Review Article

A Bibliographic Review and Meta-Analysis of the Effect of Teaching Behaviors on Learners' Performance in Videotaped Lectures

Qinghong Zhang ⬚ **and Zhonggen Yu** ⬚

Beijing Language and Culture University, Beijing, China

Correspondence should be addressed to Zhonggen Yu; 401373742@qq.com

Although the discussion about the influence of the instructor-present videos has become a hot issue in recent years, the potential moderators on the effectiveness of an on-screen instructor have not been thoroughly synthesized. The present review systematically retrieves 47 empirical studies on how the instructors' behaviors moderated online education quality as measured by learning performance via a bibliographic study using VOSviewer and meta-analysis using Stata/MP 14.0. The bibliographic networks illustrate instructors' eye gaze, gestures, and facial expressions attract more researchers' attention. The meta-analysis results further reveal that better learning performance can be realized by integrating the instructor's gestures, eye guidance, and expressive faces with their speech in video lectures. Future studies can further explore the impact of instructors' other characteristics on learning perception and visual attention including voice, gender, age, etc. The underlying neural mechanism should also be considered via more objective technologies.

1. Introduction

Online learning has obtained prominent popularity in the educational field because of its accessibility, comfort, and remote opportunities [1]. However, the challenges of online learning urged researchers to design efficient instructional courses, indispensable components of online courses, to optimize online learning outcomes [2]. Some empirical studies demonstrated that instructor-present videos could provide individualized instruction, thus leading to better online learning performance [3, 4]. At the same time, some researchers claimed that instructors' image may interfere learners or distract their attention [5]. The inconclusive findings have encouraged explorers to give reasonable explanations from different perspectives including variables related to students or instructors. Kizilcec et al. [6] emphasized that the learners' learning preferences could moderate the effectiveness of instructors' presence in videotaped lectures. Teachers' various behaviors also have potential to influence whether instructors should be present in instructional videos.

The literature suggests that whether the teacher image in videotaped lectures is beneficial to students' learning depends on teachers' various behaviors, especially the following three behaviors [7–9]. Some empirical studies devoted themselves to investigating how the instructor's gestures exert influence on students' learning [10, 11]. Instructor's gestures are deemed as hand movements that can be integrated with spoken language to convey meaningful information to learners [12]. Eye gaze, as another critical human cue in videotaped lectures, has the potential ability to trigger better learning performance [7]. Specifically, when recording a teaching video, the teacher looks directly at the camera, teaching materials, or switches his fixations between the camera and the teaching materials to guide students' attention intentionally [13]. According to multimedia learning theory, eye gaze is beneficial for reducing the competition of cognitive resources and unnecessary cognitive load, thus improving learners' performance [14]. Besides the effect of instructors' gestures and eye gaze in videos, a lack of investigations studies the on-screen teachers' emotional faces in the multimedia learning environment. In fact, learners'
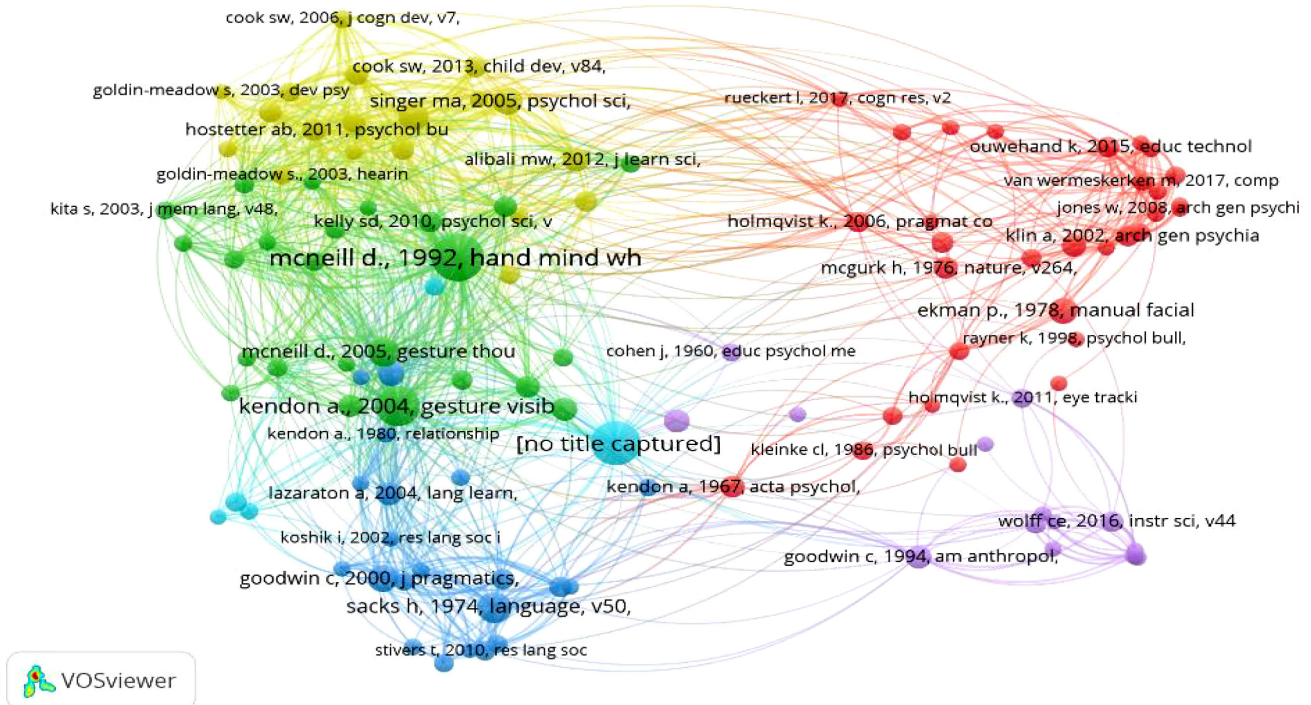
FIGURE 1: Cocitation mapping based on cited reference.

emotional perception during learning strengthens their memory for learning contents, which is beneficial for learning performance [9].

Although various empirical investigations have examined or systematically reviewed the influence of the instructor-present videos on cognitive, affective, social, and learning aspects of learners, the potential moderators on the effectiveness of an on-screen instructor have not been thoroughly synthesized [15, 16]. Furthermore, scanty researchers have conducted the meta-analysis and bibliographic review to systematically summarize and compare the moderation of various teachers' behaviors, such as the instructor's gestures, eye gaze, and emotional face, on the learning outcomes. Therefore, the current review first constructs cocitation and occurrence network maps using VOSviewer, which can present a bibliometric literature review on some hot variables related to instructors' online teaching performance in videotaped lectures. Then, meta-analyses can justify the proposed null hypotheses by quantifying and summarizing existing inconclusive research results regarding the effectiveness of instructors' behaviors in online courses. In conclusion, based on the hot topics identified by VOSviewer, the review continues to implement meta-analyses to explore how these features of instructors exert influence on learners' learning in an online environment.

2. Bibliographic Review of the Effect of Teaching Behaviors in Videotaped Lectures

2.1. Data Collection for the Bibliographic Review. We performed a comprehensive search to retrieve literature from the Web of Science Core Collection, a reliable citation index for academic research worldwide [1]. It boasts extensive resources on online databases such as Social Sciences Citation Index, Science Citation Index Expanded, and Emerging Sources Citation Index. We selected the effect of teacher, instructor, speaker, and video lecture as the term for joint search on April 15, 2022. In order to avoid search bias caused by daily database updates, we manipulated the search on a single day and obtained 781 results. Based on these data, we adopted VOSviewer to create the cocitation network based on reference by selecting full counting as the counting method, reference as the unit of analysis, and cocitation as the type of analysis. Cocitation refers to the interaction where the same literature cites two items. Of the 23,681 cited references, 115 references appeared at least eight times. We chose to show all documents, even if they were not connected with each other.

2.2. Data Analysis. Cocitation analysis enabled professionals to identify the popularity of a certain research theme. If the frequency that two publications are cited by other publications is relatively higher, the focus of the highly cited literature will be regarded as a central theme of an intellectual structure [17]. In the visualization presented in illustrations, colors mean groups of terms relatively strongly associated with each other. Each cluster is represented by a color that indicates the group to which the cluster was assigned. The lines between different clusters also represent the relatedness of clusters and the thickness of a curved line reflects the number of citations between different clusters.

