

# How students perceive lecturers' gestures? An exploration in gesture-meaning matching toward embodied pedagogical agent design

Wei, Lai<sup>a</sup>; Chow, Kenny K. N.\*<sup>a</sup>

<sup>a</sup> School of Design, The Hong Kong Polytechnic University, Hong Kong, China

\* 19085848r@connect.polyu.hk

doi.org/10.21606/iasdr.2023.115

Many studies have demonstrated that meaningful gestures, such as iconic and metaphoric gestures, can significantly aid students' comprehension and enhance their learning engagement. Additionally, virtual lecturers have had a positive impact on students' online learning experiences. However, there is currently no concrete evidence on how students recognize and interpret these gestures. Investigating students' comprehension of natural gestures used by lecturers can provide valuable insights for creating embodied gestural pedagogical agents (PAs). This paper conducted empirical studies to investigate how students recognized and interpreted gestures used by lecturers in real-life design classes. Through the analysis of four sets of paired cases, including observations and semi-structured interviews, we propose a matching model that illustrates nonverbal communication between lecturers and students. Our results indicate that the meanings of meaningful gestures commonly used by lecturers are reasonably equivalent to students' understanding of those gestures. However, there exists a significant misunderstanding between lecturers' specific gestures and students' comprehension. This research explores patterns of gesture-meaning matching for multimodal communication, which holds promise in creating a student-centered gestural PA design.

**Keywords:** *gesture, pedagogical agent, embodied learning, empirical study*

## 1 Introduction

Many researchers have investigated the credibility, appearance, emotions, gestures, and other characteristics of pedagogical agents (PAs; Castro-Alonso et al., 2021; Chen & Wu, 2015; Davis, 2018; Noetel et al., 2021). The credibility and appearance of agents have received the most attention (Kantharaju et al., 2018; Natarajan & Gombolay, 2020; Schroeder et al., 2021; Song & Luximon, 2020). Based on the existing literature, research on the trustworthiness and appearance of agents indicates that there is still great potential for exploration in the study of embodied pedagogical agents (EPAs; Grivokostopoulou et al., 2020; Petersen et al., 2021). EPAs have been found to increase students' sense of immersion in their virtual learning experience and improve learning engagement and performance (Fitton et al., 2020; W. Li et al., n.d., 2019; Wang et al., 2018). Natural anthropomorphic expressions are particularly considered to aid in developing trust in the learning process (Ratcliffe & Tokarchuk, 2020; Wu et al., 2021). However, the design of PAs depends on students' knowledge, and



students' comprehension of embodied teaching agents' gestures is less frequently advocated. Moreover, studies on embodied agent design rarely mention the relationship between lecturers' natural expression and students' understanding.

The impact of lecturer gestures on student learning outcomes has been extensively researched. Several studies suggest that lecturers' gestures with deictic, iconic, and metaphoric meanings can assist students in reinforcing learning engagement and effectiveness (Beege et al., 2020; Goldin-Meadow, 1999; Wu & Coulson, 2005). Some scholars have claimed that gestures that can potentially express meaning, such as iconic and metaphoric gestures (McNeill, 2011), applied in an educational context positively influence students' comprehension of knowledge (Goldin-Meadow, 2003; Macedonia & Kriegstein, 2012). These studies have shown that using meaningful gestures to convey instructional information can significantly improve students' learning experience and performance. Furthermore, considerable research has focused on students' self-comprehension through gestures, with the results indicating that students' usage of gestures can help reduce their cognitive load and benefit their comprehension of verbal information (Dargue et al., 2019; Goldin-Meadow, 2000, 2003; Goldin-Meadow et al., 2001). Yet, only a limited amount of research has been performed on the coherence of students' comprehension of the lecturers' gestures with the original meanings of the lecturers' meaningful gestures.

This study provides a gesture-meaning matching model from two perspectives, including lecturers and students, to understand the process of gesture expression from production to reception. Our research employed a qualitative approach to develop triangulation research through observations, interviews, and field notes to understand nonverbal communication comprehension between lecturers and students.

### **1.1 Lecturers' gestures**

Bodily movements that convey meaning have the potential to be highly effective in teaching contexts. Research into gestures suggests that iconic and metaphoric gestures are particularly useful in presenting meaningful nonverbal information (McNeill, 2011). In addition to verbal communication, pedagogical agents utilize nonverbal language to assist their students in the learning process. Lecturers' nonverbal language, educational materials, and other factors all contribute to classroom instruction and knowledge transfer (Kress, 2009; Lim, 2020; Lim et al., 2021). During lesson delivery, teachers unconsciously use gestures to help explain information and facilitate comprehension.

In particular, teachers heavily rely on cohesive and beat gestures (McNeill, 2011) to refine their language and alleviate cognitive load (Cook et al., 2008; Goldin-Meadow, 2003). However, while gestures can potentially convey social participation, they often lack meaningful information (Kress, 2009; Li & Jeong, 2020). To achieve more meaningful learning outcomes, students require a greater amount of intentional visual cues that teachers purposefully convey. Evaluating natural and meaningful gestures made by human lecturers is expected to positively impact students' embodied learning (Xu et al., 2022).

Through video observation and analysis, this study aims to identify the morphology of meaningful nonverbal movements among lecturers during teaching. The study applies multimodal discourse analysis (Lim, 2011; O'Halloran, 2004) to extract information from these meaningful gestures. The meaningful gestures are classified into two categories: common gestures and specific gestures.

Common gestures refer to lecturers' extensive use of meaningful gestures, which correspond with spoken sentences using common language to describe the teaching process. Specific gestures involve meaningful bodily movement that is infrequently used, corresponding with spoken language that explicitly details specific concepts not often mentioned in lectures. Finally, this study examines the effects of such gestures on students' comprehension within different teaching content contexts, thereby evaluating the impact of these meaningful movements on learning outcomes.

