Pointing Gestures as an External Compensatory Strategy: Mechanisms Underlying Their Role in Mitigating the Negative Effects of Virtual Teachers' Smiling Expressions

Xinrui Song*

School of Design and Architecture, Zhejiang University of Technology, Hangzhou, Zhejiang, 310023, China

*songsxr@163.com

Abstract. While a virtual teacher's smiling expression can convey positive emotions, its lack of realism may lead to cognitive dissonance and social discomfort. This study explores an external compensation strategy, examining the mitigating effects of indicative gestures. The experiment employed a 2 (Gesture Direction: Slide/Student) × 2 (Facial Expression: Smile/Neutral) within-subjects design with 52 college students. Results indicate that smiling expression significantly attracted learners' visual attention to the teacher's face, but did not significantly reduce focus on instructional content nor directly impair performance. Gesture direction exhibited clear functional differentiation: gestures pointing at slides served cognitive guidance, while those pointing at students oriented toward social interaction. When the virtual teacher displayed a smiling expression, gestures pointing at slides redirected learners' attention from the face back to the content and enhanced intrinsic motivation. These findings suggest that achieving overall optimization through functional configuration of multiple cues is more effective than simply pursuing high realism.

Keywords: Virtual Teacher, Nonverbal behavior, Pointing gestures, facial expressions, Eye tracking

1 Introduction

Virtual teachers offer great potential for personalized learning. Current research mainly follows a "feature transfer" paradigm, deconstructing real teachers' behaviors [1]. However, not all human traits transfer directly. While real teachers' smiles enhance motivation, virtual teachers' smiles often appear insincere due to their formulaic and frequent nature, leading to two negative effects: cognitive distraction and social dissonance [2,3].

© The Author(s) 2026
S. Shamsudin et al. (eds.), *Proceedings of the 4th International Conference on Language and Cultural Communication (ICLCC 2026)*, Advances in Social Science, Education and Humanities Research 1025,
https://doi.org/10.2991/978-2-38476-597-3_56

1.1 The Problem: Dual Negative Effects of Virtual Teachers' Smiles

Virtual learning environments reshape facial expression functions. Unlike physical classrooms where spatial distance reduces attention to teachers' expressions, virtual settings present a relatively static on-screen image, making expressions the primary emotional cue. Media richness theory suggests that when other information forms diminish, the remaining information gains impact [4]. Thus, a virtual teacher's smile may carry greater weight—but this amplifies problems when the smile lacks authenticity. Negative effects arise in two ways. First, frequent, formulaic virtual smiles fail to meet students' standards for sincerity; students distinguish genuine from polite smiles, and perceived insincerity follows [5]. Second, as a salient social cue, smiles automatically attract visual attention, diverting focus from learning content [6]. Most existing studies pursue "internal correction" (e.g., enhancing smile realism through technology). Given technical constraints, we propose an "external compensation" strategy: introducing directive gestures to systematically offset these negative effects, rather than relying solely on costly smile optimization.

1.2 The Solution: Functional Differentiation of Pointing Gestures

Pointing gestures represent an effective teaching modality. These gestures can automatically guide students' gaze, help establish connections between information elements, and reinforce memory [7]. Critically, the target of the pointing gesture determines its primary function. When teachers point at slides, the gesture serves cognitive guidance—directing attention to learning content. When teachers point at students, the gesture primarily serves social interaction—conveying "I am communicating with you" [8].

1.3 The Mechanism: Functional Synergy Condition Framework

The interaction of multiple visual social cues can yield competition or synergy. Visual working memory is limited [9]; simultaneous smiles and gestures without clear functional differentiation compete for resources. Conversely, distinct functional roles aligned to a common instructional goal enable synergistic gains [10]. The direction of interaction depends on functional clarity and compensatory alignment [11]: ambiguity or overlap yields competition; complementary, deficit-targeting roles yield synergy. We propose a three-part functional synergy framework—functional differentiation, compensatory alignment, goal consistency. Applied here, smiling with slide-directed gestures satisfies all three (emotional vs. cognitive functions; gesture compensates for attentional drift; both serve content learning), predicting synergy. Smiling with student-directed gestures entails functional overlap and weak compensatory fit, predicting competition. This study thus tests whether directive gestures offset the costs of virtual teacher smiles across learning and perceptual outcomes.

1.4 Research Hypotheses

Drawing from Sections 1.1–1.3, we predict: H1a—smiling increases gaze to the teacher's face; H1b—smiling reduces gaze to slides; H1c—smiling impairs performance and motivation. H2a—slide-directed gestures increase content gaze; H2b—student-directed gestures increase social presence. Regarding interactions, H3a—smiling with slide pointing redirects attention to content (compensating attentional drift); H3b—this combination yields better outcomes than smiling with student pointing; H3c—under neutral expression, gesture type has negligible differential effects.

