Gelar: Jurnal Seni Budaya
Vol 24, No. 1, June 2026 pp. 167-181
https://jurnal.isi-ska.ac.id/index.php/gelar/index
E-ISSN 2655-9153

Video-based visual micro-skills for correcting trumpet embouchure and breathing errors: a convergent mixed-methods study

Heri Murbiyantoro [a,1,*], Ibnu Syahruroji Fiqriannur [a,2], Harpang Yudha Karyawanto [a,3]; Grace Natalie [b,4]

[a] Faculty of Language and Arts, Universitas Negeri Surabaya, Indonesia
[b] Berklee College of Music, Boston, Massachusetts 02215, United States
[1] herimurbiantoro@unesa.ac.id*; [2] 240208650@mhs.unesa.ac.id; [3] harpangkaryawanto@unesa.ac.id; [4] gnatalie@berklee.edu
* Corresponding Author

ABSTRACT

Digital media research in music education has largely evaluated outcomes, engagement, and motivation, leaving unexplained how visual representation actually reconstructs technical understanding at the level of fine motor coordination. This study tests and explains the effect of visual micro skills the explicit video representation of rapid, partly unobservable coordinations of embouchure, air-stream direction, and respiratory support on the correction of technical errors in trumpet playing. A convergent mixed-methods design was applied to 25 purposively selected participants: undergraduate wind-instrument students and community drum-band players in Surabaya, aged 18–24 years, with 1–3 years of trumpet experience and no conservatory-level formal training. The intervention consisted of a single 120-minute structured session organised into five stages (misconception framing, global viewing, focused observation, guided practice with pause, and reflection), followed by a 14-day monitored independent-practice phase. Technique was scored with a 12-indicator analytic rubric covering embouchure control, airflow control, and respiratory support, rated by two blinded independent assessors. Quantitative and qualitative strands were collected in parallel and merged through a joint display. Scores rose from 62.00 ($SD = 2.94$) to 84.72 ($SD = 2.62$), $t(24) = 25.03$, $p < 0.001$, $d_z = 5.01$, 95% CI [20.85, 24.59]. Thematic analysis produced four themes: visual referencing replacing kinaesthetic guessing, decomposition of a fused action into observable components, relocation of corrective authority to the learner, and persistence of compensatory tension under performance pressure. Findings position visual micro-skills as an active pedagogical mechanism rather than a delivery medium. The single-group design and small purposive sample restrict causal and population inference.

OPEN ACCESS

Article History
Received 2026-04-18
Revised 2026-06-22
Accepted 2026-06-28

Keywords
Visual micro skills
Digital pedagogy
Trumpet technique
Embouchure
Mixed methods

1. Introduction

Trumpet playing is a fine motor skill in which sound is generated by a coupling that the player cannot see: vibrating lips inside a mouthpiece, an air stream whose direction and pressure are regulated by the abdominal and thoracic musculature, and a facial muscle configuration that must remain stable while the air column changes. Because sound production in brass instruments depends on this coupling, small deviations at the micro level produce disproportionately large consequences for pitch stability, timbre, and endurance [1], [2]. Motor habits formed in the earliest stages tend to be internalised as movement patterns that resist later correction, and sustained practice under a faulty configuration is associated not only with tone inconsistency and premature fatigue but with orofacial and respiratory strain that in the long term may develop into performance-related disorders [3]–[5]. The pedagogical problem is therefore not the absence of practice but the invisibility of the thing being practised.

The problem has escalated in severity. Instrumental learning has rapidly evolved into asynchronous learning via video. Consequently, a large number of students (those outside of conservatories) have begun to piece together their trumpet techniques by watching videos in the absence of a teacher [6], [7]. As such, learners have unprecedented access to more visual resources than ever. Although the dynamic of the teacher-student interaction has been dramatically altered, the teacher's diagnostic function has failed to migrate to the digital format. If a large number of self-motivated learners are using visual material to conduct instruction, then the methodology of visual media and technical understanding is no longer a methodological concern. Rather, it has become a necessary condition to ensure the safe and effective use of visual media. This study aims to explore this concern further. The study focuses specifically on self-motivated video learning in a semi-structured Indonesian learning environment. This is the model that typifies most self-motivated video learning.