As shown in Figure 1, cluster 1 (red), as the largest cluster, presents intensely cocited publications which aggregated the relatively more cutting-edge literature. Based on cluster 1, we can learn more hot and new things about the use of
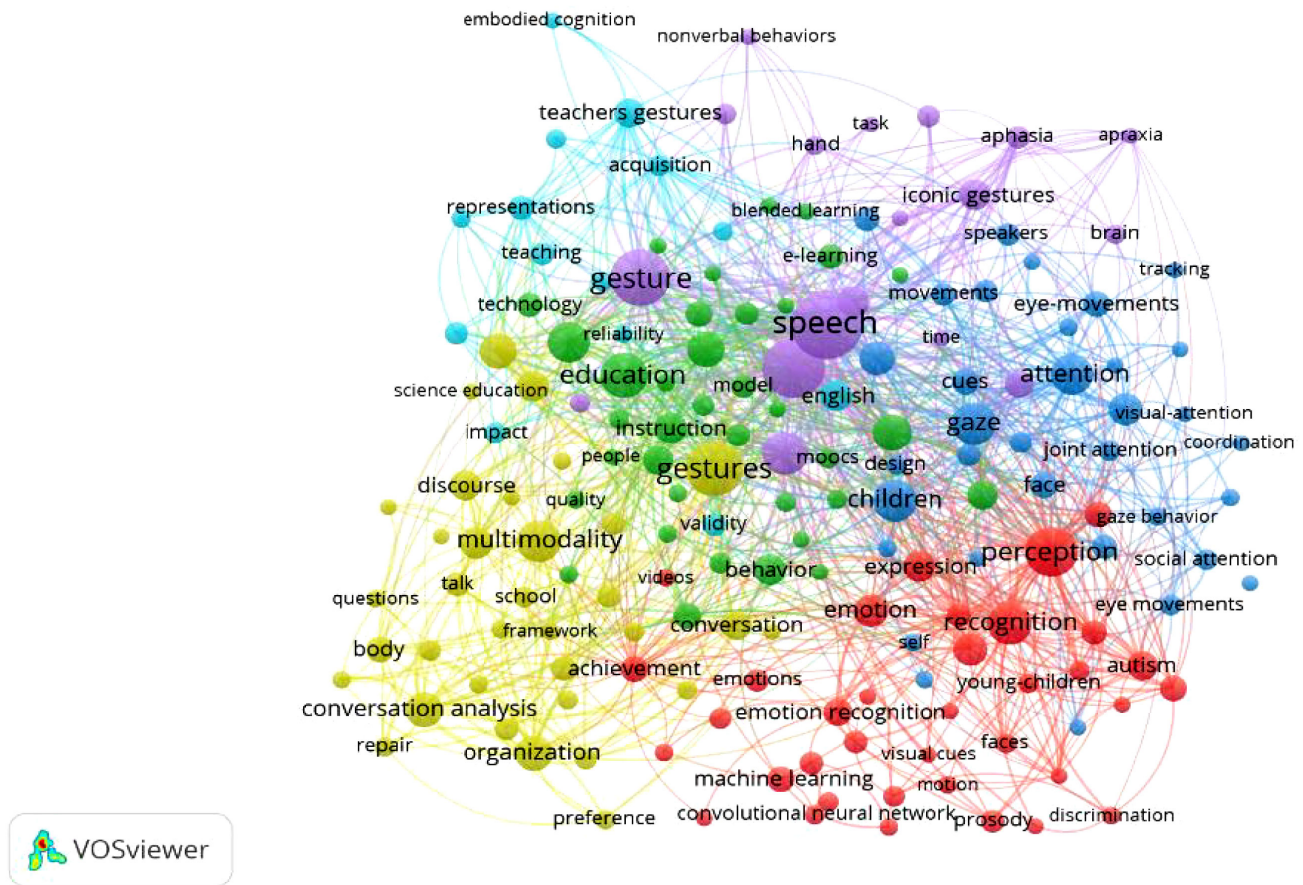
FIGURE 2: Clustering networks based on keywords.

instructors' behaviors in video lectures. To be more specific, cluster 1 involves 10 documents with relatively high citations in recent years. Eight of the 10 articles are empirical research on the instructor's nonverbal behaviors. For example, *Designing effective video-based modeling examples using gaze and gesture cues* (citations = 13) [18]; *Seeing the instructor's face and gaze in demonstration video examples affects attention allocation but not learning* (citations = 11) [19]; and *The Instructor's Face in Video Instruction: Evidence from Two Large-Scale Field Studies* (citations = 10) [6]. Therefore, when it comes to instructors' behaviors in videos, their gestures, eye gaze, and face appear to attract more researchers' attention.

Furthermore, keyword occurrence analysis in VOSviewer is also helpful in detecting hot topics, which means that the larger node represents a higher frequency, which is useful for detecting hot topics via cluster analysis, created the bibliographic map of keyword co-occurrence by selecting full counting as the counting method, co-occurrence as the type of analysis, and all keywords as the unit of analysis. In addition, we chose 5 as the threshold of the occurrence of a keyword. Of the 3,093 keywords, 172 met the threshold. Figure 2 shows the co-occurrence map of keywords that appeared over five times.

As mentioned above, the larger node represents the higher frequency which could reflect its importance [20]. As shown in Figure 2, except the teachers' speech (N = 81), their

behaviors including gestures (N = 113), eye gaze (N = 51), and emotions (N = 64) definitely a relatively higher frequency of occurrence, which is consistent with the cocitation analysis illustrated by Figure 1. Therefore, we retrieved existing studies on these hot topics which closely correlated with teachers' behaviors in videotaped lectures.

2.3. Instructor's Gestures in Video Lectures. Instructor's gestures, as essential components of teacher behavior, exert a significant impact in designing effective videos, thus becoming one of the popular research areas of video lectures [21]. Gestures refer to hand movements that can be integrated with spoken language to convey meaningful information to learners [12]. In instructional videos, gestures involve three major types [22]: deictic (pointing) gestures, depictive gestures, and beat gestures. To be more specific, deictic gestures are defined as a guiding method in which teachers point to specific teaching content with fingers or hands. Depictive/representative gestures literally or metaphorically indicate semantic content in videos through hand movements, which are beneficial for evoking a mental schema existing in listeners' memory. It contains iconic gestures representing the concreteness of things (e.g., introducing apples by drawing them in the air) and metaphoric gestures emphasizing abstract semantic aspects (e.g., describing a geometric shape in the air). Unlike the first three gestures, beat gestures cannot

convey any semantic content and are simple, up-and-down rhythmic hand movements that align with speech prosody.

A large body of research has provided positive evidence that adopting instructors' gestures could improve students' learning outcomes [23, 24]. So far, the positive effect of their gestures on students' learning has been demonstrated in various instructional videos, such as language acquisition [25], biology [21], narrative understanding [26], and statistics [27], etc. Furthermore, three major gestures were reported that their adoption was significantly and positively correlated with learners' better performance in videos [28, 29]. However, the positive effect of the instructor's gestures on enhanced learning performance could not be replicated in some empirical evidence [30, 31]. To investigate the effect of instructor's gestures in video lectures, we, therefore, identify the following questions:

(H1) The instructor's gestures in videotaped lectures may play an essential part in students' learning

(H2) The nature of learning materials and the type of instructors' gestures may positively moderate the effect of gestures on students' learning performance

2.4. Instructor's Eye Gaze in Video Lectures. The previous literature assumed that the instructor's eye gaze, another human cue in instructional videos, has the potential ability to trigger better learning performance [7]. Recently, more and more empirical efforts put emphasis on directly detecting the effect of the instructor's eye gaze on learning [18]. The visible instructor-enabled participants to become more successful in learning, presumably resulting from the guidance of the teacher's eye gaze [32]. Based on the assumption, van Wermeskerken and van Gog [19] devoted themselves to replicating the research findings with different tasks and exploring whether the visible instructors in videos facilitate learning due to their eye gaze guidance. Accordingly, they divided 69 psychology participants into three different videos concerning the composition of glutamine: no face, visible face with gaze, and visible face without eye gaze. The eye-tracking analysis reported that the instructor's gaze adjusted the attention pattern of learners in video lectures, which further failed to induce their better learning performance. Considering the contradictory effect of the instructor's eye gaze, we proposed the following hypothesis:

(H3) The instructor's eye gaze in videotaped lectures may play an essential part in students' learning

2.5. Instructor's Facial Expressions in Video Lectures. Beyond the effect of teachers' gestures and eye gaze in videos, few investigations study the teachers' emotional faces in the multimedia learning environment. The social presence theory supports that facial expression, as another nonverbal behavior, can shorten the distance learners perceive in a virtual learning environment, which is closely related to their learning experience and satisfaction [33]. Wang et al. [31] selected 69 participants to watch three kinds of videos (instructor

with a heightened, conventional expression, and instructor without any expression). The results demonstrated that instructors' emotional faces were crucial in improving students' online learning. However, another opposite prediction has been supported by the split-attention effect, which occurs if learners need to process multiple pieces of novel information [34]. In videotaped lectures, learning materials coupled with a teacher's face, particularly an emotional face, can lead to visual split attention because students must choose the more prominent content to focus on [35]. The stronger split-attention phenomenon has harmful effects on learners' performance. Thus, it is necessary to explore the following hypothesis:

(H4) The teacher's facial expressions in videotaped lectures may play an essential part in students' learning

3. Methodology

According to the hot topics identified by bibliographic maps, we further implemented meta-analyses to comprehensive review whether these variables have potential to influence online learners' performance, which is discussed in detail through the above research questions. The current review focused on how do characteristics of teachers in video lectures affect students' learning outcomes, so the meta-analysis is most suitable for quantifying and retrieving the previous research findings. The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) (see *Supplementary 1*) provided a reporting guideline for our current meta-analyses [36].

3.1. Identifying Research. The present review adopted a four-step design to explore how do characteristics of teachers in videotaped lectures affect students' learning outcomes (Figure 3). First, the visualizing results provided insight into the search terms of the current review. According to the bibliographic networks of keywords occurrence and cocitation analysis, we can summarize instructors' eye gaze, gestures, and facial expressions as three key variables that significantly impact students' learning performance. Furthermore, empirical research was pursued by more influential authors, which indicated that our review should give more priority to the empirical direction of instructors in videotaped lectures. Therefore, we selected instructor, gesture, eye gaze, facial expressions, emotion, videos, and learning performance as our primary search terms to retrieve as many publications as possible. Another complementary search included a closer examination of the selected articles' references ($N = 728$). We continued to rough filter the search results by narrowing the research to the education domain and removing editorial collections, book chapters, and reports ($N = 307$). Third, sufficient statistical information is necessary for the meta-analysis. Hence, the researcher screened available articles with full texts ($N = 92$). Finally, the researchers made a further refinement following the framework of STARLITE and other inclusion and exclusion criteria ($N = 47$).

