, and linked to competence development within the continuous learning paradigm. The working of intelligent technology was described both from the viewpoints of student and staff as well as describing implications for the higher education institutions at large.

The following were the most elaborated among the resulting twenty-one scenarios:

- Supporting career planning: supporting decision making throughout career, promoting available career services based on situational information
- Enhancing interaction in counselling: matching students and counsellors, collating previous guidance discussions to a knowledge base

- Recognizing and verbalizing existing skills: creating a competence portfolio, recognizing generic competences from work experience
- Comparing competences to goals and needs from working life: offering self-assessment tools for competence mapping, inferring competence gaps based on profile data
- Anticipating guidance needs and case management: collating information on the student for staff to see at a glance, prioritizing tasks for staff
- Recognizing networks: enabling access to up-to-date information sources on career services, leveraging existing contacts for employment opportunities

In scenarios, staff envisioned services they could use as aids when delivering guidance interventions, such as automated Q&A solutions, scheduling aids and analytics dashboards. For analytics, early warning detection systems were planned, but also systems that would highlight student successes. Staff indicated that they would like to collaborate on AI-enabled platforms, sharing information between guidance professionals.

For their students, staff described potential AI assistants for fact checking information online in social media and for recognizing skills via self-reflection. When discussing the potential of AI for student use, staff assumed that it might depend on the individuals and their career planning needs and capabilities.

Staff included mentions of the information and services available and necessary for implementing the scenario. The scenarios raised concerns linked to student privacy (personal data access, sensitive data) and the potential to enhance existing bad practices or biases. Questions were raised about data quality, both regarding the data available for AI and the AI outputs. The threat of being replaced by AI was voiced, accompanied by arguments on the irreplaceability of human effort in guidance. Staff expected AI to "outperform" them in consistency of interventions and recognizing underlying patterns in data and interactions. The necessary competences in organizations for acquiring, developing and running AI-enabled services were a concern.

Practical Trials of AI in Career Guidance

The first practical trial was conducted with an AI-enabled application that recommended courses and employment based on student's current study records and enrolment information. Information about available jobs and work placements came from a public database of employment offices. Notifications were sent via a mobile student app at their institution.

Giving feedback on the trial, students reacted positively to the idea of receiving suggestions and assistance ("I have so far only received one set of recommendations. They were appropriate, encouraging and rather timely"). They gave positive feedback on receiving information on the advancement of their studies ("it was encouraging to see that my studies are progressing, as I want to graduate soon"), sometimes linked to their wellbeing ("it was nice to get a message reminding me to also take rest").

Students appreciated the features of the app but found the accuracy of the AI-powered suggestions lacking. The usefulness of the app was rated at a median of 6 (out of 10, n=101) in an online survey a month into the trial. Out of the 63 students who responded they had received course suggestions 56% indicated that these courses were relevant to them. Some however noted that they were already attending the recommended courses. Lack of relevance was attributed to, for example, location ("I cannot take this course as distance learning"), schedule (“the courses might be interesting but do not currently fit my schedule”) or study field (“the suggestions are not within my major Tourism and service business, but in the field of health care”). Out of the 82 students who responded they had received job suggestions 62% indicated that these jobs were not relevant to them. Lack of relevance was attributed to, for example, location (“of course the AI does not know that I do not live in Finland”), field of work (“I am not interested in work in financial administration”) or career stage (“jobs available were not for beginners”).

Students wished both the information used and the delivery of the messages would be developed into more personalized direction. Students were curious to understand why certain courses or jobs were suggested and how they could provide more information for the AI on for example interests not (yet) reflected in their study or career data. Students expressed frustration when receiving multiple similar messages on the advancement of their studies, requesting more varied communication and frequently updating information.

Students hoped that they could utilize services like these when enrolment was timely and that the offering would also cover continuous education opportunities. They asked for practical features for setting up reminders and receiving study technique tips. Students envisaged that AI would mine and manage “important information” from online learning platforms, portfolios, personal email and student services serving to collate their data, enable them to control deadlines and locate appropriate services. Possible social features arose during the co-design process and via feedback, supporting networking with other students, building communality and keeping in touch with students and staff. Students recommended giving the AI “some personality”– selecting a random persona was suggested to make it more approachable. There were ideas to increase engagement and playfulness of the interaction, visualizing achievements and encouraging progress.

The second trial was conducted with a web application that utilized labour market information from commercial sources and made use of skills data more extensively. Before courses or jobs were recommended, students in two groups (n = 5 & n = 3) created their personal competence profile by compiling documents (e.g. CV) and inputting skills terms via a dedicated user interface. This resulted in a skills profile students could update when browsing courses or job ads, further refining the matches. The idea was to make the skills profile explicit, increase the visibility of the underlying matching, and enable more accurate matching.

Qualitative feedback indicates that students experienced value in verbalizing their skills and searching job ads and courses via the application. In addition to personal profiles, users experimented with creating general skills profiles for common professions in their domain, for example, sales, accounting, human resources, project management. Junior students in particular explained they are not familiar with the skills requirements of certain jobs and appreciated having a tool to explore these.

Discussion

Role of AI in Career Guidance

In focus groups, students and staff envisioned similar roles for the AI in information delivery, case management and intelligent analytics. Students tended to form concrete service ideas and describe interactions with the AI tool. Students cast AI into roles that ranged from discovery tool to pedagogical companion in their education, extending the uses to self-management tasks

(Sampson et al., 2003). Staff described AI-enabled guidance processes where the AI was an assistant to staff, rather than directly to the student.

Scenarios prompted guidance professionals to conceptualize (Tsekleves et al., 2017) new career services, where AI could be leveraged for the benefit of the student, staff and institutions. Staff even named these services and described what functionalities they should have, what data they would run on and how they would be used. When discussing the potential of AI, staff estimated that would depend on the student and their needs. References were made to matching the type and level of guidance to individual needs and types of students (Sampson et al., 2003).

Staff raised discussion about the respective roles of humans and AI. They voiced concerns, stressing the importance and role of human interventions even when facilitated by the AI process. This might reflect cautionary attitudes towards new technology and the need to mediate the interaction of students with technology. Participants envisioned a career guidance process where human and artificial effort and competences would be combined, similarly to Khare et al. (2018) who argue for a synergistic integration of human and AI support for student success. In addition to maximizing benefits, an integrated approach also moderates the risks of technology use (Fusco et al., 2020).

By Bandura's (2006) criteria, AI is not an agent as it lacks moral agency. However, the concept of proxy agency can be employed for the joint agency that users and tools possess (Neff & Nagy, 2018). This is indeed how participants described the process of developing and using AI-powered tools in guidance: extending their own competences and resources with the tools, wanting to "outsource" or "delegate" tasks to their envisioned AI collaborators capable in information retrieval, optimization, and visualization.

The construction of agency in AI-enabled guidance can be seen as an interactive process where agency can manifest via multiple modes. Table 1 details the potential role of AI in guidance on a continuum, giving examples from the study. The role of AI moves along a continuum from tool to assistant, then collaborator and eventually to coach (Kantharaju et al., 2018). This echoes the reality-virtuality continuum posited by Milgram and Kishino (1994) and the concept of augmenting human capabilities with technology (Raisamo et al., 2019).

The agency construed along the continuum expands from direct personal agency exercised with the aid of AI, to proxy agency mediated through the AI, to collective agency created together with AI and possibly even the type of symbiotic or artificial agency. Symbiotic agency is agency constructed within the human-technology interaction, where technology mediates human experiences, perceptions and behaviour, and human agency affects the uses of technology (Neff & Nagy, 2018). Kuijer and Giaccardi (2019) argue that conceptions of "artificial agency" should not focus on autonomy but the process of learning, situated and sustained in interaction. In both these constructs and along the agency continuum, the respective roles of human and machine adapt in interaction to perform optimally together.

Leveraging AI in career services may affect existing structures. Further study on the effect of AI-enabled interventions on the construction and mediation of agency in guidance is necessary in order to develop services that leverage the affordances of students, staff and technology. The construction and modes of agency should be made visible as this would further students' self-reflection and self-regulation as well as the development of guidance practices (Jääskelä et al., 2020).

Table 1
Modes of AI in Career Guidance

Mode	Role of AI	Role of human	Examples from study	Agency
AI as coach	AI acts as an interactive virtual career coach with its goal, developing career guidance practices and processes	Human guides the development of AI and data environment	Virtual career coach mentors students throughout life on career and education choices. Personal learning aid proposes competence development methods based on previous performance and preferences.	Symbiotic/ Artificial
AI as collaborator	AI learns and performs career guidance practices in real-time together with staff for a shared goal	Human works together with AI, teaching it and validating its working regularly	Virtual online counsellor delivers 24/7 guidance alongside staff. Automated weak signals detection combines with staff interventions for dropout prevention.	Collective
AI as assistant	AI assists humans in their career guidance practices in chosen areas with well-defined goals	Human assigns tasks to AI and accepts its results by case	Virtual assistant schedules meetings as needed between students and staff. Smart calendar app creates a study schedule based on enrolments and personal preferences.	Proxy
AI as tool	AI is used by humans in career guidance practices in singular tasks with set goals	Human uses AI-based tools and brings context to its results	Discovery tool maps ads against a fixed skills profile for job recommendations. Dashboard collates labour market data for analysing future competence requirements for redesigning curricula.	Personal

Expectations for AI in Career Guidance

Results from the trials reflect requirements for AI in career guidance. Students expected personalized suggestions according to location, interests, and schedule. This reflects student-centred guidance but also overall expectations towards digital services and underlying personal data. Across trials the need for a holistic learner profile became evident, ranging beyond education into personal information (e.g., interests) and informal learning (e.g., hobbies).

AI-enabled guidance services were considered inherently more accessible due to digital delivery. It was noted that their utilization requires digital devices and competences. This creates a potential conflict if students are unable to benefit from digital services, expected to enhance accessibility in temporal and spatial aspects. In order to account for accessibility, we need to pay attention to factors in the socio-technical system design, underlying algorithms and the interplay between automated and human actions (Holmes et al., 2021).

In focus groups and scenarios AI was expected to analyse vast amounts of data, mine patterns and enable proactive interventions. The trials delivered straightforward suggestions as decision-making aids but more holistic career information “wizards” were envisioned that follow student progress over longer periods. Students experimented with creating future skills profiles for professions, which could be useful for guidance staff in communicating requirements of career options.

Staff expected AI to enable them to redirect resources to more complex cases (Martiniello et al., 2020) and students benefitting from personal guidance (Sampson et al, 2003). Staff stressed the need for human connection for encounters and communication in career services, arguing for a strategic combination of human and artificial effort. Such an integrated approach (Hooley et al., 2015) could combine benefits from various technologies and moderate risks. It should be further investigated what are the areas of guidance where AI could bring most added value and how value is co-created in guidance interaction with humans and technology.

AI was seen also as a tool for guidance staff. This type of adoption could facilitate the rollout of new technology and enable development of AI-enabled services without immediate disruptions for students. Supporting staff collaboration was also a potential application of AI, echoing co-careering conceptualized by Kettunen et al. (2013) in the use of social media.

Staff wanted to clarify the responsibilities regarding AI, stressing trust and transparency within the guidance process. It was not obvious under what domain this might fall. In other words, AI in guidance might be a pedagogical (teachers), digital (IT department), or even a management (leadership) issue. This connects to the finding that the role of AI in career guidance is seen dual-fold (to support the student and to support the staff supporting the student) and that the respective roles of guidance staff and AI are diverse (staff in the foreground and AI in the background or vice versa).

Utilizing technology in guidance places novel demands on staff competences and attitudes towards technology. The particular competence requirements arising from AI have yet to be studied. Competences are required not only to use the tools in providing guidance but also for developing the services. When dealing with AI, competence for developing services would include both developing teaching data sets as well as validating the AI models. These comply with Sampson et al.’s (2020) suggestions that guidance staff should actively participate in the design, use, and evaluation of technological interventions. This evaluation should be holistic, covering the integration of AI into the process and goals of guidance in order to avoid problems and maximize the effectiveness of services.

Data for AI in Guidance

AI was described as an enabler for extending the career guidance information environment and collating data for guidance. Currently the largest bottleneck in leveraging AI technologies in education is a lack of data, especially domain-specific data (Tuomi, 2018). Intelligent tools can be useful in career guidance only if we have meaningful data for them to process. In trials the lack of location and scheduling information rendered AI-powered suggestions of studies or employment irrelevant. Data are needed first for initially training the systems, and then in a continuous and dynamic manner for delivering services. The data used to train future AI models will also shape future services.

New technologies channel greater amounts of information to individuals (Bakke et al., 2018). Further research is needed on which data are useful for the goals of career guidance. Models and solutions on data sharing across institutions should also be developed as existing data mostly covers formal learning within a single institution. Career guidance extends this into highly dynamic labour market information. Ensuring access and interoperability in the various information sources should be a priority.

Participants referred to various information relevant to career guidance. Figure 1 categorizes these into overlapping personal, career and education information. In continuous guidance, learners should be able to control their own data, submitting it to the platforms and service providers they use for developing competences and accessing career services. This type of "MyData" approach to competence and career information would support learners' agency. Shared data would enable service providers to develop and deliver services accessible throughout an individual's education and work life, not only tied to a specific enrolment or employment.

Figure 1
Career Guidance Information Environment

Ethics and Risks for AI in Career Guidance

The interaction with AI scenarios expectedly (Tsekleves et al., 2017) prompted guidance professionals also to identify a number of unwanted and negative consequences of the service concepts described by them and their colleagues. The risks of overusing or misusing AI (Floridi et al., 2018) were present in these comments. None of the participants raised concern over potential underuse, which results in opportunity costs as the benefits offered by technology would not be realised (Floridi et al., 2018).

The concerns raised by staff related to issues of quality, control, changing roles, confidentiality, privacy and equality in career guidance services as well as necessary competences and resources for providing the AI-enabled services. Students made less explicit remarks about ethical issues. When faced with AI-powered suggestions in trials, student feedback included questions about the data used and how the algorithms work, reflecting a need for transparency.

Recently, policy efforts have been directed towards sustainable development and risks associated with AI in the education domain (Pedró et al., 2019; European Commission, 2021). These do not yet address the specific dynamics and potential of AI in career guidance. The recent AI ethics guidelines by the European commission (n.d.) may serve as a basis for elaborating shared guidelines for AI developers and guidance staff.

Methodological Issues and Implications

The convergent parallel multimethod approach aided in clarifying and enriching findings across methods. The focus groups served to mine existing user needs, scenarios extended them further into possible futures and trials provided a space to experiment on specific implementations. The contribution of various methods, including creative ones, was beneficial to the study on the role of AI. Research on human-technology interaction should be multidisciplinary, involving viewpoints and methodology from such varied fields as education, cognitive psychology, human-computer interaction, artificial intelligence, social psychology, and communications (Kim & Baylor, 2006).

The limitations of this study include biases in the participating groups and context-specificity in the national environment. The case studies relied on volunteers, which may bias participants towards technology enthusiasts. Care was taken to ensure multiple ways of engagement in workshops (voice, chat and anonymous commenting) to enable wide participation. Student age varied but most (86%) were in their first or second year of studies during trials. Wider trials are planned for advanced application.

Finland has a harmonised and effective education system, considered one of the most successful in the world (Laukkanen, 2007). The existing infrastructure made it possible to utilize certain data sources in trials, which may be unavailable elsewhere. The AI applications built for this study were mature but not cutting-edge technology. They were in Finnish for usability. There are known issues in Finnish language models for AI, caused by compound words and suffixes. Further training of the model and work on cross-language interoperability is needed.

The findings of this study can be applied when planning and designing technology-assisted career guidance services, as well as monitoring the uptake and results of AI interventions. The authors apply the maturity model of Saari et al. (2018) to the use of AI into career guidance, mapping out maturity levels in organizations (Table 2). This model supports planning of intelligent technology use and its continuous assessment. The development should tie into the digitalization strategies in higher education as well as artificial intelligence roadmaps.

Conclusion

The current technological advances and their implications for society pose a manifold challenge for education. There is a need to prepare students, staff and organizations for AI-enabled education, as well as to develop AI to better understand the education domain.

This study investigated the possibilities and requirements for using AI in career guidance, including mapping out future research considerations. Technology is not simply a tool in guidance, as it has the potential to extend and transform services and practices. AI can serve students and staff in the various modes within career guidance services, depending on user needs, staff competences and organizational capability for leveraging technology.

Table 2
AI Maturity in Career Guidance

Maturity level	Data	Technology	Processes	Services	Competences
AI-trans-formed guidance	A holistic structured data ecosystem of external and internal information in real time	Interactive AI used seamlessly in career guidance practice	Human–AI cooperation in career guidance is planned and daily	AI scaled into career guidance, creating measurable value	Active contribution by organization to AI applications in career guidance
AI-integrated guidance	Data management designed for AI and integrated into guidance processes at organization level	Real-time AI models utilized for various career guidance activities	AI integrated into career guidance at a process level	AI is an integral part of career guidance services	Networked AI team cooperates with external stakeholders and reports to management
AI-informed guidance	Valid data available in a structured format in discrete subdomains	Separate tools for AI utilized in career guidance tasks in batch mode	Automation enhanced with AI in individual use cases within career guidance	A roadmap for AI in career guidance exists, and standalone implementations are initiated	AI experts work as a team, bridging guidance with other processes such as research
AI-aware guidance	Career guidance data legally validated for use in AI	Traditional analytics tools used in career guidance	AI opportunities identified from career guidance processes	User needs for AI in career guidance surveyed and evaluated	Individual resources and competences for AI exist or are available through partnerships

The authors have provided suggestions for the use of AI in career guidance processes. Artificially augmented guidance is already becoming technologically accessible. However, the visions in this article remain largely unrealized or of low maturity. Further research and development are needed to develop AI-related competences, design AI career guidance solutions that add value to student and staff, integrate AI into guidance processes and roles sustainably, enrich career data ecosystems, and ensure trustworthiness of artificial intelligence technology.

Acknowledgements
The authors appreciate the financial support from European Social Fund for the projects Artificial intelligence-aided study and career counselling (S21743) and the student's counselling for the work of the future (S21722).

The authors would like to acknowledge graphic designer Nea Vaskelainen for illustrating Figure 1 as well as Erkki Räsänen and Saija Silvennoinen for commenting on early manuscript versions.

References

Attwell, G., Bekiaridis, G., Deitmer, L., Perini, M., Roppertz, S., & Tutlys, V. (2020). *Artificial intelligence in policies, processes and practices of vocational rducation and training*. Institut Technik und Bildung, Universität Bremen.

Bakke, I.B., Hagaseth Haug, E., & Hooley, T. (2018). Moving from information provision to co-careering: Integrated guidance as a new approach to e-guidance in Norway. *Journal of the National Institute for Career Education and Counselling*, *4*(1), 48–55. https://doi.org/10.20856/jnicec.4108

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, *1*(2), 164–180. https://doi.org/10.1111/j.1745-6916.2006.00011.x

Barnes, S.-A., Bimrose, J., Brown, A., Kettunen, J., & Vuorinen, R. (2020). *Lifelong guidance policy and practice in the EU: trends, challenges and opportunities. Final report*. Publications Office of the European Union. https://doi.org/10.2767/91185

Cedefop, European Commission, ETF, ICCDPP, ILO, OECD, & UNESCO (2020). *Career guidance policy and practice in the pandemic: results of a joint international survey – June to August 2020.* Publications Office of the European Union. http://data.europa.eu/doi/10.2801/318103

Cox, A. M. (2021). Exploring the impact of Artificial Intelligence and robots on higher education through literature-based design fictions. *International Journal of Educational Technology in Higher Education, 18*(3). https://doi.org/10.1186/s41239-020-00237-8

Council of the European Union (2008). *Council Resolution on better integrating lifelong guidance into lifelong learning strategies.* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:42008X1213(02)

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage.

Eteläpelto, A., Vähäsantanen, K., Hökkä, P., & Paloniemi, S. (2013). What is agency? Conceptualizing professional agency at work. *Educational Research Review*, *10*, 45–65. https://doi.org/10.1016/j.edurev.2013.05.001

European Commission (n.d.). *Ethics guidelines for trustworthy AI.* https://ec.europa.eu/futurium/en/node/6945.

European Commission (2021). *Proposal for a Regulation laying down harmonised rules on Artificial Intelligence*. COM (2021) 206. https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

European Commission (2019). *Key Messages on Artificial Intelligence in Education Education & Training 2020 Working Group on Digital Education: Learning, Teaching and Assessment.* Accessed from https://ec.europa.eu/social/vocational-skills week/sites/evsw/files/ai_key_messages_final_and_es.pdf

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Charezand, P.,Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vavena, E. (2018).

AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds & Machines, 28*, 689–707. https://doi.org/10.1007/s11023-018-9482-5

Fusco, L., Parola, A., & Sica, L. S. (2020) Designing meaningful career tools: a proposal for an optimal use of technology in career guidance. *Proceedings of the Second Symposium on Psychology-Based Technologies.*

Greene, J. C. (2007). Mixed methods in social inquiry. San Francisco, CA: Jossey-Bass.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Igbert Bittencourt, I., & Koedinger, K. R. (2021). Ethics of AI in Education: Towards a Community-Wide Framework. International *Journal of Artificial Intelligence in Education*. https://doi.org/10.1007/s40593-021-00239-1

Hooley, T., Shepherd, C., & Dodd, V. (2015). *Get yourself connected: Conceptualising the role of digital technologies in Norwegian career guidance*. International Centre for Guidance Studies, University of Derby.

Jääskelä, P., Heilala, V., Kärkkäinen, T., & Häkkinen, P. (2020). Student agency analytics: Learning analytics as a tool for analysing student agency in higher education. *Behaviour & information technology*, *40*(8). https://doi.org/10.1080/0144929X.2020.1725130

Kantharaju, R., De Franco, D., Pease, A., & Pelachaud, C. (2018). Is two better than one?: Effects of multiple agents on user persuasion. *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, 255–262. https://doi.org/10.1145/3267851.3267890

Kettunen, J., & Sampson, J. P. (2019). Challenges in implementing ICT in career services: Perspectives from career development experts. *International Journal for Educational and Vocational Guidance*, 9, 1-18. https://doi.org/10.1007/s10775-018-9365-6

Kettunen, J., Vuorinen, R., & Sampson, J. P. (2013). Career practitioners' conceptions of social media in career services. *British Journal of Guidance & Counselling*, 41(3), 302–317. https://doi.org/10.1080/03069885.2013.781572

Khare, K., Stewart, B., & Khare, A. (2018). Artificial intelligence and the student experience: An institutional perspective. *IAFOR Journal of Education, 6(*3), 63–78. https://doi.org/10.22492/ije.6.3.04

Kim, Y., & Baylor, A.L. (2006). A social-cognitive framework for pedagogical agents as learning companions. *Education Technology Research and Development,* 54, 569–596. https://doi.org/10.1007/s11423-006-0637-3

Laukkanen R. (2008). Finnish strategy for high-level education for all. In N. C. Soguel (Eds.) *Governance and Performance of Education Systems*. Springer. https://doi.org/10.1007/978-1-4020-6446-3_14

Kuijer, L., & Giaccardi, E. (2018). Co-performance: Conceptualizing the role of artificial agency in the design of everyday life. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3173574.3173699

Martiniello, N., Asuncion, J., Fichten, C., Jorgensen, M., Havel, A., Harvison, M., Legault, A., Lussier, A., & Vo, C. (2021). Artificial intelligence for students in postsecondary education: A world of opportunity. *AI Matters*, *6*(3), 17–29. https://doi.org/10.1145/3446243.3446250

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 77(12), 1321–1329.

Moreno-Guerrero, A,-J., López-Belmonte, J., Marín-Marín, J.-A., & Soler-Costa, R. (2020). Scientific development of educational artificial entelligence in web of science. *Future Internet*, *12*(8), 124. https://doi.org/10.3390/fi12080124

Neff, G., & Nagy, P. (2018). Agency in the digital age: Using symbiotic agency to explain human–technology interaction. In Z. Papacharissi (Eds.) *A networked self and human augmentics, artificial intelligence, sentience* (pp. 113–123). Routledge

Organisation for Economic Co-operation and Development (OECD). (2004). *Career Guidance and Public Policy: Bridging the Gap*. https://doi.org/10.1787/9789264105669-en

Peavy, R. V. (2000). A sociodynamic perspective for counselling. *Australian Journal of Career Development*, *9*(1), 17–24. https://doi.org/10.1177/103841620000900105

Pedró, F., Subosa, M., Rivas, A., & Valverde, P. (2019). *Artificial Intelligence in education: Challenges and opportunities for sustainable development.* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000366994

Raisamo, R., Rakkolainen, I., Majaranta, P., Salminen, K., Rantala, J., & Farooq, A. (2019). Human augmentation: Past, present and future. *International Journal of Human Computer Studies*, 131, 131-143. https://doi.org/10.1016/j.ijhcs.2019.05.008

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, *26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

Russell, S., & Norvig, P. (2016) Artificial Intelligence: A Modern Approach (4th ed.). Prentice Hall.

Saari, L., Kuusisto, O., & Pirttikangas, S. (2019). AI Maturity Web Tool Helps Organisations Proceed with AI [White paper]. VTT Technical Research Centre of Finland. https://doi.org/10.32040/Whitepaper.2019.AIMaturity

Sampson, J. P. Jr., Peterson, G. W., Reardon, R. C., & Lenz, J. G. (2003). *Key elements to the CIP approach to designing career services*. Center for the Study of Technology in Counseling and Career Development, Florida State University.