Conventional instrumental pedagogy addresses the invisibility problem through two instruments: verbal instruction and live demonstration, and both are structurally limited for micro-level coordination. Verbal instruction encodes movement in language that is necessarily general; expressions such as "support the air" or "relax the corners" designate outcomes rather than the configurations that produce them, and learners resolve the underdetermination by importing idiosyncratic interpretations [8], [9]. Live demonstration is more concrete but is delivered at performance tempo, from a single viewing angle, and once: the learner sees the result of the coordination, not its components. Studies of coaching practice show that demonstrations are frequently supplied without the attentional scaffolding that would make the critical feature salient, so observers extract surface kinematics rather than the controlling variable [10]. The consequence in trumpet learning is a delayed and unreliable correction loop, in which the learner practises an unverified interpretation until it stabilises as habit [6].

Digital media appear to remedy exactly this, and a substantial literature now evaluates them. That literature, however, is dominated by a single evaluative logic. Systematic reviews of digital tools in music education aggregate effects on achievement, motivation, and engagement, and report them as indicators of success without opening the instructional process that produced them [11]. Reviews of game-based approaches similarly foreground participation and learning experience while treating technical acquisition as an undifferentiated outcome [12], [13]. Reviews of blended and project-based designs map trends and intellectual structure rather than mechanisms [14], [15]. Where visualisation itself is the object, the emphasis falls on classifying available technologies and their potential applications rather than on demonstrating how a specific visual affordance alters a specific technical component [7]. Even the more mechanism-oriented work in adjacent fields stops short of the level of resolution needed here: video feedback in physical education and surgical training is shown to improve performance, but the analysed skills are externally observable in full, so the question of representing a partly internal coordination never arises [16], [17]. The literature is thus not merely thin at the micro level; it is systematically organised around a unit of analysis, the learning outcome, that cannot answer a process question.

Existing theory is likewise informative but incomplete for this case. Social cognitive theory establishes that observation of a model supports skill acquisition through attention, retention, reproduction, and motivation, and predicts that fidelity of the model governs fidelity of the copy [18], [19]. It does not, however, specify what happens when the determining features of the model are not visible in the model. Stage accounts of motor learning describe a progression from cognitive to associative to autonomous control, and schema theory explains how repeated trials with knowledge of results build generalised motor programmes; both presuppose that the learner can identify the parameter to be varied, which is precisely what fails when the parameter is intra-oral [20], [21]. Research on attentional focus shows that directing attention externally, to the effect of movement rather than the body, yields superior learning a finding that reframes but does not resolve the problem, since in trumpet playing the proximate effect is itself an inaudible aerodynamic event [22].

The cognitive theory of multimedia learning supplies the processing constraints under which visual material is assimilated, including segmentation and signalling as means of managing intrinsic load, yet it is a theory of representation in general, indifferent to whether the

 ISSN 2655-9153

represented content is a diagram or a muscle [23]. Finally, accounts of deliberate practice specify that improvement requires informative feedback on a well-defined task; the trumpet case is one in which the task cannot be well defined for the learner because it cannot be perceived [24].

The concept of visual micro skills proposed here occupies the space these theories leave open. It designates a representational operation, not a medium: the deliberate decomposition of a fused instrumental action into its constituent micro coordinations lip contact and aperture, air-stream direction and continuity, diaphragmatic engagement and its synchronisation with sound onset each rendered separately observable through close-up framing, temporal dilation, freeze, and correct/incorrect juxtaposition. Three commitments distinguish it from the theories above. First, its unit of analysis is the micro coordination rather than the movement or the outcome, which makes it applicable precisely where a skill is only partly externalised. Second, it treats the visual field as something that must be constructed to make the controlling variable perceptible, rather than merely presented; the pedagogical work lies in the editing, not the recording. Third, it locates the mechanism of change in a perception-cognition-action loop that the learner can run independently, so that visualisation functions as a diagnostic instrument in the learner's own hands rather than as a substitute teacher. On this account, visualisation is an active pedagogical mechanism, an epistemological bridge from sensorimotor experience to explicit technical knowledge rather than an instructional aid.