3.2. Data Credibility. To further retrieve high-quality literature, we adopted STARLITE to assess the quality of obtained results
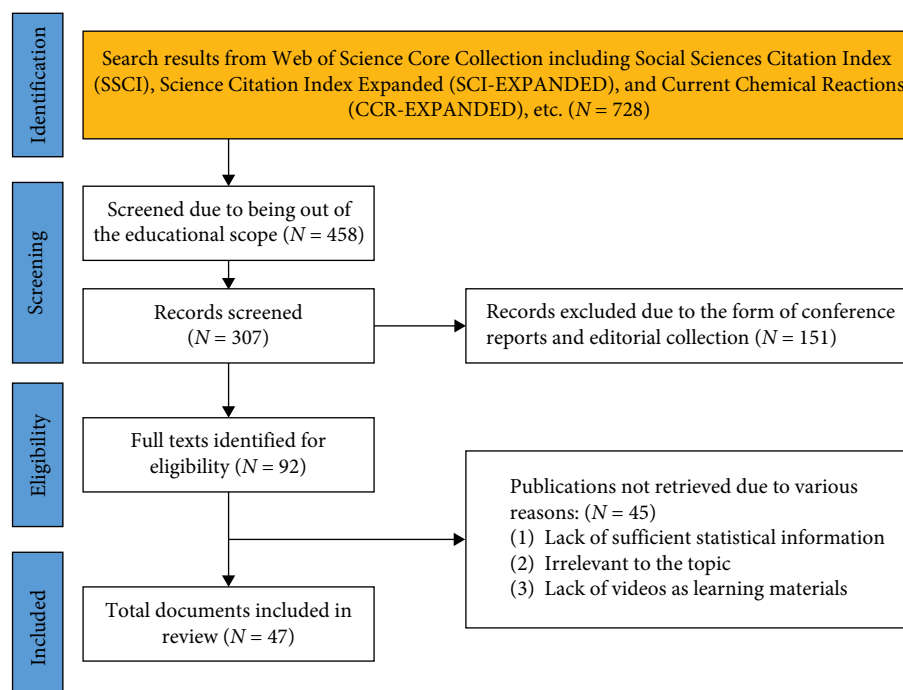
FIGURE 3: A flowchart based on PRISMA.

[37]. STARLITE refers to a systematic criterion including sample strategy (comprehensive, selective, or purposive strategy), types of studies (fully reported or partially reported), approaches, range of years (fully reported or partially reported), limits, inclusion and exclusion criteria, the term used, and electronic sources (*Supplementary 2*) [38].

3.2.1. Sampling Strategy. We carried out a purposive sampling method to search as many documents as possible among various studies. Having retrieved the previous records, we narrow the literature to how teachers' behavior characteristics in video lectures affect students' learning outcomes.

3.2.2. Type of Studies. Quantitative, qualitative, or mixed research designs were adopted to analyze reported studies.

3.2.3. Approach. We performed a hand search method to retrieve literature from the Web of Science Core Collection. It boasts extensive resources on the online databases such as Social Sciences Citation Index (SSCI), Science Citation Index Expanded (SCI-EXPANDED), Emerging Sources Citation Index (ESCI), etc. The visualizing results provided insight into the research terms in the current review. To be more specific, we conducted a joint search by keying in speaker's gesture* OR instructor's gesture* OR teacher gesture* OR speaker's eye gaze OR instructor's eye gaze OR teacher eye gaze video OR speaker's expression* OR instructor's expression* OR teacher expression* and video*. Finally, we obtained 728 documents by selecting all texts as the filter. In order to avoid search bias caused by daily database updates, we manipulated the search on a single day.

3.2.4. Range of Years. We filtered the previous literature by limiting the specific publication years ranging from 2008 to

2022. Year 2008 witnessed the commencement date of studies on the teachers' role in instructional videos. Relevant publications in 2022 failed to be thoroughly revealed because we conducted the search on April 15, 2022.

3.2.5. Limits. The review may be limited to a single database. Although the Web of Science boasts high reliability and extensive resources, we do not ensure an exhaustive retrieve in the field. The study can be improved because most documents within our scope are written in English. Naturally, some crucial publications in other languages are out of scope. We pay more attention to rigidly quantitative or mixed-designed studies at the expense of the form of conference reports, editorial articles, book chapters, etc.

3.2.6. Inclusions and Exclusions. In the present review, we included the original articles if the publications: (1) aim to examine the effect of teachers' characteristics, such as the instructor's gestures, eye gaze, and facial expressions, on students' performance in video lectures; (2) report the learners' learning performance as measured by retention, transfer, or comprehensive tests; (3) contain at least an experimental condition and a control condition; (4) must be high-quality journal articles rigidly evaluated by the framework of STARLITE. We excluded the original articles if the publications: (1) fail to shed light on teachers themselves rather than their effect on learners' performance; (2) do not select videotaped lectures as learning materials; (3) cannot provide necessary statistical conditions; (4) are poorly designed according to the guideline of STARLITE.

3.2.7. Term Used. In order to obtain as many documents as possible, we tried various terms on the Web of Science. The

search terms include instructor, gesture, eye gaze, facial expressions, emotion, videos, learning performance, etc.

3.2.8. Electronic Sources. Electronic sources are retrieved from the Web of Science Core Collection. It boasts extensive resources on online databases such as SSCI, A&HCI, CPCI-SSH, ESCI, SCI-Expanded, CPCI-S, IC, and CCR-EXPANDED.

3.3. Coding of Studies. Numerous researchers devoted themselves to reporting how the teacher's gestures in videos exert a positive influence on students' learning. The obtained documents indicated that the story is more complex than the teachers' gestures being beneficial simply. To further explore the potential boundary conditions for the effect of gestures, we recorded the obtained literature based on moderators, such as the type of gestures, the nature of knowledge (descriptive and procedural knowledge), learning measurements, and effect size (*Supplementary 3*). Some potential moderators, such as the difficulty level of learning materials, learners' prior knowledge, the age of participants, and video length, failed to be included in the meta-analysis due to limit-related literature. The studies within our scope reported various dimensions of learning achievements, including retention, transfer, and comprehensive tests. Furthermore, an empirical study often incorporates multiple experiments. Therefore, if the experimental variable falls into our current scope, it will be divided into different effect sizes. Otherwise, it will be combined into the same effect size.

3.4. Statistical Analysis. In the present review, researchers implemented a meta-analysis to retrieve data from multiple independent studies using Stata MP/14.0. We first reported meta-analyses results of the effect of instructors' gestures, eye gaze, and facial expressions on learners' performance, including calculations of effect size, 95% confidence intervals, weights of individual studies, I^2, p values, heterogeneity, Q statistics, and pooled effect sizes. Numerous documents indicated that the story is more complex than the teachers' gestures being beneficial simply. Subsequently, we continued to explore the moderating effect of the type of gestures and knowledge by regarding them as moderating variables. We also adopted Begg's and Egger's tests to examine the publication bias at the 0.05 level through effect size and standard errors. The sensitivity analysis was also adopted to determine whether the meta-analytical results were robust and stable.

The researchers also calculated the value of I^2 to identify whether the percentage of variation among individual studies was induced by heterogeneity other than random errors. Higgins et al. [39] set the standard to determine the importance of heterogeneity and the special model for the meta-analysis. Specifically, if the value of I^2 remains between 75% and 100%, the heterogeneity will be deemed as considerable, substantial from 50% to 90%, moderate from 30% to 60%, and unimportant less than 40%. Accordingly, we perform a more rigid meta-analysis via the fixed-effect model due to lower heterogeneity ($I^2 < 50\%$). Otherwise, a higher heterogeneity ($I^2 > 50\%$) asks us to adopt a random-effect modal and sensitivity analysis in the meta-analysis.
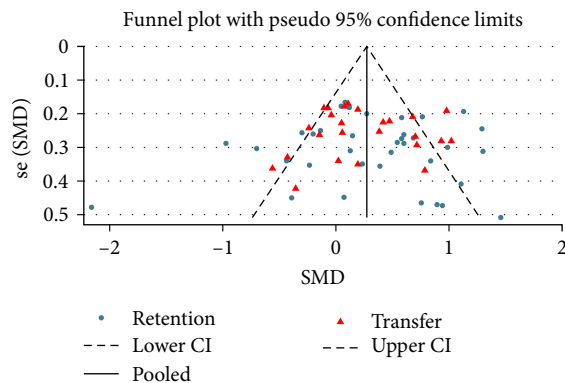

FIGURE 4: Funnel plots of publication bias analysis.

4. Results

In this section, we summarized and analyzed the meta-analytical results from three aspects including gestures, eye gaze, and facial expressions. We aimed to improve the reliability of research results by performing tests of sensitivity and publication bias analysis. Additionally, the results were reported according to the sequence of the proposed five null hypotheses.

4.1. Gestures

4.1.1. Detection of Publication Bias. To make our results more robust, we adopted Begg's and Egger's tests to estimate publication bias as precisely as possible. Figure 4 illustrates the analysis results of publication bias. Specifically, the *X*-axis means standardized mean difference and the *Y*-axis represents standard errors of the standardized mean difference. Each dot indicates an independent document, and the middle line serves as the no-effect line. If the dots appear to be symmetrically distributed along the no-effect line, we can infer that no publication bias can be found in these documents. The result (Figure 4) reflects the absence of publication bias due to the symmetrical distribution (coefficient = 0.115, std. error = 0.907, $t = 0.13$, $p = 0.899$). Similarly, the researchers also reported the absence of publication bias by using Begg's test (Kendall's score $(P - Q) = 114$, std. dev. of score = 172.6 (corrected for ties), and number of studies = 64, $z = -0.66$ Pr > $|z| = 0.513$).

4.1.2. The Sensitivity Analysis. The researchers also implemented the influence analysis using Stata 14.0 to measure the robustness of meta-analytical results (Figure 5). Each dot stands for a specific publication. When all dots fall within the scope between two lines representing the lower and upper confidence interval limits, the obtained results of the meta-analysis will be robust and stable. As illustrated in Figure 6, all documents are distributed within 95% confidence intervals, indicating that we performed a reliable meta-analysis.