Sampson, J. P., Kettunen, J., & Vuorinen, R. (2020). The role of practitioners in helping persons make effective use of information and communication technology in career interventions. *International Journal for Educational and Vocational Guidance*, *20* (1), 191-208. https://doi.org/10.1007/s10775-019-09399-y

Strategy for Lifelong Guidance 2020–2023. National Lifelong Guidance Forum. (2020). Publications of the Finnish Government 2020:34. http://urn.fi/URN:ISBN:978-952-383-536-8

Toni, A., & Vuorinen, R. (2020). Lifelong guidance in Finland: Key policies and practices. In E. Hagaseth Haug, T. Hooley, J. Kettunen, & R. Thomsen (Eds.) *Career and career guidance in the Nordic countries*, *career development series* (vol. 9) (pp. 127–143). Brill. https://doi.org/10.1163/9789004428096_009

Tsekleves, E., Darby, A., Whicher, A., & Swiatek, P. (2017). Co-designing design fictions: A new approach for debating and priming future healthcare technologies and services. *Archives of Design Research*, *30*(2), 5–21. https://doi.org/10.15187/adr.2017.05.30.2.5

Tuomi, I. (2018). *The impact of artificial intelligence on learning, teaching, and education: Policies for the future*. Publications Office of the European Union. https://doi.org/10.2760/12297

Zawacki-Richter, O., Marín, V.I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Education Technology in Higher Education,* 16, 39. https://doi.org/10.1186/s41239-019-0171-0

Ågerfalk, P.J. (2020). Artificial intelligence as digital agency. *European Journal of Information Systems,* 29(1) 1-8. https://doi.org/10.1080/0960085X.2020.1721947

Corresponding author: Stina Westman
Email: stina.westman@xamk.fi

Five Tips from Filmmakers: An Online Instructional Module for Documentary Film Research

Patsy Y. Iwasaki
University of Hawai'i at Mānoa
USA

Abstract

Information about the overall documentary filmmaking process is available; however, there is a lack of literature and educational resources about how to conduct subject research and data collection. As documentary filmmaking becomes an increasingly democratic endeavor due to technology, and information distribution and education use increases, there is a corresponding need for quality resources to support this essential step. In this study, a content-rich, technology-enhanced, online instructional module that was designed and developed in another study to guide and assist beginner documentary filmmakers with subject research and data collection, was implemented and evaluated by the target audience. This module featured five filmmaking tips summarizing professional documentary filmmakers' wisdom and expertise with subject research and data collection. Motivational and instructional models served as frameworks to inform and guide the study's learning design process. The quantitative and qualitative findings, field notes and observations provided data triangulation. After analysis and interpretation were completed, the results significantly confirmed the module had a positive, educational impact on the target audience and accomplished its purpose. This module addressed the lack of resources and utilized consultation of experts in content design and development to improve the creativity and production of beginner documentary filmmakers. This project successfully merged learning sciences theory and instructional design with humanities and arts research. It will contribute to the literature of documentary film research studies, the fields of instructional design and education, and the humanities. It has significant potential to influence and impact the broad possibilities of innovative, interdisciplinary research design and collaboration.

Keywords: instructional design and content development, documentary filmmaking, research and data collection, documentary film instructional module, ADDIE design model, ARCS motivational design concepts

The documentary film genre is a dynamic and effective information and communication medium that educates, inspires, and motivates its audience. It is a nonfictional documentation of fact-based reality, and its purpose is to share knowledge, increase understanding, or preserve historical records. Aufderheide (2007) defines a documentary as a film or video that "tells a story about real life, with claims to truthfulness" (p. 2). Documentary film's increasing utilization in educational settings, as well as its advantages in distributing information to extensive audiences, is well-timed with the field's technological advances. What was an expensive undertaking can now be created with inexpensive equipment and software applications (Loustaunau & Shaw, 2018; Winston, Vanstone & Chi, 2017). Democratizing the industry allows more people, from diverse populations, to tell meaningful stories globally via documentary film.

Although information about the overall process is available, literature on the research and data collection step of documentary filmmaking, which is essential to the production process, is limited. There is a lack of information and quality educational resources about *how* to conduct subject research and collect data for documentary film in either scholarly or popular sources (Adorama Learning Center, 2018; Desktop Documentaries, 2018). This is a disadvantage for many beginners and students exploring documentary film production who might not know how to begin subject research and data collection for documentary film. There is a clear need for informative and user-friendly educational resources to fill this gap.

An educational resource about subject research and data collection, generated from the knowledge and experiences of professional filmmakers, would be extremely helpful to student filmmakers. Receiving a head start in the challenging filmmaking process, beginner filmmakers could share their important stories with the world more quickly and with improved creativity and production. This expert guidance would not only benefit beginner filmmakers, but also audiences and society at large (Leavy, 2015; Loustaunau & Shaw, 2018). An increased availability of timely documentaries would give audiences expanded learning opportunities.

Thus, this study's purpose was to implement and evaluate the educational value of an existing resource about subject research and data collection for documentary film, one that was designed and developed in another study (Iwasaki, 2021; Iwasaki, in progress). This content-rich, online, interactive instructional module, titled *Five Tips from Filmmakers for Documentary Film Research and Data Collection,* is presented on a website. The five tips featured the wisdom and recommendations of professional filmmakers from actual filmmaking experience. The module utilized technology and consultation of experts in content design and development to support and improve the overall learning, creativity and production of student filmmakers. This article is comprised of the following sections: literature review, methodology, results, discussion and conclusion, and references.

Literature Review

Steps to a Documentary

Generally, traditional documentary film takes a highly subjective approach. The filmmaker directs the entire process, from subject selection and research, to creative approaches (Bell, 2011; Friend & Caruthers, 2016). Filmmakers claim it is an extremely rewarding experience, but also one of the most challenging and demanding creative endeavors. It is often a long and arduous process with many steps including budgeting, planning, script writing, production, editing and distribution. Since information or instruction is limited or undocumented, and producing documentary films can be such a creative, artistic process, many filmmakers learn by intuitively moving the project to completion (Adorama Learning Center, 2018; Desktop Documentaries, 2018).

The first step in creating a documentary film is to find a subject that is important to the filmmaker and is of interest to others. Since the documentary journey is often daunting and complicated, the subject needs to both energize and sustain the filmmaker for the lengthy work, and the filmmaker needs to feel compelled to share the story through film.

The second step is "research and data collection," an essential task to find resources and conduct interviews that includes background, history and context of the subject, as well as the interesting, credible, emotional and inspiring material that will resonate with the audience. This step is crucial because it determines the content of the film (Aufderheide, 2007; Bell, 2011; Frank 2013; Studio Binder, 2018; Winston et al., 2017).

Documentary Film and Education

How documentary films are used in education. Along with technological advances and lower film production costs, documentary film's information distribution and social commentary use opened the doors for its increasing applications in education (Aufderheide, 2007; Bell, 2011; Nash, Hight, & Summerhayes, 2014; Winston et al., 2017), and social science research and methodology (Frank, 2013; Goldman, Pea, Barron, & Derry, 2007). The anthropology field embraces documentary film using terms such as ethnographic film and ethnocinema (Harris, 2012; Leavy, 2015; Sjöberg, 2008). Frank (2013) expanded the educational significance, practice and application of the genre and Bell (2011) emphasized the importance of its historiographical research and scholarship. Whiteman (2004) discussed documentary films' political impact upon audiences, and Fonda (2014) combined art therapy and filmmaking.

Documentary film is used in varied research and teaching contexts and approaches from elementary education onward (Aufderheide, 2007; Frank, 2013). They can range from roughly coordinated projects to planned and scripted professional productions with a cinematographer, crew, and actors. Some may also highlight the researcher(s), participants, and others (Leavy, 2015; Leavy & Chilton, 2014). Documentary film is popular for researchers and instructors hoping to inspire and promote awareness of diverse issues such as climate change and social justice (Friend & Caruthers, 2016; Hanley, Noblit, Sheppard, & Barone, 2013); migration (Loustaunau & Shaw, 2018); to the environmental concerns of agricultural chemicals (The Monsanto Papers, 2018).

How students learn about documentary filmmaking. Just as there are steps in the documentary filmmaking process, a filmmaker's path necessitates specific actions. The journey can include obtaining a degree in film studies. Some believe a successful career requires enrolling at an elite film program at the American Film Institute or University of Southern California, while others believe that experiential learning is more important. Tuition at these schools is approximately $60,000 per year for undergraduates and $65,000 for graduate students (Learn How, 2020; Galuppo & Chuba, 2020).

Although access to state-of-the-art equipment and the networking opportunities that may come with attending these schools are helpful, there are no requisite rules. Many leading filmmakers say their passion for filmmaking began while they were young, making short videos with friends and family, and that experience, matters the most. "Even films created on a smartphone and edited on a laptop can convey your raw talent and eye for cinematography." (Learn How, 2020, para. 4).

Quality online educational resources about the documentary filmmaking process would further democratize the industry for beginner filmmakers. The overall goal of this study was to add an original, informative, and free educational resource to help student filmmakers make their goals a reality.

The Current Study

The purpose of this study was to evaluate the educational value and impact of an instructional module about subject research and data collection for documentary film that was designed and developed in another study (Iwasaki, in progress) to address the lack of information and inadequate educational resources about this topic.

The module featured beneficial recommendations by professional filmmakers from their actual filmmaking experiences (Iwasaki, 2021) condensed into five major tips with explanatory information: 1) do the research, 2) tell the story visually, 3) find strong characters, 4) support universal themes, and 5) relate to your audience. This study implemented and evaluated the module to answer the research question:

> RQ: What is the impact of an instructional module for documentary film subject research and data collection upon student filmmakers learning about and exploring documentary film production?

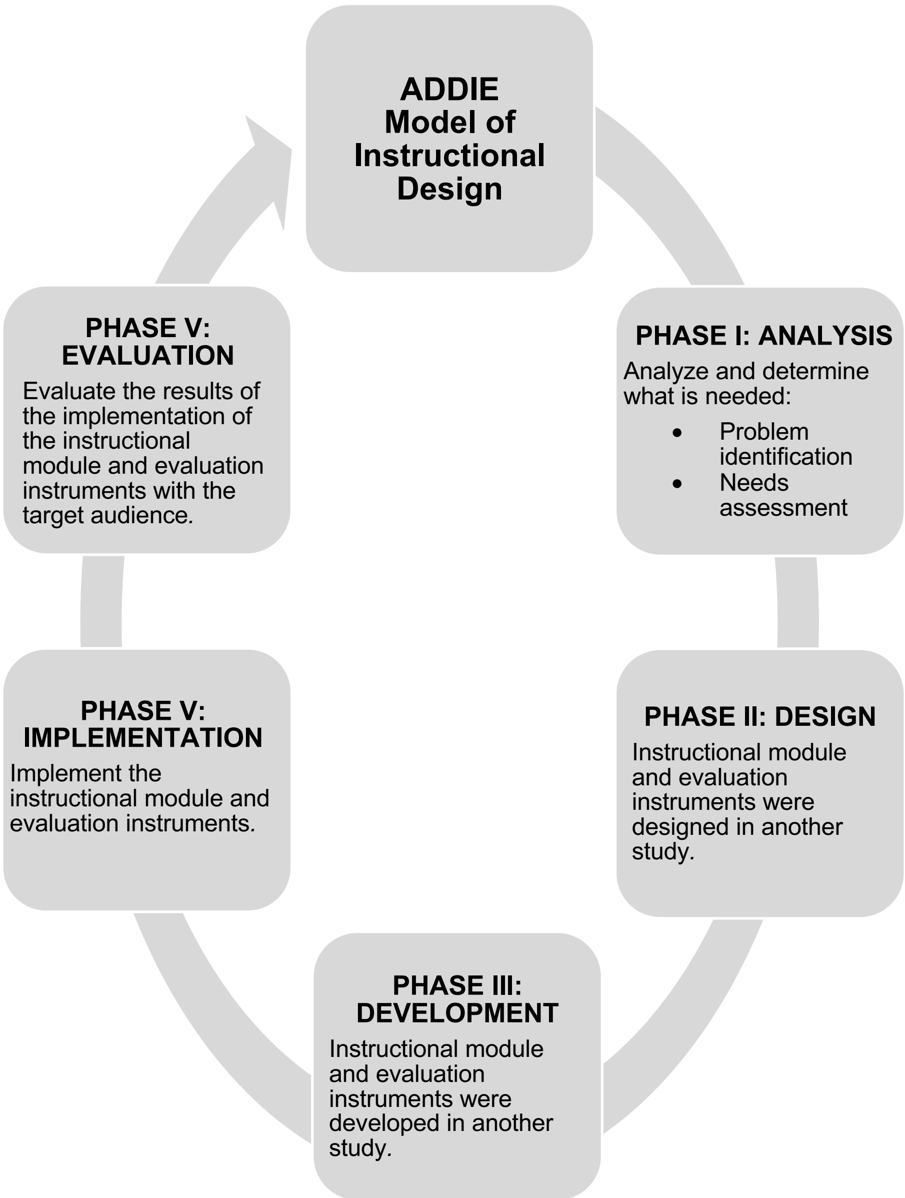
Methodology

Research Design

A convergent mixed methods design approach was used (Creswell, 2009, 2015, 2018) to gather quantitative and qualitative data from the target audience during the same time frame, rigorously strengthening the data collection, analysis and interpretation of the study. Along with the quantitative and qualitative instruments, field notes and observations were recorded and used in analysis and interpretation, applying triangulation to strengthen the study and increase credibility and validity (Glesne & Peshkin, 1992; Yin, 2016). A pilot test (Bryman, 2012) of the instructional module and evaluation instruments was conducted with a student and confirmed that the instruments worked well to collect the necessary data. Integration of the study's mixed methods approach ensured feasibility and the data was recorded, documented, and validated as much as possible (Bryman, 2012; Maykut & Morehouse, 1994; Yin, 2016).

ADDIE model of instructional design. The ADDIE model of instructional design, with its iterative phases of Analysis, Design, Development, Implementation and Evaluation, systematically guided and organized this study's learning design process (see Figure 1) (Dick, Carey, & Carey, 2001; Gagne, Wager, Golas, & Keller, 2005; McGriff, 2000; Molenda, 2015; Molenda, Pershing, & Reigeluth, 1996; Serhat, 2017). The process included formative assessments by experienced reviewers that informed and guided the strategic design modifications involving multiple iterations of the instructional module (Iwasaki, in progress). The final format of the module was important to this study. A basic, single-page infographic progressively evolved to become an engaging, easy-to-navigate, multimedia module on a website. The ADDIE process ensured that the study focused on the best educational practices for the target audience, beginner filmmakers.

Figure 1
Representation of the study using the ADDIE model



Analysis: Problem identification. A lack of information and educational resources about subject research and data collection for documentary film in either scholarly or popular sources (Adorama Learning Center, 2018; Desktop Documentaries, 2018) is a problem as many students exploring documentary film production might need assistance on how to begin research and data collection.

Analysis: Needs assessment. Since the target audience is student filmmakers in higher education, addressing their academic, cognitive, and physical needs were essential to the study's design. U.S. college student generalizations include: diverse population at 42% non-white, 62% work part- or full-time, demand for digital technology, self-paced learning, multimedia integration, online learning's flexibility and convenience, need for financial aid and affordable housing, and focus on the cognitive knowledge, exploration, discovery and meaning needs (Bill & Melinda Gates Foundation, 2020; Education Dive, 2017; McGraw-Hill, 2016; National Center for Education Statistics, 2020). This study was also informed by the students at a four-year comprehensive, regional university in the western United States who participated in this study. The learner profile includes a diverse, mostly full-time undergraduate and graduate student population.

Design and development. The design and development of the module and evaluation instruments for this study were completed in other studies (Iwasaki, 2021; Iwasaki, in progress).

Implementation. The implementation of this study will be reported in the Methodology section.

Evaluation. Evaluation of this study will be reported in the Results section.

Participants

There were 17 participants out of a pool of 20; students taking a free, 11-week non-credit class on film and video production at a four-year university in the western United States, and other students interested in documentary film production at the same institution. Taught by two professionals, the course covered pre-production, production and post-production, including educational and documentary filmmaking. Some of the students taking the course were also associated with the university's video production program. An exempt status IRB approval was secured for the study.

The majority of students, seven students (44%), were sophomores. College majors of the participants varied widely. Two participants declared Administration of Justice, two Business, two Education, two Kinesiology, and the rest were individually different: Accounting, Biology, Communication, English, Environmental Science, Hawaiian Studies, Nursing, Political Science, and Psychology.

Instructional Module

Table 1 below presents the content in the module: five major themes aligned with the ARCS model concepts of attention, relevance, confidence and satisfaction (Keller, 1983, 2010, 2017), and organized as five beneficial "tips," summarizing the wisdom and knowledge of professional filmmakers from their actual filmmaking experiences. This information was derived and condensed from the results of a previous study (Iwasaki, 2021). The module included photos supplementing each tip, and short audio clips of filmmakers elaborating highlights of each tip (Iwasaki, in progress).

Table 1
Five tips from filmmakers and their relationship to the ARCS concepts

Tip 1. Do the Research (Confidence) • Complete an exhaustive resource search. • Collect existing material about your subject. • Identify key characters who can tell the story. • Determine experts who can add legitimacy. • Pinpoint a gap in the story, or a lack of the story. • Fill that void with your documentary film.
Tip 2. Tell the Story Visually (Attention) • Collect interesting interviews, historical documents, material, photos, videos and supplemental footage. • Answer why this story needs to be presented visually. • Determine if the sources and materials are accessible. • Establish an organized system for all of the data. • "Show" the audience, not just tell the audience.
Tip 3. Find Strong "Characters" (Attention) • Focus on the strength of your interview sources. • Feature genuine interview characters who are engaging, fascinating, vulnerable, revealing, and who feel true. • Create an emotional and impactful audience connection. • Generate affinity and empathy with the audience. • Guide the audience on a storytelling journey.
Tip 4. Support Universal Themes (Relevance) • Focus on all-embracing topics such as love, joy, peace, family, survival, pain, suffering, equity, or the striving and struggling one takes to reach a goal. • Unravel the universal human stories and relationships. • Shed light on the shared and collective human experience. • Select topics that entertain and move audiences.
Tip 5. Relate to your audience (Relevance and Satisfaction) • Create a meaningful, relevant story that resonates, informs, educates, inspires and empowers audiences to action. • Help the audience apply the story to the real world, current issues, and to their own lives and circumstances. • Encourage viewers to insert their own stories, experiences and struggles into what they're seeing.

The module began with a welcome page thanking participants for reviewing and evaluating the module, followed by: an overview page with instructions, the five filmmaking tips on separate webpages, embedded evaluation surveys, a webpage featuring all five tips serving as a review, and a final page thanking participants (Iwasaki, in progress). Here is the link to the instructional module website: http://www2.hawaii.edu/~piwasaki/Five_Tips/

Figure 2
Screenshot of "Tip 1. Do the Research" webpage

Instruments

The evaluation instruments that were designed and developed in another study (Iwasaki, in progress) were embedded into the module: the pre-module (pretest) and post-module (posttest) surveys, and five short in-module surveys assessing each of the five filmmaking tips. There were seven surveys in total. A student filmmaker completed a pilot test of the module and instruments, successfully navigating through the module and answering all seven surveys smoothly "with no problems." He said the module was interesting, inspiring and very helpful, noting that the audio clips of filmmakers were especially enlightening and useful. He did not provide any suggestions for improvement.

ARCS model of motivational design. This study sought to evaluate the educational value of an existing instructional module (Iwasaki, in progress) about subject research and data collection for documentary film. To increase the likelihood that the module and evaluation instruments would resonate with the target audience, beginner filmmakers, in addition to applying educational best practices, a well-established model in motivation and instructional design, Keller's ARCS model (1983, 2010, 2017), served as the framework that informed and guided the design and development of all three types of evaluation surveys: pre-module, post-module and in-module. The framework also guided the analysis and interpretation of the data collected. Two other purposes in using ARCS was to: 1) motivate students to become interested in subject research and data collection for documentary film, and 2) inform, guide and instruct students about the process of subject research and data collection for documentary film.

The ARCS motivational factors of attention, relevance, confidence and satisfaction (Pappas, 2015; Peterson, 2003) are strongly applicable to the field of documentary film with similar motivational goals for filmmakers, whether they are veterans or beginners (educators/instructional designers), and their audience (learners/students) (Astleitner & Lintner, 2004; Keller & Suzuki, 2004; West, Hoffman, & Costello, 2017). While the goals for filmmakers may not be referred to as "instructional design," their educational goals are very similar; they use relevant, motivational methods of film and video production (Frank, 2013; Nash et al., 2014; Winston, et al., 2017).

Pre-module survey. The pre-module survey was designed to determine participants' baseline interest in and familiarity with documentary film, and their knowledge of the process of subject research and data collection. To capture participants' voices, one qualitative question was asked: explain why they were interested in film and video production. To measure the degree of participants' feelings, six quantitative questions were designed using a 5-point Likert psychometric scale with appropriate qualifiers (Bryman, 2012; Gorard, 2003).

In-module surveys on the five tips. To draw out the rich and thick descriptions of qualitative data and hear the opinions of participants, five two-question surveys evaluating each of the five tips immediately followed each tip. Question 1 asked: "What stood out to you most?"; and Question 2 asked: "What was the most helpful element? Why?" These short surveys were implemented to allow a deeper exploration of each of the five tips. The "Tip 5. Relate to your Audience" survey also asked which of the five tips was the most helpful overall.

Post-module survey. The post-module survey was designed to determine the amount of learning participants had acquired from reviewing the module. Some of the post-module survey questions paralleled the pre-module survey questions in order to discover the impact of the module and if their knowledge had changed. To capture participants' voices, two qualitative questions were asked: to explain if they felt the module was helpful for research and data collection for documentary film, and if they had anything else to share. To measure the degree of participants' feelings, the 14 quantitative questions were designed using appropriate qualifiers that could be answered using a 5-point Likert psychometric scale (Bryman, 2012; Gorard, 2003).

Procedure

The researcher worked with the two instructors of the film production course and visited the class in January 2020 to introduce herself and provide an overview of the research project with a recruitment presentation, emphasizing that participation was entirely voluntary and would not affect their class participation at all. The researcher informed them she would be returning in March, after Spring break, to present the hard copy educational resource (it had not yet evolved to become an online, technology-enhanced instructional module) (Iwasaki, in progress) for them to review and evaluate. However, due to the COVID-19 pandemic, the university moved entirely to online learning and face-to-face classes were cancelled. The researcher was unable to return to the class as planned.

Thus, the researcher was motivated to develop and implement the resource and the evaluation instruments online using appropriate technology (Iwasaki, in progress). Previous findings by Mayer and other researchers report that multimedia online delivery can increase learning and outcomes (Alessi & Trollip, 2001; Chiu & Churchill, 2016; Clark & Mayer, 2011; Levonen, Biardeau, & Rouet, 2001; Mayer, 2001, 2009, 2017). Self-paced learning, multimedia, technology integration, and online learning were important academic needs for the target audience as discussed in the Analysis section; thus, the move to online implementation was a positive, constructive development (McGraw-Hill, 2016; National Center for Education Statistics, 2020).

The researcher received participants' university email addresses from the instructors, and was able to move forward with the online implementation and evaluation of the resource after its progression into a digital, multimedia instructional module (Iwasaki, in progress). In late April 2020 participants received an email with a link to the module, recruitment letter, and a consent form. They were asked to return signed consent forms or respond affirmatively to the email, review the module, and complete the evaluation surveys via the link within two weeks.

Participants were expected to progress through the module in the following manner: welcome page, overview page with instructions, pre-module survey, review of the five filmmaking tips with a two-question survey following each tip, and a review of all five tips on a single webpage served as a reminder of the overall purpose of the module. Participants were then instructed to complete the post-module survey. Lastly, a webpage thanked the participants for their time and effort in completing the module.

During the two-week period of data collection, 15 students (88%) did not have any problems with viewing and completing the instructional module. After completing the module, some of the students emailed the researcher with positive comments such as: "This is great information!" "The next time I create an educational video or short documentary, I'll remember those five tips." "I was able to complete the module with ease." "I listened to all of the five tips and finished all of the surveys."

Results

Pre-Module Survey

Table 2 below features the results of the 5-point Likert psychometric scale (Bryman, 2012; Gorard, 2003) items for the pre-module survey and their alignment with the appropriate ARCS model concepts of attention, relevance and confidence. When asked to rate their level of interest in film and video production, the 17 participants reported an average rating of 4.12 (*SD* = 0.78). Four participants (24%) selected 5.00 (Strongly Agree). Their familiarity with the documentary film genre reported an average rating of 4.00 (*SD* = 0.71). The participants' familiarity and interest in film and video production were quite strong. However, their familiarity and knowledge about research and data collection for documentary film were weaker. Participants reported an average rating of 2.88 (*SD* = 1.11) when asked about their familiarity with the process of creating documentary film; 2.82 (*SD* = 1.01) when asked about their familiarity with the tools of documentary film; and 2.76 (*SD* = 1.09) when asked about their knowledge of research and data collection. This evidence aligns with the researcher's theory that students need more information and resources in this area.

Table 2
Average interest and familiarity ratings from the pre-module survey (n=17)

ARCS Categories	Interest and/or familiarity with:	Mean	SD	Min	Max
Attention	Film and video production	4.12	0.78	3.00	5.00
Relevance	Documentary film genre	4.00	0.71	2.00	5.00
	Documentary films	3.65	0.61	2.00	4.00
Confidence	Process of creating documentary film	2.88	1.11	1.00	4.00
	Tools of documentary film	2.82	1.01	1.00	4.00
	Research and data collection for documentary film	2.76	1.09	1.00	4.00

Participants answered one open-ended question explaining their interest in film and video production. This question produced thick and rich, narrative type responses. An inductive approach (Bryman, 2012; Maykut & Morehouse, 1994; Yin, 2016) was implemented and responses were analyzed, interpreted, and grouped according to their relationship with the ARCS model concepts of attention, relevance, confidence and satisfaction (Astleitner & Lintner, 2004; Keller, 1983, 2010, 2017).