## **1.2 Students' perceptions**

Intrinsic and extrinsic factors influence students' perception of learning in class (Cachia et al., 2018; Ferrer et al., 2020). The physical teaching environment also affects students' perception of the course format (Anagnostopoulos et al., 2005; Beckers et al., 2016; Tondeur et al., 2015). This study primarily focuses on the teaching environment in the form of lectures, which involves one-way information dissemination. Students' intrinsic factors are particularly relevant to the context of this research, as the learning modality involves limited interaction with lecturers. Numerous studies have shown that students' perceptions of the learning process are significantly related to personal factors (Ahmed et al., 2018; Amir et al., 2020; Martin et al., 2018; Scheer et al., 2019).

To investigate the relationship between students' self-comprehension cognition and the effectiveness of in-class gestures on learning outcomes, this study extends its analysis to include various factors that may influence students' interpretation of these gestures. Factors such as course evaluation, prior learning experiences, learning habits, visual focus, cognitive load, and perceptions of the teaching environment were considered as primary research directions. It is critical to note that students' recognition and comprehension of gestures may be influenced by their learning style, visual attention, and experiences in prior learning contexts.

Semi-structured interviews were conducted to capture student perceptions of these factors, enabling the researchers to gather rich and nuanced student data. Importantly, since this observational study was conducted within authentic classroom settings, it was deemed necessary to integrate information on students' personal and environmental conditions into the interview protocol design. By doing so, the study ensures that contextual factors are suitably considered and incorporated into the analysis.

## **2 Method**

This study adopted a qualitative case study approach, employing gestural analysis (McNeill, 2011) and content analysis (Neuendorf, 2017) to understand lecturers' gestures and students' perceptions of them. Multimodal discourse analysis (O'Halloran, 2004) was utilized to comprehend video content. Field notes were taken by a researcher present in the classes to contextualize and triangulate the observations made from the video analysis, providing an alternative viewpoint on the lecturers' gestures that complemented the interpretations made in the subsequent student interviews. During the multimodal discourse analysis, the researchers familiarized themselves with the data by reviewing the video recordings and listening to the interviews multiple times.

### **2.1 Subject and research site**

This study investigated four paired cases in real-life teaching activities at the University, focusing on the natural behaviors of lecturers in a genuine teaching situation. The participants included three male lecturers (Allen, Brian, and Carl) and one female lecturer (Della), all with approximately ten years of

teaching experience at the undergraduate level and specialized in a design discipline. Additionally, one male student (Willis) and three female students (Tracy, Una, and Viola), all first-year undergraduate students in the design discipline, were recruited for the study. The four lecturers conducted four design-related lectures, and the corresponding students attended these lectures as part of their regular academic requirements.

## **2.2 Procedures**

We informed the lecturers and students about our research, and before the teaching sessions began, the teaching process was video recorded. A camera was positioned approximately three meters away from the lecturers. During the class, one of our researchers sat to observe the lecturers' and students' actions and consistently monitored the camera without interrupting the regular teaching process. We observed and recorded four one-hour-long lectures. Within a week after each lecture, we conducted approximately one-hour-long interviews with each student in a separate room. The interview processes were video recorded.

## **2.3 Materials**

The data analyzed in this empirical study consists of lecture observations and interviews with both the lecturer and student participants during real-life educational activities. Approximately 60 minutes of observation was conducted for each lecture, where we closely observed the lecturers' and students' speech, non-verbal movements, and teaching process. The lecturers' verbal language was transcribed from video recordings using an online transcription service and manually reviewed for accuracy. From the video recordings, we identified three typical meaningful gestures and two specific gestures from each lecturer participant's gestural movements, based on their verbal language and the teaching context.

During the interviews, we asked questions to the corresponding students about the three typical meaningful gestures and two specific gestures from the video recordings. Employing a semi-structured interview method, we designed the questions for the students into five categories, including inquiring about their learning behaviors and habits, memories of the observed lectures, interpretations of the lecturers' gestures, gestural performance, and gesture categorizations.

# **3 Data and analysis**

## **3.1 Tracy's understanding of Allen's gestures**

### **3.1.1 Allen's gestures**

In Allen's lecture, he primarily introduced the methods of collecting information from target users to facilitate communication between designers and customers. Allen gestured frequently and made changes to his gestures throughout the lecture. He walked around the classroom, focusing mainly on interactions between students and instructional materials. Figure 1 displays a matching model of the three common gestures and the two specific gestures extracted from Allen's 1-hour lecture video recording, along with Tracy's interpretation.

The first common gesture was shown through Allen drawing a vertical rectangle in the air with his index finger (upper left) when he said, "we will learn to design wireframe" the second common gesture (upper middle) was demonstrated that Allen raised his left hand and moved five fingers slightly together in front of his chest when he stated, "when you research on the activity, do not only focus

on one thing” the third common gesture (upper right) was stroked that Allen stretched out his left palm to move around in the air when he mentioned “form different clusters in the process” The two specific gestures one (lower left) that Allen’s left hand raised to his chest, and slightly crossed his index finger and thumb together when he spoke, “you want something more expensive or cheaper, cheaper right? ”, another one (lower right) appeared through Allen clenched his hands in front of his chest, and moved slightly up and down, when he said, “imaging you are driving a car.”