4.1.3. The Effect of Teacher's Gestures on Learning in Videos. We calculated 39 and 24 effect sizes to investigate teacher gestures' effect on students' learning, including retention and transfer scores. The meta-analysis results (Table 1) prove that heterogeneity induces the variance of effect sizes rather than random errors in terms of retention and transfer tests ($Q = 173.19$, $I^2 = 78.1\%$, $p < 0.01$; $Q = 71.74$, $I^2 = 66.5\%$,
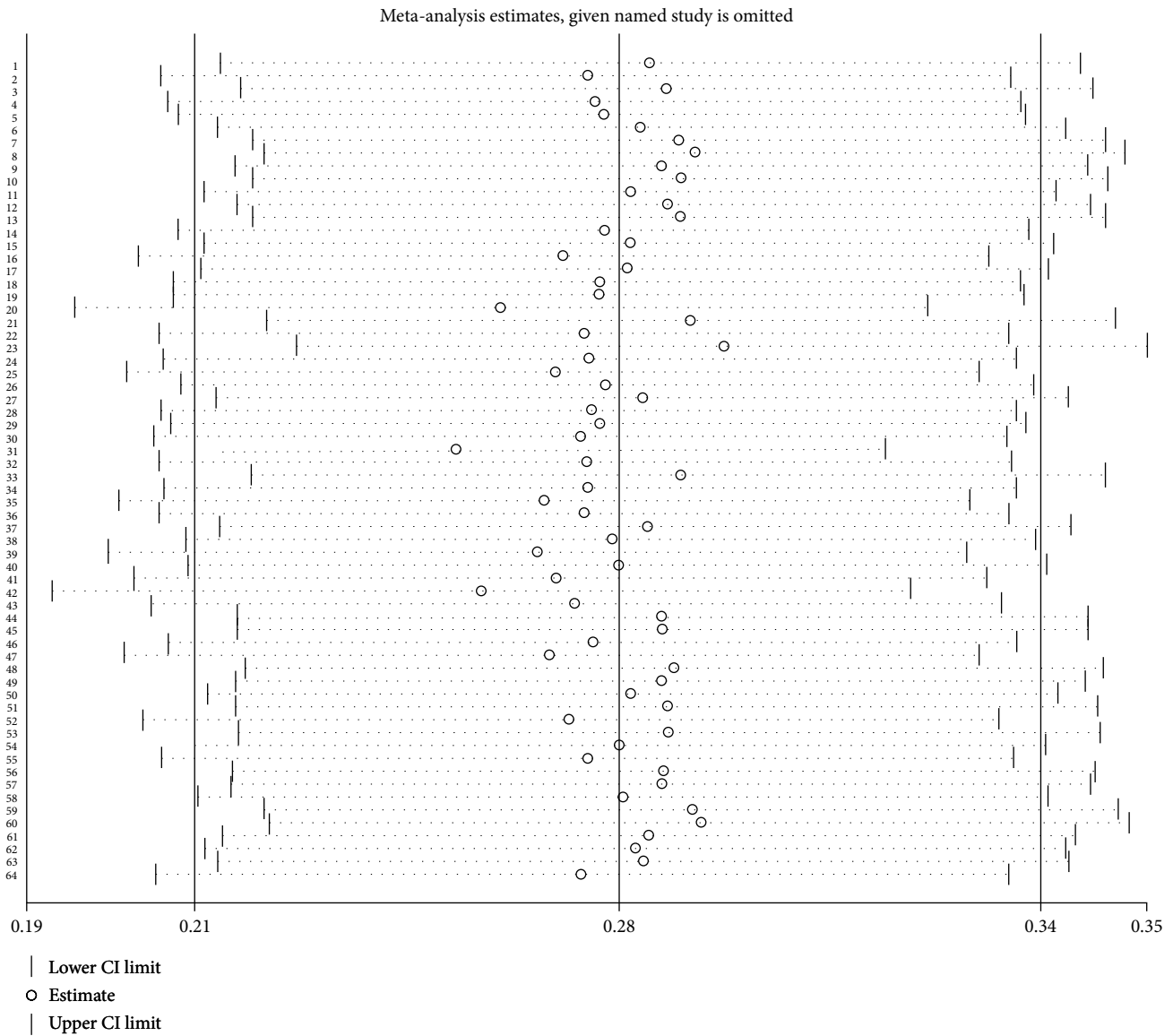
FIGURE 5: Results of the influence analysis.
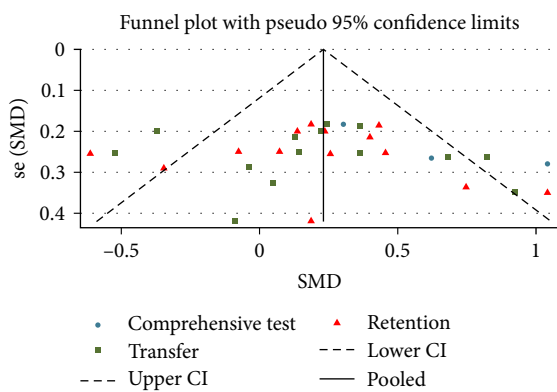
FIGURE 6: Funnel plots of publication bias analysis.

$p < 0.01$). As a result, we selected a random-effect model to examine the effect sizes. The meta-analysis results (Table 1) reveal that there is a significant difference between videos with and without gestures group based on retention performance ($d = 0.93$, 95% CI = 0.38, 0.49) and transfer scores ($d = 0.24$, 95% CI = 0.08, 0.41). Therefore, researchers, accepting the null hypothesis, indicate that students in instructional videos with teacher's gestures outperform the control group based on retention and transfer tests.

4.1.4. Meta-Analyses for Different Gestures. We calculated 24, 26, and 12 effect sizes to separately investigate the moderating effect of the type of gestures used by instructors in videotaped lectures in terms of pointing gesture (PG), depictive gesture (DG), and beat gesture (BG). The meta-analysis results (Table 2) prove that heterogeneity induces the variance of effect sizes rather than random errors according to pointing gesture and depictive gesture ($Q = 103.88$, $I^2 = 76.9\%$, $p < 0.01$; $Q = 102.15$, $I^2 = 74.5\%$, $p < 0.01$). As a result, we selected a random-effect model to examine the effect sizes. The meta-analysis results (Table 2) reveal that there is a significant

TABLE 1: Meta-analyses on the main effect of instructor's gestures for learning.

Learning performance	SMD	95% CI		wt%	Heterogeneity statistic	df	p	I^2 (%)	z	p
R	0.93	0.38	1.49	58.6	173.19	38	<0.01	78.1	3.16	0.002
T	0.24	0.08	0.41	41.4	71.74	24	<0.01	66.5	2.92	0.004
Overall	0.28	0.15	0.41	100	246.06	63	<0.01	74.4	4.27	<0.01

Note: R: retention; T: transfer.

TABLE 2: Meta-analyses on the moderation effect of instructor's gestures for learning.

Learning performance	SMD	95% CI		wt%	Heterogeneity statistic	df	p	I^2 (%)	z	p
PG	0.38	0.17	0.59	40.18	103.88	24	<0.01	76.9	3.54	<0.01
DG	0.28	0.08	0.48	41.89	102.15	26	<0.01	74.5	2.79	0.005
BG	0.02	0.2	0.23	17.93	20.17	11	0.043	45.5	0.07	0.946
DK	0.34	0.18	0.51	73.6	208.62	47	<0.01	77.5	4.09	<0.01
PK	0.11	0.04	0.26	26.4	25	15	0.05	40	1.94	0.052
Overall	0.28	0.15	0.41	100	246.06	63	<0.01	74.4	4.27	<0.01

Note: PG, pointing gesture; DG, depictive gesture; BG, beat gesture; DK, declarative knowledge; PK, procedural knowledge.

difference between the experimental group and the control group based on pointing gestures ($d = 0.38$, 95% CI = 0.17, 0.59) and depictive gestures ($d = 0.28$, 95% CI = 0.08, 0.48).

However, the effect sizes of beat gestures are considered not significantly heterogeneous ($Q = 20.17$, $I^2 = 45.5\%$, $p = 0.043$). We, thus, chose a fixed model to explore the effect sizes. Students in the beat gestures group do not significantly perform better than the no-gesture group ($d = -0.01$, 95% CI = −0.16, 0.15). Nevertheless, researchers, partially accepting the null hypothesis, indicated that students harvested greater academic achievements in the experimental group where teachers adopted pointing and depictive gestures to teach in instructional videos.

4.1.5. Meta-Analyses for Different Knowledge. The results (Table 2) summarize 48 and 16 effect sizes to respectively reveal the moderating effect of knowledge type in terms of declarative knowledge (DK) and procedural knowledge (PG). The meta-analytically findings of test scores are heterogeneous at the type of knowledge level based on declarative knowledge ($Q = 208.62$, $I^2 = 77.5\%$, $p < 0.01$). As a result, we chose a random-effect model to examine the effect sizes. As shown in Table 2, learners' performance in the experimental group was significantly better than in the control group when teachers narrated declarative knowledge in instructional knowledge ($d = 0.34$, 95% CI = 0.18, 0.51). Since $I^2 = 40\%$ in the procedural gestures ($Q = 25$), the effect sizes tend to be not significantly heterogeneous ($p = 0.052$). Thus, a fixed model appeared to be more appropriate for conducting the meta-analysis regarding procedural knowledge. As shown in Table 1, not enough evidence can prove that students' performance would be better if an instructor teaches procedural knowledge coupled with gestures in video lectures.

4.2. Eye Gaze

4.2.1. Detection of Publication Bias. As mentioned above, we also adopted Begg's and Egger's tests to estimate publication bias (Figure 6). The results illustrate that all dots seem to be nearly symmetrical (coefficient = 0.931, std. error = 1.293, $t = 0.72$, $p = 0.477$). Similarly, the researchers also reported the absence of publication bias by using Begg's test (Kendall's score $(P - Q) = 47$, std. dev. of score = 58.84 (corrected for ties), number of studies = 31, $z = 0.8$, $\Pr > |z| = 0.434$).