The gap is therefore specific and statable. It is not known (i) whether a visual intervention constructed on micro-coordination principles produces measurable change in the specific technical components it targets, as distinct from global performance gains reported in outcome studies; (ii) through what pedagogical mechanism such change occurs, since existing studies do not observe the correction process; and (iii) which components resist correction, a question the outcome literature cannot pose at all because it does not disaggregate. Accordingly, this study pursues three objectives: (1) RO1. To determine whether a 120-minute visual micro-skill intervention followed by a monitored independent-practice phase produces a statistically significant change in rubric-assessed trumpet technique scores across embouchure control, airflow control, and respiratory support (H1: post-test > pre-test, $\alpha = 0.05$), and to quantify the magnitude of that change; (2) RO2. To explain, from parallel observational and reflective data, the pedagogical mechanism by which video-based visual representation reorganises learners' correction of embouchure and breathing errors; (3) RO3. To identify which technical components remain resistant to correction after the intervention, and under what conditions they recur. The contribution is twofold. Theoretically, the study advances visual micro-skills as a construct that specifies the representational conditions under which a partly invisible coordination becomes learnable, thereby extending digital pedagogy from a discourse about access and engagement toward one about diagnostic resolution. Practically, it yields an operationalised rubric and a staged intervention protocol that instrumental teachers can apply and audit. The scope is deliberately bounded: the study concerns two fundamental error domains in early-stage trumpet playing within one semi-formal Indonesian setting, and makes no claim to generalise across instruments, expertise levels, or populations.

2. Method

2.1. Research design

This study utilized a convergent mixed-methods design [25]. Quantitative and qualitative data were collected and analyzed simultaneously, and, at the level of interpretation, data were integrated using a joint display [26]. The design was convergent rather than explanatory sequential, because the two strands address complementary rather than sequential questions: the quantitative strand focuses on the degree of technique changes (RO1), while the qualitative strand focuses on the description of the correction process, which is not possible to identify after the fact. If the qualitative strand had occurred after the quantitative results, the participants' process accounts would be biased by the predicted gain in test score. Equal priority was given to each strand (QUAN + QUAL), and the point of interface was at the interpretation stage [27]. The intervention used a design-based educational inquiry. This involves a learning sequence that is both treatment and way of seeing learning dynamics [28]. The quantitative aspect consists of a single-group pre-test and post-test. The limitations for drawing causal relationships are presented in Section 3.8 and are the basis for the constraints of the claims discussed.

2.2. Participants, sampling, and ethics

Twenty-five purposively sampled participants, comprising of Wind Instrument pupils at Universitas Negeri Surabaya and members of community drum bands in Surabaya, formed part of this study. The selected participants must have had at least 6 months of continuous trumpet practice, have the ability to perform at least one octave of a sustained tone, and must not have had conservatory-level music training. Participants must also have had the availability to engage in the study sessions and the independent practice construct. Respiratory and orofacial health issues as well as engagement in other trumpet-structured programs, were exclusion criteria.

These selection criteria were established since the focus of the present study is the early-stage internalized error. For this to successfully take place, the participants must have had experience in trumpet playing for a period long enough for misconceptions to have developed and consolidated, but, at the same time, not too long for such misconceptions to have been formally and essentially corrected. Participant characteristics were: n men and n women; mean age M years (SD = SD, range 18–24); mean trumpet experience M years (SD = SD); n university students and n drum-band members. Sample size was determined by the accessible population rather than by power analysis; a post-hoc sensitivity analysis indicates that $n = 25$ with $\alpha = 0.05$ and power = 0.80 detects a paired effect of $d_z \geq 0.58$, comfortably below the observed effect. All participants gave written informed consent, including specific consent for video recording and for the use of still frames in publication. Participation was voluntary and unpaid, withdrawal was possible at any point without consequence, and all recordings were stored under coded identifiers. Ethical clearance was granted by an institutional body (approval no. number).

2.3. Intervention: content, structure, and duration

The intervention comprised two components with a total contact time of 120 minutes and a subsequent 14-day independent-practice phase. The instructional material consisted of two purpose-produced video modules addressing the two error domains identified at baseline: Module A Air Stream Direction and Embouchure Coupling and Module B Diaphragmatic Respiratory Support. Each module presents the target coordination four times: at performance tempo, in close-up at 0.25× speed, as a freeze-frame sequence with on-screen annotation of the controlling feature, and as a side-by-side juxtaposition of a correct and a faulty execution. These four presentations instantiate, respectively, the global reference, temporal dilation, signalling, and contrastive framing operations that define visual micro-skills as a representational construct [16], [23], [29]. The 120-minute session was delivered in five stages (Table 1). During the independent-practice phase, participants retained access to both modules and were asked to complete n practice sessions of n minutes, logging each session in a structured practice diary. Compliance was verified through the diaries and through weekly check-ins; participants completing fewer than n logged sessions were to be excluded, and n were excluded on this basis. The post-test was administered at the end of the phase.