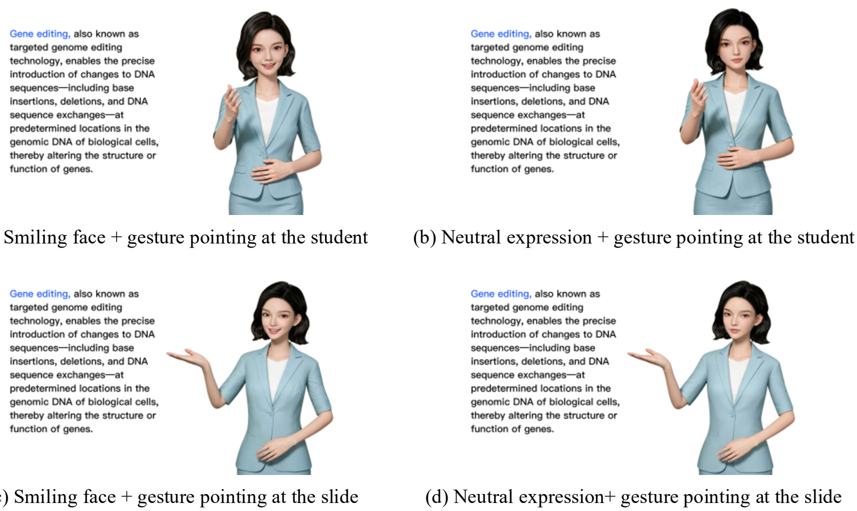
2 Method

2.1 Virtual Teacher Design

The virtual teacher was modeled in Poser and MetaHuman, animated with MetaHuman's presets, voiced by iFlytek TTS, and lip-synced via Audio2Face with a 4–6 frame offset.

2.2 Experimental Design

This study employed a 2 (facial expression: neutral vs. smiling) × 2 (gesture: pointing at slides vs. pointing at students) design. Dependent measures comprised academic performance (retention, transfer, intrinsic motivation, cognitive load) and eye-tracking indices of attentional focus (fixation duration and count on slides, teacher face, and hands). Four 10-minute video lectures were created, each containing 25 instances of the assigned gesture–expression combination (Figure 1).



(a) Smiling face + gesture pointing at the student (b) Neutral expression + gesture pointing at the student

(c) Smiling face + gesture pointing at the slide (d) Neutral expression+ gesture pointing at the slide

Fig. 1. Screenshot examples of each lecture

2.3 Participants

A priori power analysis required ≥52 participants. Fifty-seven students (27 male, 30 female) aged 19–25 (M = 22.04, SD = 2.00) completed a 2×2 within-subjects design: (a) smiling facial expression + pointing gesture toward the slide; (b) smiling expression + gesture pointing at the slide; (c) neutral expression + gesture pointing at the student; (d) neutral expression + gesture pointing at the student.

2.4 Equipment

Using the Tobii Pro Glasses 3 head-mounted eye tracker, with a sampling rate of 100 Hz, a 17-inch LCD monitor ,1920×1080 px, and a viewing distance of approximately 65 cm.

2.5 Experimental Procedure

Participants provide consent, wear Tobii Pro Glasses 3, complete calibration, and are randomly assigned. After a 10-min instructional video, eye-tracking data are saved and the device removed. Questionnaires, a 15–20-min post-test, and a one-on-one interview follow. Total duration: 40–45 min; withdrawal permitted anytime.

3 Results

To test the interactive effect of the teacher's gesture target and facial expression, we conducted a series of 2 (gesture target: pointing at a slide vs. pointing at a student) × 2 (facial expression: smiling vs. neutral) repeated measures analyses of variance . This analysis examined students' percentage of time spent looking at the slides and the virtual teacher's body, as well as their scores on learning retention and transfer, learning experience, intrinsic motivation, and cognitive load.

3.1 Attention allocation.

Repeated measures ANOVAs revealed: for slide gaze (duration and count proportions), only the main effect of gesture was significant. For facial gaze duration, both main effects and the interaction were significant; simple effects showed slide pointing reduced facial gaze only under smiling. For facial gaze count proportion, smiling and gesture main effects and the interaction were significant; simple effects showed student pointing increased facial gaze only under neutral expression. For hand gaze, no effects were significant for duration, but for count proportion, smiling and gesture main effects and the interaction were significant; simple effects showed student pointing increased hand gaze only under neutral expression.

3.2 Learning Outcomes and Subjective Perceptions

For retention scores, only gesture type showed a significant main effect; neither smiling expression nor the interaction reached significance. Transfer scores revealed no significant main or interaction effects. Regarding intrinsic motivation, both main effects and the interaction were significant, with simple effects analysis indicating that smiling expression mattered only when the gesture pointed at materials, not at the student. The learning experience measure yielded significant main effects for both smiling expression and pointing gesture but no interaction. Finally, pointing gestures significantly affected internal cognitive load, whereas smiling expression and the interaction did not.

