

Leveraging AI to Design Comprehensible Instructional Videos



Ryan Rumpf

WCEPS & Missouri Western University, Missouri, USA // rrumpf5@gmail.com

ABSTRACT: This article examines the integration of artificial intelligence (AI) into the design of instructional videos specifically targeting newcomer multilingual learners within a mastery-based educational framework. Drawing from the author’s experience as the leader of the newcomer program, the article outlines the challenges of using traditional methods in a mastery-based system. The article details how AI can bridge the gap between effective instructional programming and the needs of newcomers who are at the beginning stages of English language acquisition. Situating these practices within Cognitive Load Theory and current research on video-integrated instruction, the author highlights both pedagogical benefits and practical limitations as well as the tools needed to do the work. Implications include the need for professional development in AI integration, potential enhancements to multilingual instructional design, and calls for further research on AI’s impact on newcomer education.

Keywords: artificial intelligence, instructional video, newcomers, mastery-based learning, cognitive load theory

1. Introduction

Between 2021 and 2024, I was the administrator of a newcomer program in the Midwest that served recently-arrived immigrants and refugees. This program was the first phase of the district’s Language Instruction Educational Program (LIEP) and was complemented by comprehensive and centralized wraparound services. At any given time, we served between 100-140 students from around the world. The most common first languages in our program were Spanish, Swahili, Pashto, Dari and Kinyarwanda. Our students came to us with a wide range of experiences and educational backgrounds and consequently, required that our team think beyond traditional approaches to school.

As a team, we collaboratively developed a program where students were in self-contained classrooms, those comprised exclusively of newcomers, for core academic classes while also immersed with grade-level peers during electives, interventions and more social time periods, such as lunch. Rather than replicating the district’s standard approach to curriculum, instruction and assessment, the team and I developed a mastery-based system, or a system with a clear learning progression where students advance upon demonstrating mastery. Our goal was to accelerate learning, not keep pace. The team spent time in the summer planning a model that infused language development into the content-driven, mastery-based system. As a team, we methodically developed learning progressions, a coding system, multilingual assessments and several other elements to ensure the system functioned properly. The final component was developing a learning cycle that complemented our approach which started with instructional video.

2. The Use of Video to Enhance Instructional Design

An essential component of a mastery-based system is the continual development of virtual resources for each learning target (Marzano & Abbott, 2022). The team favored the approach due to the clear and obvious popularity of video in the modern world; many kids and adults around them spent an exceedingly amount of time glued to the videos. In 2023, the worldwide average of short video consumption was 19 hours per week (Ceci, 2024). In 2025, YouTube marked 20 years and had grown beyond entertainment to become a popular tool for learning (Brown, 2025). Naturally, video has grown increasingly popular in K-12 and higher education classrooms as well (Walsh & Henderson, 2022).

While the research base supporting instructional video is inconsistent, there are promising results. A study on students at the university level showed that replacing existing teaching methods with video showed minimal benefits, but incorporating video into existing teaching led to improved student learning (Noetel et al., 2021). This important finding highlights how instructional videos are complementary, not a replacement for classroom instruction. In our setting, we intended to use video as the initial exposure to a learning target, tutorials and a potential medium for assessment.

Instructional video impacts working memory and this concept can be viewed through the lens of Cognitive Load Theory (CLT). A study on the effectiveness of instructional videos concluded that the use of video is enhanced when instructors consider its place in active learning, student engagement and cognitive load (Brame, 2016). CLT proposes three distinct categories that compete for space in our working memory: intrinsic load, extraneous load and germane load (Strohmaier et al., 2023). CLT defines intrinsic load as the inherent difficulty or complexity of a task, extraneous load as distractions that hinder learning, and germane load as instructional support that facilitates the development of meaningful schemata stored in long-term memory (Reese, Pawluk & Taylor, 2016). Continued studies of the effects of video on cognitive load showed similar results. In a recent study on the impact of instructional videos on cognitive load in a K-12 setting, researchers found video tutorials decreased students' intrinsic load, increased germane load, but did not affect extraneous load (Fan, Bower & Sieman, 2024). In other words, video was shown to reduce the inherent difficulty of a concept or task and eased the mental effort needed for learning academic content.

Research on the structure and format of effective video is scarce, but it is emerging. In a meta-analysis of instructional video technologies, researchers found that in-class video clips enhanced learning more than outside videos and video games while impacting motivation and practical skill learning more than knowledge acquisition (Lin & Yu, 2024). Additionally, the authors found that K-12 students and adults experienced greater benefits than undergraduates (Lin & Yu, 2024).