2.4. Quantitative instrument and its psychometric properties

Technique was assessed with the Trumpet Micro-Technique Rubric (TMTR), developed for this study because no existing instrument disaggregates trumpet technique at the level of micro-coordination. The TMTR comprises three dimensions with four indicators each (Table 2), scored on a five-point behaviourally anchored scale (1 = error present throughout; 5 = target configuration maintained throughout). Raw scores (12–60) were linearly transformed to a 0–100 scale to aid interpretation. Each participant performed a standardised task a sustained tone at three dynamic levels, a five-note slurred ascending pattern, and an eight-bar excerpt specify excerpt recorded frontally and laterally under identical conditions at both time points. Content validity was established by three independent expert judges (a trumpet performance specialist, a wind-instrument pedagogue, and a music-education measurement specialist), each rating every indicator for relevance and clarity. Aiken's V ranged from V to V (mean V = V), exceeding the critical value of 0.80 for three raters on a five-point scale; n indicators were reworded following judge comments [30]. Internal consistency on the pre-test data was Cronbach's α = α for the total scale and α–α across the three dimensions. Inter-rater reliability was computed on all 50 recordings independently double-rated by two assessors, yielding a two-way random-effects, absolute-agreement, single-measures intraclass correlation coefficient of ICC(2,1) = ICC, 95% CI [lo, hi], interpretable as good/excellent agreement [31]. Analyses used the mean of the two raters' scores.

2.5. Qualitative data collection

Three qualitative sources were used, all generated within the study; no literature was treated as qualitative data. First, structured non-participant observation of every session and of n sampled practice periods, recorded on a protocol with four pre-specified fields (observed technical action, visual feature attended to, self-correction attempted, outcome) plus an open field for unanticipated events. Second, structured practice diaries in which participants recorded, per session, what they attempted, what they compared their playing against, and what they changed. Third, semi-structured individual reflective interviews of 20–30 minutes conducted at the end of the independent-practice phase, following a guide with four core prompts (append as supplementary material) and non-directive probes. Interviews were audio-recorded and transcribed verbatim. Data collection continued until no new codes emerged from three consecutive transcripts.

2.6. Control of subjective bias

Six procedures were used to constrain assessor and researcher subjectivity. (i) Assessor blinding: pre- and post-test recordings were stripped of identifying metadata, assigned random codes, and presented to raters in randomised order, so that raters could not know which recording preceded the intervention. (ii) Assessor independence: the two raters were qualifications who had no role in delivering the intervention and no prior contact with participants; they rated independently and were not permitted to confer. (iii) Calibration: before scoring, raters jointly scored five training recordings not included in the dataset until discrepancies of more than one scale point were eliminated. (iv) Discrepancy resolution: indicator-level differences exceeding one scale point (n instances, n% of ratings) were adjudicated by a third rater whose score was final. (v) Separation of roles: the researcher who facilitated the session did not code qualitative data. (vi) Analytic transparency: a reflexive journal recorded interpretive decisions, and an audit trail of raw data, codes, and category revisions was maintained throughout.

2.7. Data analysis

Quantitative data were analysed in SPSS v.__ / JASP v.__. Normality of the difference scores was checked with the Shapiro–Wilk test (W = W, p = p) and inspection of Q–Q plots; the assumption was met. A paired-samples t-test compared pre- and post-test composite scores, with the effect size reported as Cohen's dz together with the 95% confidence interval of the mean difference [32]. Dimension-level scores were analysed separately with Bonferroni-adjusted paired tests ($\alpha = 0.017$) to identify differential change across the three components, addressing RO3. Qualitative data were analysed by reflexive thematic analysis following the six-phase procedure, an approach with established precedent in instrumental music research: familiarisation, generation of initial codes, construction of candidate themes, review against the coded extracts and the full dataset, definition and naming, and reporting [33], [34]. Coding was inductive at the semantic level and latent at the interpretive stage. Two coders independently coded 30% of the corpus; intercoder agreement was Krippendorff's α = α, and remaining disagreements were resolved discursively [35]. Draft themes were returned to n participants for member checking, and their responses were incorporated as a further layer of interpretation rather than as verification of a single correct reading [36].