4.2.2. The Sensitivity Analysis. The researchers also implemented the influence analysis using Stata 14.0 to measure the robustness of meta-analytical results. As illustrated in Figure 7, all dots belong to the scope defined by two lines representing the lower and upper confidence interval limit. Thus, the obtained results of the meta-analysis are distributed within 95% confidence intervals, which indicates that we performed a reliable meta-analysis.

4.2.3. The Effect of the Teacher's Eye Gaze on Learning in Videos. We calculated 3, 14, and 13 effect sizes to respectively investigate the effect of the teacher's eye gaze on students' learning, including comprehensive, retention, and transfer scores. The meta-analysis results (Table 3) prove that heterogeneity induces the variance of effect sizes rather than random errors in terms of comprehensive, retention, and transfer tests ($Q = 4.96$, $I^2 = 59.7\%$, $p = 0.084$; $Q = 27.09$, $I^2 = 52\%$, $p = 0.012$; $Q = 32.49$, $I^2 = 60\%$, $p = 0.002$). As a result, we selected a random-effect model to examine the effect sizes. The meta-analysis results (Table 3) reveal that there is a significant difference between videos with and without eye gaze group based on comprehensive performance ($d = 0.62$, 95% CI = 0.19, 1.05) and retention scores ($d = 0.2$, 95% CI = 0.02, 0.39). However, transfer performance had neither detrimental nor beneficial change when teachers added the eye gaze cues in videos ($d = 0.2$, 95% CI = 0.00, 0.41, $p = 0.053$). Therefore, researchers, partially accepting the null hypothesis, indicate that students in instructional videos with the teacher's eye gaze outperform the control group based on comprehensive and retention tests.
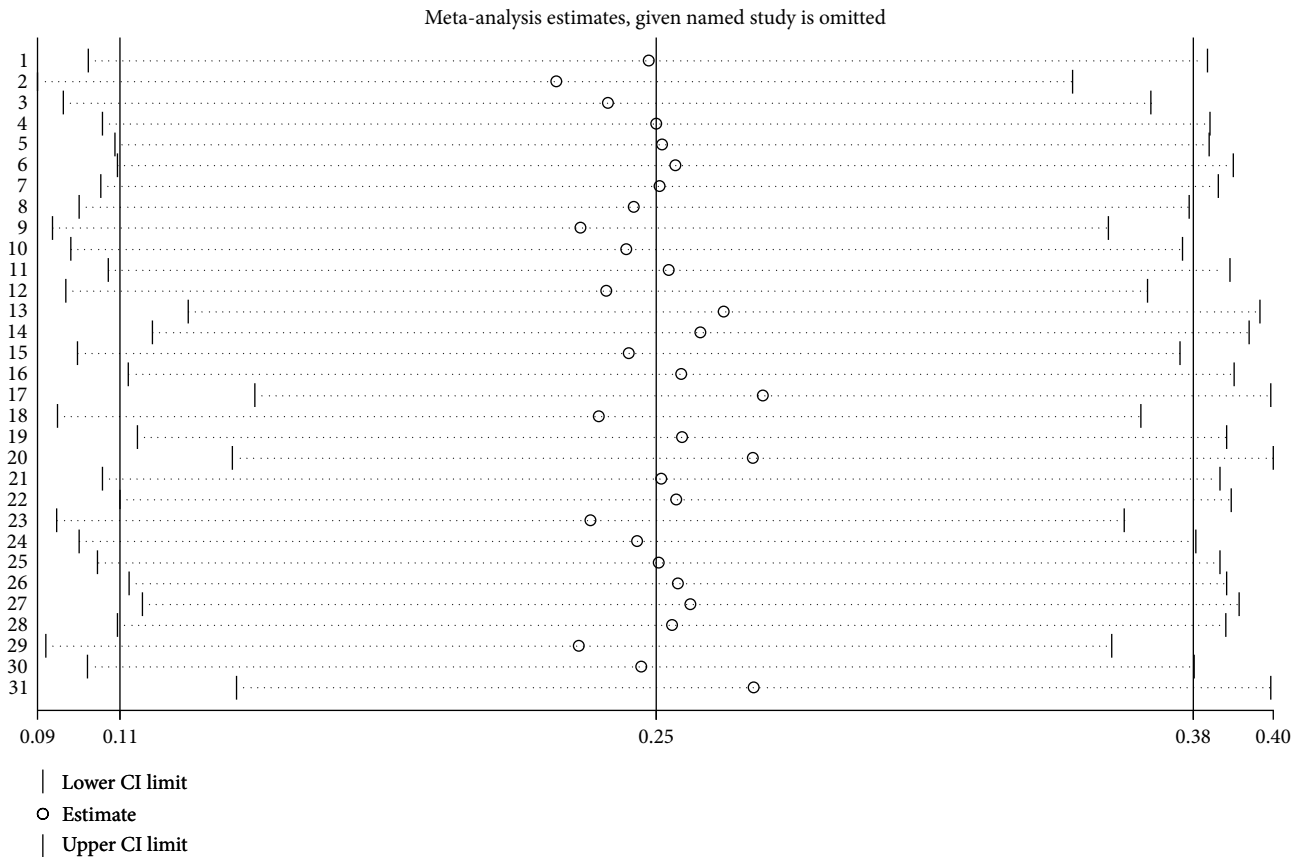
Figure 7: Results of the influence analysis.

Table 3: Meta-analyses of instructor's eye gaze for learning.

Learning performance	SMD	95% CI		wt%	Heterogeneity statistic	df	p	I^2 (%)	z	p
Comprehensive test	0.62	0.19	1.05	10.1	4.96	2	0.084	59.7	2.82	0.005
Retention test	0.2	0.02	0.39	44.94	27.09	13	0.012	52	2.15	0.032
Transfer test	0.2	0.00	0.41	44.95	32.49	13	0.002	60	1.94	0.053
Overall	0.25	0.11	0.38	100	70.87	30	<0.01	57.7	3.62	<0.01

4.3. Facial Expressions

4.3.1. The Detection of Publication Bias and Sensitivity. In the current review, we also adopted Begg's test and Egger's test to estimate publication bias (Figure 8). The results illustrate that all dots seem nearly symmetrical (coefficient = 7.189, std. error = 5.046, $t = 1.42$, $p = 0.214$). Similarly, the researchers also reported the absence of publication bias by using Begg's test (Kendall's score $(P - Q) = 11$, std. dev. of score = 6.66 (corrected for ties), number of studies = 7, $z = 1.65$ $\Pr > |z| = 0.133$). Concerning the sensitivity analysis, all dots belong to the scope defined by two lines representing the lower and upper confidence interval limit (see Figure 9). Thus, the obtained results of the meta-analysis are distributed within 95% confidence intervals, which indicate that we performed a reliable meta-analysis.

4.3.2. The Effect of Teacher's Facial Expressions on Learning in Videos. We calculated seven effect sizes to respectively investigate the effect of the teacher's facial expressions on students' learning achievement. Since $I^2 = 45.9\%$ in the effect
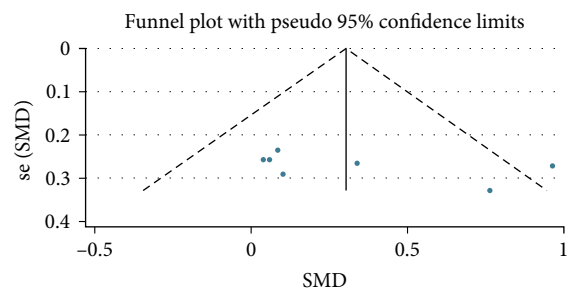

Figure 8: Funnel plots of publication bias analysis.

sizes of teacher's eye gaze ($Q = 11.08$), the effect sizes tend to be not significantly heterogeneous ($p = 0.086$). Thus, a fixed model appeared to be more appropriate for conducting the meta-analysis regarding teachers' facial expressions. The meta-analysis results (Figure 10) reveal a significant difference between videos with emotional faces and without

Meta-analysis estimates, given named study is omitted
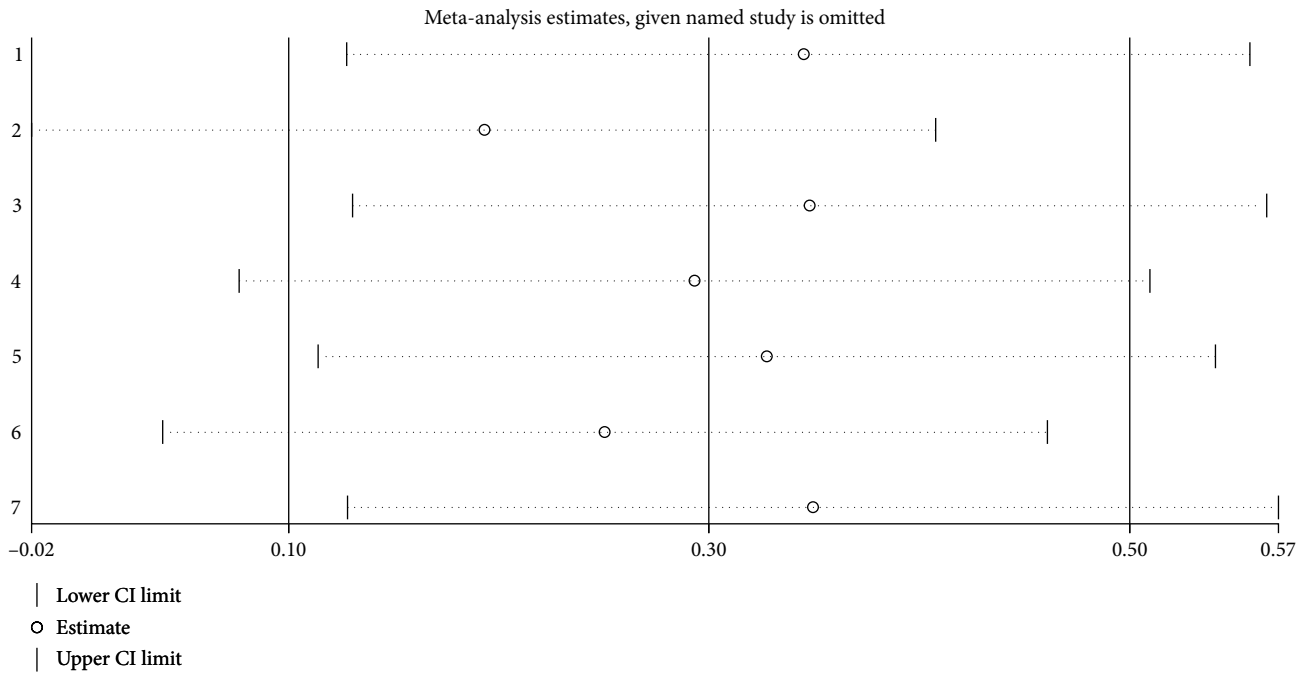