In an empirical study, Walsh and Henderson (2022) summarized the benefits of video that were consistently found in studies across time and place:

- Positive affective and cognitive attitudes toward learning—students believed they learned more as a direct result of viewing the videos.
- Greater control over the learning process including what needed to be learned, when to learn it, where to learn and the pace of learning.
- Study habits improved and that led to greater independence, self-reflection, and test preparation.
- Higher scores on tests and improved technology skills led to increased overall learning performance (p. 18).

Our team bought into the use of video; however, some were skeptical about making them. Some team members were much more comfortable creating videos than others. As the leader, I took the lead in helping certain team members understand how to use the tools and technology. Once the teachers were comfortable using cameras and the software, the structure and format of the video became the issue. Unfortunately, there is little guidance or research on the structure and format of an effective instructional video. In fact, terminology in the research community is not consistent (Walsh & Henderson, 2022). The

team and I decided to align the videos with the flow of the gradual release of responsibility model: I do, we do, and you do (Badwan, 2018).

Using a well-crafted prompt to ChatGPT, the team used the AI tool to produce an initial draft of the script, then collaborated on edits to ensure the video clearly conveyed the intended message. When the script was ready, the team used a document camera, screencast software or a camcorder to record each scene. Some scenes were combined using video editing software while others were made in their entirety through a single recording.

When the students showed up, the team was eager to get started. They had spent a lot of time and energy designing the structure and were cautiously optimistic about implementation; it was new after all. As it turned out, even the most thought-out video that was slowed down and captioned was not accessible to the newcomers. The students could not understand the videos. While the team carried on and continued working through the routines and procedures with the students, I started looking for a solution that tapped into the students' native languages.

3. Supporting Newcomers through the L1

The value of the native languages was not lost on our team. We had collectively adopted the six guiding principles published by The Understanding Language District Engagement Subcommittee at Stanford University. One of the principles included the use of the home language in connecting prior knowledge to new knowledge and providing comprehensible instruction (Stanford University, 2013). The power of the home language is widely recognized across the field of TESOL and is commonly advocated for by groups such as TESOL International, Sheltered Instruction Observation Protocol (SIOP), the Common Core State Standards Initiative and the Next Generation Science Standards (Wright, 2019). In fact, prominent scholars have advocated for native language use in classrooms since at least as early as 1837 when Calvin Stowe submitted a report to the Ohio General Assembly urging state leaders to develop specialized programs for the growing immigrant population (Stowe, 1837). We discovered that students' difficulty in grasping the learning objectives often hindered their engagement with grade-level material. We believed that starting the learning process in the native language would streamline instruction and maximize time spent on academic work.

An added benefit of offering translations was the increased sense of belonging by affirming students' home languages. A qualitative study on Hmong and Thai newcomers reported that the use of the home language improved confidence-levels and learning efficiency (Bigelow, Basford & Smidt, 2008). In a mixed-methods descriptive case study on recently arrived immigrant adolescents, the use of the home language helped students feel included, accepted and valued (McInerney, 2022). Recognizing the social and academic value of the home language, it was clear that the language barrier in our instructional video needed to be addressed.

Initially, I asked the district's family liaisons to help recreate the videos in the students' home languages, but the process quickly became too burdensome and time-consuming. The family liaisons were excellent cultural brokers and strong at building relationships with families. However, they lacked the educational background to create instructional videos and the time required to produce them. As the search for a viable solution progressed, I began investigating technology-based approaches and found an AI-powered video dubbing platform called Maestra (Maestra, n.d.).

4. Using AI to Bridge Language Gaps

Like many others, I had faced difficulties with technology-based translation tools in the past, but I was intrigued by advancements in artificial intelligence and curious about their translation capabilities. In a comparative study between human translators and AI-powered translation services focused on legal documents, the AI-powered service scored slightly lower (89.1) than the human translation (92.7) (Moneus

& Sahari, 2024). The researchers concluded that while the human translator was more accurate and was able to address cultural context and ambiguity, the AI tools excelled at speed and scalability as well as consistency (Moneus & Sahari, 2024).

In a comparative study between two popular AI bots, ChatGPT slightly outperformed DeepL in accuracy rates while demonstrating more sophisticated intelligence (Sun, 2024). AI-powered translation services are constantly growing and learning though, with an increasing amount of options and continually improved accuracy rates across languages (Hyvarinen, 2024). More recently, AI translations have increased in precision and speed raising questions about whether humans or artificial intelligence are more efficient or accurate (Pederson, 2025).

[Maestra.ai](#) offered us an opportunity to upload the videos to their program and the artificial intelligence replaced the original speech with a natural-sounding voice in the target language. It was also cost effective, offering an hour of video dubbing for roughly \$15 plus a monthly platform fee which was offset by an educator discount. I scheduled a meeting with Maestra's managing partner Mike Falis who shared his vision of leveraging the nascent artificial intelligence technology to bridge language gaps in an increasingly interconnected world (personal communication, May 24, 2025). Rather than subscribing to a monthly package, he arranged a yearly deal that was better suited for our setting. At first, we used federal funds for the program, but as we moved forward, we purchased the license through community partners and local grants.