In-Module Surveys on the Five Tips

The five in-module surveys about each tip featured two open-ended questions to collect rich, qualitative answers and an inductive approach (Bryman, 2012; Maykut & Morehouse, 1994; Yin, 2016) was implemented to analyze, interpret, and organize the data. Participants were asked what stood out most, was the most helpful, and will "take away." Frequent response themes were: collect important and relevant information and images about your subject during the research portion determines whether you have a story worth telling visually; feature characters that will emotionally move and resonate with the audience; and help the audience see themselves in the film and relate to the struggle or conflict.

One example of the depth of the responses was a participant who associated Tip 3. Find Strong "Characters" to the Hawaiian art of hula, saying "Similar to hula, filmmaking is the art of capturing the audience and making them feel a part of the story. Just as hula needs the right *wahine* (female) or *kane* (male), the characters for the film need to be a perfect match."

The five tips' effectiveness was demonstrated by the participants using the exact words used in the tips in their own responses. For example, for Tip 2: Tell the story visually: "'Show' the audience, not just tell the audience," or they paraphrased the bullet points. Each page also featured just one photo that supplemented the content. The instructional module was designed and developed using Mayer's 12 principles of multimedia learning (2001, 2009, 2017; Clark & Mayer, 2011) in another study (Iwasaki, in progress) and one of the principles states: extraneous words, pictures and sounds are excluded rather than included. The data successfully collected in this study supports this principle.

The audio clips, ranging from 34 seconds to 2:05 minutes, featuring professional filmmakers sharing their vast insight and experience were the highlights of the module. The majority of responses for the two questions "what stood out most" and "was most helpful" included the audio clips.

Participants also expressed appreciation of the multimedia and technology integration within the module. One participant said, "Having the audio file while reading through the tips and seeing the photo on the page was helpful. It made me understand how important the visual is." Another said "The text helped me understand what the filmmaker was talking about." This supports Mayer's multimedia design principles of combining graphics, narration and text and that narration is spoken in a friendly human voice. The researcher believes the real-world wisdom and experience from the professional filmmakers was key to the module's success. The audio clips featuring real voices supported the validity and truth of the information. One of the participants said, "Having the audio file to explain did help."

In answering which tip was the most helpful overall, participants selected each of the five tips almost equally. One student said, "In order to develop an astonishing documentary, all tips should be considered. Being able to generate ideas into film is not an easy task, so taking these words of recommendation can really help with the process of creating a documentary."

Post-Module Survey

In the post-module survey, 16 participants answered 16 Likert response items and one qualitative question. Table 3 features the results of the 5-point Likert psychometric scale (Bryman, 2012; Gorard, 2003) and their alignment with the ARCS model concepts.

Table 3

Average interest and familiarity ratings from the post-module survey (n=16)

ARCS Categories	The module/module's:	Mean	SD	Min	Max
Attention	Captured interest and attention	4.38	0.81	2.00	5.00
	Layout and design are appealing	4.44	0.63	3.00	5.00
	Colors are appealing	4.63	0.50	4.00	5.00
Attention/ Confidence	Read and understand information	4.81	0.40	4.00	5.00
Attention/ Relevance	Graphics are interesting and appropriate	4.50	0.52	4.00	5.00
	Sound clips are engaging and interesting	4.19	0.91	2.00	5.00
Confidence/ Satisfaction	Ease of navigation	4.69	0.79	2.00	5.00
	Hear sound clips	4.13	1.36	1.00	5.00
	Access surveys	4.50	1.21	1.00	5.00

Attention/ Relevance	Increased interest in watching documentary films	4.50	0.89	2.00	5.00
	Increased interest in research and data collection for documentary film	4.44	0.73	3.00	5.00
Confidence/ Satisfaction	Increased familiarity with documentary filmmaking	4.56	0.51	4.00	5.00
	Increased familiarity with research and data collection for documentary film	4.56	0.51	4.00	5.00
	Increased knowledge of how to begin the process of creating a documentary film	4.56	0.51	4.00	5.00
Relevance/ Satisfaction	Increased knowledge of how to conduct research and data collection for documentary film	4.50	0.73	3.00	5.00
	Helpful for research and data collection for documentary film	4.44	0.63	3.00	5.00

Participants found the instructional module captured their interest and attention; the layout and design, and colors were appealing; graphics were interesting and appropriate; and the sound clips engaging and interesting. When asked to rate their ability to read and understand information, the 16 participants reported an average rating of 4.81 (*SD* = 0.40); 13 participants selected 5.00 (Strongly Agree). Another high average rating of 4.69 (*SD* = .79) was reported in the participants' ability to navigate the module, which also received 13 "strongly agree" responses. These two items received the most 5.00 (Strongly Agree) selections. These findings validate the successful application of Keller's ARCS motivational model (1983, 2010, 2017) and Mayer's 12 principles for multimedia learning (2001, 2009, 2017; Clark & Mayer, 2011) as frameworks in the design and development of the module (Iwasaki, in progress), and the application of best practices. The lowest average rating of 4.13 (*SD* = 1.36) was reported when participants were asked about their ability to listen to the sound clips. This could be accounted for by two students who had difficulty accessing and reviewing the module. They contacted the researcher, and after troubleshooting with the researcher, the two participants were able to successfully view the module and complete the surveys.

As a posttest, the post-module survey provided positive results of the module's effectiveness. Participants added substantial and important new information to their knowledge base, applying constructivism. After reviewing the module, when participants were asked to rate their level of familiarity with documentary filmmaking, they reported an average rating of 4.56 (*SD* = 0.51), an increase of 0.56 from the pre-module 4.00 survey average. Marked rating increases were seen especially with familiarity about the process of creating a documentary film, and familiarity and knowledge of research and data collection.

When asked to rate their familiarity with the process of creating a documentary film, participants reported an average rating of 4.56 (*SD* = .0.51), a significant increase of 1.68 from the 2.88 pre-module average. When rating their familiarity with research and data collection for documentary film, participants reported an average rating of 4.56 (*SD* = 0.51); and when asked about knowledge of research and data collection for documentary film, participants reported an average rating of 4.50 (*SD* = 0.73), a substantial increase of 1.74 from the 2.76 pre-module average. In the post-module survey, 10 participants selected 5.00 (Strongly Agree), four selected 4.00 (Agree), and two selected 3.00 (Neutral). There were no “Disagree” or “Strongly Disagree” responses. Since the original number is zero, this represents a 100% increase in the “strongly agree” category. In the pre-module survey, 0 participants selected 5.00 (Strongly Agree), five selected 4.00 (Agree), six selected 3.00 (Neutral), three selected 2.00 (Disagree), and three selected 1.00 (Strongly Disagree). Clearly, participants were not very familiar with subject research and data collection for documentary film before reviewing the module, and it was a positive and successful learning experience for the participants.

The module substantially increased participants’ familiarity with documentary film, knowledge of the process of creating a documentary film, and interest in and knowledge on how to conduct research and data collection for a documentary film after deciding upon a subject.

The results also revealed that applying the ARCS model (Keller, 1983, 2010, 2017) as a design and development (Iwasaki, in progress) framework helped motivate students to become interested in subject research and data collection for documentary film. It also helped successfully inform and guide students about the process of subject research and data collection for documentary film, and increased their understanding and knowledge. These were important additional purposes in utilizing ARCS and the data confirmed that the module accomplished this.

The final item was “This module is helpful for research and data collection for documentary film,” and participants reported an average rating of 4.44 (*SD* = 0.63). Eight participants selected 5.00 (Strongly Agree), seven selected 4.00 (Agree), and one selected 3.00 (Neutral). There were no “Disagree” or “Strongly Disagree” responses. This demonstrates the module is a valuable and useful resource to guide and assist beginner filmmakers with documentary film research and data collection.

Participants answered one post-module survey qualitative question: Why is the instructional module helpful for research and data collection for documentary film? An inductive approach (Bryman, 2012; Maykut & Morehouse, 1994; Yin, 2016) was implemented and responses were analyzed, interpreted, and aligned with the ARCS concepts (Keller, 1983, 2010, 2017; Pappas, 2015). Participants affirmed that the module was helpful to guide and assist beginner and student filmmakers. “It's a very useful and helpful module for filmmakers beginning their projects,” said a participant. Table 4 below lists a sampling of participant answers and alignment of ARCS concepts.

Table 4
Examples of post-module qualitative answers

The instructional module is helpful for research and data collection for documentary film. Please explain why.
Attention
"Really encouraged connecting with others using visuals, personal stories, and being relatable." – Participant #7
Relevance
The module gave a lot of useful information for people who are beginning to start their filmmaking career. – Participants #8, #10, #13
Confidence
"Very thorough, easy-to-digest and informative." – Participant #4
Satisfaction
"This survey was well done! It was effective in providing organized and easy to understand information." – Participant #12

These significant quantitative and qualitative findings, as well as the field notes and observations that utilized triangulation, (Bryman, 2012; Glesne & Peshkin, 1992; Yin, 2016) confirmed that the instructional module did indeed have a positive educational impact on the participants. The data substantiates the educational quality, value and significance of the module and this study.

Discussion and Conclusion

The results from the quantitative and qualitative survey instruments, along with documented field notes and observations, validated the educational value of the instructional module about subject research and data collection for documentary film. The findings revealed that content and format contributed to the instructional module's success. Participants applied constructivism, the building of new knowledge upon their pre-existing knowledge base.

According to the data, the highlight of the module that positively resonated with the participants were the professional filmmakers' audio clips discussing key content points. The filmmakers' actual voices added honesty, validity and legitimacy to the material. This study successfully utilized consultation of experts in content design and development to improve the understanding, creativity, and production of beginner documentary filmmakers.

Due to the COVID-19 pandemic, the researcher was unable to implement the module in person as planned; however, it motivated the researcher to design and develop an online format which contributed to its increased success with participants. Utilizing the ADDIE model of instructional design that included multiple formative assessments by experienced reviewers and modified iterations (Iwasaki, in progress), what began as a basic, single page infographic turned into a content-rich, engaging, technology-enhanced, multimedia module on a website.

This development was extremely important because self-paced learning, multimedia and technology integration, and online learning were important academic needs for the target audience (Education Dive, 2017; McGraw-Hill, 2016). The data collected in this study validates the successful online implementation of the module and that multimedia online delivery can increase learning potential and outcomes (Alessi & Trollip, 2001; Chiu & Churchill, 2016; Clark & Mayer, 2011; Levonen et al., 2001; Mayer, 2001, 2009, 2017).

The module addressed the lack of educational resources about subject research and data collection, an important step in the documentary filmmaking process that has been challenging to document. *Five Tips from Filmmakers for Documentary Film Research and Data Collection* is a much-needed, relevant, easy-to-use, quality educational resource that will help guide and assist beginner and student filmmakers exploring documentary film production.

The wisdom and recommendations from veteran filmmakers provide student filmmakers a head start in the complex documentary filmmaking process, enabling them to share important stories more quickly and with improved outcomes. This professional guidance in the film process benefits beginner filmmakers, as well as society at large (Leavy, 2015; Loustaunau & Shaw, 2018), exposing audiences to increased learning opportunities through additional timely and important documentaries.

Looking forward, since documentary film is such an integral part of learning in a thriving, multicultural society, future research would add to the educational resources in the field and practice of documentary film research and data collection, film studies, and education, adding knowledge to this increasingly democratic, effective, and ever-expanding communication medium. Further research could also delve deeper into the need and importance of constructing knowledge and information from authentic, knowledgeable experts and voices in the design and development of instructional content to support, increase and improve overall learning.

This study successfully and innovatively applied and integrated learning sciences theory, methodology, technology, and instructional design with humanities and arts research. This study and its effective and dynamic approach to designing and developing educational content and resources has made an important and significant contribution to the fields of instructional design, the learning sciences, creative arts and the humanities. This study has significant potential to impact the broad possibilities of contemporary, interdisciplinary research design and collaboration.

References

Adorama Learning Center. (2018, February 26). How to make a documentary. Adorama earning Center how to. Accessed from https://www.adorama.com/alc/how-to-make-a-documentary

Alessi, S., & Trollip, Stanley R. (2001). *Multimedia for learning: Methods and development*. Boston, MA: Allyn and Bacon.

Astleitner, H. & Lintner, P. (2004). The effects of ARCS-strategies on self-regulated learning with instructional texts. *E-journal of instructional science and technology, 7*(1).

Aufderheide, P. (2007). *Documentary film: A very short introduction*. Oxford, UK: Oxford University Press.

Bell, D. (2011). Documentary film and the poetics of history. *Journal of Media Practice, 12*(1), 3-25. https://doi.org/10.1386/jmpr.12.1.3_1

Bill & Melinda Gates Foundation. (2020). *Postsecondary Success: Today's college students.* Retrieved from https://postsecondary.gatesfoundation.org/what-were-learning/todays-college-students/

Bryman, A. (2012). *Social research methods*. Oxford, UK: Oxford University Press.

Chiu, T. K. F. & Churchill, D. (2016). Design of learning objects for concept learning: Effects of multimedia learning principles and an instructional approach. *Interactive Learning Environments, 24*(6), 1355–1370. https://doi.org/10.1080/10494820.2015.1006237

Clark, R. E. & Mayer, Richard E. (2011). *E-learning and the science of instruction: Proven guidelines consumers and designers of multimedia learning*. San Francisco, CA: Pfeiffer.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Creswell, J. W. (2015). *A concise introduction to mixed methods research.* Thousand Oaks, CA: Sage Publications.

Creswell, J. W. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Thousand Oaks, CA: Sage Publications.

Desktop Documentaries. (2018). *Making documentaries: Step-by-step, how to make a documentary.* Accessed from https://www.desktop-documentaries.com/making-documentaries.html

Dick, W., Carey, L., & Carey, J. O. (2001). *The systematic design instruction.* New York: Longman.

Education Dive. (2017, November 2). *Report: 87% of students prefer tech-savvy schools.* Accessed from https://www.educationdive.com/news/report-87-of-students-prefer-tech-savvy-schools/509829/

Frank, J. (2013). The claims of documentary: Expanding the educational significance of documentary film. *Educational Philosophy and Theory, 45*(10), 1018–1027. https://doi.org/10.1111/j.1469-5812.2012.00849.x

Friend, J., & Caruthers, L. (2016). Documentary film: The next step in qualitative research to illuminate issues of social justice in urban education. *International Journal of Learning, Teaching and Educational Research, 15*(6), 33–47.

Gagne, R. M., Wager, W. W., Golas, K. C., & Keller, J. M. (2005). *Principles of instructional design* (5th ed.). Belmont CA: Wadsworth/Thomson Learning, Inc.

Galuppo, M., & Chuba, K. (2020, August 24). 2020's top 25 American film schools, ranked. *The Hollywood Reporter.* Accessed from https://www.hollywoodreporter.com/lists/2020s-top-25-american-film-schools-ranked

Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction.* White Plains, NY: Longman.

Goldman, R., Pea, R., Barron, B., & Derry, S. J. (Eds.). (2007). *Video Research in the Learning Sciences.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc. Publishers.

Gorard, S. (2003). Quantitative methods in social science. New York: Continuum.

Hanley, M. S., Noblit, G. W., Sheppard, G. L., & Barone, T. (Eds.). (2013). *Culturally relevant arts education for social justice: A way out of no way.* New York, NY: Routledge.

Harris, A. (2012). *Ethnocinema: Intercultural arts education.* New York: Springer.

Iwasaki, P. Y. (2021). Gathering the story: Documentary film research and data collection. *The 11th Asian Conference on Media, Communication & Film 2020 Official Conference Proceedings.* (pp. 21-37). The International Academic Forum (IAFOR). ISSN: 2186-5906.

Iwasaki, P. Y. (in progress). From Concept to Website: Designing an Instructional Module for Documentary Film Research. *In Progress.*

Keller, J. M. (1983). Motivational design of instruction. In Reigeluth, C. M. (Ed). *Instructional design theories and models.* (383-434). Hillsdale, NJ: Lawrence Erlbaum Associates.

Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach.* New York, NY: Springer.

Keller, J. M. (2017). The MVP model: Overview and application. *New Directions for Teaching and Learning, 2017*(152), 13–26. https://doi.org/10.1002/tl.20265

Keller, J. M., & Suzuki, K. (2004). Learner motivation and e-learning design: a multinationally validated process. *Journal of Educational Media, 29*(3), 229–239. https://doi.org/10.1080/1358165042000283084

Learn How to Become (2020). *Film school and careers in film: How to become a filmmaker.* Accessed from https://www.learnhowtobecome.org/arts-humanities-careers/film/

Leavy, P. (2015). *Method meets art: Arts-based research practice*. New York, NY: The Guilford Press.

Leavy, P., & Chilton, G. (2014). Art-based research practice: Merging social research and the creative arts. *Oxford handbook of qualitative research.* Oxford University Press. Accessed from https:ebookcentral.proquest.com

Levonen, J., Biardeau, A., & Rouet, J. F. (2001). *Multimedia learning: Cognitive and instructional issues*. New York, NY: Pergamon.

Loustaunau, E. E., & Shaw, L. E. (Eds.). (2018). *Telling migrant stories: Latin American diaspora in documentary film.* Gainesville, FL: University Press of Florida. Accessed from: https://ebookcentral.proquest.com/lib/uhm/detail.action?docID=5264844

Mayer, R. (2001). *Multimedia learning*. New York, NY: Cambridge University Press.

Mayer, R. (2009). *Multimedia learning*. Cambridge, UK: Cambridge University Press.

Mayer, R (2017). Psychological & Brain Sciences. *University of California, Santa Barbara.* Accessed from https://www.psych.ucsb.edu/people/faculty/mayer

Maykut, P., & Morehouse, R. (1994). *Beginning qualitative research: A philosophical and practical guide.* London, UK: The Falmer Press.

McGraw-Hill. (2016, October 17). *New survey data: Four out of five college students say digital learning technology helps improve their grades.* Accessed from https://www.mheducation.com/news-media/press-releases/2016-digital-study-trends-survey.html

McGriff, S. J. (2000). Instructional system design (ISD): Using the ADDIE model. Accessed from https://www.lib.purdue.edu/sites/default/files/directory/butler38/ADDIE.pdf

Molenda, M. (2015). In search of the elusive ADDIE model. *Performance Improvement, 54*(2), 34–36. https://doi.org/10.1002/pfi.21461

Molenda, M., Pershing, J. A., & Reigeluth, C. M. (1996). Designing instructional systems. *The ASTD training and development handbook.* New York, NY: McGraw-Hill.

Nash, K., Hight, C., & Summerhayes, C. (Eds.) (2014). *New documentary ecologies: emerging platforms, practices and discourse.* New York, NY: Palgrave Macmillan.

National Center for Education Statistics. (2020). *Fast facts: Enrollment.* Accessed from https://nces.ed.gov/fastfacts/display.asp?id=98

Pappas, C. (2015). *Instructional design models and theories: Keller's ARCS model of motivation.* Accessed from https://elearningindustry.com/arcs-model-of-motivation

Peterson, C. (2003). Bringing ADDIE to life: Instructional design at its best. *Journal of Educational Multimedia and Hypermedia, 12*(3), 227–241.

Serhat, K. (2017). ADDIE model: Instructional design. Retrieved from https://educationaltechnology.net/the-addie-model-instructional-design/

Sjöberg, J. (2008). Ethnofiction: Drama as a creative research practice in ethnographic film. *Journal of Media Practice, 9*(3), 229–242.

The Monsanto Papers. (2018). Accessed from https://topdocumentaryfilms.com/monsanto-papers/

West, B., Hoffman, K., & Costello, M. (Eds.). (2017). *Creative instructional design: Practical application for libraries.* Chicago, IL: Association of College and Research Libraries.

Winston, B., Vanstone, G., Chi, W. (2017). *The act of documenting: Documentary film in the21st century.* New York, NY: Bloomsbury Academic & Professional. Accessed from https://ebookcentral.proquest.com/lib/uhm/detail.action?docID=4742336.

Yin, R. K. (2016). *Qualitative research from start to finish*. (2nd ed.). Guilford Publications. ProQuest Ebook Central. Accessed from https://ebookcentral.proquest.com/lib/uhm/detail.action?docID=2008479.

Corresponding author: Patsy Y. Iwasaki
Email: piwasaki@hawaii.edu

The Effects of Task Selection Approaches to Emphasis Manipulation on Cognitive Load and Knowledge Transfer

Seohyun Choi
Hanyang University
South Korea

Dongsik Kim
Hanyang University
South Korea

Jaewon Jung
Korean Educational Development Institute
South Korea

Abstract

Emphasis manipulation is a way to help learners by directing their attention to particular subcomponents of a learning task. This study investigated the effects of different approaches to emphasis manipulation on knowledge transfer and cognitive load. This was done by examining the impact of three task selection strategies: system-controlled, learner-controlled, and shared-controlled. Forty-five students (n = 45) in the first or second year of high school were randomly assigned to three groups and each group used a different type of task selection to manipulate emphasis in a complex learning context. The system-controlled group carried out learning tasks that were identified as essential by the system. The learner-controlled group selected and carried out learning tasks they needed to learn. The shared-controlled group chose and carried out learning tasks that they wanted to learn from a list of suggested learning tasks. The tasks had four learning phases: pre-test, training, mental-effort rating, and transfer test. After participants completed the training, their cognitive load was measured. One week after the training, a transfer test was conducted to measure the constituent skill acquisition. The findings revealed that the system-controlled task selection strategy was the most effective in optimizing cognitive load and enhancing knowledge transfer. In addition, learners benefited from personalized guidance on learning task selection based on their expertise. Given that the shared-controlled task selection method was more effective than the learner-controlled task selection, this study's results indicate that learners should be provided with information about how to select learning tasks when they are allowed to do so.

Keywords: cognitive load, complex learning, emphasis manipulation, task selection

Modern society requires individuals to solve real-life problems. In education, there is more and more emphasis on using complex tasks to help learners integrate the knowledge, skills, and attitudes essential for effective task performance in real-life applications (Merrill, 2002; van Meeuwen et al., 2018; Frerejean et al., 2019). Although previous studies have shown complex tasks effectively enhance learners' cognitive skills and help them achieve higher-quality solutions (Beers et al., 2005; Slof et al., 2011; Shin et al., 2018), learners can have difficulty carrying out such tasks. One reason complex tasks are difficult is they can impose a high cognitive load on learners due to the fact that they are loosely structured problems composed of diverse subcomponents (van Merriënboer & Kirschner, 2018; Jung et al., 2019).

To overcome this difficulty, researchers have suggested using whole-task sequencing strategies such as simplifying conditions, knowledge progression and emphasis manipulation (van Merriënboer & Kirschner, 2018). The simple to complex sequencing of whole tasks facilitates the learning experience by encouraging learners to coordinate and integrate the constituent skills that make up a whole task (van Merriënboer & Kirschner, 2018). One of the advantages of whole-task sequencing is that it enables learners to carry out whole tasks without segmenting the elements into individual tasks. For example, emphasis manipulation is a whole-task approach that can help learners see the "big picture", as it allows them to practice constituent skills within the context of a whole task (Gopher, 2007; van Merriënboer & Kirschner, 2018; Frerejean et al., 2021). Emphasis manipulation directs learner focus to specific subcomponents of a task, rather than having them perform the whole task all at once (Gopher et al., 1989; Frerejean et al., 2021). As a result, emphasis manipulation can help learners coordinate constituent skills and process whole tasks while focusing on learning component skills.

Although many studies have indicated emphasis manipulation is an effective way to direct learner attention within a whole-task module, some researchers have argued that it is ineffective because the whole task is repeated many times throughout the process (Gopher et al., 2007). Repetitive whole-task performance can increase cognitive load by increasing the amount of redundant information, especially as learners' expertise increases (Chen, 2008). This can cause the expert reversal effect (Sweller, 1994; Kalyuga et al., 2003), particularly for more experienced learners who are already proficient in selecting, controlling and monitoring their learning processes (Ertmer & Newby, 1996). To reduce unnecessary cognitive load and promote cognitive skill acquisition, previous researchers have recommended a *personalised* approach to emphasis manipulation based on learners' prior knowledge (Corbalan et al., 2006; Kalyuga & Sweller, 2014).

Researchers have focused on task selection strategies as an effective personalised approach (Salden et al., 2004; Salden et al., 2006). The rationale for focusing on task selection strategies is they can enable learners to learn interrelated constituent skills more efficiently. Such strategies facilitate the cognitive learning process by ensuring that the task classes match the learners' level of expertise (Corbalan et al., 2006). During task selection, learners can be provided with a personalised sequence of tasks based on their specific proficiency and current learning status. This sequence of learning tasks can be chosen by a system, an instructor or a learner (Paas et al., 2011). Ideally, more personalised learning sequences, informed by input from learners, will allow them to perform learning tasks that are most suitable for their current level of expertise. For this reason, a personalised task selection approach to emphasis manipulation is expected to prevent the expert reversal effect and promote constituent skill acquisition.

Some studies have shown that the system-controlled approach to task selection is effective at complex learning (e.g., Camp et al., 2001; Kalyuga, 2006). In contrast, others have shown that high levels of system control may negatively affect learners' motivation (e.g., Corbalan et al., 2006), suggesting that giving learners control over task selection is a more effective approach to complex learning (e.g., Salden et al., 2006). Despite the advantages of task selection, research has produced inconclusive results about the effectiveness of various task selection strategies. In addition, few studies have explored the relationships between task selection, whole-task sequencing, and emphasis manipulation.

To address these issues, the current study explored effective instructional strategies for whole-task sequencing based on personalised approaches to complex learning. Specifically, the study examined three types of task selection strategies for emphasis manipulation. These strategies were categorised according to the *agent* who selected the learning tasks: system, learner, and shared (system and learner).