| Lower CI limit
○ Estimate
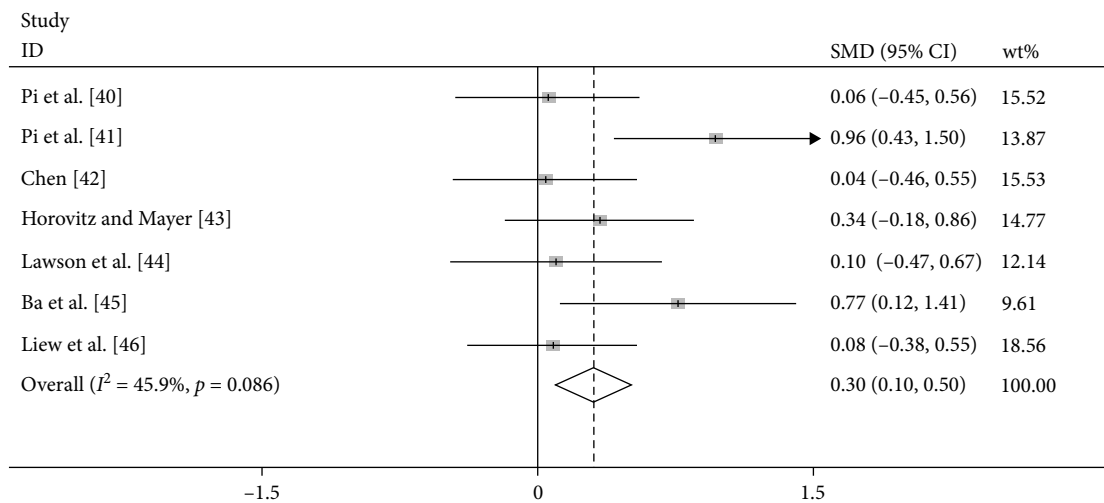| Upper CI limit

FIGURE 9: Results of the influence analysis.

FIGURE 10: A forest plot concerning the effect of teacher's facial expressions [40–46].

expressive instructors based on learning performance ($d = 0.3$, 95% CI = 0.10, 0.5). Therefore, researchers, rejecting the null hypothesis, indicate that students in instructional videos with emotional faces outperform the control group based on comprehensive tests.

5. Discussion

Consistent with Hypothesis 1, the finding indicated that better learning performance could be realized by integrating the instructor's gestures with their speech in video lectures as measured by retention and transfer tests. Pi et al. [21] attempted to explain the cognitive and perceptive mechanism on how teachers' gestures work in instructional videos. On

this basis, we further explained from another perspective why teachers' gestures in instructional videos exert a positive impact on students' learning and illustrated intuitively these three paths by building a more comprehensive model (see Figure 11).

Cognitively, teachers can use appropriate gestures to guide learners' effective attention to more crucial learning content in selecting information. According to Mayer [47], multimedia learning contains three cognitive processes: selecting, organizing, and integrating. Selecting is deemed the process engaged by students selectively to distribute their attention to relevant verbal and visual information. Then, learners focus on organizing information into meaningful representations and integrating them with prior knowledge
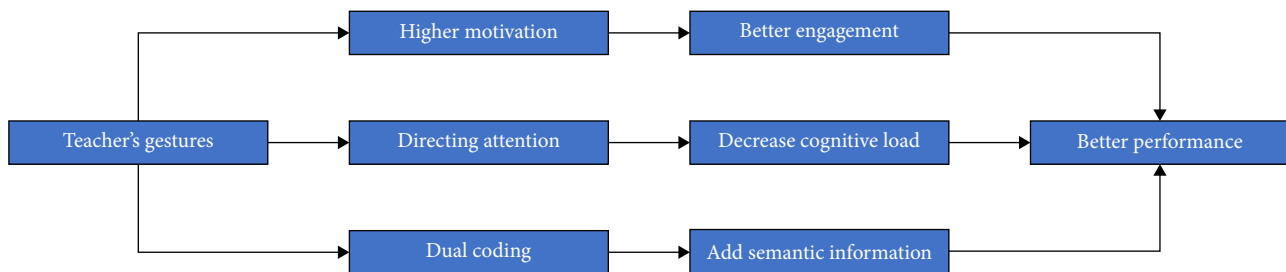
FIGURE 11: A model of how teacher gestures work in instructional videos.

stored in learners' long-term memory. Therefore, the instructor's gestures help more cognitive resources allocate to the organization and integration stage of learning content, thus improving learning performance. The research on teachers' gestures from the cognitive perspective can be evidenced by Yang et al. [13] and So et al. [30], who found that the teacher's pointing or beat gestures were beneficial for better learning performance.

From the perception perspective, Mayer's social cues principle in multimedia learning provided an insight into why better learning outcomes can be realized by adopting the gestures in teacher-present videos [47]. According to the principle, gesture, a typical social cue provided by instructors, may enable learners to perceive a more intimate and familiar social relation with their virtual instructor. Naturally, the closer connection means a higher level of satisfaction and interest, thus leading to learners' higher motivation [48]. Using appropriate gestures during lectures encourages the learners to participate in the video lectures more actively and reach higher academic achievements [49]. The analysis is aligned with empirical findings investigated by Tian et al. [50]. They found that teacher's gesture guidance induced higher HBO of the right prefrontal and right temporal lobes among participants as measured by fNIRS, which proved teachers' gesture was an essential consideration in deepening learners' cognitive processing and improving recall performance.

Furthermore, the third path should also be included in the model of how gestures work in videotaped lectures. The dual channels principle revealed that learners process novel information in separate channels: visual and verbal systems [51]. The cognitive theory of multimedia learning further indicated that instructional designers should sufficiently use these two channels to ensure active cognitive processing while preventing information overloading [49]. Therefore, instructors' gestures served as another important way to provide semantic information besides speech. Specifically, teachers in videos use gestures to make their verbal expressions more explicit through the combination of different channels when gestures and speech describe congruent information. Additionally, if the gestures express information absent in speech, instructors' gestures are able to make up for the incompleteness of semantic information, thus improving students' performance. In sum, gestures, a visual stimulus, cannot only clarify the expression of verbal information but also present semantic information that verbal expressions cannot.

Additionally, the results related to eye gaze are congruent with the claims of social presence theory [52] and Mayer's [53] multimedia learning theory. The social presence theory explains one possible underlying mechanism for the teacher's gaze benefits learning. Learners' perception of interaction with others in an unreal setting positively predicts learning outcomes by stimulating their learning motivation and engagement [32]. According to the multimedia learning theory, the teacher's eye gaze might not only provide a sense of social interaction and intimate connection between the learner and instructor but also guide the learner's effective attention toward what they should follow. Consequently, as an attentional cue, eye gaze reduces the competition of cognitive resources and unnecessary cognitive load [14]. Learners' higher social presence and attention distribution level may contribute to better learning performance.

As for the effect of instructors' facial expressions, meta-analyses results align with the emotional enhancement effect, which elaborates that learners' emotional perception during learning strengthens their memory for learning contents, such as pictures, faces, and words [9]. Furthermore, Wang et al. [31] analyzed the facial expressions produced by a female teacher during teaching activities in videotaped lectures. The final summarized six kinds of facial expressions including the neutral face (55.9%), happy face (12.9%), sad face (10.2%), scared face (9.2%), surprised face (8.7%), and other faces (3.1%). Our review concluded that more expressive emotions produced by instructors provided greater benefits for students' learning. Future researchers should specifically compare the effect of instructors' various expressions on learning, thus shedding light on the design of effective videos.

Despite the effectiveness of the teacher's gestures, the story on transfer tests ($d = 0.93$) seems to be more complicated than gestures being beneficial to retention tests ($d = 0.24$). Therefore, we can infer that the instructor's gestures appear more conducive to elementary memorization and superficial recall of learning content. The transfer test needs more cognitive resources to apply novel knowledge in new situations [49]. Naturally, it requires learners to integrate various elements of the learning materials to construct more sophisticated mental models. It is acceptable why the instructor's gestures in videos induce more satisfactory recall outcomes than transfer performance.

While instructors accompany different gestures to conduct instructional activities, significant differences in the type

of gestures are reported in gesture-assisted learning performance. Specifically, students tend to achieve more success learning when they watch videos with the instructor's pointing gestures and depictive gestures than beat gestures. The conclusion is consistent with the empirical findings of Macoun and Sweller [26]. They believed that gestures, such as depictive gestures, provided more incredible benefits to learning if they could complement additional information for assisting learners in understanding speech. However, beat gestures failed to convey semantic information [22]. In addition, we could not find direct and robust evidence to support the effect of beat gestures' attention guidance. As one path mentioned above, teacher's pointing gestures guide learners' effective visual attention in instructional videos, thus providing more excellent benefits to learning. Furthermore, beat gestures seemed to convey no nonredundant information, thus leading to lower expectations to recall unless they can perceive it in the visual channel [26].