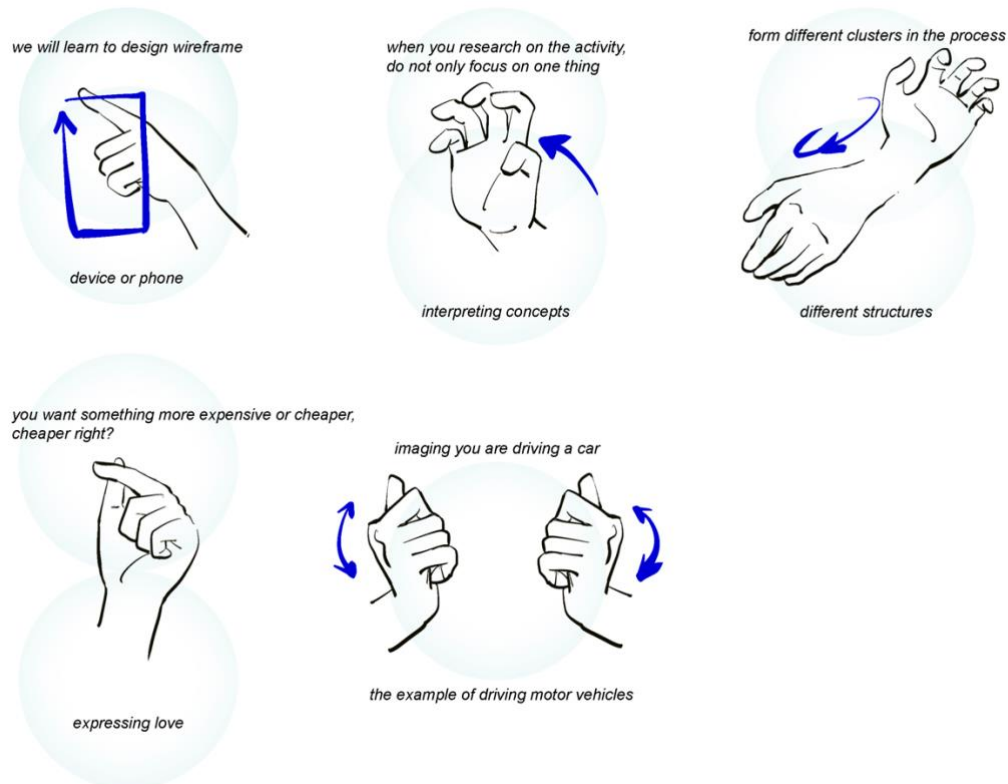


Figure 1. A matching model between Allen’s three meaningful gestures, two specific gestures, and Tracy’s interpretation. There are two circles behind each gesture. The upper one is Allen’s oral language when the gesture is expressed, and the lower is Tracy’s interpretation. The degree of proximity is expressed by the positional relationship between these two circles.

### 3.1.2 Tracy’s understanding

According to Tracy’s interview data, she believes she is an independent and eager-to-learn student. Tracy evaluates her learning behaviors as active and self-controllable. During lectures, she asserts that she can focus her attention on the lecturers and teaching materials effectively. Tracy also believes that lecturers’ gestures can significantly aid in students’ deeper understanding (e.g., gestures that illustrate the shapes of objects). She emphasizes that lecturers use gestures more frequently when verbalizing adjectives and creating cohesion in their speech. When asked about the gestures performed by Allen that she can remember, Tracy recalls some general gestures like stretching her arms flat and drawing circles, but she finds it difficult to describe the meanings behind them.

Then we presented Allen’s video clips without sound to help Tracy recall her memories. She could immediately explain the three common gestures, and the two specific gestures are device or phone, interpreting concepts, different structures, expressing love, and the example of driving motor vehicles,

respectively. There has been a slight discrepancy in Tracy's interpretations of Allen's meaningful gestures and specific gestures. However, Tracy's explanations on general gestures, like meaningful gestures one and three that are frequently applied during the subject, are assumed that Tracy's understanding of these gestures is enhanced due to the course content and Allen's repetition of the gestures. As for the deviation in the understanding of specific gestures, we suspect it is related to her previous experience due to the influence of popular culture.

After performing gesture tasks, which involved using hands to convey the meanings of importance, introducing concepts, interacting with students, and making conclusions, Tracy demonstrated the meaning of importance by pointing her index fingers up and down. However, she found it challenging to imagine any specific gestures to represent the meanings of introducing concepts, interacting with students, and making conclusions. Tracy considered that gestures could be categorized into three functions: guiding, describing, and emphasizing. To illustrate guiding and emphasizing functions, she used her index fingers to point and move vertically in the air. When describing a particular operation, she raised both hands together and formed a sphere in front of her chest.

## **3.2 Una's understanding of Brian's gestures**

### **3.2.1 Brian's gestures**

Brian's lecture focused on the history and characteristics of typography. Brian made fewer gestures and sat more steadily; he taught online and spent much of his lecture interacting with instructional materials, with very little interaction with students. Figure 2 shows a matching model on the three common gestures and the two specific gestures taken from Brian's one-hour lecture video footage and Una's interpretation.

The first common gesture was exhibited through Brian raising two hands two together and forming a sphere (upper left) when he said, "*What is typography*" the second common gesture (upper middle) was exposed when Brian's two hands together and moved from one place to another when he said, "*you put on the outside of something that's on the inside*" the third common gesture (upper right) was performed that Brian spread his hands forward in front of his chest when he mentioned "*we as designers, we have to be able to change that shape*" The two specific gestures one (lower left) that Brian stretched his index finger and thumb to make a small square when he said, "*a stone or a wall that you can build up with the slack,*" another one (lower right) is shown through Brian's hands were close. His fingers were clasped together, and he moves his hands horizontally when he said, "*typefaces that are designed with those little legs.*"

### **3.2.2 Una's understanding**

Una's interview data revealed that she seeks an active learning environment with support from teachers and peers. Although she enjoys an interactive atmosphere, she feels apprehensive about actively participating in class. Una describes her learning behaviors as eager to learn and receptive to the teacher's instructions. During lectures, she can maintain focus on the lecturers and instructional materials, but her attentiveness can be impacted by the classroom surroundings. Una prefers offline classes over online ones, and she frequently emphasizes the importance of the learning atmosphere. Una finds that live demonstrations by the teacher and storytelling help her better understand the learning content. When asked about the gestures made by Brian, she could remember some aspects. While she couldn't recognize any specific gestures that the lecturer used, she recalled that he moved the mouse, which led her to guess that he might have been drawing circles on the desk.