This study's research questions were formulated to identify effective emphasis-manipulating task selection strategies for learners. The first question was, "What are the effects of different task selection strategies on learners' cognitive load during emphasis manipulation sequencing?" The second question was, "What are the effects of task selection strategies on knowledge transfer during emphasis manipulation sequencing?" By addressing these questions, this study was conducted to identify the most effective task selection approaches to emphasis manipulation for complex learning.

Theoretical Background

Emphasis Manipulation Sequencing

Complex learning requires learners to integrate knowledge, skills and attitudes. Researchers have proposed many methods of promoting cognitive skill acquisition during complex learning (van Merriënboer & Kirschner 2018; Jung et al., 2019). Although the methods differ, they share a focus on learning experiences based on authentic, real-life tasks (Kirschner et al., 2006; Wang et al., 2017). The cognitive tasks used in complex learning have been shown to improve learners' cognitive skill acquisition and facilitate knowledge transfer (Kester & Kirschner, 2012; van Merriënboer & Kirschner, 2018).

Studies have recommended whole-task sequencing as an effective way to help learners coordinate and integrate constituent skills and promote knowledge transfer (van Merriënboer & Kirschner, 2018). In particular, emphasis manipulation, which is a whole-task sequencing strategy that emphasises or de-emphasises constituent skills within a whole-task project, can help learners to acquire complex cognitive skills (van Merriënboer & Kirschner, 2018). In emphasis manipulation sequencing, the relative emphasis on selected subcomponents is manipulated, but the whole task remains intact (Gopher et al., 1989; van Merriënboer & Kirschner, 2018). The essence of emphasis manipulation is that learning occurs continuously as the subcomponent priorities are varied. Emphasis manipulation allows learners to work on all of the constituent skills from the beginning of the learning process while focusing learners' attention on the most significant subcomponents. It is most effective if priorities and trade-offs are established and appropriate attention-allocation learning strategies identified (Gopher, 1993).

Despite emphasis manipulation's ability to help learners focus on important subcomponents of effectively learning the target material, it can create cognitive difficulties if too much redundant information is generated. This can happen when the whole learning process is repeated from beginning to the end while emphasizing different subcomponents. As a result, instructional strategies that prevent unnecessary cognitive load and promote learning performance should be applied when using emphasis manipulation.

Task Selection Strategy

Researchers studying whole-task sequencing have proposed task selection strategies for personalised learning that can effectively provide learning content that matches the characteristics and differences of individual learners (van Merriënboer & Kirschner, 2018). Some researchers have categorised task selection strategies according to the agent who determines (controls) which learning tasks will be emphasised: system control, learner control and shared control (Corbalan et al., 2006; Corbalan et al., 2008; Paas et al., 2011). In a system-controlled condition, learning tasks are selected by an instructional agent such as the computer system or teacher (Tennyson & Buttery, 1980; Corbalan et al., 2006). The system-controlled approach is used in some electronic learning environments, providing personalised learning by selecting tasks based on learners' current stage of learning (Camp et al, 2001). However, high levels of system control may negatively affect learners' interest and task involvement in learning (Corbalan et al., 2006). To prevent this, there is a need for learner control over some aspects of the learning process.

Salden and colleagues (2006) reported that giving learners control over task selection promotes learners' motivation and helps them engage in self-regulated learning. As the aim of complex learning is to promote complex cognitive skills and self-regulation skills, learner-controlled instruction is believed to lead to learning success. Giving learners task selection opportunities assumes that they are able to identify the learning tasks that are the most suitable to their needs (van Merriënboer, 2002; Bergamin & Hirt, 2018). In turn, this may increase learner motivation and strengthen their belief that they can accomplish their learning goals, ultimately leading to successful complex learning. In addition, learner control avoids unnecessary cognitive load by eliminating non-essential learning tasks and increasing germane cognitive load by increasing learning engagement (Vandewaeter & Clarebout, 2013; Lange, 2018). However, learner-controlled task selection is not always effective. Task selection may overburden novices and learners may omit essential parts of learning tasks when selecting one learning task from a large number of learning tasks (Merrill, 2002; Schwartz, 2004).

Another task selection strategy, shared-controlled, has been developed to compensate to address the limitations of learner-controlled selection (Corbalan et al., 2006). In the shared-controlled approach, learners select tasks while referring to personalised information about the most essential learning tasks based on their expertise. Some researchers have reported that shared control is more efficient than either system control or learner control alone, as it helps learners to make the right selections and to eliminate redundant learning tasks (van Meeuwen et al., 2018).

Cognitive Load and Expert Reversal Effect

Cognitive load theory (CLT) developed by Sweller (1988) has suggested that cognitive load is a critical factor in the process of complex tasks learning (Salden et al., 2006). Cognitive load can be divided into three elements: Intrinsic cognitive load, Extraneous cognitive load, and Germane cognitive load (Sweller, 1994). Intrinsic load is determined by the complexity of learning elements that are related to performing tasks (Gerjets et al., 2006). Extraneous load is imposed learning methods, information presentation methods, and learning strategies (Corbalan et al., 2008). Intrinsic load and extraneous load do not promote learning while germane load does (Moreno & Park, 2010). Germane load is imposed as a result of the cognitive efforts required to form schemas during learning. Germane cognitive load occurs by learning methods designed to promote automation, and researchers suggest securing the space for germane load by reducing extraneous load (Jung et al., 2016). Studies exploring CLT have suggested that intrinsic load and extraneous load must be reduced and germane load promoted for successful complex learning (Sweller, 1994; Moreno & Park, 2010).

Although emphasis manipulation sequencing is effective for teaching complex cognitive skills, repeated learning processes can lead to the expert reversal effect (Sweller, 1994; Kalyuga et al., 2003). Many studies of this effect have shown that educational approaches that are successful for novice learners are often less effective for more experienced learners (Jung et al., 2016; Kalyuga & Renkl, 2010; van Merriënboer & Sweller, 2010). Repetitive cognitive processing and learning materials constitute extraneous cognitive load, which may hinder learning. Thus, appropriate learning tasks and methods that consider learners' prior knowledge are needed to prevent the expert reversal effect as learner expertise increases over the learning process (Kalyuga & Renkl, 2010; Jung et al., 2016; Choi et al., 2019).

Individualised instruction that considers learners' expertise can be an effective strategy, as it may provide learners with appropriate learning tasks and reduce unnecessary cognitive load. Individual learners' cognitive load is determined by interactions between the learners' expertise and the difficulty of the learning tasks (Paas et al., 2003); thus, redundant educational materials and support can be removed as learner proficiency in a specific learning task improves. In addition, to optimise cognitive load and promote learning efficiency, educational guides must be provided at the time they best suit learners' needs. Using a personalised learning strategy to guide emphasis manipulation is likely to be an effective approach to reducing cognitive load and achieving learning objectives because it is challenging to identify the level of learning difficulty and particular subcomponent skills that should be emphasised for each individual learner. Thus, this study examined how applying three task selection approaches (system-controlled, learner-controlled, shared-controlled) to emphasis manipulation affected the success of complex learning.

Research Methodology

Participants

Forty-five students (n = 45) in the first or second year of high school in South Korea enrolled in a career development class were participants in this study. Permission to conduct this study was initially obtained from the school, as well as the students and their parents. Thus, there were no significant ethical conflicts. The students were 17 or 18 years old and 31 students (70%) were male and 14 (30%) were female. Participants were randomly assigned to one of three groups: a system-controlled group (SC; n=15), a learner-controlled group (LC; n=15), and a shared-controlled group (SHC; n=15).

Description of Task Selection Learning Environment

The task selection learning environment developed for this study consisted of three categories of constituent skills related to some advanced features in PowerPoint: animation effects, chart and graph effects, and multimedia effects. Each category included five constituent skills (see Table 1). Each of the three groups was presented with a different set of learning tasks on the main page of a website (see Figure 1). The system-controlled instruction condition provided a list of learning tasks in which the participants were weak; the learner-controlled instruction condition provided a list of all of the learning tasks; and the shared-controlled instruction condition provided a list of suggested learning tasks. When the participants clicked on a particular learning task displayed on the main page, a short video lecture played. Brief videos (3 – 5 minutes) were provided to help participants acquire the constituent skills. None of the learning phases for constituent skill acquisition were time limited. Each group performed different learning tasks based on the type of task selection control.

Table 1
List of Constituent Skills

Animation effects	Chart and graph effects	Multimedia effects
1. Screen transitions	1. Insert a table	1. Word Art
2. Add animation effects	2. Insert a chart	2. Smart Art
3. Customise animation effects	3. Chart layout	3. Insert an audio file
4. Modify animation effects	4. Edit data in Excel	4. Insert a video file
5. Animation Pane	5. Change chart type	5. Hyperlink

Figure 1
Structure of Task Selection Learning Environment

Pre-Test

Fifteen self-rated problems were used to measure participants' prior knowledge of the target skills. A computer-based task was developed to measure the learners' level of competence in PowerPoint. The participants were asked to make PowerPoint slides so that their weakest constituent skills could be identified. Responses to pre-test items were categorized by three evaluators as either poor and awarded a score of 0 or good and awarded a score of 1. Thus, the minimum pre-test score was 0 and the maximum was 15. There was no difference between the groups' pre-test scores [$F(2, 42) = .097, p > .05$].

Experimental Conditions

The system-controlled (SC) instruction condition provided learners with the learning tasks that they were the weakest in, which was determined by learners input. The SC group only carried out learning tasks that were identified as essential by the system based on their pre-test results. The SC group was not given the opportunity to carry out other learning tasks (see Figure 2).

The learner-controlled (LC) instruction condition provided learners with a list of all of the learning tasks. The learners in the LC group selected the learning tasks that they needed to learn and then carried out these learning tasks. The LC group was not provided with guidance on what was essential information or feedback to help with their task selection (see Figure 3).

The shared-controlled (SHC) instruction condition provided learners with a list of suggested learning tasks based on their pre-test results. However, unlike those in the SC group, the learners in the SHC group could choose any learning task that they wanted to learn (see Figure 3).

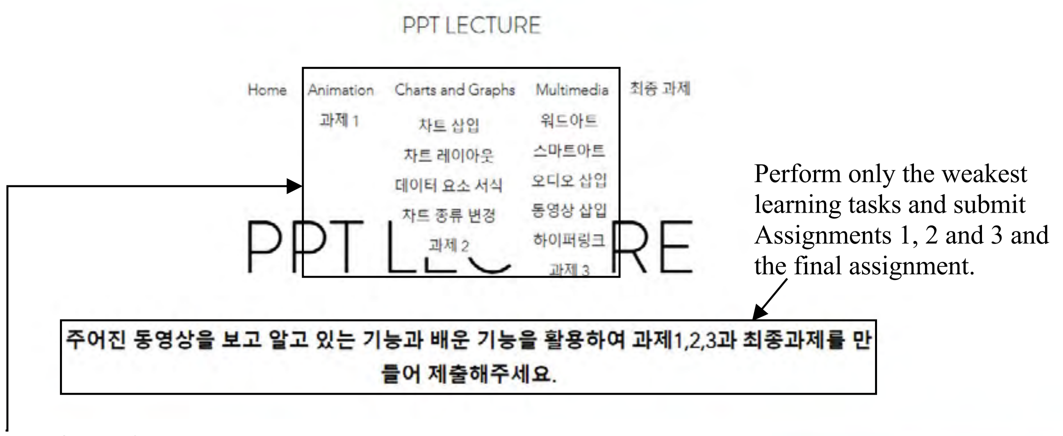
Table 2

Three Types of Selection Control and Learning Tasks

Group	Type of control	Learning tasks
SC	System	Weakest learning tasks
LC	Learner	All the learning tasks without information about which are the essential learning tasks
SHC	System and Learner	All the learning tasks with suggestions about essential learning tasks

Figure 2

Weakest Learning Tasks for the SC Group

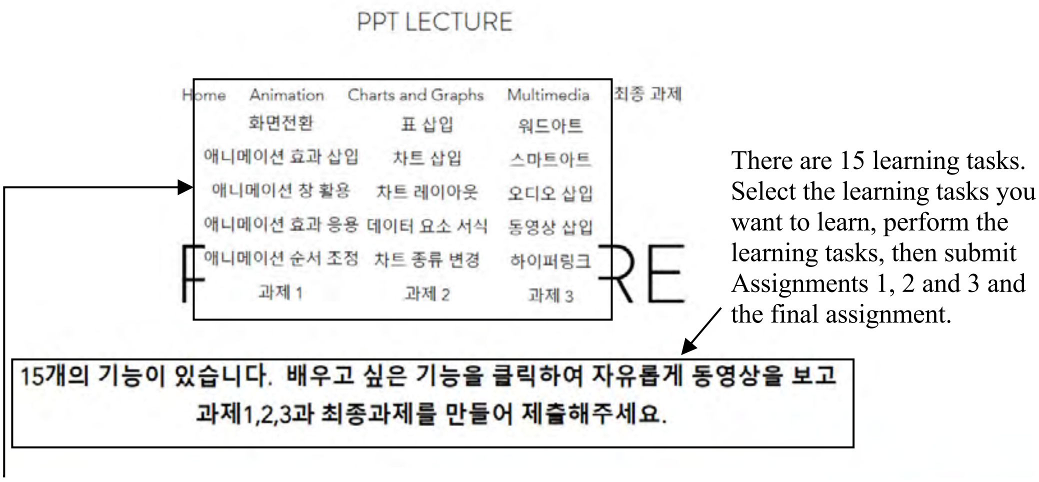


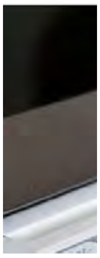
Weakest sets of learning tasks

Animation	Charts and graphs	Multimedia
Assignment 1	Insert a chart	Word Art
-	Chart layout	Smart Art
-	Edit data in Excel	Insert an audio file
-	Change chart type	Insert a video file
-	Assignment 2	Hyperlink
		Assignment 3

** Functions to perform specific learning tasks were open to learners in the system-controlled group.*

Figure 3
All of the Learning Tasks for the LC group and the SHC Group





All of the learning tasks

Animation	Charts and graphs	Multimedia
Screen transitions	Insert a table	Word Art
Add animation	Insert a chart	Smart Art
Customise animation	Chart layout	Insert an audio file
Modify animation	Edit data in Excel	Insert a video file
Animation Pane	Change chart type	Hyperlink
Assignment 1	Assignment 2	Assignment 3

** Functions to perform whole learning tasks were open to learners in the learner-controlled group.*

Cognitive Load Measures

A 10-point Likert-scale was used to measure cognitive load (Leppink et al., 2013). Previous studies have used this instrument to measure cognitive load (e.g., Becker, Klein, Gößling, & Kuhn, 2020; Thees, Kapp, Strzys, Beil, Lukowicz, & Kuhn, 2020). The scale ranged from "*not at all the case*" (0) to "*completely the case*" (10). The measurements consisted of three intrinsic load items, three extraneous load items, and four germane load items (see Table 3). All of the participants were asked to rate their cognitive load after the training phase. The reliability analysis revealed a Cronbach's alpha value of 0.807 for the intrinsic load items, 0.895 for the extraneous load items, and 0.911 for the germane load items.

Table 3
Sample Questions to Measure Cognitive Load

Type of load	Questionnaires
Intrinsic load	The topic/topics covered in the activity was/were very complex.
Extraneous load	The instructions and/or explanations during the activity were very unclear.
Germane load	The activity really enhanced my understanding of the topic(s) covered.

Transfer Measures

Learning outcomes were measured using a transfer test. One week after the training session, the participants were asked to make PowerPoint slides related to their major and recent trends in their disciplines. The purpose was to use the constituent skills acquired during the training phase. The transfer tests were recorded in the task selection learning environment to determine whether the participants had used the constituent skills properly. Three evaluators were chosen to rate the submitted PowerPoint slides using two scale values: *poor* (0) or *good* (1). The minimum transfer test score was 0 and the maximum was 5. The measurements consisted of three items (see Table 4). Interrater reliability analysis revealed a Cohen's Kappa value of 0.765.

Table 4
Sample Indicators to Measure the Level of Knowledge Transfer

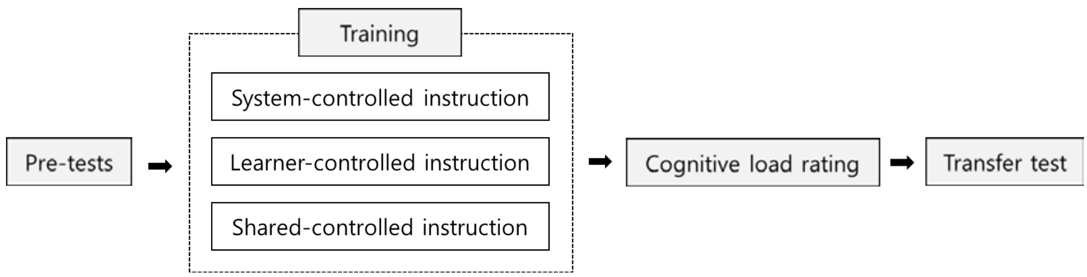
Item #	Questionnaires	Categories
1	Student made PowerPoint slides using screen transitions.	Animation effects
2	Student made PowerPoint slides using a table.	Charts and graph effects
3	Student made PowerPoint slides using a video file.	Multimedia effects

Procedure

The study was conducted over two weeks in an online learning environment. The participants were assigned to one of three groups and asked to perform tasks across four learning phases: pre-tests, training, mental-effort rating, and transfer tests (see Table 5). After the pre-tests, each group was provided with learning tasks associated with their assigned task selection condition. They then carried out the learning tasks for constituent skill acquisition. The participants were required to watch the short video lectures and then make three presentation slides that were identical to the presented samples. After the participants completed the training, their cognitive load was measured. One week after the training phase, a transfer test was administered to measure the acquisition of constituent skills.

In the transfer test, participants were asked to make PowerPoint presentation slides on a specific topic (My dream major at university). Participants were asked to use all of the functions about PowerPoint they had learned during the training phase. The PowerPoint presentation slides made in the transfer test were evaluated by one visual communication expert and two educational experts. All of the procedures and responses were recorded using a video recording program so the exact performance of the participants could be evaluated (see Figure 4).

Figure 4
Learning Processes in the Task Selection Learning Environment



Data Collection and Analysis

This study used a one-way factorial design. The independent variable was type of task selection control and the dependent variable was cognitive load, which consisted of intrinsic load, extraneous load, and germane load, as well as learning success. Analysis of variance was conducted to compare the cognitive load and learning between the three groups. The significance level was set at 0.05. The Statistical Procedures for Social Sciences version 24.0 was used to code and analyze the data.

Results

Effects of Task Selection on Learners' Cognitive Load

The SC group had the highest intrinsic cognitive load (M = 10.73; SD = 3.10), and the SHC group had the lowest intrinsic cognitive load (M = 9.40; SD = 2.26). The LC group had the greatest extraneous cognitive load (M = 10.33; SD = 2.47), and the SC group had the lowest extraneous cognitive load (M = 8.33; SD = 2.13). The SC group had the highest germane cognitive load (M = 14.13; SD = 5.09) (see Table 5).

Table 5

Descriptive Statistics for Cognitive Load and Knowledge Transfer (N=45)

Types of control	*n*	Intrinsic load		Extraneous load		Germane load		Transfer	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
System control	15	10.73	3.10	8.33	2.13	14.13	5.09	4.93	0.70
Shared control	15	9.40	2.26	9.07	3.22	9.60	4.01	4.80	0.49
Learner control	15	10.33	2.47	14.07	4.79	9.20	2.14	4.30	0.77

ANOVA for cognitive load revealed that the types of task selection control had a statistically significant effect on extraneous cognitive load [F (2, 42) = 11.59, $p < .001$, = .36] and germane cognitive load [F (2, 42) = 3.35, $p < .05$, η^2 = .26], but not intrinsic load [F (2, 42) = 1.01, $p > .05$] (see Table 6).

A *post hoc* Tukey test showed that the differences between the extraneous cognitive loads of the SC and LC groups (p = .000) and between the SHC and LC groups (p = .001) were statistically significant. A post hoc Tukey test also showed that the differences in the germane cognitive loads of the SC and LC groups (p = .005) and SHC and LC groups (p = .011) were statistically significant. These results mean that the SC and SHC groups experienced statistically significantly less extraneous cognitive load than the LC group. However, there was no statistically significant difference in the extraneous cognitive loads experienced by the SC and SHC groups. Furthermore, the SC and SHC groups experienced statistically significantly similar germane cognitive load that was statistically significantly more than the LC group.

Table 6
ANOVA of Cognitive Load

Type of load	Sources	*SS*	*df*	*MS*	*F*	*p*	η^2
Intrinsic load	Between groups	14.04	2	7.02	1.01	.373	
	Within groups	291.87	42	6.95			
	Total	305.91	44				
Extraneous load	Between groups	292.04	2	146.02	11.59	.000*	.356
	Within groups	529.20	42	12.60			
	Total	821.24	44				
Germane load	Between groups	225.24	2	112.62	7.24	.000*	.256
	Within groups	653.73	42	15.57			
	Total	878.98	44				

$*p < .05$

Table 7
Post hoc Tukey Test Results for Cognitive Load

Cognitive Load	**Sources**	**Type of control** **System control**	**Shared control**	**Learner control**
Intrinsic Cognitive Load	System control		1.33 (.358)	.40 (.909)
	Shared control	-1.33(.358)		-.93 (.60)
	Learner control	-.40 (.909)	.93 (.60)	
Extraneous Cognitive Load	System control		-.73 (.839)	-5.73 (.000**, -1.61)
	Shared control	.73(.839)		-5.0 (.001*, -1.41)
	Learner control	5.73 (.000**, 1.61)	5.0 (.001*, 1.41)	
Germane Cognitive Load	System control		.40 (.964)	5.13 (.005*, 1.21)
	Shared control	-.40 (.964)		4.73 (.011*, 1.11)
	Learner control	-5.13 (.005*, -1.21)	-4.73 (.011*, -1.11)	

*p<.05 **p<.001

Effects of Task Selection on Knowledge Transfer

The SC group had the highest transfer success (M = 4.93; SD = 0.70), followed in decreasing order by the SHC group (M = 4.80; SD = 0.49) and the LC group (M = 4.30; SD = 0.77) (see Table 5). The ANOVA test for transfer success showed that task selection type statistically significantly affected knowledge transfer success, F (2, 42) = 3.35, p < .05 (see Table 8). A *post hoc* Tukey test showed that the differences in knowledge transfer success between the SC and LC groups (p = .005) and the SHC and LC groups (p = .011) were statistically significant. Both the SC and SHC groups had statistically significantly greater knowledge transfer success than the LC group.

Table 8
ANOVA of Knowledge Transfer

Sources	*SS*	*df*	*MS*	F	*p*	η^2
Between groups	2.98	2	1.49	3.35	.045*	.138
Within groups	18.67	42	0.44			
Total	21.64	44				

*p < .05

Table 9
Post hoc Tukey Test Results for Knowledge Transfer

Sources	Types of control System control	 Shared control	 Learner control
System control		.20 (.950)	2.20 (.005*, 1.23)
Shared control	-.20 (.950)		2.0 (.011*, 1.12)
Learner control	-2.20 (.005*, -1.23)	-2.0(.011*, -1.12)	

Discussion

This study was conducted to investigate the effect of task selection approach type for emphasis manipulation on knowledge transfer success and cognitive load. Therefore, this study was conducted to answer the following research questions: "What are the effects of different task selection strategies on learners' cognitive load during emphasis manipulation sequencing?" and, "What are the effects of task selection strategies on knowledge transfer during emphasis manipulation sequencing?"

Effects of Task Selection on Learners' Cognitive Load

The findings indicated that there were no statistically significant differences in intrinsic cognitive load between the SC, LC and SHC groups. Intrinsic cognitive load is determined by the complexity of the content being learned, such as the number of components to be learned and how they interact (Jung, et al., 2016). This finding indicates that there was no statistically significant difference in the groups' prior knowledge or learning experiences and that task selection strategies did not have statistically significantly different effects on intrinsic cognitive load.

The results revealed that the system-controlled approach to emphasis manipulation was the most effective in reducing extraneous cognitive load and enhancing germane cognitive load. In contrast, the learner-controlled approach to emphasis manipulation produced the highest extraneous cognitive load and the lowest germane cognitive load. Because working memory capacity is limited (Sweller, 1988), reducing extraneous cognitive load is an important part of increasing germane cognitive load. Related research has shown increasing germane cognitive load can help learners secure more cognitive space for acquiring complex skills (Jung et al., 2019; Paas & Sweller, 2012; Renkl & Atkinson, 2003). The findings suggested that the system-controlled approach to emphasis manipulation effectively decreased unnecessary cognitive load by providing learners essential learning tasks based on their prior knowledge. In addition, it can be assumed that the learners in the SC group did not experience the expert reversal effect, as redundant information would have been eliminated by this approach (Sweller, 2010). However, the learner-controlled approach, which allowed learners to select learning tasks but gave them no guidance on necessary learning tasks, seemed to increase cognitive overload. These results imply that learners in the LC group experienced difficulty in accurate task selection and performed unnecessary learning tasks, potentially leading to extra cognitive work. Thus, even though learner control is typically perceived as an instructional approach that enhances learners' motivation (Corbalan et al., 2008), if learners attempt to learn all possible constituent components without any guidance about what is most essential, they may experience unnecessary cognitive load.