It is also acceptable to find that students who learned declarative knowledge in video lectures outperformed those who participated in the learning of procedural knowledge. Declarative knowledge, mainly including the abstract understanding of facts and concepts, is more challenging to construct mental representations of graphic and textual information, which is not conducive to learners' stable memory. Therefore, nonverbal cues, such as gestures, have the potential to improve students' learning [54]. However, the finding of van Gog et al. [32] was different from our results. Procedural knowledge mainly introduced the knowledge of how to operate. Each step of the teacher's operation will induce subtle changes in the interface, which requires the instructor to direct learners' focus toward key points, thus improving learning outcomes. Therefore, the teacher's gestures can trigger a faster visual search by precisely locating the key issues. The possible elaboration for the discrepancy may be the different gestures used in videos. In addition, it remains an open question if procedural knowledge has the same positive effects on learning in video lectures coupled with other gestures due to the limited application of procedural materials in previous empirical research.

6. Conclusion

6.1. Major Findings. The current review is committed to implementing a bibliographic study and meta-analysis concerning how instructors' characteristics exert influence on learners' learning in a multimedia environment. The visualizing results provided insight into the search terms of the current review. According to the bibliographic networks of keywords occurrence and cocitation analysis, we can summarize instructors' eye gaze, gestures, and facial expressions as three key variables that significantly impact students' learning performance. Furthermore, empirical research was pursued by more influential authors, which indicated that our review should give more priority to the empirical direction of instructors in videotaped lectures. Through the meta-analysis via Stata/MP 14.0, the review revealed that better learning performance could be realized by integrating the instructor's gestures, eye gaze, and expressive faces with their speech in video lectures as measured by retention, transfer, and comprehensive tests. Furthermore, the authors found that the type of gestures and the topics of learning materials may moderate the effect of teachers' gestures.

6.2. Limitations. We would inevitably ignore some articles due to the limited availability of resources. Although we tried various terms on the Web of Science, a database with high reliability and extensive resources, to obtain as many documents as possible, we do not ensure an exhaustive retrieve in the field. Another important influencing factor lies in a biased selection of publications concerning the language. Most articles in our current review were written in English, although we conducted a complementary search by examining the references of selected articles. Admittedly, the small effect size of instructors' eye gaze and facial expressions may be closely related to the small sample size due to the limitations of empirical evidence. The insufficient empirical evidence could also lead to our failure to conduct meta-analyses concerning the effect of critical moderators on the relationship between teaching behaviors and learning performance.

6.3. Future Research Directions. The current review presented the potential of teachers' gestures in videotaped lectures which could give new insights into instructional practices and students' learning. Specifically, instructors with deictic gestures tend to point to specific teaching content with fingers or hands. Teachers should make good use of deictic gestures which contribute to attracting more prominent attention to key information in education. By gesturing themselves, instructors can optimize learners' attention distribution, thus helping learners to invest more cognitive resources in the effective organization and integration of learning content. In addition to the guidance of learners' attention, teachers' gestures can scaffold their verbal information. Teachers with depictive gestures usually choose to indicate semantic content in videos literally or metaphorically. For example, teachers can use gestures in place of physical objects that are too cumbersome to adopt in lectures. Depictive gestures can also be used to demonstrate molecules that are challenging to represent because of their small sizes. In sum, gestures, as teaching tools, have the potential to improve learning contexts.

Another insight for the practice of onscreen teachers is to make further efforts to focus on gaze shifts between the relevant content of the board and the audience while lecturing. According to our review, learners perform better in instructional videos with the teacher's eye gaze in terms of comprehensive and retention tests. These findings shed light on the design of video lectures that should involve an instructor who looks at the learners while talking and occasionally shifts eye gaze to the materials of the board rather than looking directly at the audience. It is intended to reduce learners' cognitive burden and guide their attention with gaze shift by suggesting the relevant portion of learning materials. In addition, the review demonstrates that video designers should be aware of the power of expressional faces displayed by teachers. Onscreen teachers in instructional videos also need to attach more importance to showing their

expressional faces, especially positive expressions, as much as possible during lectures.

Regarding the beat gestures, we concluded that their effect size appeared to be much lower than other gestures, including pointing and depictive gestures, in terms of retention and transfer tests. However, a few studies still reported that beat gestures are beneficial for learners' performance [55]. Future research could further conduct moderator analyses to explore whether other influencing factors in instructional videos moderate the effect of instructors' beat gestures, thus proving the rationality of their existence. Notably, the small sample size concerning the instructor's emotional face is mainly attributed to the fact that there are not many empirical data included in the meta-analysis. According to the bibliographic results, facial expressions attracted researchers' increasing attention. But the empirical studies on facial expressions need more researchers to examine their influence in the future.

In addition, we put the impact of teachers' characteristics on learning performance at the top of our research. Therefore, it remains an open question due to the limitation of empirical research if instructors' features have congruent positive effects on learners' visual attention distribution and learning perceptions in multimedia learning. In the future, researchers can devote themselves to exploring these issues. In the current review, we have discussed the moderate effect of the type of gestures and knowledge on the relationship between teaching behaviors and learning outcomes. The studies should emphasize learners' differences to enhance the quality of instructional videos [56]. Further works still need to investigate whether the learners' individual differences, such as age, prior knowledge, and learning preference, mediate the effect of teaching behaviors.

With further research on teachers' behaviors in videos, researchers have tried to shift their research focus from the effect of teaching behaviors on learning performance to revealing the underlying cognitive neural mechanism. However, previous studies have only focused on how instructors' hand movements affect learners' performance by their cognitive neural mechanism. Little attention has been paid to revealing the underlying cognitive activity concerning the effect of other teaching behaviors, such as eye gaze, facial expressions, and voice. Furthermore, the researchers devoted themselves to observing the learners' EEG components, neural oscillation strength, and brain activation when the instructor's gestures accompanied their speech. Future research should strengthen the link between cognitive neuroscience evidence and behavioral data. In other words, researchers can investigate the relationship between students' brain area activation and their learning effects when teaching videos with or without teacher gestures, eye gaze, or other teaching behaviors.

In accordance with the social presence theory, the multimedia instruction should trigger learners' social interaction schema in minds, which encourages learners to assume the virtual image as a social partner [53]. Whether an instructor is appropriate as a social partner mainly depends on learners' perceptions concerning the instructor's social quality [53]. The instructor's voice, as one of the vital instructional factors, affects the learners' perception of social qualities. Liew et al. [57] explored if a speaker with an enthusiastic voice will differ from a speaker with low enthusiasm based on perceived social presence, learning outcomes, and the cognitive load. The researchers found that the instructor's enthusiastic voice could improve social agency compared to a calm voice, thus leading to better transfer performance. They needed to make additional efforts to examine which aspects of voice serve as the most critical factors in promoting deep learning. Future research can try to explore other characteristics of instructors, such as their accent (foreign vs. human accent), gender (female vs. male voice), mechanization (machine vs. human voice), and dialect (standard dialect vs. regional speech).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Funding

Supplementary Materials

Supplementary 1. PRISMA guideline.

Supplementary 2. Appendix A: The STARLITE Evaluation.

Supplementary 3. Appendix B: The coding of studies.

References

[1] Z. Yu, "The effects of gender, educational level, and personality on online learning outcomes during the COVID-19 pandemic," *International Journal of Educational Technology in Higher Education*, vol. 18, Article ID 14, 2021.

[2] H. Polat, "Instructors' presence in instructional videos: a systematic review," *Education and Information Technologies*, 2022.

[3] Z. Yu, "The effect of teacher presence in videos on intrinsic cognitive loads and academic achievements," *Innovations in Education and Teaching International*, vol. 59, no. 5, pp. 574–585, 2022.

[4] T. Colliot and É. Jamet, "Understanding the effects of a teacher video on learning from a multimedia document: an eye-tracking study," *Educational Technology Research and Development*, vol. 66, pp. 1415–1433, 2018.

[5] M. van Wermeskerken, S. Ravensbergen, and T. van Gog, "Effects of instructor presence in video modeling examples on attention and learning," *Computers in Human Behavior*, vol. 89, pp. 430–438, 2018.

[6] R. F. Kizilcec, J. N. Bailenson, and C. J. Gomez, "The instructor's face in video instruction: evidence from two large-scale field studies," *Journal of Educational Psychology*, vol. 107, no. 3, pp. 724–739, 2015.

[7] J. Wang and P. D. Antonenko, "Instructor presence in instructional video: effects on visual attention, recall, and perceived learning," *Computers in Human Behavior*, vol. 71, pp. 79–89, 2017.

[8] J. Llanes-Coromina, I. Vilà-Giménez, O. Kushch, J. Borràs-Comes, and P. Prieto, "Beat gestures help preschoolers recall and comprehend discourse information," *Journal of Experimental Child Psychology*, vol. 172, pp. 168–188, 2018.

[9] Ł. Okruszek, A. Bala, M. Dziekan et al., "Gaze matters! The effect of gaze direction on emotional enhancement of memory for faces in patients with mesial temporal lobe epilepsy," *Epilepsy & Behavior*, vol. 72, pp. 35–38, 2017.

[10] O. Kushch, A. Igualada, and P. Prieto, "Prominence in speech and gesture favour second language novel word learning," *Language, Cognition and Neuroscience*, vol. 33, no. 8, pp. 992–1004, 2018.

[11] Z. Pi and J. Hong, "Learning process and learning outcomes of video podcasts including the instructor and PPT slides: a Chinese case," *Innovations in Education and Teaching International*, vol. 53, no. 2, pp. 135–144, 2016.

[12] S. Goldin-Meadow, "How gesture works to change our minds," *Trends in Neuroscience and Education*, vol. 3, no. 1, pp. 4–6, 2014.

[13] Yang, Zhang, Li, and Pi, "The interaction of an instructor's guided behaviors and learners' prior knowledge in predicting learning from video," *China Educational Technology*, no. 7, Article ID 8, 2019.

[14] E. Ozcelik, I. Arslan-Ari, and K. Cagiltay, "Why does signaling enhance multimedia learning? Evidence from eye movements," *Computers in Human Behavior*, vol. 26, no. 1, pp. 110–117, 2010.