Una found it challenging to understand the three common gestures and the two specific gestures even though we showed her Brian's video clips without sound. She interpreted the three common gestures as a system, a concept different from others, and a step-by-step process, respectively. She scarcely understood the two specific gestures. Una thought she could grasp the topic of Brian's speech, but she found it challenging to comprehend fully. The result shows that the meanings of Brian's gestures according to the class context significantly differ from the students' understanding. However, Una's interpretations of three meaningful gestures are relatively close to the original meanings, albeit imperfect. Her instructional environment influences Una's knowledge of gestures, yet she can still discover comparable explanations for regularly used gestures. Regarding her skewed perception of specific gestures, we believe she lacks a strong recall of the teaching content and pays little attention to Brian throughout the lecture.

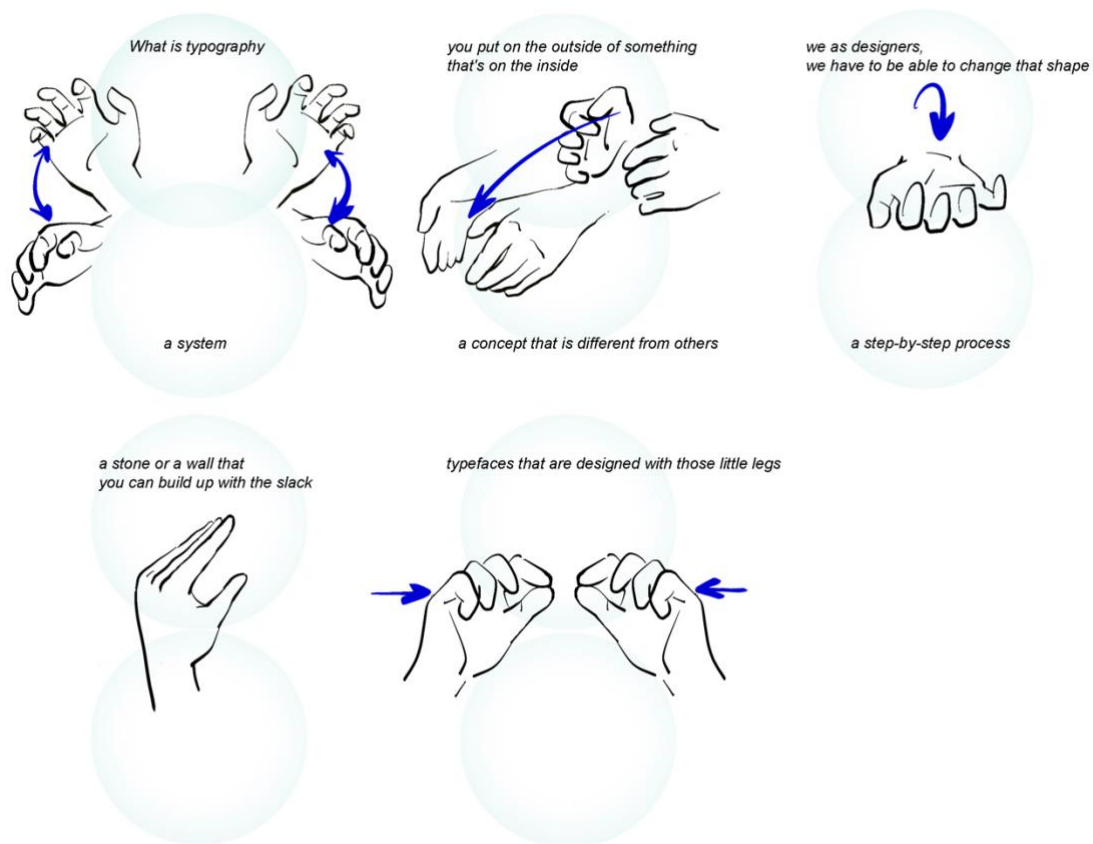


Figure 2. There are two circles behind each gesture. The upper one is Brian's oral language when the gesture is expressed, and the lower one is Una's interpretation. The degree of proximity is defined by the positional relationship between these two circles.

During the gesture performance tasks, where Una used her hands to convey the meanings of importance, introducing concepts, interacting with students, and making conclusions, she pointed her index fingers up and down to demonstrate the meaning of importance. She explained that these gestures indicated telling and repetitive movements. To represent the meanings of introducing concepts, she clenched one hand into a fist and hammered the other to mark the beginning or end of a chapter. Una believed that lecturers prefer to walk around without using many gestures when interacting with students, and they often put their hands down to make conclusions. Una classified gestures into three categories: processing, describing, and spatial functions. As she explained the

classification, she raised her hands and made different shapes in the air. It appeared that she was particularly focused on the connection between objects and space, leading her to use more spatial-related motions while viewing and executing gestures. However, Una's attention to Brian's gestures was greatly diminished as they were small and limited to the screen, resulting in mismatches in the information she perceived.

### 3.3 Viola's understanding of Carl's gestures

#### 3.3.1 Carl's gestures

Carl's lecture focused on the principles of typography and layout design for visual communication. Throughout his speech, Carl employed several different gestures and utilized the space effectively by walking around. Video clips of three meaningful gestures and two specific gestures were extracted from Carl's one-hour lecture video recording, and Viola's interpretation illustrated a matching model shown in Figure 3.