Although previous studies have shown that personalised task selection with shared control decreases extraneous cognitive load (Kostons et al., 2020) and increases germane cognitive load (Corbalan et al., 2006), the findings of this study were partially consistent with previous studies. The results demonstrated that the shared-controlled approach to emphasis manipulation is less effective than the system-controlled approach in optimising cognitive load, whereas it was more effective than the learner-controlled approach in diminishing extraneous cognitive load and increasing germane cognitive load. This study's results showed that *personalised* advice can help learners with task selection. This effect would likely be stronger for novice learners who are less able to accurately assess their own performance and choose learning tasks that fit their learning needs than experienced learners (Kostons, van Gog, & Paas, 2009).

In the meantime, the shared-controlled approach overcomes the difficulty of learner task selection compared to the learner-controlled approach by providing information about necessary learning tasks (Corbalan et al., 2006). The findings of this study provide evidence that a shared-control task selection strategy provides learning tasks that exclude repetitive learning tasks based on learner expertise, which may be more effective in controlling cognitive load compared to not providing such information. Therefore, this study suggests providing personalized learning information should be implemented to support accurate task selection by learners in order to support learning success.

Effects of Task Selection on Knowledge Transfer

In this study, the SC group had the best scores on the knowledge transfer test. The system-controlled approach emphasised essential learning tasks and excluded redundant learning. The findings imply that the learners in the SC group could effectively coordinate and integrate their prior knowledge and new information by focusing on learning the tasks in which they were weakest. Previous studies have shown that personalised learning based on learners' expertise leads to successful learning (Kalyuga & Sweller, 2005; Salden et al., 2004). This study provides empirical evidence that the personalised task sequencing of complex tasks leads to better knowledge transfer outcomes. A different study found that giving learners task selection opportunities increased their motivation and helped them engage in self-regulated learning (Salden et al., 2006). However, in this study, the SHC and LC groups that had control over task selection had *worse* learning outcomes than the SC group, which only had access to the learning tasks in which they were the weakest. This result indicates that learners did not have sufficient knowledge to select the learning tasks that were the most suitable to their needs.

The findings suggest that a task selection strategy that only provides essential learning tasks in consideration of learners' prior knowledge can effectively lead to success in learning by allowing learners to learn only the learning tasks in which they were weakest. Previous researchers have recommended using shared-controlled instruction to overcome the limitations of learner-controlled instruction (Corbalan et al., 2006). As shown by the SHC group's higher scores on the knowledge transfer test, providing learners with a list of suggested learning tasks based on their expertise appears to eliminate repetitive and redundant learning tasks and thereby promote knowledge transfer. This result is partially consistent with the empirical evidence that shared control is more efficient than either system control or learner control alone (van Meeuwen et al., 2018). Although the findings of this study indicated that the SC group had higher scores than the SHC group, given that the SHC group had better transfer scores than the LC group, it can be assumed that the task selection strategy of providing learners with a list of suggested learning tasks is more effective than not providing any learning information. Thus, the findings of this study provide empirical evidence that providing information about essential learning tasks can help learners select necessary learning tasks and improves knowledge transfer acquisition in contexts that use emphasis manipulation.

Conclusion and Limitations

This study investigated the effects of task selection strategies for emphasis manipulation on cognitive load and knowledge transfer. Three task selection strategies for emphasis manipulation were tested: system control, learner control and shared control. The system-controlled task selection approach to emphasis manipulation was found to be the most effective in optimising cognitive load and enhancing knowledge transfer. The results were consistent with previous studies of how personalised task sequencing focusing on learners' weaknesses can improve learning (e.g. Camp et al., 2001; van Merriënboer et al., 2002). In addition, the findings demonstrated that the system-controlled task selection strategy which provided necessary learning tasks based on learners' prior knowledge was effective in eliminating redundant information and preventing the expert reversal effect. The shared-controlled task selection strategy for emphasis manipulation was also found to be more effective than the learner-controlled task selection strategy. These findings revealed that providing learners with individualised information about essential learning tasks can help them select necessary learning tasks and achieve their learning goals (Corbalan et al., 2006).

In this study, it was hypothesized that allowing learners to select their own learning tasks would increase their motivation and improve learning outcomes as a result. However, the LC group had the most control over their task selection strategy, but learners did not have sufficient expertise to accurately select the most suitable tasks and, as a result, had the worst learning outcomes. Taken together, the results indicate that learners should be provided with information about how to select learning tasks when they are allowed to do so. Learners should be allowed to select learning tasks according to their prior knowledge level. In addition, learners should be presented essential information regarding how to select meaningful learning tasks.

This study has some limitations. First, the study focused on task selection strategies for emphasis manipulation in complex learning contexts. However, there is a need for more research on investigating task selection strategies suitable for novice and more experienced learners. Second, in this study, we investigated the effects of task selection on cognitive load and knowledge transfer. Future research should diversify the dependent variables to include motivation, engagement and self-efficacy. Third, this study only had 45 participants in total. More data is needed to examine the effects of task selection strategies in emphasis manipulation with various domains of expertise.

References

Beers, P. J., Boshuizen, H. P. E., Kirschner, P. A., & Gijselaers, W. H. (2005). Computer support for knowledge construction in collaborative learning environments. *Computers in Human Behavior, 21*(4), 623–643. https://doi.org/10.1016/j.chb.2004.10.036

Becker, S., Klein, P., Gößling, A., & Kuhn, J. (2020). Using mobile devices to enhance inquiry-based learning processes. *Learning and Instruction*, *69*, 101350. https://doi.org/10.1016/j.learninstruc.2020.101350

Bergamin, P., & Hirt, F. S. (2018). Who's in Charge?—Dealing with the Self-regulation Dilemma in Digital Learning Environments. In K. North, R. Maier, & O. Haas (Eds.), Knowledge Management in Digital Change (pp. 227–245). Springer. https://doi.org/10.1007/978-3-319-73546-7_14

Camp, G., Paas, F., Rikers, R., & van Merriënboer, J. J. (2001). Dynamic problem selection in air traffic control training: A comparison between performance, mental effort and mental efficiency. *Computers in Human Behavior, 17*(5-6), 575–595. https://doi.org/10.1016/S0747-5632(01)00028-0

Chen, C. M. (2008). Intelligent web-based learning system with personalized learning path guidance. *Computers & Education, 51*(2), 787–814. https://doi.org/10.1016/j.compedu.2007.08.004

Choi, S., Kim, N., Choi, S., & Kim, D. (2019). Emphasis manipulation effect in terms of the least-abled sets on cognitive load, transfer, and instructional efficiency. *Problems of Education in the 21st Century, 77*(2), 228–243. https://doi.org/10.33225/pec/19.77.228

Corbalan, G., Kester, L., & van Merriënboer, J. J. (2006). Towards a personalized task selection model with shared instructional control. *Instructional Science, 34*(5), 399–422. https://doi.org/10.1007/s11251-005-5774-2

Corbalan, G., Kester, L., & van Merriënboer, J. J. (2008). Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement. *Contemporary Educational Psychology, 33*(4), 733–756. https://doi.org/10.1016/j.cedpsych.2008.02.003

Ertmer, P. A., & Newby, T. J. (1996). The expert learner: Strategic, self-regulated, and reflective. *Instructional Science, 24*(1), 1–24. https://doi.org/10.1007/BF00156001

Frerejean, J., van Merriënboer, J. J., Kirschner, P. A., Roex, A., Aertgeerts, B., & Marcellis, M. (2019). Designing instruction for complex learning: 4C/ID in higher education. *European Journal of Education, 54*(4), 513–524. https://doi.org/10.1111/ejed.12363

Frerejean, J., van Geel, M., Keuning, T., Dolmans, D., van Merriënboer, J. J., & Visscher, A. J. (2021). Ten steps to 4C/ID: training differentiation skills in a professional development program for teachers. *Instructional Science*, 1–24. https://doi.org/10.1007/s11251-021-09540-x

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction, 16*(2), 104–121. https://doi.org/10.1016/j.learninstruc.2006.02.007

Gopher, D. (1993). The skill of attention control: Acquisition and execution of attention strategies. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience* (pp. 299–322). Cambridge, MA: MIT Press.

Gopher, D. (2007). Emphasis change as a training protocol for high-demand tasks. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice* (pp. 209–224). New York: Oxford University Press.

Gopher, D., Weil, M., & Siegel, D. (1989). Practice under changing priorities: An approach to the training of complex skills. *Acta Psychologica, 71*(1-3), 147–177. https://doi.org/10.1016/0001-6918(89)90007-3

Jung, J., Kim, D., & Na, C. (2016). Effects of WOE Presentation Types Used in Pre-training on the Cognitive Load and Comprehension of Content in Animation-Based Learning Environments. *Educational Technology & Society, 19*(4), 75–86. https://www.jstor.org/stable/10.2307/jeductechsoci.19.4.75

Jung, J., Shin, Y., & Zumbach, J. (2019). The effects of pre-training types on cognitive load, collaborative knowledge construction and deep learning in a computer-supported collaborative learning environment. *Interactive Learning Environmen*ts, 1–13. https://doi.org/10.1080/10494820.2019.1619592

Kalyuga, S. (2006). Assessment of learners' organised knowledge structures in adaptive learning environments. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 20*(3), 333–342. https://doi.org/10.1002/acp.1249

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4

Kalyuga, S., & Renkl, A. (2010). Expertise reversal effect and its instructional implications: Introduction to the special issue. *Instructional Science, 38*(3), 209–215. https://doi.org/10.1007/s11251-009-9102-0

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development, 53*(3), 83–93. https://doi.org/10.1007/BF02504800

Kalyuga, S., & Sweller, J. (2014). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (2nd revised ed.). New York: Cambridge University Press.

Kester, L., & Kirschner, P. A. (2012). Cognitive tasks and learning. In N. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 619–622). New York: Springer

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1

Kostons, D., Van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction, 22*(2), 121-132. https://doi.org/10.1016/j.learninstruc.2011.08.004

Lange, C. (2018). The relationship between system-provided learner control and maintained situational interest within e-learning courses. *Interactive Technology and Smart Education, 15*(3), 205–219. https://doi.org/10.1108/ITSE-12- 2017-0062

Leppink, J., Paas, F., van der Vleuten, C. P., van Gog, T., & van Merriënboer, J. J. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods, 45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development, 50*(3), 43–59. https://doi.org/10.1007/BF02505024

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist, 38*(1), 1–4. https://doi.org/10.1207/S15326985EP3801_1

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review, 24*(1), 27–45. https://doi.org/10.1007/s10648-011-9179-2.

Paas, F., van Merriënboer, J. J. G., & van Gog, T. (2011). Designing instruction for the contemporary learning landscape. In K. Harris, S. Graham, & T. Urdan (Eds.), *APA educational psychology handbook: Vol. 3. Application to learning and teaching* (pp. 335–357). Washington: American Psychological Association.

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38*(1), 15–22. https://doi.org/10.1207/S15326985EP3801_3

Salden, R. J., Paas, F., Broers, N. J., & van Merriënboer, J. J. G. (2004). Mental effort and performance as determinants for the dynamic selection of learning tasks in air traffic control training. *Instructional Science, 32*(1-2), 153–172. https://doi.org/10.1023/B:TRUC.0000021814.03996.ff

Salden, R. J. C. M., Paas, F., & van Merriënboer, J. J. G. (2006). A comparison of approaches to learning task selection in the training of complex cognitive skills. *Computers in Human Behavior, 22*(3), 321–333. https://doi.org/10.1016/j.chb.2004.06.003

Schwartz, B. (2004). *The paradox of choice: Why more is less.* New York, NY: HarperCollins.

Shin, Y., Kim, D., & Jung, J. (2018). The effects of representation tool (Visible-annotation) types to support knowledge building in computer-supported collaborative learning. *Journal of Educational Technology & Society, 21*(2), 98–110. https://www.jstor.org/stable/26388383

Slof, B., Erkens, G., & Kirschner, P. A. (2011). Constructing part-task congruent representations to support coordination of collaborative problem-solving tasks. In *Proceedings of the International Conference on CSCL 2011* (pp. 264–271).

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2010). Cognitive load theory: Recent theoretical advances. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 29–47). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511844744.004

Tennyson, R. D., & Buttery, T. (1980). Advisement and management strategies as design variables in computer-assisted instruction. *Educational Communication and Technology Journal, 28*(3), 169–176. https://doi.org/10.1007/BF02765363

Thees, M., Kapp, S., Strzys, M. P., Beil, F., Lukowicz, P., & Kuhn, J. (2020). Effects of augmented reality on learning and cognitive load in university physics laboratory courses. *Computers in Human Behavior, 108*, 106316. https://doi.org/10.1016/j. chb.2020.106316.

Vandewaetere, M., & Clarebout, G. (2013). Cognitive load of learner control: Extraneous or germane load? *Education Research International, 2013*(1), 1–11. https://doi.org/10.1155/2013/902809

van Meeuwen, L. W., Brand-Gruwel, S., Kirschner, P. A., de Bock, J. J., & van Merriënboer, J. J. (2018). Fostering self-regulation in training complex cognitive tasks. *Educational Technology Research and Development, 66*(1), 53–73. https://doi.org/10.1007/s11423-017-9539-9

van Merriënboer, J. J. G., Schuurman, J. G., De Croock, M. B. M., & Paas, F. (2002). Redirecting learners' attention during training: Effects on cognitive load, transfer test performance and training efficiency. *Learning and Instruction, 12*(1), 11–37. https://doi.org/10.1016/S0959-4752(01)00020-2

van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: design principles and strategies. *Medical Education, 44*(1), 85–93. https://doi.org/10.1111/j.1365-2923.2009.03498.x

van Merriënboer, J. J. G., & Kirschner, P. A. (2018). *Ten steps to complex learning: a systematic approach to four-component instructional design*. New York: Routledge.

Wang, M., Derry, S., & Ge, X. (2017). Guest editorial: Fostering deep learning in problem-solving contexts with the support of technology. *Journal of Educational Technology & Society, 20*(4), 162–165. http://www.jstor.org/stable/26229214

Corresponding author: Jaewon Jung
Email: jungj5@gmail.com

Machine Translation in the Language Classroom: Turkish EFL Learners' and Instructors' Perceptions and Use

Murat Ata
Middle East Technical University
Northern Cyprus Campus

Emre Debreli
European University of Lefke
Northern Cyprus

Abstract

Online machine translation (OMT) tools are not exclusively designed for language learners; however, these tools are popular among them. This quantitative study investigated the perceptions and attitudes of Turkish speaking EFL learners and instructors in a university English program regarding the use of OMT tools. Two online questionnaires were administered to 462 Turkish-speaking learners and 34 instructors. The results revealed that 94% of the learner participants reported using OMT tools for their language learning studies. The learners predominantly used these tools for single-word or phrase translations. Reading and writing assignments were the main areas where the learners most frequently referenced to OMT tools. The learner participants thought the accuracy of the tools was not high, and the ethicality of using them depended on how they were used. Three-quarters of the instructor participants reported using OMT tools, and their judgements concerning the accuracy of these tools were more positive than the learners'. The results also revealed a mismatch between learners' and instructors' perceptions and attitudes regarding OMT tools in foreign language learning. Accordingly, the instructors often overestimated how much learners use OMT tools, while learners underestimated the instructors' interest in them. These findings suggest policies should be developed within language learning institutions to guide students' use of OMT tools, as well as improve the mutual understanding between students and teachers in terms of their ethicality.

Keywords: attitudes, EFL, language learning, Google Translate, machine translation, perceptions

In today's age, students and instructors easily and freely benefit from a variety of online tools. Machine translators are among the most frequently referred online tools by learners. Free online machine translation (OMT) tools such as Google Translate offer written, voice and other types of translations between many languages. Practicality, ease of use, and free access to such websites and apps have made these tools very popular, especially among language learners. Scholars and educators, on the other hand, have varying reflections regarding the use of these tools in language learning. While some institutions do not allow students to use these tools, and some educators have reservations about their use for classwork or assignments on academic integrity grounds (Correa, 2011; Harris, 2010), others have looked for effective ways to make use of OMT tools (Benda, 2014; Chandra & Yuyun, 2018; Garcia & Pena, 2011).

Despite increasing interest, many institutions have not produced clear-cut policies regarding the use of OMT tools by students, nor have they specified the possible beneficial applications of these tools to aid language learning and teaching (Glendinning, 2014). Language teachers have also been struggling with the appropriate ways to approach the issue of their students' increasing use of machine translation. Besides, empirical data regarding the use of OMT tools in foreign language (FL) and second language (SL) education is very limited. Some studies have tried to describe the use of these tools by learners and teachers (Briggs, 2018; Clifford et al., 2013; Jolley & Maimone, 2015; Niño, 2009; O'Neill, 2019). Some other researchers have looked into possible ways to make use of these tools as Computer Assisted Language Learning (CALL) tools (Benda, 2014; Chandra & Yuyun, 2018; Garcia & Pena, 2011; Knowles, 2016; Lee & Briggs, 2021; Tuzcu, 2021). Others have focused on the issue from the perspective of academic misconduct (Correa, 2011; Groves & Mundt, 2015; Harris, 2010). Only a handful of these studies were exclusive to the context of English as a foreign (EFL) or second language (ESL). This suggests a significant lack of literature regarding the use of OMT tools in EFL, especially in the Turkish context.

The aim of this study was to address the lack of literature regarding the use of OMT tools by learners of English in the Turkish context. Conducted in a Turkish university EFL context, the purpose of the study was three-fold. First, the study aimed to describe the attitudes and perceptions of English language learners regarding OMT tools in terms of frequency of use, effectiveness, and ethicality for learning English. Second, the study aimed to explore the same issues from the perspective of instructors. Finally, the study aimed to document the thoughts of learners and instructors about each other's perceptions of OMT tools. To accomplish these aims, the present study was designed to address the following research questions:

1. How often and for what purposes do EFL learners use machine translation tools, and what are their perceptions and attitudes regarding the effectiveness and ethicality of these tools in learning English?
2. How often and for what purposes do English instructors use machine translation tools, and what are their perceptions and attitudes regarding the effectiveness and ethicality of these tools in EFL education?
3. What are students' and instructor's beliefs regarding each other's views on OMT use?

The results of this study would help reveal how OMT tools are used and perceived by students and instructors. In turn, such information would be instrumental for language teachers and administrators in the process of making policy about OMT tool use in EFL education. Together, the results may also prove useful to material designers, test developers, and educational software designers.

Literature Review

Translation has been one of the oldest means of language teaching. For centuries, language teachers have used the grammar-translation method, teaching their students how to analyse the grammatical structure of a target language and translate texts. The introduction of other teaching methods, such as the Direct Method and the Communicative Language Teaching, has emphasised communicative proficiency in foreign language learning. Thus, making use of students' first language has become a rather overlooked tool to present or explain new language, and it is commonly referred to as mother tongue facilitated teaching (Richards, 2015).

Printed dictionaries have always been a necessity for language learners, but they have started to lose their position as a primary resource of the target language. This is partially explained by continuous developments in the field of OMT and the introduction of smartphones with internet capabilities. Because of these developments, language learners have started to enjoy the practicality of online dictionaries and user-friendly software applications that support online OMT. Since the 1990s, the pedagogical implications of OMT were studied for FL education, especially in the area of computer-assisted language learning (CALL) (Benda, 2014; Chandra & Yuyun, 2018; Garcia & Pena, 2011) and their practical and ethical uses (Clifford et al., 2013; Correa, 2011; Groves & Mundt, 2015; Jolley & Maimone, 2015; Knowles, 2016; Lee & Briggs, 2021; McCarthy, 2004; Niño, 2009; Tuzcu, 2021).

Previous Research on OMT Use in Language Education

Niño (2009) attempted to group previous research on machine translation (MT) use in FL teaching and learning into four areas: "1. Use of MT as a bad model, 2. Use of MT as a good model, 3. Vocational use: Translation quality assessment, pre-editing and post-editing, and 4. MT as a "CALL" tool" (p. 242). Niño stated the strengths of MT as a CALL tool by highlighting features such as "wide availability, immediacy, multilingualism, good lexical translation, and good simple structure translation" (p. 245). On the other hand, Niño (2009) listed "literal translations, grammatical inaccuracies, discursive inaccuracies, spelling errors, missing cultural references, and unnatural writing" (p. 245 - 246) as drawbacks of MT use in FL education. It is important to bear in mind at this point that with the advancing MT algorithms and especially with the introduction of neural machine translation (NMT) in current MT tools (Briggs, 2018; Ducar & Schocket, 2018), some of the weaknesses identified by Niño have been substantially improved since 2009. NMT eliminated major errors involving the translation of proper nouns, literal translations of idioms, archaic vocabulary suggestions, and discursive inaccuracies, which has led more students to use OMT in their language studies (Ducar & Schocket, 2018).

Machine translation as a CALL tool. OMT tools offer many features from speech recognition to pronunciation; however, the most widely used feature has been translation of written text between languages. In order to aid students with writing more efficiently, some educators and researchers have made attempts to employ OMT as a CALL tool. To this end, Benda (2014) explored the possible benefits of using Google Translate (GT) in his English writing classes. Working in a Taiwanese context, Benda (2014) concluded that his undergraduate university students, and language learners in general, are more "…motivated by the need to obtain some kind of credentialing or certification" (p. 323) rather than learning for communicative purposes. Therefore, the learners used OMT to achieve higher scores in their writing without properly checking for the accuracy of the results. Chandra and Yuyun (2018) investigated English learners' habits of using GT during essay writing in the Indonesian context. Findings suggested that students used GT for three main purposes: vocabulary, grammar and spelling. Data on vocabulary use indicated that students used GT mostly for translating individual words. According to Chandra and Yuyun (2018), translating phrases and full sentences was less common. Because most previous research on OMT tool use in foreign language writing centred on students with a high level of proficiency, Garcia and Pena (2011) decided to investigate whether OMT can be considered a CALL tool for beginners and early intermediate level students. Their findings suggested that with the help of OMT, participants were able to produce a higher number of words in their paragraphs, proportional to their language proficiency level. Garcia and Pena (2011) stated that "… the lower their mastery, the greater the help provided by the MT draft…" (p. 478).

Student and teacher attitudes and perceptions regarding machine translation. Several scholars explored the attitudes and perceptions of FL students and instructors. Clifford et al. (2013) investigated the perceptions of language students and instructors regarding OMT use in FL in the United States. The results indicated that only a small proportion of the 905 participants *never* used OMT for language studies. In contrast, the majority of OMT users reported that they preferred Google Translate as their tool of choice, and they used OMT for individual words. For the next phase of the research, 43 instructors of Romance languages (Spanish, French, Italian, and Portuguese) were surveyed, and nearly half of them reported considering OMT use to be equal to cheating. More than half of the faculty regarded OMT as "*not useful"* or "*somewhat not useful"* for elementary and intermediate level students. In their conclusion, Clifford et al. (2013) recommended foreign language teaching policies evolve to be "proactive and pedagogically forward thinking to develop the best language learning experience possible" (p. 116).

Jolley and Maimone (2015) set out to explore the perceptions, attitudes, and beliefs of Spanish learners and instructors regarding OMT tools by evaluating their quality and ethicality. The results suggested that almost all of the students reported using OMT tools for their language studies but at varying frequencies. The study found that 65.08% of students reported using OMT tools for the purpose of translating individual words. The authors posited that many of the students considered Free Online Machine Translation (FOMT) as "…having a positive impact on their language learning and want instructors to cover strategies for effective use" (Jolley & Maimone, 2015, p. 192). A survey of instructors revealed the majority of them used OMT tools for teaching or personal needs. Like the students, the instructors judged individual words to be more suitable for accurate translation via OMT tools. In addition, over 60% of the instructors felt the accuracy of longer text translations to be ineffective. In terms of ethicality, instructors' perceptions deviated from that of their students. More than 85% of the instructors considered OMT use for texts longer than individual words to be "*unethical"* or "*equal to cheating"*. Similarly, Baskin and Mumcu (2018) found that higher-level students used GT less for sentence translations as they considered these tools ineffective for effective writing.

Academic misconduct and misuse issues of OMT in language learning. The literature on OMT use in FL education reveals that while students regard OMT use as relatively ethical, it is common for language educators to be disapproving of their use on ethical grounds (Clifford et al., 2013; Correa, 2011; Groves & Mundt, 2015; Jolley & Maimone, 2015; Knowles, 2016; Niño, 2009). Niño (2009) found that nearly half of participating instructors considered OMT use as cheating, and more than two-thirds disapproved of their use in language learning. Similarly, Jolley and Maimone (2015) found that 85% of instructor participants thought OMT use for assignments with texts longer than individual words was cheating, while a much lower percentage of students considered such use as cheating. Correa (2011) surveyed SL instructors to find out their opinion about academic dishonesty. One questionnaire item about the use of OMT (*Using an online translator for one or more sentences (if use of dictionaries is permitted*) was marked as academic dishonesty by more than half of the instructors. Similarly, Knowles (2016) surveyed 20 Romance language instructors for their perception of OMT use. Their findings showed that nearly half of the participants agreed or strongly agreed that OMT use equated to cheating, while only a small proportion disagreed or strongly disagreed with the notion.

In light of previous research, the goal of the present study is to better understand the perceptions and attitudes of Turkish language learners and instructors regarding the use of OMT tools in EFL education.

Method

The present study adopted a quantitative research design. Two separate online questionnaires (described below) were administered via Google Forms to collect data from student and instructor participants. The questionnaires were piloted first in four different universities with volunteering students and instructors. After the implementation of suggested modifications, the questionnaires were administered in the target setting.

Participants

In the study, two groups of participants, namely student participants and instructor participants, were involved.

Student participants. The student participants were recruited from an English Preparatory Programme (EPP) at an English-medium university and were not proficient in English. Out of 631 eligible Turkish speaking students enrolled in the EPP during the data collection period, 462 students participated in this study. In terms of language proficiency level, 138 (29.8%) participants were repeating students, and they were placed in Repeat (A2+) and B1 groups. The remaining 324 (70.2%) first-year students were placed in one of four levels within the EPP (Beginner, Elementary, Pre-Intermediate, and Intermediate). Male students accounted for 267 participants, and 174 participants were female. Twenty-one participants did not specify their sex. Student participants ranged in age from 17 to 34, with a mean of 20.