2.8. Integration procedure

The two strands were merged through a joint display in which each dimension-level quantitative result is placed alongside the themes bearing on that dimension, and each pairing is classified as convergent, complementary, or discordant [26], [27]. Discordant pairings were retained and interrogated rather than reconciled, since divergence between measured gain and observed process is itself evidence about the mechanism. Trustworthiness was addressed

 ISSN 2655-9153

through triangulation across the three qualitative sources (credibility), thick description of the setting and the intervention (transferability), the audit trail (dependability), and the reflexive journal (confirmability).

3. Results and Discussion

3.1. Baseline error domains: airflow–embouchure coupling and respiratory support

Pre-test analysis located participants' technical errors in two domains, which are named here as analytic constructs rather than instructional labels: airflow embouchure coupling, concerning the relationship between lip configuration and the direction and continuity of the air-stream entering the mouthpiece; and respiratory support regulation, concerning the recruitment of diaphragmatic rather than thoracic mechanisms to supply that air-stream. In the first domain, errors appeared as excessive labial pressure against the mouthpiece rim and an asymmetric aperture, which redistributed the air-stream away from the mouthpiece throat. In the second, they appeared as clavicular and upper-thoracic inhalation through the nose with minimal abdominal displacement, producing an air supply that was both limited in volume and unstable in pressure. As can be seen in Fig. 1.

Fig 1. Baseline manifestations of the two error domains: mouthpiece-contact panels showing labial in-folding and forward placement, and the inhalation panel showing thoracic-dominant, nasally initiated breathing without abdominal displacement.

Fig. 1 has two configurations that serve as empirical bases for the subsequent arguments and both can be interpreted mechanistically. The mouthpiece contact panels show the mouthpiece sitting on the lower lip, which has now folded inwards and is situated along the lower lip's labial surface. According to the regenerative model of sound production in brass, the lips act as a valve whose pressure is controlled and whose oscillation is determined by the coupling of the lip's tissue to the air column [2], [37]. Increasing stiffness and decreasing the effective diameter of the lip folds results in the need for the brass player to create greater air pressure in order to start the oscillation. The fact that brass players often report the feeling of "holding in" the air as a means of pressure control is an objective measure of the mechanism that opposes the air release control. The tradeoff is limited control of the pitch while the air pressure is sustained in the lips. An electromyographical analysis of student and elite trumpet players reveals that for the same sound production, students demonstrate a greater engagement of the muscles governing lip pressure. The negative effects of this increase in pressure has been linked to an increase in the fatigue and strain of the muscles in the mouth [4], [38].

The inhalation panel demonstrates the respiratory correlate. Control of pressure during slow, regulated exhalation will be weakened if thoracic-dominant inhalation is practiced, as the inspiratory reserve will also shorten. Research into the control of exhalation conducted on expert trumpeters shows that the economy of the larger systems of the body was more vital to the performance than the control of the muscular systems [39]–[41]. If coordination of abdominal pressure is absent, then adequate pressure control will be implemented at the embouchure. This is why in performance the two areas are fundamentally interconnected. This is to say that a respiratory control error will be manifested at the embouchure. The analytical system employed in this research was impacted directly by this phenomenon. The reason the intervention needed to include both areas, while the analysis viewed the two areas separately, was because the two areas were interconnected so that a respiratory control error would be

 ISSN 2655-9153

manifested at the embouchure. Reading Fig. 1 against the theoretical position set out in Section 1 clarifies what the outcome-oriented literature cannot supply. Reviews of digital tools in music education report gains without identifying the configuration that produced them, the present analysis identifies a specific, mechanistically explicable configuration, which is what makes a targeted representational intervention possible [11], [12]. It also shows why conventional instruction fails here: verbal correction ("relax", "breathe deeply") addresses the compensation rather than the valve condition that necessitates it, while a real-time demonstration displays a correct configuration without making the *difference* between it and the learner's own configuration perceptible [8], [10]. The intervention was therefore designed around contrast and dilation rather than around exposure, which is the operational meaning of the visual micro-skills construct and the direct link between the baseline finding and RO1–RO2.