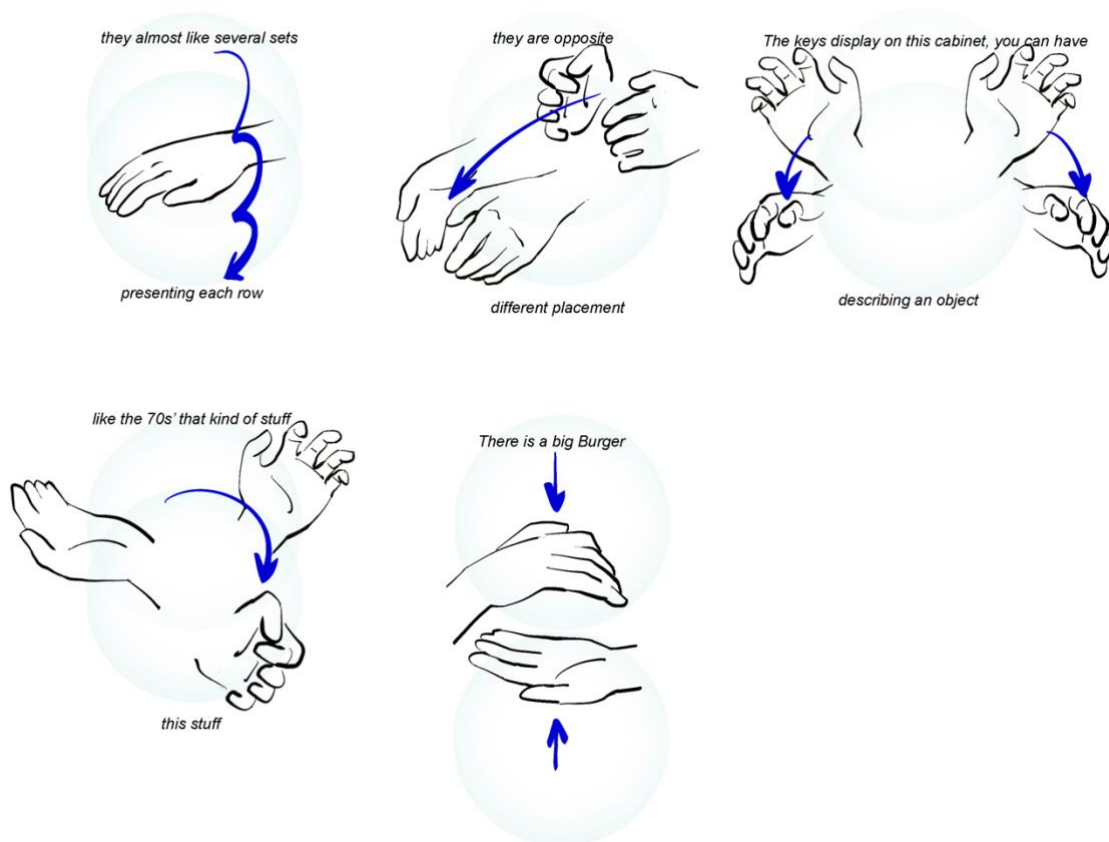


Figure 3. A matching model between Carl's three meaningful gestures and two specific gestures and Viola's interpretation. There are two circles behind each gesture. The upper one is Carl's oral language when the gesture is expressed, and the lower one is Viola's interpretation. The degree of proximity is expressed by the positional relationship between these two circles.

Carl demonstrated various common gestures during his speech. He moved his palms from top to bottom, remarking, "they almost like several sets," indicating the first common gesture (upper left). The second common one, Carl put his hands together and moved them from one place to another, mentioning "they are opposite," which signified the upper-middle common gesture. The third common gesture, Carl raised his hands and formed a huge sphere in front of his chest while discussing



the cabinet keys(upper right). He performed two specific gestures, one being shaking his hands in front of his chest while stating, "like the 70s' that kind of stuff" (lower left). Another one (lower right) appeared through Carl performing the movement of his palms vertically together when he mentioned, "There is a big Burger."

### 3.3.2 Viola's understanding

Viola's interview data revealed that she paid more attention to the slides than to the lecturers because she believed it would prevent distractions during the lecture. She described her learning habits as being curious and practicing repeatedly. Viola emphasized that the design of instructional tactics and the visual information on the slides were crucial for her understanding of the course content. While she could remember pointing gestures that highlighted important knowledge, she also recalled certain specific gestures, but she had difficulty understanding the majority of them. According to Viola, Carl tended to move around during the lecture but rarely made gestures. She felt that specific gestures, like the example of the burger gesture, would help her maintain focus and concentration.

Even when we showed Viola Carl's video clips without sound, she struggled to understand the three common gestures and the two specific gestures. She interpreted the three common gestures and the first specific gesture as presenting each row, different placement, describing an object, and this stuff, respectively. Her understanding of the last specific gesture was inadequate. Viola worried she hadn't paid enough attention to the lecturer's gestures and was unsure of her comprehension. Yet, the results demonstrate that her interpretations are closely connected to the original meanings of Carl's gestures. Viola is clearly aware of the general meaningful gestures, but she didn't intentionally memorize them. And she couldn't give a precise meaning for specific gestures when she wasn't provided context.

During the gesture performance tasks, which involved using hands to convey the meanings of importance, introducing concepts, interacting with students, and making conclusions, Viola demonstrated her understanding by opening her palms to represent introducing concepts and using crosscutting movements from top to bottom with her palms to make conclusions. She also pointed her index fingers up and down to show the meanings of importance. However, she felt that no gestures could effectively describe interacting with students. Viola believed that gestures are not easily categorizable and are influenced by individual habits. Despite Viola's unique perspective on gestures, we assume that she can correctly comprehend Carl's general meaningful gestures. Additionally, the level of understanding of gestures and the co-occurrence of contextual verbal language and gestures are interconnected.

## 3.4 Willis's understanding of Della's gestures

### 3.4.1 Della's gestures

Della's lecture focused on interface design and layout design classification basics. Throughout her lecture, she made minor gestures and remained in a steady position in front of the students, near her teaching materials. Figure 4 illustrates a matching model of the three common gestures and the two specific gestures extracted from the video clips of Della's one-hour lecture video recording, along with Willis' interpretation.