I started dubbing each video into nearly all the students' languages: Spanish, Swahili, Arabic, Haitian-Creole, Pashto and several others. Mr. Falis stated that the accuracy analyses differ by language, but the average accuracy rate was around 95% across the platform (personal communication, May 24, 2025). To double-check the accuracy of the translation, I showed the videos to the team of family liaisons, who were native speakers of many of the district's languages. They were all amazed that the videos were in their languages and were curious how I made them. Their approval was enough to move forward.

I experienced a small learning curve using the software. Uploading the original video to the platform was straightforward: choose the file, choose the target language and click the upload button. Since our videos ranged in length from 3-5 minutes, Maestra.ai transcribed and translated the script in less than a minute. From there, I could review the transcription for accuracy and make any necessary edits. I could delete filler words, such as *so* and *like*, and ensure content area words were translated accurately. I could also review the Characters Per Second (CPS) rates to ensure the speed of the voice was natural-sounding. Once the editing was finished, I chose a voice and downloaded the video as an MP4. The whole process took ten minutes or less. Once the first video was complete, I was able to translate the video into other languages maintaining the edits I made to the original translation.

Once the videos were dubbed, they were downloaded and placed into the same shareable files as the original videos. The teachers made QR codes and students were able to quickly access them on their Chromebooks. Now that the video challenge was solved, the team was able to continue their work implementing the mastery-based system. Each unit had a learning progression and the teachers gradually developed videos for each learning target. The students started the learning process in the native language via the video, then proceeded to work in small groups as well as with the classroom teacher to develop the target language associated with the skill. If the student was confused or made a mistake, the video was available to watch again. Additionally, the parents received the QR code through the multilingual communication tool, Talking Points, so they were aware of what their child was learning. In theory, the students could watch the video and practice the skill anywhere, at any time.

At first, the system was not perfect. There is a lot more that goes into being an effective learner than mastering academic skills. Many newcomers who are transitioning to life and school in the United States must also learn how to be a student as well as the routines and procedures of the school. We had to address that as well. After the instructional videos were developed, the focus turned to the development of learner dispositions.

Over a three-year period, we created a video for every learning target. We also expanded the video project to include learner dispositions, routines and procedures, important components of the Student Code of

Contact, and explanations of the various emergency drills throughout the school year. Video became a part of the culture, and it worked. Our newcomer program's exit criterion was an overall 2.0 on the WIDA MODEL. Each student started at the bottom of the scale, a 1.0, and could stay in our program for a full calendar year. Students spent an average of 8.7 months in the program. Additionally, the average score on the WIDA MODEL was a 2.3 overall. On average, we exceeded expectations in both time and language growth.

5. Conclusion

Newcomers, who are beginning their journeys in English, are also taking their first steps in a long, difficult journey to acquiring English and mastering academic content. The technology boom, especially with the growing power and availability of artificial intelligence, can provide access to rigorous, grade-level content. Innovation in this area could provide many benefits to students and the teachers that serve them, including removing a language barrier that prevents equal access to the opportunities afforded to students who understand the language of instruction.

The value of AI-powered translation technology was a successful piece of our instructional program. Potential innovations in the education field are more robust AI-powered multilingual instructional design platforms or integrating AI technology into existing educational software, such as popular Learning Management Systems, personalized learning platforms and assessment systems. Future technology-based educational software should allow for multilingual scripting, dubbing and distribution of video as well as allow for real-time feedback in a language students can understand. Such a technology, used to complement English development, can potentially provide clarity on not only the content being learned, but also a student's performance and next steps in their learning journeys. Importantly, professional development for educators and decision-makers are necessary to ensure proper use of the tool.

Finally, rigorous studies are needed to examine the impact of AI-translated instructional video on multilingual learner outcomes, including comprehension, engagement, sense of belonging and long-term academic growth. While AI has found its way into many different industries, there is still much to learn about its potential benefits and drawbacks. Empirical research on AI-driven instruction will clarify its value to educators, students and their families.

The Author

Dr. Ryan Rumpf is an educator, author and consultant with 22 years' experience as an ELD teacher and leader in Missouri, Colorado, and Saigon. He formerly served as the Director of English Language Development and World Languages at the Missouri Department of Education as well as the principal of an innovative newcomer center in the Midwest. He currently serves as an adjunct professor at Missouri Western State University, a consultant with the Wisconsin Center for Education Products and Services and as an advisor with several other organizations.