[15] E. Alemdag, "Effects of instructor-present videos on learning, cognitive load, motivation, and social presence: a meta-analysis," *Education and Information Technologies*, vol. 27, pp. 12713–12742, 2022.

[16] M. L. Henderson and L. Schroeder, "A systematic review of instructor presence in instructional videos: effects on learning and affect," *Computers and Education Open*, vol. 2, Article ID 100059, 2021.

[17] S. P. Nerur, A. A. Rasheed, and V. Natarajan, "The intellectual structure of the strategic management field: an author co-citation analysis," *Strategic Management Journal*, vol. 29, no. 3, pp. 319–336, 2008.

[18] K. Ouwehand, T. van Gog, and F. Paas, "Designing effective video-based modeling examples using gaze and gesture cues," *Journal of Educational Technology & Society*, vol. 18, no. 4, pp. 78–88, 2015.

[19] M. van Wermeskerken and T. van Gog, "Seeing the instructor's face and gaze in demonstration video examples affects attention allocation but not learning," *Computers & Education*, vol. 113, pp. 98–107, 2017.

[20] Z. Yu, "Visualizing artificial intelligence used in education over two decades," *Journal of Information Technology Research*, vol. 13, no. 4, pp. 32–46, 2020.

[21] Z. Pi, Y. Zhang, and J. Yang, "The influence of instructor's gestures on video learning and its cognitive neural mechanism," *China Educational Technology*, vol. 4, Article ID 9, 2019.

[22] R. Arnheim, "Hand and mind: what gestures reveal about thought," *Leonardo*, vol. 27, no. 4, p. 358, 1994.

[23] S. Kang, G. L. Hallman, L. K. Son, and J. B. Black, "The different benefits from different gestures in understanding a concept," *Journal of Science Education and Technology*, vol. 22, pp. 825–837, 2013.

[24] A. Igualada, N. Esteve-Gibert, and P. Prieto, "Beat gestures improve word recall in 3- to 5-year-old children," *Journal of Experimental Child Psychology*, vol. 156, pp. 99–112, 2017.

[25] N. Sweller, A. Shinooka-Phelan, and E. Austin, "The effects of observing and producing gestures on Japanese word learning," *Acta Psychologica*, vol. 207, Article ID 103079, 2020.

[26] A. Macoun and N. Sweller, "Listening and watching: the effects of observing gesture on preschoolers' narrative comprehension," *Cognitive Development*, vol. 40, pp. 68–81, 2016.

[27] L. Rueckert, R. B. Church, A. Avila, and T. Trejo, "Gesture enhances learning of a complex statistical concept," *Cognitive Research Principles and Implications*, vol. 2, Article ID 2, 2017.

[28] S. D. Kelly, T. McDevitt, and M. Esch, "Brief training with co-speech gesture lends a hand to word learning in a foreign language," *Language and Cognitive Processes*, vol. 24, no. 2, pp. 313–334, 2009.

[29] Z. Pi, F. Zhu, Y. Zhang, L. Chen, and J. Yang, "Complexity of visual learning material moderates the effects of instructor's beat gestures and head nods in video lectures," *Learning and Instruction*, vol. 77, Article ID 101520, 2022.

[30] W. C. So, C. S. Chen-Hui, and J. L. Wei-Shan, "Mnemonic effect of iconic gesture and beat gesture in adults and children: is meaning in gesture important for memory recall?" *Language and Cognitive Processes*, vol. 27, no. 5, pp. 665–681, 2012.

[31] Y. Wang, Q. Liu, W. Chen, Q. Wang, and S. David, "Effects of instructor's facial expressions on students' learning with video lectures," *British Journal of Educational Technology*, vol. 50, no. 3, pp. 1381–1395, 2019.

[32] T. van Gog, I. Verveer, and L. Verveer, "Learning from video modeling examples: effects of seeing the human model's face," *Computers & Education*, vol. 72, pp. 323–327, 2014.

[33] J. Borup, R. E. West, and C. R. Graham, "Improving online social presence through asynchronous video," *The Internet and Higher Education*, vol. 15, no. 3, pp. 195–203, 2012.

[34] P. Ayres and J. Sweller, "The split-attention principle in multimedia learning," in *The Cambridge Handbook of Multimedia Learning*, R. Mayer, Ed., pp. 206–226, Cambridge University Press, Cambridge, 2014.

[35] J. Zhang, M.-L. Bourguet, and G. Venture, "The effects of video instructor's body language on students' distribution of visual attention: an eye-tracking study," in *Proceedings of the 32nd International BCS Human Computer Interaction Conference (HCI)*, pp. 1–5, BCS Learning and Development Ltd., Belfast, UK, July 2018.

[36] D. Moher, A. Liberati, J. Tetzlaffff, D. G. Altman, and PRISMA Group, "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement," *Annals of Internal Medicine*, vol. 151, no. 4, pp. 264–269, 2009.

[37] Z. Yu, M. Gao, and L. Wang, "The effect of educational games on learning outcomes, student motivation, engagement and satisfaction," *Journal of Educational Computing Research*, vol. 59, no. 3, pp. 522–546, 2021.

[38] A. Booth and MCLIP, "Brimful of STARLITE": toward standards for reporting literature searches," *Journal of the Medical Library Association: JMLA*, vol. 94, no. 4, pp. 421–429, e05, 2006.

[39] J. Higgins, S. G. Thompson, J. J. Deeks, and D. G. Altman, "Cochrane handbook for systematic reviews of interventions version 5.1.0. the Cochrane collaboration," *Naunyn-Schmiedebergs*

Archiv für Experimentelle Pathologie und Pharmakologie, vol. 5, no. 2, p. S38, 2008.

[40] Z. Pi, M. Chen, F. Zhu, J. Yang, and W. Hu, "Modulation of instructor's eye gaze by facial expression in video lectures," *Innovations in Education and Teaching International*, vol. 59, no. 1, pp. 15–23, 2022.

[41] Z. Pi, Y. Zhang, F. Zhu, L. Chen, X. Guo, and J. Yang, "The mutual influence of an instructor's eye gaze and facial expression in video lectures," *Interactive Learning Environments*, pp. 1–18, 2021.

[42] M. Chen, *The influence of teachers' facial expressions and eye gaze on learners' learning in teaching videos*, Master's Theory, Huazhong Normal University, 2020.

[43] T. Horovitz and R. E. Mayer, "Learning with human and virtual instructors who display happy or bored emotions in video lectures," *Computers in Human Behavior*, vol. 119, Article ID 106724, 2021.

[44] A. P. Lawson, R. E. Mayer, N. Adamo-Villani, B. Benes, X. Lei, and J. Cheng, "The positivity principle: do positive instructors improve learning from video lectures?" *Educational Technology Research and Development*, vol. 69, pp. 3101–3129, 2021.

[45] S. Ba, D. Stein, Q. Liu, T. Long, K. Xie, and L. Wu, "Examining the effects of a pedagogical agent with dual-channel emotional cues on learner emotions, cognitive load, and knowledge transfer performance," *Journal of Educational Computing Research*, vol. 59, no. 6, pp. 1114–1134, 2021.

[46] T. W. Liew, N. A. Mat Zin, and N. Sahari, "Exploring the affective, motivational and cognitive effects of pedagogical agent enthusiasm in a multimedia learning environment," *Human-centric Computing and Information Sciences*, vol. 7, Article ID 9, 2017.

[47] R. E. Mayer, *The Cambridge Handbook of Multimedia Learning*, Cambridge University Press, 2nd edition, 2014.

[48] C. Frechette and R. Moreno, "The roles of animated pedagogical agents' presence and nonverbal communication in multimedia learning environments," *Journal of Media Psychology: Theories Methods, & Applications*, vol. 22, no. 2, pp. 61–72, 2010.

[49] R. E. Mayer, "Evidence-based principles for how to design effective instructional videos," *Journal of Applied Research in Memory and Cognition*, vol. 10, no. 2, pp. 229–240, 2021.

[50] Y. Tian, Z. Qi, X. Huang, H. Xiang, and Y. Wang, *Cognitive Neural Mechanism of Social Cues Promoting Online Learning*, China Educational Technology, 2021.

[51] A. Paivio, *Mind and its Evolution: A Dual Coding Theoretical Approach*, Psychology Press, New York, 1st edition, 2014.

[52] C. N. Gunawardena, "Social presence theory and implications for interaction and collaborative learning in computer conferences," *International Journal of Educational Telecommunications*, vol. 1, no. 2, pp. 147–166, 1995.

[53] R. E. Mayer, "Principles of multimedia learning based on social cues: personalization, voice, and image principles," in *The Cambridge Handbook of Multimedia Learning*, R. E. Mayer, Ed., pp. 201–212, Cambridge University Press, 2005.

[54] H. Xie, F. Wang, Z. Zhou, and P. Wu, "Cueing effect in multimedia learning: a meta-analysis," *Acta Psychologica Sinica*, vol. 48, no. 5, pp. 540–555, 2016.

[55] E. E. Austin, N. Sweller, and P. Van Bergen, "Pointing the way forward: gesture and adults' recall of route direction information," *Journal of Experimental Psychology: Applied*, vol. 24, no. 4, pp. 490–508, 2018.

[56] Q. Zhang, X. Chen, Y. Duan, and X. Yan, "Relationships between teacher presence and learning outcomes, learning perceptions, and visual attention distribution in videotaped lectures," *International Journal of Technology-Enhanced Education (IJTEE)*, vol. 1, no. 1, Article ID 79, 2022.

[57] T. W. Liew, S.-M. Tan, T. M. Tan, and S. N. Kew, "Does speaker's voice enthusiasm affect social cue, cognitive load and transfer in multimedia learning? Voice enthusiasm," *Information and Learning Sciences*, vol. 121, no. 3/4, pp. 117–135, 2020.