The first common gesture (upper left) was made through Della's left palm moving vertically in the air from top to lower when she uttered, "*type two or type three*" the second common gesture (upper

middle) was shown that Della’s raised left hand and moved five fingers slightly together in front of her chest when she said, “*you know this number*” the third common gesture (upper right) was carried out that Della’s left-hand swung sharply in the air when she started with a new sentence “*All of those things*” The two specific gestures one (lower left) that Della’s palms were lifted vertically when she spoke, “*desktop, tablet, and mobile,*” another one (lower right) appeared through Della stretched out her palm and moved her five fingers upwards slightly, when she said, “*the downloaded data and living in the device.*”

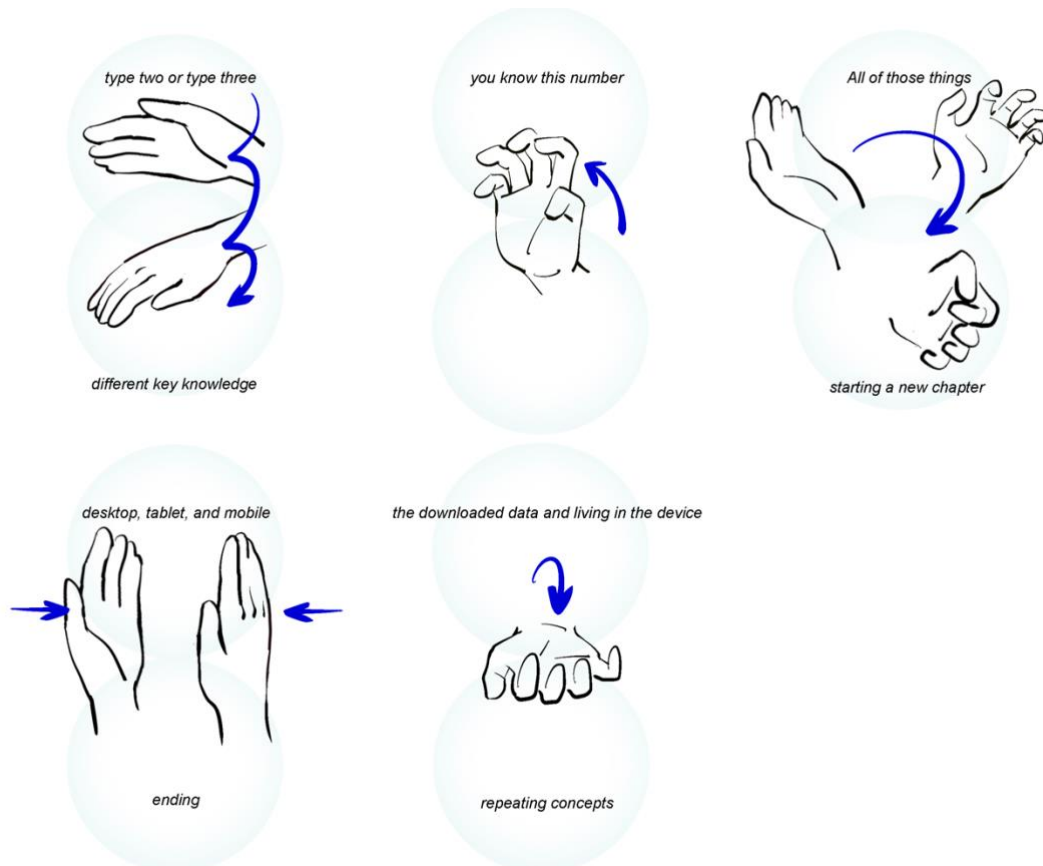


Figure 4. A matching model between Della’s three meaningful gestures, two specific gestures, and Willis’s interpretation. There are two circles behind each gesture. The upper one is Della’s oral language when the gesture is expressed, and the lower one is Willis’s interpretation. The degree of proximity is expressed by the positional relationship between these two circles.

### 3.4.2 Willis’s understanding

Willis' interview data indicated that he could recall Della employing gestures for interface navigation since gestures were a regular part of her teaching. He described his study habits as paying close attention to what the teacher said and seeking additional information online to enhance his understanding. Willis felt that having multiple perspectives to explain the same concept helped him grasp the knowledge better. Therefore, he believed that the teacher's voice and tone, the information on the slides, and the availability of devices for self-acquisition of information were beneficial for his learning. Willis also observed that due to the demands of the course, Della would occasionally hold her phone during class and produce gestures based on it.

When we showed him Della's video clips without sound, Willis confidently predicted the three common and the two specific gestures. The first and third meaningful gestures, as well as the two specific gestures, were regarded by him as different fundamental knowledge, starting a new chapter, ending, and repeating concepts, respectively. He reported that it was hard to imagine the second gesture. Willis' interpretations of the gestures differed significantly from the gestures' original meanings. Even if the substances are similar, there is no significant overlap. We considered Della's repertoire of gestures was limited, and she made no body movements or rotations that may acutely impact the level of Willis's understanding of gestures.

During the gesture performance tasks, which included using hands to express important meanings, introducing concepts, interacting with students, and making conclusions, Willis frequently used handwaving to indicate interaction and summarization. He found that pointing with his index finger was an effective way to accentuate and introduce information. Additionally, Willis believed that drawing certain shapes in the air with his hands could enhance the description of an object. According to Willis, gestures can be categorized into two functions: pointing and description. We assume that Willis is familiar with Della's gestures to some extent. While he can provide a general description of the gestures, he lacks detailed explanations due to the absence of a visualizable scenario.

#### **4 Discussion**

Tracy can accurately interpret Allen's gestures, possibly because he uses a wide range of gestures and maintains a facing position towards the pupils. However, her interpretation of gestures could be influenced by contemporary culture, leading to mismatches or miscommunications. For instance, she misread the first specific gesture and perceived it as a stroke to draw a heart shape. Una misunderstands several of Brian's meaningful gestures, possibly due to the limitations of cyberspace and minimal teacher-student interaction. Although she has a better grasp of Brian's general gestures, her awareness of gestures diminishes when Brian uses specific gestures to indicate concrete things. We anticipate that explicit and participatory gestures could assist students in comprehending information, but if specific gestures are not clearly articulated, it is difficult to establish a consensus between lecturers and students. Viola has a strong sense of identification with Carl's gestural meaning and can relatively accurately repeat the teaching content conveyed by Carl's gestures. This could be because Carl's gestures are more defined and rhythmic. He does not overuse gestures, but he employs their magnitude to convey information effectively. On the other hand, Willis can roughly describe the meaning of Della's gestures without providing detailed statements. It is possible that Della's gestures may not be as evident or as rich as they could be, and this might not be Willis' primary concern.