Instructor participants. Thirty-six instructors were employed at the EPP. Three spoke English as their native language, and 33 were native speakers of Turkish. Thirteen instructors held a Master's degree, and two held a PhD. Out of the 36 eligible instructors, 34 (94.4%) took the online questionnaire. Twelve of the participating instructors were male, and the remaining 22 were female. The instructor participants ranged in age from 26 to 63, with a mean of 36.

Online Questionnaires

The questionnaires used in the present study were modified versions of those employed by Jolley and Maimone (2015). Permission to use and modify the questionnaire items were kindly granted by the authors. Accordingly, items 7-22 and 30-33 were modified to suit EFL education in the Turkish context. This was necessary since the original questionnaires were designed for Spanish learners in the United States. One other reason for modifications was to gather more accurate frequency data. This was necessary because the frequency options used by previous studies (always – often – sometimes – rarely – never) were found to be misunderstood by participants. For example, the option "*sometimes*" for one participant may mean a unit of frequency that covers a day or, for another, a week. To address this issue and gather more accurate overall usage data, the present study employed the following options: *Never - a couple of times a year - a couple of times a month - a couple of times a week - once a day - multiple times a day*.

The online questionnaire for student participants included 27 items across four sections. The first section addressed the participants' habits of OMT use and their expectations from these tools. The five items in the second section addressed student participants' perception of output quality of OMT tools and ethicality or appropriateness of using OMT tools for their English language studies. The multi-part items in this section focused on the perception of the participants regarding the length of the translated segments and the language activities for which they use OMT tools. The third section included items regarding student perception of the instructor views about OMT. The last section was designed to gather data about the demographic and background information.

The online questionnaire for instructor participants included 30 items in four sections. The first section was about their habits of OMT use and teaching practices involving OMT tools. The second section addressed the participants' perception of the output quality of OMT tools. The third section included items addressing the participants' perception of their students' use of OMT tools for different language activities in terms of the length of translated segments, output quality, and appropriateness. The fourth section was designed to gather demographic information.

Procedures

In the present study, all of the participants were informed of the objectives of the study, and what they were expected to do, through online information forms. This information included the purpose of the study, a brief description, assurance of anonymity and confidentiality, and how their data would be used. It was also explained that they had the right to withdraw from the study at any stage or time. To maintain anonymity and confidentiality, names of the participants were not requested. In addition, the name of the university where the study was conducted was not provided in any part of this study to make the participants unidentifiable to persons from the same context and university.

Both the student and instructor questionnaires were administered at the end of the semester when students were taken to a computer lab to complete a course evaluation. The students stayed in the computer lab for an extra 15 to 20 minutes to respond to the online questionnaire. The students who needed more time to complete the questionnaire could do so by clicking on the link they received via email after pausing the questionnaire. The instructor participants received their online questionnaire through personalised emails. They were requested to complete the questionnaire within two weeks.

The Google Forms platform was used to create and administer the questionnaires. Each participant was presented with a consent form together with the questionnaire. Since the present study is descriptive in nature, the frequencies for the items were produced for descriptive analysis. Coded data provided by the survey platform were analysed in IBM SPSS Statistics 25 software.

Findings

The following sections present and discuss the key findings obtained from analysing the data collected via the two questionnaires.

OMT Tools and Features Used by EFL Students

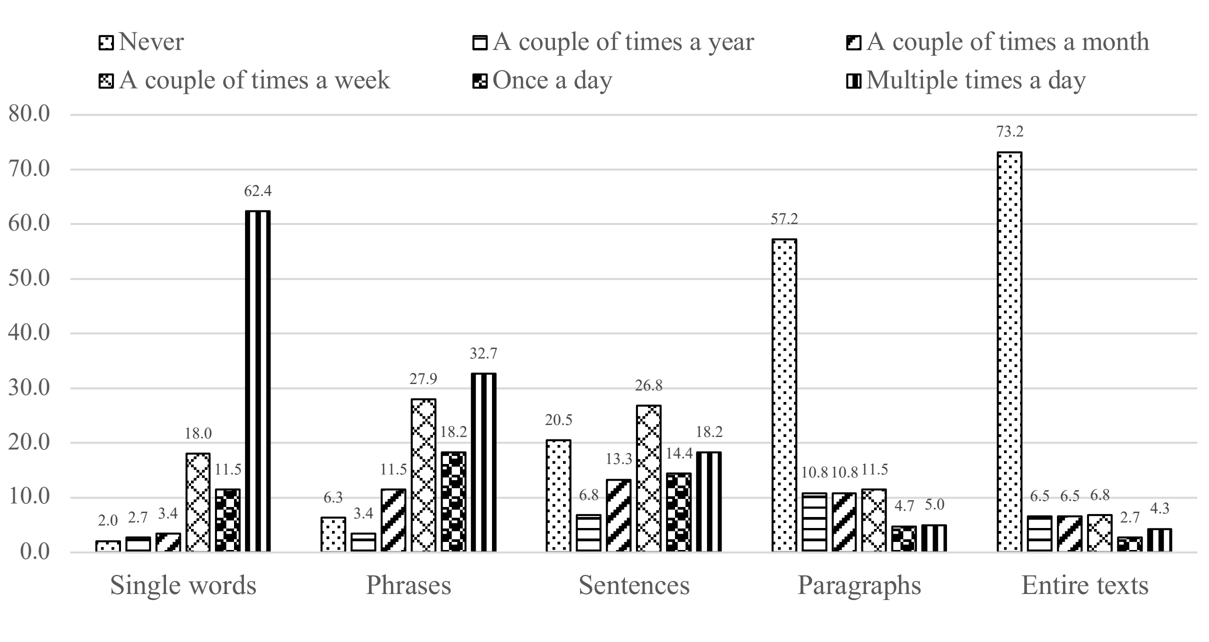
The results indicated that the majority of student participants (94.4%) reported using OMT tools for language studies. Google Translate was found to be the most commonly used online OMT tool, used by 82.2% of participants. This tool was followed by Yandex Translate, which was selected by 51 (11.5%) participants, and Microsoft Translator, which was chosen by 17 (3.8%) participants. When asked about the features of OMT tools student participants used, 425 (95.7%) participants selected *written translation,* making this the most commonly used feature. Nearly 200 participants (44.1%) indicated that they used the *pronunciation* feature in these tools. The *voice translations* feature was selected by 142 participants (31.9%). Forty-three participants (9.6%) selected *visual/image translation*; 20 participants (4.5%) chose *handwriting translation*; 15 participants (3.4%) chose *translation of conversations*, and nine participants (2.2%) chose *translation of uploaded documents*.

Frequency of EFL students' use of OMT tools. Frequency of use is one of the most critical aspects of the presence of OMT tools in language learning. Determining how frequently and with what purposes these tools are used by learners provides language educators with some perspective as to how students use OMT tools. These findings are presented in the following sections.

Frequency of ELF students' overall use of OMT tools. A vast majority of student participants (94.4%) reported using OMT tools. A breakdown of the overall frequency of use indicates that more than half of the students (52.4%) reported using them *multiple times a day*. In contrast, only 3.9% of students stated that they *never* used these tools. Additionally, when the responses were grouped as less frequent and more frequent, the results indicated a very high frequency of OMT tool use among participants (see Figure 4).

Responses indicated a decreasing trend in OMT tool use based on the length of translated segments. The majority of participants (91.9%) stated that they used these tools for *single words* a couple of times a week or more. For *phrases*, the percentage of frequent users was substantial at 78.8%. For *sentences*, more than half of the participants (59.4%) stated they used OMT tools a couple of times a week or more. When it came to translating *paragraphs*, the percentage of participants who *infrequently* or *never* used OMT tools accounted for 78.8%. For *entire texts*, only 13.7% of participants reported using OMT tools (see Figure 1).

Figure 1
Frequency of EFL Student Participants' OMT Tool use by Segment Length, in Percentages



Frequency of EFL students' use of OMT tools for reading and writing assignments. The responses by student participants revealed the frequency of their OMT tool use for different stages of reading (pre-reading, while reading, post-reading, reading assignments). A descriptive analysis of the frequencies showed that students were more inclined to use OMT tools for reading assignments rather than in-class reading activities (46.4% vs. 40.7%, respectively). *Never* was the most frequently chosen option for both situations. As for responses regarding the different stages of reading, students tended to use OMT tools most frequently during the post-reading stage (66.2%) and least frequently during the pre-reading stage (47.7%).

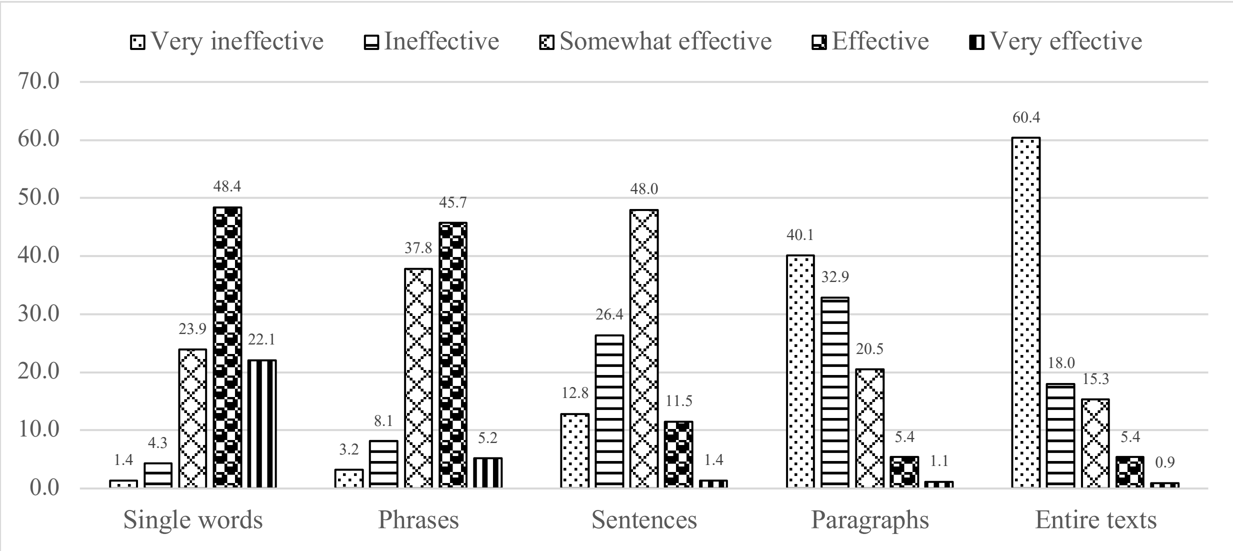
Comparable to the trend in reading activities, the use of OMT tools for writing assignments was higher than that of in-class writing activities (53.4% and 44.4%, respectively). More than half of the respondents (54.7%) reported using OMT tools for editing, and a similar percentage (54.1%) stated they used OMT tools for while-writing activities. Planning was the stage where students made use of the OMT tools the least (45.3%).

EFL Students' overall perception of the effectiveness of OMT tools. As for the students' perception of the effectiveness of OMT tools, 29.5% of the participants found OMT tools *effective* or *very effective* overall for English to Turkish translations, while 23.6% found them *ineffective* or *very ineffective*. A majority of the participants considered the results *somewhat effective,* with a percentage of 46.8. For Turkish to English overall translation effectiveness, the tendency shifted toward ineffective with 27.5% of participants choosing *ineffective* or *very ineffective* against 22.3% choosing *effective* or *very effective*.

These results suggest a decreasing trend in the perceived effectiveness of OMT tools as text segments get longer. For single words, 70.5% of participants found OMT tools *effective* or *very effective*. For phrases, 50.9% deemed OMT tools *effective* or *very effective*. For sentences, the percentage of participants who found OMT tool *effective* or *very effective* fell dramatically to

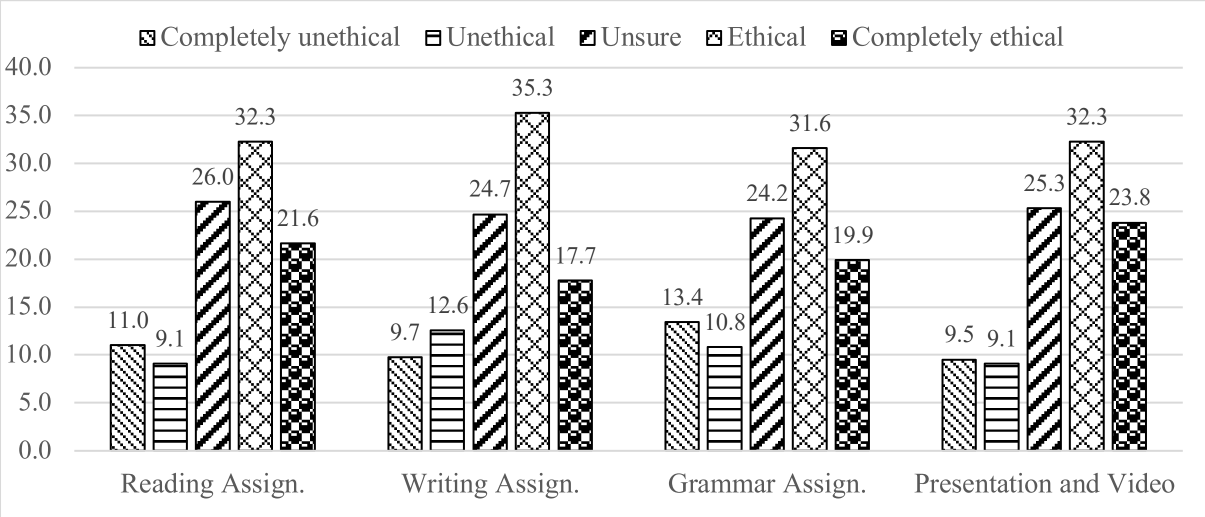
12.8%, while 39.2% of students thought they were *ineffective* or *very ineffective*. For paragraphs, 73% found them *ineffective* or *very ineffective*. For entire texts, the number of participants who found OMT tools *very effective* dropped below one percent. In total, 78.4% found them *ineffective* or *very ineffective* (see Figure 2).

Figure 2
EFL Student Participants' Perceptions of the Effectiveness of OMT Tools Based on Segment Length, in Percentages



EFL students' perceptions of the ethicality of using OMT tools for assignments. Along with the frequency of use and the perceived effectiveness of OMT tools, the concept of ethicality for students was investigated. Accordingly, for reading and writing assignments, 53.9% and 53% of participants considered OMT use *ethical* or *completely ethical*. Around 20% of the participants deemed the use of these tools as *unethical* or *completely unethical* for both types of assignments. For grammar assignments, 51.5% thought OMT tool use was *ethical* or *completely ethical*, while 24% considered it *unethical* or *completely unethical*. For the presentation and video assignments, 56.1% considered OMT tool use *ethical* or *completely ethical*, while 18.6% considered it *unethical* or *completely unethical* (see Figure 3).

Figure 3
Student Participants' Perception of the Ethicality of Using OMT Tools for Assignments in Percentages



The student questionnaire included a section about student participants' perception of the ethicality of using OMT tools to translate language units of different lengths. Translating single words using OMT tools received the highest positive ethicality rating (81.6%). Phrase translations showed a similar trend, with 71.4% of participants considering OMT use *ethical* or *completely ethical*. However, as the segments got longer, the trend started to reverse. For sentence translations, 46.8% considered OMT use *ethical* or *completely ethical*, while 26.8% thought it was *unethical* or *completely unethical*. A more dramatic change in the perception of ethicality manifested itself for even longer segments. For paragraph translations, 45.9% of participants considered OMT use *unethical* or *completely unethical*. For entire text translations, 51.1% of participants deemed it *unethical* or *completely unethical*.

OMT Tools and Features used by EFL Instructors

Written translation was the most frequently used feature of OMT tools, which was selected by all 25 participants. For ten participants (40%), pronunciation was the second most common feature. Five participants (20%) chose translation of uploaded documents; 4 (16%) participants chose voice translation, and 3 participants (12%) chose visual/image translation. One participant added "*translation of an entire web page*" which was not an option on the questionnaire.

Frequency of EFL instructors' overall use of OMT tools. About a quarter (26.5%) of the teacher participants indicated that they *never* used OMT tools for personal and teaching purposes. The majority (82.4%) of the remaining participants reported using OMT tools *a couple of times a month or less*. Only 5.9% reported using them *once a day* or *multiple times a day*. The results reveal that the instructors' use of OMT tools is not nearly as frequent as that of students (See Figure 4).

Figure 4
Frequency of OMT Tool use by EFL Instructors and Students

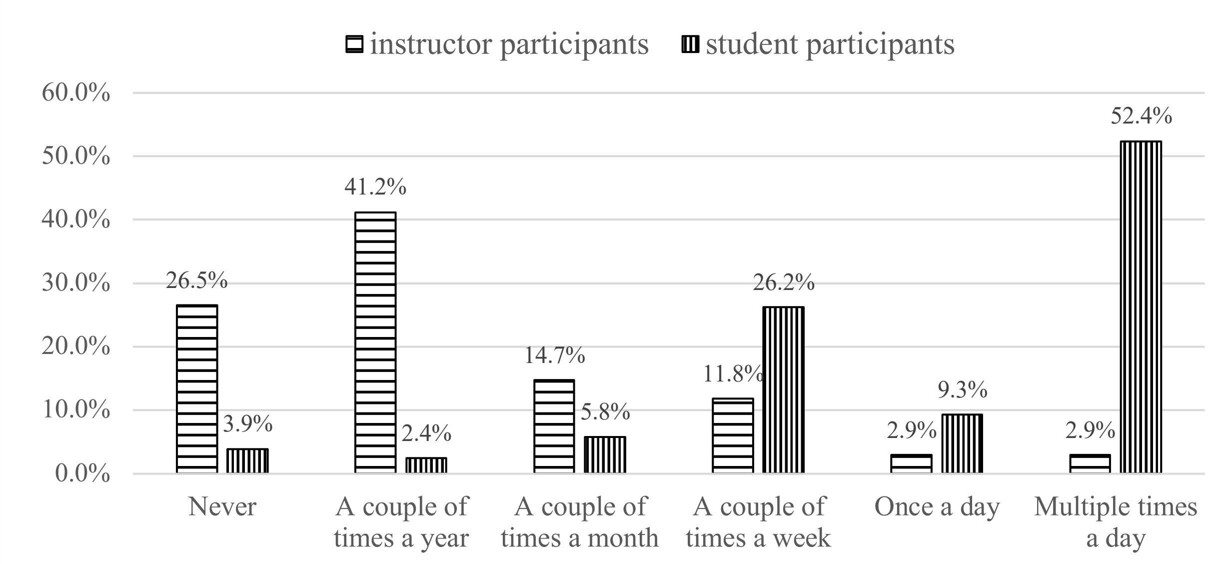


Figure 4: Frequency of OMT tool use by EFL instructors and students

EFL instructors' perceptions of the effectiveness of OMT tools. Participants refrained from reporting OMT tools as *very ineffective* or *very effective*. No participant considered translations produced by OMT tools for Turkish to English or English to Turkish as *very ineffective*, and only 2.9% of the participants deemed English to Turkish translations as *very effective*. For English to Turkish translations 56% of participants reported OMT tools as *effective* or *very effective* overall, while 16% found them *ineffective*. For Turkish to English translation, 36% of participants found OMT tools *effective* overall, while 28% found them *ineffective*. In total, around 30% of the participants rated OMT tool as *somewhat effective* for translating in both directions.

The results from the instructor questionnaire bear both similarities and differences to those of students. Like students, instructors found OMT translation results from English to Turkish more effective. However, while the majority of students chose the midpoint option (*somewhat effective*) for both translation directions, instructors had a more positive perspective regarding the quality of OMT translations by leaning more to the *effective* side.

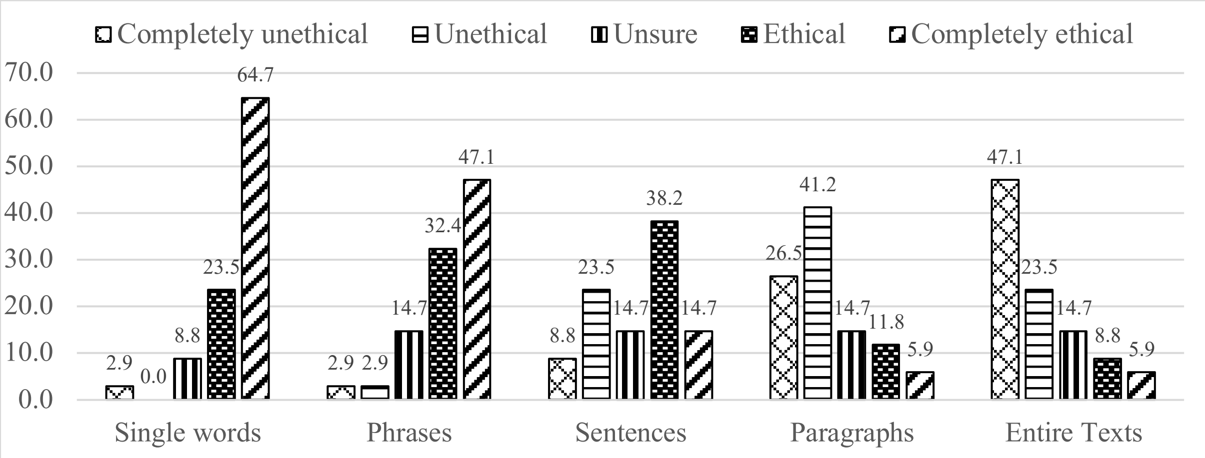
The data gathered from the participants regarding their perceived effectiveness of OMT tools revealed that 60% of the participants considered the tools *effective* or *very effective* for single word translations. The perceived effectiveness of the OMT tools fell to 44% for phrases, and 12% thought the results were *ineffective* for phrases. For sentence translations, the participants on the *effective* or *very effective* side accounted for 48%, while those who considered them as *ineffective* accounted for 24%. For paragraph and entire text translations, 36% and 24% of the participants considered OMT as an *effective* tool, respectively.

These findings were important as they displayed the only increasing trend when it came to the perceived effectiveness of segments of different lengths translated in OMT tools. That is, instructors rated the effectiveness of sentence translations as higher compared to phrase translations (48% vs. 44%). For longer segments, the perceived effectiveness fell again. This fluctuation was not observed in the student data.

EFL instructors' perception of overall ethicality of students' use of OMT tools for English assignments. With regard to the ethicality of students using of OMT tools for English assignments, 2.9% of the instructor participants reported considering it *ethical*, while no participants considered OMT use *unethical*. Seventy percent of the participants reported that it *depends on how MT tools are used for English assignments*, while 26.5% of the instructor participants were *not sure if MT use for assignments were ethical or not*. This finding suggests that more than a quarter of the instructor participants did not have enough experience or exposure to OMT tools to form a judgment concerning their ethicality. Considering the finding that 87.9% of students use these tools more than a couple of times a week, this lack of an ethicality judgement on the part of instructors was an important finding. While more than half of the students considered using OMT for their language assignments ethical, more than a quarter of the instructors had yet to form their judgment on the issue, probably due to their lack of interest in the subject. Similar to the student participants, instructors considered using OMT tools for reading assignments mostly ethical. Nearly 60% of the instructor participants considered OMT use *ethical* or *completely ethical*. In contrast, 70.6% of instructors thought using such tools for writing assignments as *unethical* or *completely unethical*. Only 11.8% of instructors considered the use of OMT tools for writing assignments *ethical* or *completely ethical*. For grammar assignments, the majority of instructors again leaned to the ethical side at 47.1%. However, the highest number of unsure participants regarding the ethicality of using OMT tools for assignments was found in this category with 23.5%. Regarding presentation assignments, 47.1% of the instructors considered OMT tool use as *unethical* or *completely unethical*. On the other hand, 38.2% considered this kind of use as *ethical* or *completely ethical*.

EFL instructors' perception of the ethicality of students' use of OMT tools based on segments of different lengths. Like the student participants, instructor participants thought that OMT use to translate single words was tolerable in terms of ethicality. Specifically, 88.2% considered OMT use *ethical* or *completely ethical*, while only 2.9% considered it *completely unethical*. For phrase translations, 47.1% considered OMT use *ethical* or *completely ethical*, while 5.9% considered it *unethical* or *completely unethical*. When it comes to translating sentences, more than half of the participants (52.9%) considered OMT use *ethical* or *completely ethical*, while 32.4% considered it *unethical* or *completely unethical*. Again, like the student participants, when it comes to longer segments, more instructors leaned towards the unethical side. For paragraph translations, 67.6% considered OMT use *unethical* or *completely unethical*. For entire text translations, 70.6% considered it *unethical* or *completely unethical*. For the entirety of this section, around 15% of the participants chose the option *unsure*. This may suggest that instructor participants need further clarification on the ethicality of OMT tool use for FL education (see Figure 5).

Figure 5
EFL Instructor Participants' Perceptions of the Ethicality of Students Using OMT Tools Based on Segment Length



EFL Student Instructor Beliefs Regarding Each Other's Views on OMT Use

In order to reveal how accurately the two participant groups evaluated each other's perception of OMT use, they were presented with a series of statements. Accordingly, when asked about how often student participants believed their instructors used OMT tools, 25.5% chose *never*, and around 40% chose *a couple of times a month* or less. This suggests 65% of students thought their instructors used OMT tools infrequently. When compared to data from the instructor questionnaire, however, it is striking that students guessed the percentage of instructors who *never* used OMT tools quite accurately (25.5% vs. 26.5%). However, in terms of overall frequency, students overestimated the number of instructors who used OMT tools *a couple of times a week* or *more* (34.9% vs. 17.6%). On the other hand, instructors overestimated the daily use of OMT tools by students. Self-reported total use of OMT tools *once a day* or *multiple times a day* by students added up to 61.7%; however, the instructors' guessed 85.3%.