References

Badwan, K. (2018). Effective classroom strategies. In Liontas, J.I. (Ed.) *The TESOL Encyclopedia of English Language Teaching*. Hoboken, NJ: John Wiley & Sons, Inc.

Bigelow, M., Basford, L., & Smidt, E. Supporting Hmong newcomers academic and social transition to elementary school. *Journal of Southeast Asian American Education and Advancement*, 3(1),1-22. <http://doi.org/10.7771/2153-8999.1098>

- Brame, C.J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE Life Science Magazine*, 15(4), 1-6. <https://doi.org/10.1187/cbe.16-03-0125>
- Brown, C. (2025). 20 years after its first upload, YouTube is changing the way people learn. Retrieved from <https://why.org/segments/20-years-after-its-first-upload-youtube-is-changing-the-way-people-learn/>
- Ceci, L. (2024). Weekly time spent with online video worldwide 2018-2023. Retrieved from <https://www.statista.com/statistics/611707/online-video-time-spent/>
- Fan, E., Bower, M. & Siemon, J. (2024). Video tutorials in the traditional classroom: The effects on different types of cognitive load. *Tech Know Learn* 29, 2017-2036 <https://doi.org/10.1007/s10758-024-09754-1>
- Hyvarinen, J. (2024). 8 best AI translators: Ranked compared for accuracy & features. Retrieved from <https://medium.com/@jussi-hyvarinen/8-best-ai-translators-ranked-compared-for-accuracy-features-3e9ee49b4f46>
- Lin, Y., Yu, Z. (2024). A Meta-analysis Evaluating the Effectiveness of Instructional Video Technologies. *Tech Know Learn*, 29, 2081-2115. <https://doi.org/10.1007/s10758-023-09669-3>
- Maestra. (n.d.). *Maestra: Voice-over, subtitles & transcription powered by AI*. <https://maestra.ai/>
- Marzano, R. & Abbott, S. (2022). *Teaching in a competency-based elementary school*. Bloomington, IN: Marzano Resources.
- McInerney, K. (2022). Perceptions from newcomer multilingual adolescents: Predictors and experiences of sense of belonging in high school. *Child & Youth Care Forum*, 52, 1041-1072. <https://doi.org/10.1007/s10566-022-09723-8>
- Moneus, A.M. & Sahari, Y. (2024). Artificial intelligence and human translation: A contrastive study based on legal texts. *Heliyon*, 10(6). <https://doi.org/10.1016/j.heliyon.2024.e28106>
- Noetel, M., Griffith, S., Delaney, O., Sanders, T., Parker, P., del Pozo Cruz, B., & Lonsdale, C. (2021). Video improves learning in higher education: A systematic review. *Review of Educational Research*, 91(2), 204-236. <https://doi.org/10.3102/0034654321990713>
- Pederson, F.R. (2025). *Why AI translation is held to a higher standard than human translators*. Retrieved from <https://www.forbes.com/councils/forbestechcouncil/2025/01/29/why-ai-translation-is-held-to-higher-standards-than-human-translators/>
- Reese, D. D., Pawluk, D. T. V., & Taylor, C. R. (2016). Engaging learners through rational design of multisensory effects. In S. Y. Tettegah & S. U. Noble (Eds.), *Emotions and technology: Emotions, technology, and design* (pp. 103-127). Academic Press. <https://doi.org/10.1016/B978-0-12-801872-9.00006-5>.
- Stanford University, Understanding Language District Engagement Subcommittee. (2013). *Key Principles for ELL instruction*. Retrieved 7/24/25 from <https://ul.stanford.edu/resource/key-principles-ell-instruction>.
- Stowe, C.E. (1837). *Report on elementary public instruction in Europe, made to the thirty-sixth general assembly of the state of Ohio, December 19, 1837*. Harrisburg: Thompson & Clark Printers.
- Strohmaier, A.R., Ehmke, T., Hartig, H., & Leiss, D. (2023). On the role of linguistic features for comprehension and learning from STEM texts. A meta-analysis. *Educational Research Review*, 39, 100533. <https://doi.org/10.1016/j.edurev.2023.100533>.
- Sun, R. (2024). Evaluating the translation accuracy of ChatGPT and DeepL through the lens of applied subjects. *AWEJ for Translation and Literacy Studies*, 8(4), 41-53. <http://dx.doi.org/10.24093/awejtls/vol8no4.5>.
- Walsh, T. & Henderson, M. (2022). The impact of video in the K-12 classroom: A scoping study of instructional video research. *Monash University Report*. <https://doi.org/10.26180/19614093>.
- Wright, W.E. (2019). *Foundations for teaching English language learners: Research, theory, policy and practice*. Philadelphia: Calson, Inc.