Based on the recognition and classification of gestures, students' understanding of instructional gestures implies that gestures with guiding and descriptive roles can substantially support learning. The findings from human lecturers' gestures and students' perceptions of them indicate three gesture-meaning matching patterns. First, the students' interpretation of the lecturer's common meaningful gestures and the meaning indicated by the gestures themselves are essentially equivalent, with slight deviations due to the degree and range of the gesture. Second, the use of specific gestures in the classroom is accurately understood by students when provided with context. Additionally, students' personal experiences and popular culture play a crucial role in their perceptions of teacher gestures. In designing gestural PAs, we consider the student's perspective. The lecturer's expression of

meaningful gestures can reach maximum consensus with the students under conditions of clear visibility and rhythm. We propose that the types, scale, and rhythm of the design agent's gestures and body orientation have considerable potential to enhance the effectiveness of embodied PAs.

## 5 Conclusion

Our study conducted a qualitative analysis of four paired cases to understand how students interpret lecture gestures in real-life classrooms, resulting in a gesture-meaning matching model. The research involved gestural and content analysis, along with multimodal discourse analysis, to comprehend meaningful gestures. The study's unique aspect is the evaluation of students' cognition of lecturers' gestures by assessing their comprehension and interpretation of gesture meanings, an area seldom explored by previous studies. However, the preliminary results and limited sample size, focusing on one university, may restrict generalizability. To provide a broader perspective on gesture comprehension, data from diverse educational settings should be considered. As our research continues, we are currently investigating more paired lectures to develop a gesture-word matching model design. The analysis results highlight the importance of multimodal interaction in higher education discourse, learning perception, and PA and virtual learning design studies.

## References

- Ahmed, Y., Taha, M. H., Al-Neel, S., & Gaffar, A. M. (2018). Students' perception of the learning environment and its relation to their study year and performance in Sudan. *International Journal of Medical Education, 9*, 145–150. <https://doi.org/10.5116/ijme.5af0.1fee>
- Amir, L. R., Tanti, I., Maharani, D. A., Wimardhani, Y. S., Julia, V., Sulijaya, B., & Puspitawati, R. (2020). Student perspective of classroom and distance learning during COVID-19 pandemic in the undergraduate dental study program Universitas Indonesia. *BMC Medical Education, 20*(1), 392. <https://doi.org/10.1186/s12909-020-02312-0>
- Anagnostopoulos, D., Basmadjian, K., & Mccrory, R. (2005). The decentered teacher and the construction of social space in the virtual classroom. *Teachers College Record, 107*(8), 1699–1729.
- Beckers, R., van der Voordt, T., & Dewulf, G. (2016). Learning space preferences of higher education students. *Building and Environment, 104*, 243–252. <https://doi.org/10.1016/j.buildenv.2016.05.013>
- Beege, M., Ninaus, M., Schneider, S., Nebel, S., Schlemmel, J., Weidenmüller, J., Moeller, K., & Rey, G. D. (2020). Investigating the effects of beat and deictic gestures of a lecturer in educational videos. *Computers & Education, 156*, 103955. <https://doi.org/10.1016/j.compedu.2020.103955>
- Cachia, M., Lynam, S., & Stock, R. (2018). Academic success: Is it just about the grades? *Higher Education Pedagogies, 3*(1), 434–439. <https://doi.org/10.1080/23752696.2018.1462096>
- Castro-Alonso, J. C., Wong, R. M., Adesope, O. O., & Paas, F. (2021). Effectiveness of Multimedia Pedagogical Agents Predicted by Diverse Theories: A Meta-Analysis. *Educational Psychology Review, 33*(3), 989–1015. <https://doi.org/10.1007/s10648-020-09587-1>
- Chen, C.-M., & Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education, 80*, 108–121. <https://doi.org/10.1016/j.compedu.2014.08.015>
- Cook, S., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*. <https://doi.org/10.1016/j.cognition.2007.04.010>
- Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin, 145*(8), 765–784. <https://doi.org/10.1037/bul0000202>
- Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: A meta-analysis. *Educational Research Review, 24*, 193–209. <https://doi.org/10.1016/j.edurev.2018.05.002>