There was a considerable discrepancy between students' self-reported use of OMT tools for reading and writing assignments and the perception of their instructors. While 46.4% of the students reported using OMT tools frequently (*sometimes*, *often*, and *always*) for reading assignments and 53.4% for writing assignments, the instructors reported thinking 85.3% and 88.3% of students used them frequently, respectively.

The student questionnaire included the following statement for participants to indicate their level of agreement: *Our instructors consider these tools as helpful to the language learning process.* Nearly one-third (32.2%) of the participants reported they *agreed* or *completely agreed* with the statement, while 23.4% *disagreed* or *completely disagreed*, and 44.4% were *unsure*. According to self-reported instructor responses, 26.5% of the instructors indicated that they thought these tools were helpful in the language learning process. These similar results indicated that the students had a good grasp of their instructors' attitude about this issue.

The student questionnaire also included the following statement for participants to agree or disagree with: *I feel proficient in using these tools for language learning*. More than half (58.7%) of the students *agreed* or *completely agreed*, which was very close to the number of students who reported using these tools *once a day or more* (61.7%). The instructors thought

44.1% of the students felt that they were proficient in using these tools. Exactly half of the instructors were *unsure* about the statement.

The students were also asked to indicate their level of agreement for the following statement regarding their instructors' attitudes towards students' use of OMT tools: *Our instructors encourage us to learn to use these tools in appropriate ways.* More than one-third (34.4%) of the students *agreed* or *completely agreed* with the statement. On the other hand, 28.4% *disagreed* or *completely disagreed*, and 37.2% were *unsure*. The instructor questionnaire included the following statement regarding the issue: *I think students should be encouraged to learn to use these tools in appropriate ways.* More than half (53%) of the instructors *agreed* or *completely agreed* with the statement, while 8.8% *disagreed* or *completely disagreed*, and 38.2% of the participants were *unsure*. The small discrepancy between the percentage of the student and instructor participant groups with regard to the encouragement item might indicate that instructors thought that students should be encouraged, but they simply did not do so in the classroom themselves, or some students might be ignoring the encouragement provided by their instructors.

Discussion

In this section, the study's findings are discussed with regard to the issues raised by the research questions, which included student and instructor perceptions and attitudes about OMT use and their perceptions of each other's perceptions and attitudes.

Discussion of EFL Students' Perceptions and Attitudes

The results of the study revealed a large number of participants use OMT tools for features that go beyond simple translation tasks and are not available in traditional dictionaries. This is similar to existing findings in the literature (Briggs, 2018; Clifford et al., 2013; Jolley & Maimone, 2015; O'Neill, 2019). For example, many students prefer Google Translate as the main provider of OMT, and the most commonly preferred feature is written translation. Vocabulary was reported to be the area where students used OMT tools heavily in the previous literature (Clifford et al., 2013; Jolley & Maimone, 2015; O'Neill, 2019). The results of the present study are in line with the existing literature. In addition, the results concerning the frequency of use for text of different lengths bore similarities with the findings of Chandra and Yuyun (2018) and Jolley and Maimone's (2015), which indicated that the majority of students reported using OMT tools for single words and only a few students used them for longer texts (e.g., paragraphs or entire texts). This may be due to the language level of the students. Novice language learners in preparatory programs are often not exposed to complex language structures; therefore, there might be less need to translate longer segments of text. The need for longer text translation may emerge in more proficient stages of language learning, which should be investigated through future research.

In terms of the perceived effectiveness of OMT tools, the findings bear similarities with that of Jolley and Maimone (2015). Inversely proportional results can be observed where longer segments resulted in lower perceived effectiveness. A similar trend to frequency and effectiveness of OMT tool use for different lengths of text was observed in terms of ethicality. The shorter the text segments, the higher the positive ethicality attributions. As the length of the segments increased to sentences and paragraphs, the perceived ethicality of OMT use fell dramatically. When considered together with the frequency data, these trends can be considered further evidence that students do not approach OMT tools uncritically in terms of translation effectiveness and ethicality. Students' preference for refraining from OMT use for longer texts may show that they do not use the tools for solely pragmatic reasons. Instead, the quality of the final text-based product and the appropriateness of their work is also important. Therefore, students avoid using OMT tools when they think the results are not of high quality or might violate ethical norms.

Discussion of EFL Instructors' Perceptions and Attitudes

The findings suggest that compared to students, instructor participants referenced OMT tools much less frequently for teaching purposes or personal use. As for the quality and effectiveness of OMT results for English to Turkish and Turkish to English translations, instructors held a more positive view than students. The significance of these findings becomes more pronounced when analysed together with the overall frequency data. Accordingly, students used OMT tools much more frequently than instructors, although their overall ratings of the results were lower (nearly 30% considered the results effective or very effective). The convenience of OMT tools may be one reason why students keep referring to them despite doubting their quality. The results may also suggest that students are not very good at judging the quality of the translations produced by OMT tools, whereas instructors are more comfortable using the OMT output to express meaning in English and Turkish. This might be an opportunity for instructors to explain to students how they decide whether or not an OMT is effective.

In terms of ethicality, unlike Clifford et al.'s (2013) findings, the instructor participants in this study refrained from judging OMT use as cheating. Instead, a significant proportion thought the ethicality judgment depended on how the tools were used by students. For example, the student and instructor participants agreed that it was ethical to use OMT tools for shorter segments of written text (e.g., single words, phrases). The two groups also found OMT use ethical for reading tasks. In contrast, the majority of instructors felt OMT use was unethical for writing tasks, which was different from the student perspective. Since writing is a productive skill, usually requiring creative production from students, instructors may be less tolerant about OMT use for such tasks. To limit student use of OMT during writing assignments, instructors may need to introduce new rules restricting internet use. In this way, instructors can make sure the writing tasks are students' original work, and students can practise producing written work unaided by OMT.

Discussion of Participants' Perceptions of Each Other's OMT Use

The findings revealed that the participants tended to overestimate each other's OMT tool use. In line with Jolley and Maimone's (2015) findings, instructors reported feeling that students rely on OMT tools for their language learning activities. In addition, the majority of the instructor participants did not think OMT tools were helpful for students. They also felt students were aware of this attitude. On the other hand, instructors overestimated students' attitudes about how helpful OMT tools are for language learning. It can be argued that the more teachers know their students' learning attitudes and habits, the better they can guide them in the learning process. In turn, this may suggest that if instructors know students' real usage and perception of OMT tools, they can address issues arising from OMT use more precisely and effectively. When instructors know their students do not refer to OMT tools, regardless of the learning activity and the length of text, and that they have some reservations regarding their quality and ethicality, instructors may be able to focus on making effective use of OMT for language teaching.

Conclusions and Implications

This study provided notable insights into the use of OMT tools in EFL teaching and learning in the Turkish context. One important result was that the vast majority of the EFL learners made use of OMT tools in order to aid their learning. This was done on a frequent basis. Given the novelty of OMT tools and their swift adoption by language learners, these findings suggest that researchers, language policy makers, educators and educational technology developers need to look deeper into this topic.

Another valuable finding was that a substantial proportion of learners used OMT tools for reading and writing assignments. The amount of use for in-class activities was slightly lower. It is apparent that banning or discouraging OMT tools is of little use as they are easy to access and widely available as long as learners have devices and connectivity. Therefore, teachers and administrators should either find new ways to limit the use of OMT tools by students for graded or ungraded schoolwork or, better yet, teachers and school administration may team up to discover new ways to integrate such tools into learners' academic work. In a study with Turkish EFL students, Tuzcu (2021) found that using OMT during writing activities increased creativity and improved students' "fluency, flexibility, originality and elaboration" (p. 48). Similarly, in a study with EFL learners in the Korean context, Lee and Briggs (2021) found that after OMT revisions, student errors in writing decreased significantly. Such innovative efforts to utilise OMT for instructional purposes may yield valuable benefits beyond efforts to limit or ban their use.

Another noteworthy finding has to do with the perceived effectiveness of the translations produced by OMT tools. Accordingly, students rated the results of these translations as being less accurate than their instructors. This curious finding may signify a conflict within student thinking. One the one hand, they do not think OMT tools are totally accurate. On the other, however, they continue to use them frequently. The answer to this puzzle might be found in how they described these tools. In data from the open-ended sections of the questionnaire, the student participants described OMT tools most frequently with adjectives such as "easy to use" and "quick". It may be imperative to note that the participating students in this study were almost entirely millennials. Millennials are also considered to be digital natives, who are, according to Prensky (2001) "used to receiving information really fast. They like to parallel process and multi-task" (p. 4). Based on Prensky's description, it might be argued that the age and generational characteristics of the student participants could be a factor leading them to use OMT tools in their language studies. In other words, they value their convenience and speed. With this insight, educators, dictionary developers, and material designers should consider that, for millennials, speed and ease of access are vital aspects of engaging in learning activities.

More than a quarter of the instructor participants reported never having used OMT tools, and for several sections of the instructor questionnaire, between a quarter to half chose the *unsure* option. Considering how frequently OMT tools are used by students for classwork and assignments, the instructors' lack of exposure might be considered alarming in terms of ensuring a healthy FL learning environment. It is recommended that instructors familiarise themselves with OMT tools and find potential ways their students may benefit from using them in their studies.

Finally, the results related to the third research question suggested there is a significant mismatch between learners' and instructors' thoughts regarding each other's use of OMT tools in FL learning. In order to overcome these discrepancies, language teachers, policymakers, and school administrators may define clear policies regarding the use of OMT tools in language classes. The instructors may be briefed through seminars and workshops regarding these policies, and they should inform their students regarding the established rules.

There are a number of limitations which must be noted in the present study. First, the setting for the study was one university. That is, both participant groups were from the same university. This limited the sample size and reduced the generalisability of the findings. Second, compared to the number of student participants (n=462), the number of instructor participants was much smaller (n=34). A larger sample size with more instructors from other universities may have yielded more generalisable results.

References

Baskin, S., & Mumcu, M. (2018). Dictionary culture of university students learning English as a foreign language in Turkey. *International Education Studies*, *11*(3), 101–114. https://doi.org/10.5539/ies.v11n3p101

Benda, J. (2014). Google Translate in the EFL classroom taboo or teaching tool?. *Writing & Pedagogy*, *5*(2), 317–332.

Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *The JALT CALL Journal*, *14*(1), 3–24.

Chandra, S. O., & Yuyun, I. (2018). The use of Google Translate in EFL essay writing. *LLT Journal: A Journal on Language and Language Teaching*, 21(2), 228–238.

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *The Acquisition of Second Languages and Innovative Pedagogies*, *10*, 108–121.

Correa, M. (2011). Academic dishonesty in the second language classroom: Instructors' perspectives. *Modern Journal of Language Teaching Methods*, *1*(1), 65–79.

Ducar, C., & Schocket, D. H. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google translate. *Foreign Language Annals*, *51*(4), 779–795. https://doi.org/10.1111/flan.12366

Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: writing for beginners. *Computer Assisted Language Learning*, *24*(5), 471–487. https://doi.org/10.1080/09588221.2011.582687

Glendinning, I. (2014). Responses to student plagiarism in higher education across Europe. *International Journal for Educational Integrity*, *10*(1), 4–20. https://doi.org/10.21913/IJEI.v10i1.930

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, *37*, 112–121. https://doi.org/10.1016/j.esp.2014.09.001

Harris, H. (2010). Machine translations revisited: issues and treatment protocol. *The Language Teacher, 34*(3), 25–29.

Jolley, J. R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. In A. J. Moeller (Ed.), *Learn languages, explore cultures, transform lives* (pp. 181-200). Minneapolis: 2015 Central States Conference on the Teaching of Foreign Languages.

Knowles, C. L. (2016). *Investigating instructor perceptions of online machine translation and second language acquisition within most commonly taught language courses* (Doctoral dissertation). The University of Memphis.

Lee, S., & Briggs, N. (2021). Effects of using machine translation to mediate the revision process of Korean university students' academic writing. *ReCALL*, *33*(1), 18–33.

McCarthy, B. (2004). Does online machine translation spell the end of take-home translation assignments. *CALL-EJ Online*, *6*(1), 6–1.

Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, *21*(2), 241–258. https://doi.org/10.1017/S0958344009000172

O'Neill, E. M. (2019). Online Translator, Dictionary, and Search Engine Use Among L2 Students. *CALL-EJ*, *20*(1), 154–177.

Prensky, M. (2001). Digital natives, digital immigrants Part 1. *On the Horizon*, *9*(5), 1–6.

Richards, J. C. (2015). *Key issues in language teaching*. Cambridge University Press.

Tuzcu, A. (2021). The Impact of Google Translate on Creativity in Writing Activities. *Language Education &Technology (LET Journal)*, *1*(1), 40–52.

Corresponding author: Emre Debreli
Email: edebreli@eul.edu.tr

Next Generation Mobile Learning: Leveraging Message Design Considerations for Learning and Accessibility

Eunice Ofori
Virginia Tech
USA

Barbara B. Lockee
Virginia Tech
USA

Abstract

Access to mobile learning (mLearning) opportunities has become widespread and continues to proliferate as a means of educational continuity during the COVID-19 pandemic. Due to such proliferation, guidance is needed to inform the design of mobile learning content from both learning and accessibility perspectives. Though evidence-based recommendations for mobile learning message design do not currently exist, prior research in multimedia learning and instructional design-related areas may be used to support the planning and production of such educational programming. Design efforts for mLearning would also benefit from the incorporation of strategies to enhance the accessibility of mLearning for learners with differing needs. Taking evidence-based practices from instructional design and universal design for learning could inform the future development of mLearning toward more effective learning experiences for all learners. Employing a design and development methodology, this study focused on the creation of evidence-based guidelines for mLearning content design, informed by prior research on instructional message design combined with recognized universal design principles for media-based learning. The study resulted in a set of considerations to guide the message design of accessible and effective mLearning experiences. The resulting guidelines underwent validation by expert reviewers representing the areas of instructional design, message design, universal design, and mLearning. Their feedback informed the final version of the guidelines produced as the outcome of this study, research-based considerations which can be practically applied by those responsible for the creation of mLearning instruction.

Keywords: accessibility, instructional message design, mobile learning

Mobile devices, for the purpose of education, have become a popular way for students to learn and access information (NMC Horizon Report, 2017). Uther (2019) stated that "Mobile learning has become one of the more influential aspects in the field of educational technology given the ubiquity of modern mobile devices and proliferation of educational applications or 'apps' for mobile devices" (p. 1). Mobile screens have the ability to support classroom learning in various capacities and settings (Pegrum, 2019). Researchers see the potential of mobile learning because of its portability, cost effectiveness, and communication features (NMC Horizon Report, 2017). McQuiggan, McQuiggan, Sabourin, and Kosturko (2015) define mobile learning as, "the experience and the opportunity afforded by the evolution of the educational technologies, it is anywhere anytime learning enabled by instant on demand access to a personalized world filled with tools and resources" (p.8).

With the widespread use of mobile devices in education for teaching and learning purposes, there is a need to ensure that information presented on the mobile device screen is designed to help learners comprehend the information displayed. Effective instructional message design is crucial for learning to take place. An instructional message is, "a pattern of signs (words, pictures, gestures) produced for the purpose of modifying the psychomotor, cognitive, or affective behavior of one or more persons" (Fleming & Levie 1993, p. x). An instructional message also, "provides a setting for new information that is conveyed by a message" (Gibbons, 2014, p. 215).

With the popularity of mobile devices for learning purposes, instructional designers must adhere to sound instructional design principles to generate meaningful and effective instructional material. A design that adheres to evidence-based design principles will place images, spoken language, and printed words in proper combinations to maximize instructional effectiveness (Wang & Shen, 2012). A careful and comprehensive study of existing literature confirms earlier studies which indicate that very little empirical exploration has been conducted related to design strategies for mLearning (Haag & Berking, 2015; Saleh & Bhat, 2015; Shen, Wang, Gao, Novak, & Tang, 2009; Vincent-Layton, 2015; Wang & Shen, 2012; Wishart, 2009). As a result, there is a need for more systematic research focusing on instructional message design for mLearning in order to design engaging and accessible instructional materials for mobile devices.

Literature Review

Mobile learning, in contrast to other forms of distance education, is unique. Tereshchenko, Zagorskaya, Polyanskaya and Bobritskaya (2020) stated that learning using mobile devices offers new opportunities for students and teachers. Along with the flexibility and affordability of mobile technologies, challenges related to mLearning design are also present. Wang and Shen (2012) note the constraints of designing instruction to work effectively across a wide variety of sizes and formats. As such, Vavoula and Karagiannidis (2005) proposed that care should be taken to design mLearning content effectively from a learning standpoint.

Research-based strategies for the design and development have been noted in the literature. Vincent-Layton (2015) states:

> There are few resources published that offer comprehensive mobile lessons and concrete methods to effectively implement mobile learning into the classroom. Educators need specific guidelines and model examples of mobile lessons to fully

> understand how to create the lesson, what to consider when developing, and how to successfully integrate it into the classroom (pp. 149–150).

Other researchers have noted the need for evidence-based guidance related to mLearning design. For example, Shen, Wang, Gao, Novak and Tang (2009) also indicated that, "the best practices for using mobile devices in teaching and learning are still unknown. Systematic studies are needed to investigate student and instructor experiences with mobile learning" (p. 539). Additionally, Haag and Berking (2015) mentioned that, "Instructors, educators, and instructional designers are quickly adopting mobile technology in their learning environments, but strategic design considerations and proven pedagogical practices have not been systematically documented" (p. 42). With prolific adoption of mLearning around the world, especially in response to the COVID-19 pandemic, practical research to inform mLearning design is essential.

Mobile Devices and User Interface Design

User interface design for small screens is one of the prominent challenges in the development of mobile devices, mostly because of the small size and functionality of such mobile devices (Gong & Tarasewich, 2004). "Mobile platforms have called for attention from HCI practitioners, and, ever since 2007, touchscreens have completely changed mobile user interface and interaction design" (Punchoojit & Hongwarittorrn, 2017, p.1). The design features of the user interface on a mobile device are an important aspect of the device to ensure efficient and effective usability. Punchoojit and Hongwarittorrn (2017) further stated that "although mobile platforms are becoming an indispensable part of daily lives, true standards for mobile UI design patterns do not exist" (p. 1).

Mobile devices also lack some affordances of standard computing technologies. For example, computers present a landscape view which, while possible on a mobile device, is not the typical approach to its use. To ensure that learners who are using these devices get the most out of it, care must be taken to consider the design constraints, as well as the many possibilities. Uden (2006) states that "mobile applications must be carefully designed to justify the limitations of their size, lower processing power, and low bandwidth" (Uden 2006, p .82).

Multimedia Theory for mLearning Design

The psychological principles from multimedia learning theory, long applied in the design of computer-based instruction, hold the potential to inform mLearning message design. Multimedia learning theory is grounded in the fact that instructional messages should be take into consideration how the human mind works (Mayer, 2001, 2005, 2009). Mayer (2009) describes multimedia learning as the presentation of instructional materials in words and pictures to ensure that learning takes place. Mayer (2009) explains that by "words," he means that the material is presented in a verbal form which is printed or spoken text. By "pictures," he means pictorial forms which include using still graphics such as illustrations, graphs, photos, or maps, or dynamic graphics such as animations or video. He further suggested that the goal of multimedia learning is:

> To minimize extraneous cognitive processing during learning (i.e., cognitive processing that does not serve the instructional goal), to manage essential processing during learning (i.e. cognitive processing needed to mentally represent the essential material), and to foster generative processing during learning (i.e., cognitive processing aimed at making sense of the material). (Mayer, 2013, p. 395)

More recently, Clark and Mayer (2016) derived eLearning design principles from the earlier multimedia-related research of Mayer (2001, 2009). According to these authors, there are eight principles that instructional designers can follow to design instructionally sound eLearning materials. These principles include: multimedia, modality, contiguity, redundancy, coherence, personalization, segmentation, and pre-training principle (Clark & Mayer, 2016). Additionally, these scholars provided theory-based, concrete guidance for operationalizing these principles in e-learning environments, serving as a possible foundation for mLearning instructional development.

Universal Design of Instruction for mLearning

When considering plans for mobile learning design, it is important to also look at design strategies that are inclusive, meaning that everyone partaking in instructional content will benefit from it. According to Burgstahler (2012), universal design for instruction (UDI) centers on this inclusion to address a wide array of individual learner differences. UDI has been an important consideration for electronically delivered educational experiences since the early days of e-learning. In 1997, The Center for Universal Design at North Carolina State University developed a set of universal design principles that continue to inform accessible design strategies across disciplines and environments. These principles include: equitable use, flexibility in use, simple and intuitive use, perceptible information, tolerance and error, low physical effort, and size and space considerations. Since then, interest in universal design for learning has continued to grow, reflecting many of these early principles. The current global pandemic has accelerated the define strategies for making mobile learning more accessible to all (Taildong & Toquero, 2021).

Need for the Study

Research conducted so far shows that mLearning is a growing trend as a learning delivery modality and the momentum it has taken will keep accelerating. Hanbridge, Tin, and Sanderson (2018) state that, “it is anticipated that mLearning will grow quickly in the next few years” (p. 119). As such, McQuiggan et al. claim that “the future adoption of mobile learning and the success of such efforts requires continuous awareness and integration of new technologies and functions” (2015, p. 333). The proliferation of mLearning underscores the imperative to develop evidence-based guidance for the design and development of effective mobile learning experiences, especially in the wake of the COVID-19 pandemic (Pebriantika, Wibawa, & Paristiowati, 2021).

Purpose of the Study

The purpose of this study was to propose a set of instructional message design guidelines for mobile learning content development, based on prior research in the areas of instructional message design and universal design for learning. A design and development research methodology was used to coalesce findings from the literature into a set of evidence-based mLearning design considerations. The resulting considerations are proposed for use by content developers and instructional design professionals in the overall instructional design process for mobile learning experiences.

Research Questions

The following research questions were used to guide this study:

(1) What evidence from prior research in the areas of instructional message design and

universal design be used to inform mLearning content design?
(2) How can such evidence be translated in a set of message design guidelines for the development of mLearning content?

Methodology

This study employed a design and development method of research. Design and development, as defined by Richey and Klein (2007) is, "the systematic study of design and development and evaluation process with the aim of establishing an empirical basis for the creation of instructional products and tools and new or enhanced models that govern their development" (p. 1). Design and development research addresses two different types of inquiry: product and tool research (Type 1) and model research (Type 2). The proposed study employed product and tool research. This approach includes all of the processes leading to the production of an instructional or non-instructional tool (Richey & Klein, 2007). To address concerns regarding areas of validity, causal inferences, generalizations and interpretation, and anticipation of problems in this type of study, Richey and Klein (2007) have recommended that experts with differing areas of specialization should be used for tool review.

Instrumentation

The survey for the expert reviewers consisted of three sections and 27 questions representing the specific guideline categories and validity factors. Section One focused on the general overview of the considerations and included a yes/no question to provide consent. Section Two of the survey focused on obtaining data about specific factors associated with the considerations. These factors included general design, functionality, text, color, video and audio, and graphics. Section Three emphasized the elements of design and development research recommendations. Sections of the survey inquired about the clarity of individual guideline categories and asked for any suggestions to improve them.

Study Procedures

The study employed three phases in its overall design. The first phase (Analysis) was to identify relevant design principles to guide mLearning creation. This involved a literature review of instructional message design principles (multimedia principles), universal design principles (universal design of instruction), and best practices for mobile user interface design. The second phase (Development) involved taking the findings from the research and formulating a set of principles. This was achieved through synthesizing Clark and Mayer's (2016) multimedia principles, universal design principles, and mobile user interface design best practices. In the third and final stage (Evaluation and Revision), feedback and recommendations were gathered from five expert reviewers and revisions were made to the considerations based on their feedback. Institutional Review Board (IRB) approval was obtained before formal emails were sent to recruit potential reviewers.

Phase 1: Analysis. The guideline development process began with a comprehensive literature review of relevant topics pertaining to the research. The importance of doing a literature review is essential as it uncovers the knowledge gap and informs the study (Webster & Watson, 2002). The essential literature review areas covered for the study included instructional message design, principles for designing for mobile learning (mLearning), universal design principles, and mobile interface design.

Literature that was gathered for the study was carefully analyzed to determine relevance to the guideline development process. The articles, books, and other resources were critically evaluated for any recurring ideas. Key areas included Clark and Mayer's multimedia principles (2016) and their interpretations, universal design principles for mediated instruction, empirical studies supporting each principle, applicable mobile interface design best practices, and any other content areas needed for the considerations.

Phase 2: Development. After analyzing the relevant literature, key research findings and best practices from the articles and resources were grouped as they aligned with multimedia and/or universal design principles. The purpose of engaging in this activity was to organize ideas to ensure that resources for each classification were easily accessible once the guideline building process began, as well as to ensure that each area was supported with credible evidence.

After careful analysis, four main categories were identified. These included statements of the principle, explanations of the principle, supporting literature, and how to operationalize considerations for designing instruction for mobile phones. These four categories organized the resources for easy understanding and interpretation. The final stage of categorizing the resources was to operationalize the considerations for designing for mobile phones based on the evidence presented. At this stage, thought went into the challenges of mobile phone use and how the principles might be applicable for designing content. These were stated in simple and easy to understand terms so that content developers and instructors could easily conceptualize how the principles could be used in the mobile phone content development process. Appendix A reflects Mayer's Multimedia Principles (2016) and the correlating recommendations operationalizing these principles based on relevant literature. Appendix B represents the outcomes of the literature analysis related to the Center for Universal Design's principles for universal design for instruction (1997). These foundational principles set the initial standards for the creation of accessible design, based on input from experts across a broad array of design sciences. Since their inception, these recommendations remain widely used to inform accessible web design (Zheng, 2021).