4 Discussion

4.1 The Effects of Smiling Emoticons: Attention Capture and Limited Spillover

H1 received partial support. Smiling increased fixation duration and count on the teacher's face, confirming H1a and its attention-capturing effect. However, smiling did not reduce attention to slides (refuting H1b) nor impair academic performance (refuting H1c). Smiling's negative effects are not inevitable. Learners may maintain content attention via rapid switching or strategic back-tracking, and pointing gestures may have offset attentional shifts. Thus, the virtual teacher's smile is neither inherently harmful nor beneficial—its impact depends on the broader multimodal configuration.

4.2 Functional Differentiation of Gestural Pointing: Cognitive Guidance and Social Attention

H2a and H2b were generally supported. Slide-directed gestures increased gaze to slides while reducing looks to the instructor's face under smiling, and yielded positive main effects on performance and cognitive load—consistent with a cognitive guidance function. In contrast, student-directed gestures increased gaze to the instructor's face and hands, especially with a neutral expression, but did not direct attention toward instructional content, aligning with a social-cue enhancement rather than cognitive guidance account. This social attention shift did not, however, improve learning: student-directed gestures offered no advantage over slide-directed gestures for performance or motivation. The critical distinction lies in whether the attentional target serves the learning task. Slide-directed gestures align gaze with instructional goals; student-directed gestures keep attention on social cues, yielding a more indirect instructional effect.

4.3 Interaction Effects and Compensatory Mechanisms: Context-Dependent Functional Synergy

H3a and H3b received support, revealing context-dependent interaction patterns. At the attentional level, a significant interaction emerged for facial gaze duration: under smiling, slide-directed gestures reduced facial gaze compared to student-directed gestures;

under neutral expression, gesture type made no difference. This suggests slide-directed gestures compensate for smiling-induced attentional drift by redirecting gaze toward instructional content. For intrinsic motivation, the interaction was also significant: smiling enhanced motivation only when paired with slide-directed gestures, disappearing with student-directed gestures. Functional synergy thus appears contingent on role division—cognitive guidance from gesture combined with emotional signaling from expression—rather than overlapping social functions.

5 Conclusions

This study examined pointing gestures as external compensation for a virtual teacher's smiling expressions. Findings partially support the functional synergy framework. Smiling reliably captured attention but neither significantly reduced focus on instructional content nor directly impaired performance. Gesture function diverged clearly: slide-directed gestures guided cognition, whereas student-directed gestures activated social attention without learning gains. Under smiling, slide-directed gestures reduced excessive face gaze and yielded motivational synergy; however, this advantage remained largely confined to process measures, with no significant transfer to performance outcomes. Contributions include articulating a "functional differentiation–compensatory match–goal consistency" logic for multimodal design and showing that overall optimization hinges on cue configuration rather than isolated realism. Limitations comprise a single brief session, a restricted gesture/expression repertoire, and untested generalizability. Future work should extend training duration and diversify nonverbal behaviors.

References

1. MU X, HE J. Virtual Teacher-Aided Learning System Based on Voice Operated Character Animation[J/OL]. Applied Sciences, 2024, 14(18).
2. LIEW T W, ZIN N A M, SAHARI N, et al. The effects of a pedagogical agent's smiling expression on the learner's emotions and motivation in a virtual learning environment[J/OL]. The International Review of Research in Open and Distributed Learning, 2016, 17(5).
3. PI Z, LING H, LI X, et al. Instructors' pointing gestures and positive facial expressions hinder learning in video lectures: Insights from teachers and students in China[J/OL]. Education and Information Technologies, 2024, 29(16): 21115-21131.
4. DENNIS A R, KINNEY S T. Testing media richness theory in the new media: The effects of cues, feedback, and task equivocality[J/OL]. Information Systems Research, 1998, 9(3): 256-274.
5. ISLAM M Z, WANG G. Avatars in the educational metaverse[J/OL]. Visual Computing for Industry, Biomedicine, and Art, 2025, 8(1): 15.
6. IRFAN M, PATEL P, HASSAN B. From distraction to interaction: Investigating learner engagement challenges in virtual classrooms[C/OL]//BHATEJA V, PATEL P, SIMIC M. Singapore: Springer Nature, 2025: 605-615.

7. SHI J, CHEN Z, ZHU J, et al. Research on the impact of pointing gestures based on computer vision technology on classroom concentration[J/OL]. Neural Computing and Applications, 2025, 37(19): 13237-13249.
8. KELLY S, KIM J, CHAUDHARY P. Understanding the role of instructor gestures during virtual lectures[J/OL]. Frontiers in Computer Science, 2025, 7[2026-04-22].
9. BALI C, TASDELEN B, BANDI S, et al. Understanding the cognitive cost of multimedia learning: Effects of visual load and language proficiency[J/OL]. Cognitive Research: Principles and Implications, 2026, 11(1): 2.
10. SCHNEIDER S, BEEGE M, NEBEL S, et al. A meta-analysis of how signaling affects learning with media[J/OL]. Educational Research Review, 2018, 23: 1-24.
11. PI Z, LING H, LI X, et al. Instructors' pointing gestures and positive facial expressions hinder learning in video lectures: Insights from teachers and students in China[J/OL]. Education and Information Technologies, 2024, 29(16): 21115-21131.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.