- Ferrer, J., Ringer, A., Saville, K., A Parris, M., & Kashi, K. (2020). Students' motivation and engagement in higher education: The importance of attitude to online learning. *Higher Education*.  
<https://doi.org/10.1007/s10734-020-00657-5>
- Fitton, I. S., Finnegan, D. J., & Proulx, M. J. (2020). Immersive virtual environments and embodied agents for e-learning applications. *PeerJ Computer Science*, 6, e315. <https://doi.org/10.7717/peerj-cs.315>
- Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences*, 3(11), 419–429. [https://doi.org/10.1016/S1364-6613\(99\)01397-2](https://doi.org/10.1016/S1364-6613(99)01397-2)
- Goldin-Meadow, S. (2000). Beyond Words: The Importance of Gesture to Researchers and Learners. *Child Development*, 71(1), 231–239. <https://doi.org/10.1111/1467-8624.00138>
- Goldin-Meadow, S. (2003). Hearing Gesture: How Our Hands Help Us Think. In *Bibliovault OAI Repository, the University of Chicago Press* (Vol. 14). <https://doi.org/10.2307/j.ctv1w9m9ds>
- Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining Math: Gesturing Lightens the Load. *Psychological Science*, 12(6), 516–522. <https://doi.org/10.1111/1467-9280.00395>
- Grivokostopoulou, F., Kovas, K., & Perikos, I. (2020). The Effectiveness of Embodied Pedagogical Agents and Their Impact on Students Learning in Virtual Worlds. *Applied Sciences*, 10(5), Article 5.  
<https://doi.org/10.3390/app10051739>
- Kantharaju, R. B., 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>
- Kress, G. (2009). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge.  
<https://doi.org/10.4324/9780203970034>
- Li, P., & Jeong, H. (2020). The social brain of language: Grounding second language learning in social interaction. *Npj Science of Learning*, 5(1), Article 1. <https://doi.org/10.1038/s41539-020-0068-7>
- Li, W., Wang, F., Mayer, R. E., & Liu, H. (2019). Getting the point: Which kinds of gestures by pedagogical agents improve multimedia learning? *Journal of Educational Psychology*, 111(8), 1382–1395.  
<https://doi.org/10.1037/edu0000352>
- Li, W., Wang, F., Mayer, R. E., & Liu, T. (n.d.). Animated pedagogical agents enhance learning outcomes and brain activity during learning. *Journal of Computer Assisted Learning*, n/a(n/a).  
<https://doi.org/10.1111/jcal.12634>
- Lim, F. (2011). *A systemic functional multimodal discourse analysis approach to pedagogic discourse*.
- Lim, F. V. (2020). *Designing Learning with Embodied Teaching: Perspectives from Multimodality*. Routledge.
- Lim, F. V., Towndrow, P. A., & Min Tan, J. (2021). Unpacking the Teachers' Multimodal Pedagogies in the Primary English Language Classroom in Singapore. *RELC Journal*, 00336882211011783.  
<https://doi.org/10.1177/00336882211011783>
- Macedonia, M., & Kriegstein, K. (2012). Gestures Enhance Foreign Language Learning. *Biolinguistics*, 64, 393–416.
- Martin, F., Wang, C., & Sadaf, A. (2018). Student perception of helpfulness of facilitation strategies that enhance lecturer presence, connectedness, engagement and learning in online courses. *The Internet and Higher Education*, 37, 52–65. <https://doi.org/10.1016/j.iheduc.2018.01.003>
- McNeill, D. (2011). Hand and Mind. In *Hand and Mind* (pp. 351–374). De Gruyter Mouton.  
<https://doi.org/10.1515/9783110874259.351>
- Natarajan, M., & Gombolay, M. (2020). Effects of Anthropomorphism and Accountability on Trust in Human Robot Interaction. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 33–42). Association for Computing Machinery.  
<https://doi.org/10.1145/3319502.3374839>
- Neuendorf, K. A. (2017). *The Content Analysis Guidebook*. SAGE.
- Noetel, M., Griffith, S., Delaney, O., Harris, N. R., Sanders, T., Parker, P., del Pozo Cruz, B., & Lonsdale, C. (2021). Multimedia Design for Learning: An Overview of Reviews With Meta-Meta-Analysis. *Review of Educational Research*, 00346543211052329. <https://doi.org/10.3102/00346543211052329>
- O'Halloran, K. (2004). *Multimodal Discourse Analysis: Systemic Functional Perspectives*. A&C Black.
- Petersen, G. B., Mottelson, A., & Makransky, G. (2021). Pedagogical Agents in Educational VR: An in the Wild Study. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445760>
- Ratcliffe, J., & Tokarchuk, L. (2020). Presence, Embodied Interaction and Motivation: Distinct Learning Phenomena in an Immersive Virtual Environment. In *Proceedings of the 28th ACM International*

- Conference on Multimedia* (pp. 3661–3668). Association for Computing Machinery.  
<https://doi.org/10.1145/3394171.3413520>
- Schroeder, N. L., Chiou, E. K., & Craig, S. D. (2021). Trust influences perceptions of virtual humans, but not necessarily learning. *Computers & Education*, *160*, 104039.  
<https://doi.org/10.1016/j.compedu.2020.104039>
- Song, Y., & Luximon, Y. (2020). Trust in AI Agent: A Systematic Review of Facial Anthropomorphic Trustworthiness for Social Robot Design. *Sensors*, *20*(18), Article 18. <https://doi.org/10.3390/s20185087>
- Tondeur, J., De Bruyne, E., Van Den Driessche, M., McKenney, S., & Zandvliet, D. (2015). The physical placement of classroom technology and its influences on educational practices. *Cambridge Journal of Education*, *45*(4), 537–556.
- van der Scheer, E. A., Bijlsma, H. J. E., & Glas, C. A. W. (2019). Validity and reliability of student perceptions of teaching quality in primary education. *School Effectiveness and School Improvement*, *30*(1), 30–50.  
<https://doi.org/10.1080/09243453.2018.1539015>
- Wang, F., Li, W., Mayer, R. E., & Liu, H. (2018). Animated pedagogical agents as aids in multimedia learning: Effects on eye-fixations during learning and learning outcomes. *Journal of Educational Psychology*, *110*(2), 250. <https://doi.org/10.1037/edu0000221>
- Wu, Q., Wu, C.-J., Zhu, Y., & Joo, J. (2021). Communicative Learning with Natural Gestures for Embodied Navigation Agents with Human-in-the-Scene. *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4095–4102. <https://doi.org/10.1109/IROS51168.2021.9636208>
- Wu, Y. C., & Coulson, S. (2005). Meaningful gestures: Electrophysiological indices of iconic gesture comprehension. *Psychophysiology*, *42*(6), 654–667. <https://doi.org/10.1111/j.1469-8986.2005.00356.x>
- Xu, X., Kang, J., & Yan, L. (2022). Understanding embodied immersion in technology-enabled embodied learning environments. *Journal of Computer Assisted Learning*, *38*(1), 103–119.  
<https://doi.org/10.1111/jcal.12594>

#### **About the Authors:**

**Lai Wei:** A PhD candidate in the School of Design at The Hong Kong Polytechnic University. Her current research focuses on pedagogical agent design, gesture design, virtual learning design, and multimodal interaction. [thefuture.wei@connect.polyu.hk](mailto:thefuture.wei@connect.polyu.hk).

**Kenny K. N. Chow:** An Associate Professor in the School of Design at The Hong Kong Polytechnic University. His current research focuses on digital media, interaction design, cognitive science, and media studies. [sdknchow@polyu.edu.hk](mailto:sdknchow@polyu.edu.hk).

**Acknowledgement:** We acknowledge the University Grants Committee of Hong Kong for funding this research.