Following the identification of key design considerations from instructional message design and universal design literature, further synthesis was conducted to consolidate and organize these factors into overarching mLearning design recommendations. These guidelines were organized for application, with ease of use and functionality as key design priorities. Upon careful consideration, the following categories were chosen to frame the synthesized recommendations: general design principles for content and presentation, designing for context and function, and guidance for adding text, video and audio, animation, graphics, and color. The general design principles and specific suggestions for designing for function were stated first so that designers can easily access them for specific forms of content. Table 1 represents the culminating guidelines for mLearning message design.

Table 1
Guidelines for mLearning Message Design

General Content Design	• Design instructional text using a simple and clear writing style. • Limit concepts to one per screen. • Use speech input as a viable alternative for text entry. • Limit the use of external links. • Include a menu or table of contents for easy navigation of instruction. • Apply consistency in the use of design elements like color, font, graphics, etc. • Keep sentences short. • Be consistent with navigation functions. • Avoid the need for excessive scrolling. • Preview the content on a variety of mobile phone screens.
General Content Presentation	• Design content in small units. • Deliver content in the simplest possible formats. • Present information in multiple formats, such as a combination of text, graphics, and/or video. • Avoid small font size to ensure legibility. • Check text for readability
Design for Context	• Provide a clear and consistent way to return to the home screen. • Make buttons easy to click/use with one hand. • Navigation should provide easy access to help, both technical and instructional.
Design for Function	• Include the ability to review previously viewed content. • Provide the opportunity to stop and start module activities as desired. • Design content such that mobile users can readily view content, despite device screen size. • Avoid pop-ups, mouse-overs, or auto-refresh for mobile content. • Use cloud-computing file storage and sharing to address storage and access needs. • Explore the use of speech recognition as a plausible means of entering information.
Adding Text	• Avoid the inclusion of text that duplicates audio narration information. • Use text signaling strategies such as outlines, headings, highlighting, bolding, or pointer words (e.g., first, second, etc.) to draw attention to salient points. • Use sans-serif fonts to increase legibility. • Make textual content as concise as possible.

Adding Video and Audio	• Segment video and audio files into smaller chunks, when possible. • Add captions to video content and transcripts to audio content. Text-to-speech features can assist with this process. • Provide easy and accessible controls for video/audio playback (pause, go back, go forward).
Adding Graphics	• Add appropriate graphics to textual content to help visualize concepts. • Add alt-text descriptions to graphics.
Adding Color	• Use contrasting colors to increase legibility of text. • Use color for visual cueing. • Keep color coding consistent throughout the content design. • Use contrasting colors to highlight and draw attention to key concepts. • Use a color contrast checker to preview color selection decisions.

Phase 3: Evaluation and revision. Five expert reviewers were recruited from the areas of instructional message design, instructional design, mobile learning, and accessibility/universal design based on their expertise and scholarly reputation in these areas. These reviewers provided an evaluation of the proposed guidelines based, using a customized rubric created by the researchers to record their suggestions and feedback. Perspectives and recommendations from the reviewers were synthesized and revisions were made to the guidelines based on their collective input. Revisions that fell outside of the scope of the study were not addressed, such as a suggestion to create an app for guideline application.

Results

Overall, the expert reviewer feedback indicated that the guidelines would be helpful to instructional designers and content developers. Four out of five reviewers agreed that the proposed design strategies are likely to effectively improve the quality of mobile learning in distance learning experiences. Three of the experts also agreed that the organization of the considerations, a key design feature based on the synthesized literature, supported the purpose and use of the final product. Additionally, each of the reviewers provided feedback about the practicality and effectiveness of the proposed design guidelines, and indicated possible challenges related to their use. Some concerns were expressed about factors such as lack of designer expertise or project time constraints posing barriers for guideline usage for mLearning development. Each reviewer also provided feedback for revision of the considerations, with most recommendations focused on the technical writing style, nomenclature related to message design terms, and structuring of the recommendations to match the applicable design task. Much of the feedback provided was incorporated into the final version of the guidelines in Table 1, with some recommendations falling beyond the scope of the proposed study.

Discussion

The design and development activities in this study, framed by the following research questions, culminated in the final set of message design considerations for mLearning content:

(1) What evidence from prior research in the areas of instructional message design and universal design can be used to inform mLearning content design?
(2) How can such evidence be translated in a set of message design guidelines for the development of mLearning content?

Decades of inquiry and theory-based principles in the areas of multimedia learning, instructional message design, and universal design provided relevant insights to inform the aesthetic and functional design of mLearning content. Analysis and synthesis of these findings generated a practical collection of message design considerations for instructional content developers charged with mLearning creation that is accessible and effective.

A field test will be necessary to further validate the use of the proposed considerations (Richey & Klein, 2007). Such testing should be conducted with instructional designers, user experience design professionals, and potential faculty users, with the opportunity to apply the considerations in their natural environment. Formative evaluation should be performed at every stage, and changes made to enhance the usability and effectiveness of the considerations. Further systematic research should be conducted on the impact of instructional message design decisions, such as those proposed herein, on both student and instructor mLearning experiences.

Conclusion

The use of mobile phones for accessing online content has become a common phenomenon in online learning (Statistica, 2017) and has experienced exponential growth during the COVID-19 pandemic as a means of instructional continuity (Pebriantika, Wibawa, & Paristiowati, 2021). However, a long-standing need for research has existed to inform the design of mLearning content, given its unique aesthetic features and functional affordances (Gao, Novak, & Tang, 2009; Haag & Berking, 2015; Saleh & Bhat, 2015; Shen & Vincent-Layton, 2015; Wang & Shen, 2012; Wang & Wishart, 2009). This study offers guidance for designing mobile learning experiences, drawn from prior relevant research and best practices in multimedia learning, instructional message design, and universal design for learning. As Beirne and Romanoski (2018) state, “growing numbers of students are looking for more flexible formats for undertaking courses, certificates, and degree programs” (p. 1). Given that mLearning adoption and growth will likely continue into the future, it is even more important that evidence-based guidance, such as the message design considerations provided herein, is available to inform the design of mobile learning courses and programs.

References

Aguayo, C., Cochrane, T., & Narayan, V. (2017). Key themes in mobile learning: Prospects for learner-generated learning through AR and VR. *Australasian Journal* of *Educational Technology*, *33*(6). https://doi.org/10.14742/ajet.3671

Beck, I. L., McKeown, M. G., Sandora, C., Kucan, L., & Worthy, J. (1996). Questioning the author: A yearlong classroom implementation to engage students with text. *The Elementary School Journal, 96*(4), 385–414. https://doi.org/10.1086/461835

Beirne, E., & Romanoski, M. P. (2018). Instructional design in higher education: Defining an evolving field. *OLC outlook: An environmental scan of the digital learning landscape.*

Burgstahler, S. (2012). *Equal access: Universal design for instruction*. Seattle, WA: DO-IT, University of Washington. https://www.washington.edu/doit/equal-access-universal-design-instruction

Butcher, K. R. (2006). Learning from text with diagrams: Promoting mental model development and inference generation. *Journal of Educational Psychology, 98*(1), 182. https://doi.org/10.1037/0022-0663.98.1.182

Butcher, K. R. (2014). The multimedia principle. *The Cambridge handbook of multimedia learning* (pp. 174–205). Cambridge University Press.

Center for Universal Design (1997). *The principles of universal design*. North Carolina State University.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293–332.

Clark, R. C., & Mayer, R. E. (2016). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning.* Wiley.

Dye, K. L. (1997). *Message design: A key to effective instructional materials*. http://kathydye.com/IDEAS/MessageDesign2.pdf

Elias, T. (2011). Universal instructional design principles for mobile learning. *The International Review of Research in Open and Distributed Learning*, *12*(2), 143–156. https://doi.org/10.19173/irrodl.v12i2.965

Fleming, M. L., & Levie, W. H. (1993). *Instructional message design: Principles from the behavioral and cognitive sciences* (2nd ed.). Educational Technology Publications.

Gong, J., & Tarasewich, P. (2004). Guidelines for handheld mobile device interface design. In Proceedings of DSI 2004 Annual Meeting (pp. 3751–3756).

Gu, X., Gu, F., & Laffey, J. M. (2011). Designing a mobile system for lifelong learning on the move. Journal of Computer Assisted Learning, 27, 204–215. https://doi.org/10.1111/j.1365-2729.2010.00391.x

Haag, J., & Berking P. (2015). Design considerations for mobile learning. In Y. Zhang (Ed.).Handbook of mobile teaching and learning (pp. 41–60). Springer. https://doi.org/10.1007/978-3-642-54146-9

Hanbridge, S. A., Tin, T., & Sanderson, N. (2018). Information literacy skills on the go: Mobile learning innovation. *Journal of Information Literacy*, *12,* 118–136. https://doi.org/10.11645/12.1.2322

Hoober, S., & Berkman, E. (2011). *Designing mobile interfaces: Patterns for interaction design.* O'Reilly Media, Inc.

Huang, R., Zhang, H., Li, Y. & Yang J. (2012). A framework of designing learning activities for mobile learning. In S. K. S. Cheung, J. Fong, L. F. Kwok., K. Li. & Kwan R. (Eds.). *Hybrid Learning*. Lecture notes in computer science, vol. 7411. Springer. https://doi.org/10.1007/978-3-642-32018-7_2

Lal, R. (2013). *Digital design essentials: 100 ways to design better desktop, web, and mobile interfaces.* Rockport Publishers.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 93*(2), 377–389. https://doi.org/10.1037/0022-0663.93.2.377

Mayer, R. E. (2001). *Multimedia learning.* Cambridge University Press. https://doi.org/10.1017/CBO9781139164603

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511811678

Mayer, R. E. (2013). Multimedia learning. In J. M. Spector, M. D. Merrill, J. Ellen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed.). Springer.

Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology, 84*(4), 444–452. https://doi.org/10.1037/0022-0663.84.4.444

Mayer, R. E., & Chandler, P. (2001). When learning is just a click away: Does simple user interaction foster deeper understanding of multimedia messages? *Journal of Educational Psychology*, *93*(2), 390–397.

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology, 95*(4), 806–812.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187–198.

McQuiggan, S., McQuiggan, J., Sabourin, J., & Kosturko, L. (2015). *Mobile learning: A handbook for developers, educators, and learners*. John Wiley & Sons.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358-368.

Nilsson, E. G. (2009). Design patterns for user interface for mobile applications. *Advances in Engineering Software, 40*(12), 1318–1328. https://doi.org/10.1016/j.advengsoft.2009.01.017

NMC Horizon Report. (2017). *Higher education edition*. http://cdn.nmc.org/media/2017-nmc-horizon-report-he-EN.pdf

O'Malley, C., Vavoula, G., Glew, J., Taylor, J., Sharples, M., & Lefrere, P. (2003). Guidelines for learning/teaching/tutoring in a mobile environment. *Mobilearn Project Report,* D4.1.

Pebriantika L, Wibawa B, Paristiowati M. (2021). Adoption of mobile learning: The influence and opportunities for learning during the Covid-19 pandemic. *International Journal of Interactive Mobile Technologies, 15*(5). 222–230. https://doi.org/10.3991/ijim.v15i05.21067

Pegrum, M. (2019). Mobile devices. In *Mobile lenses on learning* (pp. 127-170). Springer. https://doi.org/10.1007/978-981-15-1240-7

Richey, R. C., Klein, J. D., & Nelson, W. A. (2004). Developmental research: Studies of instructional design and development. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 1099–1130). Lawrence Erlbaum Associates.

Richey, R., & Klein, J. D. (2007). *Design and development research: Methods, strategies, and issues*. Routledge.

Saleh, S. A., & Bhat, S. A. (2015). Mobile learning: A systematic review. *International Journal of Computer Applications*, *114*(11), 1–5. https://doi.org/10.5120/20019-1406

Shen, R., Wang, M., Gao, W., Novak, D., & Tang, L. (2009). Mobile learning in a large blended computer science classroom: System function, pedagogies, and their impact on learning. *IEEE Transactions on Education, 52*(4), 538–546. https://doi.org/10.1109/TE.2008.9307

Shneiderman, B., & Plaisant, C. (2010). *Designing the user interface: Strategies for effective human-computer interaction* (5th ed.). Addison-Wesley.

Stull, A. T., & Mayer, R. E. (2007). Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology, 99*(4), 808–820. https://doi.org/10.1037/0022-0663.99.4.808

Subramanya, S. R., & Yi, B. K. (2006). User interfaces for mobile content. *Computer*, *39*(4), 85-87. https://doi.org/10.1109/MC.2006.144

Sweller, J. (2011). Cognitive load theory. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 37–76). Elsevier Academic Press. http://dx.doi.org/10.1016/B978-0-12-387691-1.00002-8

Talidong, K. J. B., & Toquero, C. M. D. (2021). Mobile learning grounded on universal instructional design principles on the shift to emergency online education. *Journal of Digital Educational Technology, 1*(1), ep2102. https://doi.org/10.21601/jdet/10943

Tereshchenko, S., Zagorskaya, M., Polyanskaya, O., & Bobritskaya, J. (2020). Mobile learning in forestry education. n IOP Conference Series: Earth and Environmental Science (Vol. 507, No. 1, p. 012031). IOP Publishing. https://doi.org/10.1088/1755-1315/507/1/012031

Uther, M. (2019). *Mobile learning*. MDPI - Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/educsci9010033

Vavoula, G., & Karagiannidis, C. (2005, November). Designing mobile learning experiences. In *Panhellenic Conference on Informatics* (pp. 534–544). Springer.

Vincent-Layton, K. (2015). Mobile learning and engagement: Designing effective mobile lessons. In Y. Zhang (Ed.), *Handbook of mobile teaching and learning*. Springer. https://doi.org/10.1007/978-3-642-54146-9_62

Wang, M., & Shen, R. (2012). Message design for mobile learning: Learning theories, human cognition and design principles. *British Journal of Educational Technology, 43*(4), 561–575. https://doi.org/10.1111/j.1467-8535.2011.01214.x

Webster, J., & Watson, R. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly, 26*(2), Xiii-Xxiii.

Wishart, J. (2009). Use of mobile technology for teacher training. In M. Ally (Ed.), *Mobile learning: Transforming the delivery of education and training* (pp. 265–278). Athabasca University Press.

Zheng, R. (2021). Learn to create accessible websites with principles of universal design. Interaction Design Foundation. https://www.interaction-design.org/literature/article/learn-to-create-accessible-websites-with-the-principles-of-universal-design

Corresponding author: Barbara B. Lockee
Email: lockeebb@vt.edu

Appendix A

Mobile Learning Design Considerations by Multimedia Principle

Principle	Principle Explanation	How to Operationalize for mLearning Design	Supporting Literature
Multimedia Principle (Mayer 2001, 2009)	Presentation of content should be a combination of both words and images. (Mayer, 2005).	Add appropriate graphics to text content to help learners' get the most of meaning from the content. Present information in multiple formats.	People who learned from words and graphics produced between 55 percent to 121 percent more correct solutions to transfer problems than people who learned from words alone (Clark & Mayer, 2016). People developed a deeper understanding of how the human heart works from text with simple illustrations than from text alone (Butcher, 2006).
Modality principle (Mayer 2002, 2009)	Students learn better from a combination of animation and narration than from animation and on-screen text. (Mayer, 2002)	Avoid creating animation with narration and onscreen text as this could cause cognitive overload. Choose narration when possible and avoid narration with text at the same time.	Students performed better on a transfer test after receiving a narrated animation on lightning formation than after receiving the same animation with on-screen captions that contained the same words as the narration (Moreno & Mayer, 1999).
Contiguity Principle (Mayer 2002, 2009)	The effectiveness of multimedia instruction increases when words and pictures are presented near each other in time or space (Mayer	Place printed words near corresponding images or other media to bring more meaning to content.	Moreno and Mayer (1999) found that students performed better on a transfer test after viewing an animation about lightning in which printed words were placed next to the part

	& Anderson, 1992). Learners build connections between corresponding words and graphics (Mayer, 2009).		of the lightning system they described than when printed words were placed at the bottom of the screen as a caption.
Redundancy principle (Mayer 2002, 2009)	People learn better from concurrent graphics and audio than from concurrent graphics, audio, and on-screen text (Clark & Mayer, 2016). Students learn better from animation and narration than from animation, narration, and on-screen text (Mayer, 2002).	Avoid including on-screen text with a narrated graphic as that will be duplicating information.	Avoid e-learning courses that contain redundant on-screen text presented at the same time as on-screen graphics and narration (Clark & Mayer, 2016). Learning from a diagram or graph is hurt by the addition of textual information that redundantly explains with words what the diagram or graph already shows pictures (Chandler & Sweller, 1991).
Coherence principle Mayer (2002)	People learn more deeply from a multimedia message when extraneous material is excluded rather than included. Student learning is improved when unneeded words or sound are eliminated from a multimedia presentation (Mayer, 2002). Rather than extensive narrative	Avoid adding extraneous material which is not included in instructional goal or not relevant to the understanding of the content.	Consistency of "look and feel" should be the same across multiple platforms. Elements of mobile interfaces such as names, color schemes, and dialog appearances should have consistency in functionality (Gong & Tarasewich, 2004; Lal, 2013; Shneiderman, & Plaisant, 2010). When pictures are used only to decorate the page or screen, they are not likely to improve learning (Clark & Mayer,

	descriptions, stick to concise presentation of the content (Clark & Mayer, 2016).		2016).
Personalization principle (Mayer 2002, 2009)	Students learn better when words are presented in conversational style rather than formal style (Mayer, 2002). Use conversational rather than formal style. Use polite wording rather than direct wording and use human voice rather than machine voice (Clark & Mayer, 2016).	Use a conversational style of writing unless otherwise specified when creating content for students. Use the first or second person in your narration or text.	People learn better from a narrated animation on lightning formation when the speech is in conversational style rather than formal style (Moreno & Mayer, 2000b). People work harder to understand material when they feel they are in a conversation with a partner rather than simply receiving information (Beck, McKeown, Sandora, Kucan, & Worthy, 1996).
Segmentation Principle (Mayer, 2002)	People learn more deeply when a multimedia message is presented in learner-paced segments rather than as a continuous unit (Mayer, 2009).	Design mobile content in small or unit sections to enable learners better understand without any overload.	Design content as small, self-contained units (Gu, Gu, & Laffey, 2012). Learners who received segmented content presentation performed better on transfer tests than the learners who received a continuous presentation, even though identical material was presented in both conditions (Mayer & Chandler, 2001).
Pre-training Principle (Mayer 2002,	People learn more deeply from a multimedia	Include names and characteristics of main concepts at the	People performed better on problem-solving transfer tests

2009)	message when they have first learned the names and characteristics of the main concepts (Mayer, 2009).	beginning of mobile content when designing a module to help leaners gain awareness of each major component.	when a multimedia lesson was preceded by pre-training in the names and characteristics of each key component (Mayer, 2009).
Signaling Principle (Mayer 2002, 2009)	People will learn more efficiently if the lesson is designed to call attention to the most important material (Mayer, 2009). People learn better when cues are added to highlight the organization of essential material (Mayer, 2009).	Use text signaling strategies such as outline, headings, highlight, bolding, pointer words such as first, second etc. when designing mobile content to draw learners' attention to salient points.	Signaling of visual material includes arrows, flashing, and spotlighting. For example, in a narrated animation on how an airplane achieves lift, students performed better on a transfer test if the narration included an initial outline, headings, and voice emphasis on key words (Mautone & Mayer, 2001).

Appendix B

Mobile Learning Design Considerations by UDI Principle

UDI Principle	Explanation	How to Operationalize for mLearning design	Supporting Literature
Equitable use	The design is useful and marketable to people with diverse abilities. Develop content and assignments that can be accessed on a wide variety of devices. Course content should be accessible to people with diverse abilities and in diverse locations (Elias, 2011).	Deliver content in the simplest possible formats. Short Messaging Systems (SMS), or texting technology is inexpensive and given its high levels of penetration, is universally accessible (Elias, 2011). Given the small storage capacity of most smart phones and do not have a big storage capacity, using cloud-computing file storage and sharing sites may be a better option.	Cater for universal usability and recognize the diverse users (Shneiderman & Plaisant, 2010). Design for multiple and dynamic contexts, configure output to users' needs and preferences (e.g., text size, brightness) (Gong & Tarasewich, 2004).
Flexibility in use	Accommodating a wide range of individual abilities, preferences, schedules, levels of connectivity, and choices in methods of use (Elias, 2011).	Design content such that mobile users who have smaller screens and bigger screens can still access the content they need. Do not use pop-ups, mouse hover, or auto-refresh for mobile contents. Inputting text data into small devices can also present challenges for the user.	Ensure that lives take precedence because mobiles are contextual and are used alongside people's actual lives (Hoober & Berkman, 2012). Do not use pop-ups, mouse hover, or auto refresh (Lal, 2013). Should be Intuitive and easy to use. The interface should be simple enough for anyone educated or not to use (Subramanya & Yi, 2006).

Simple and intuitive use	Use design that is easy to understand, regardless of the user's experience, knowledge, language skills, or current concentration level (Center for Universal Design, 1997). Unnecessary complexity should be eliminated and course design rendered simple and intuitive (Elias, 2011).	Design the content in very simple and easy to understand format. Think about designing for the smallest mobile phone content and ensure content will fit without cognitive overload. Speech input is a viable alternative since some devices may be too small for buttons.	The interface should be simple enough for anyone to use (Subramanya & Yi, 2006). Keep learners' interfaces simple. It should be ensured that they contain only information that can fit comfortably on the smallest of screens (Elias, 2011). Allow website to scale for all mobile browser layouts for both portrait and landscape (Lal, 2013).
Perceptible information	The design communicates necessary information effectively to the user, regardless of ambient conditions or the user's sensory abilities (Center for Universal Design, 1997). Adding captions, descriptors, and transcriptions increases learners' perception of the content and reaches everybody in spite of any disability (Elias, 2011).	Add captions to video content and transcripts to audio content. Use text to speech features in authoring tools.	A video presentation should include alternative forms of the spoken work, including captions, descriptors and transcriptions (Burgstahler, 2009; Elias, 2011).
Tolerance and error	The design minimizes hazards and the adverse consequences of accidental or unintended actions (Burgstahler, 2012). Minimize hazards and adverse consequences of errors in software operation by designing learning environments	Design content such that learners can go back to review content which was previously viewed. Provide opportunity to reduce error by allowing mobile content to be stopped and started, as well as revisited.	Users must have some control over the use of the device (Shneiderman & Plaisant, 2010). Allow website to scale for all mobile browser layouts for both portrait and landscape (Lal, 2013).

	with a tolerance for error (Elias, 2011).		Allow applications to be stopped, started, and resumed with little or no effort (Gong & Tarasewich, 2004). Provide word selection instead of requiring text input (Gong & Tarasewich, 2004).
Low physical and technical effort	The design can be used efficiently, comfortably, and with a minimum of fatigue (Burgstahler, 2007; Elias, 2011).	Limit use of external links Use short messaging systems (SMS), or texting technology which is easy to use and widely available. Include menu or table of content to a module for easy navigation.	In addition to radio buttons, combo and check boxes, include spinners, sliders, and menu for easy manipulation which is usually more efficient and easier to perform than typing (Nilsson, 2009). Provide information about level of progress to make learners more patience and anticipate how long it will take to complete a module (Nilsson, 2009).
Size and space	Appropriate size and space are provided for approach, reach, manipulation, and use regardless of the user's body size, posture, or mobility (Burgstahler, 2007).	Limit to one idea per screen. Use the phone screen effectively as it is small compared with regular desktop computer. Design content in small units.	Design for small devices and provide word selection instead of requiring text input (Gong & Tarasewich, 2004). Design micro content items as small, self-contained and granular learning (Gu, Gu, & Laffey, 2012). Keep screen layout with plenty of white space and do not clutter. Keep one idea

			in one screen, and don't overload data (Lal, 2013).

Reviewers – Volume 9 – Issue 4

The editorial team would like to thank the following reviewers for their contributions to the peer review of this issue. Their dedication and assistance are greatly appreciated.

Senior Reviewers

Dr Shiao-wei Chu, National Chin-Yi University of Technology, Taiwan
Dr Mikel Garant, Beijing Institute of Technology, China
Dr Guruprasad Nagraj, Global Academy of Technology, Bangalore, India
Dr Shwadhin Sharma, California State University Monterey Bay, USA
Dr Michelle Yeo, Coventry University and University of Newcastle, Australia
Dr Elena Mishieva, Moscow State University, Russia

Reviewers

Dr Hee Seon Jang, Pyeongtaek University, Korea
Dr Debbie Meharg, Edinburgh Napier University, UK
Dr Serpil Meri-Yilan, Ibrahim Cecen University, Turkey
Mirela Ramacciotti, Independent Researcher, Brazil
Dr Leonardo Munalim, The Philippine Women's University-School of Arts & Sciences and School of Education, The Philippines
Dr Murielle El Hajj Nahas, Lusail University, Qatar
Dr Melinda Cowart, Texas Woman’s University (TWU), USA
Dr Philip Ajewole, Federal University Oye-Ekiti, Nigeria
Dr Queen Ogbomo, Tennessee Technological University, USA
Dr Urarat Parnrod, Songkhla Rajabhat University, Thailand
Dr Ana Garcia Diaz, UCJC-University Camilo José Cela, Spain
Dr Menyani Nourreddine, Tofail University Faculty of Arts and Humanities, Morocco
Michelangelo Magasic, Kochi University of Technology, Japan

the iafor

journal of education

ISSN: 2187-0594