3.2. Representational structure of the intervention

Fig. 2 and Fig. 3 show representative frames from the two modules. The pedagogical work is visible in the editing decisions rather than in the content: the framing isolates a single controlling feature, the annotation directs attention to it, and the juxtaposition supplies the comparison against which the learner's own configuration becomes interpretable.

Fig 2. Module A, *Air-Stream Direction and Embouchure Coupling*: close-up framing of aperture formation (left) and annotated air-stream trajectory relative to the mouthpiece throat (right).

Three properties distinguish this material from conventional demonstration. Temporal dilation decouples events that occur too rapidly to be resolved at tempo, converting a single perceived act into an ordered sequence; this is the segmentation principle of multimedia learning applied to a motor rather than a symbolic content [23]. Signalling through annotation performs the attentional work that unguided demonstration leaves to chance, and by directing attention to the air-stream effect rather than to the muscles producing it, it induces the external focus associated with superior motor learning [10], [22]. Contrastive juxtaposition supplies a discriminative rather than merely a normative reference: the learner is given not only the target but the dimension along which their own execution departs from it. Together these establish the conditions under which observational learning can operate on a partly invisible skill, extending rather than restating the model-fidelity account [18], [19], [42]. As shown in Fig. 3.

Fig 3. Module B, *Diaphragmatic Respiratory Support*: lateral framing of abdominal displacement during inhalation (left) and synchronisation of expiratory release with tone onset (right).

Table 1 sets out how these representational properties were sequenced across the 120-minute session.

Tabel 1. Stages, duration, and pedagogical function of the visual micro-skill intervention (total contact time 120 minutes)

Stage	Activity	Pedagogical function	Duration
Misconception framing	Facilitator identifies the two error domains and their consequences	Activates awareness of a specific discrepancy as the object of correction	10 min
Global viewing	Uninterrupted viewing of both modules at performance tempo	Establishes a holistic reference representation before decomposition	20 min
Focused observation	Viewing at 0.25× with freeze-frame and annotation	Resolves micro-coordinations that are unresolvable at tempo; signals the controlling feature	35 min
Guided practice with pause	Alternating playing and paused comparison against the model	Couples perception to action within a single cycle; enables discriminative self-correction	40 min
Reflection and consolidation	Structured discussion comparing self-performance with the model	Converts episodic correction into an explicit, restatable technical rule	15 min

The sequence is cumulative rather than additive. The first two stages construct a reference against which discrepancy can be perceived; the third makes the discrepancy perceptible; the fourth closes the perception–action loop while the visual reference is still available; the fifth externalises the resulting adjustment as a rule the learner can apply without the video. This ordering is what distinguishes a visual micro-skill intervention from repeated exposure to demonstration material, and it is the mechanism proposed in answer to RO2.

3.3. Quantitative change in technique

Technique was operationalised through the three-dimensional rubric shown in Table 2, each dimension carrying four behaviourally anchored indicators.

Tabel 2. Dimensions and indicators of the Trumpet Micro-Technique Rubric (TMTR)

Dimension	Indicators (scored 1–5 each)
Embouchure control	(1) Mouthpiece placement without excessive labial pressure; (2) aperture stability during sustained tone; (3) absence of compensatory perioral tension; (4) consistency of embouchure formation across repeated onsets
Airflow control	(1) Air-stream direction toward the mouthpiece throat; (2) continuity of flow across the phrase; (3) tone stability free of cracking or pitch drift; (4) efficiency of air use relative to phrase length
Respiratory support	(1) Abdominal rather than clavicular inhalation; (2) depth of inspiration relative to phrase demand; (3) graded control of expiration during playing; (4) synchronisation of expiratory onset with tone onset

Composite scores are reported in Table 3. The difference scores were normally distributed (Shapiro–Wilk W = W, p = p), satisfying the assumption for the paired test.

Tabel 3. Descriptive statistics and paired-samples t-test for composite technique scores (n = 25)

Measurement	n	Mean	SD	Mean difference	t(24)	p
Pre-test	25	62.00	2.94	—	—	—
Post-test	25	84.72	2.62	22.72	25.03	< 0.001

The composite score improved significantly by an average of 22.72 points, $t(24) = 25.03$, $p < 0.001$, indicating a statistically significant result with a 95% confidence level of the mean

 ISSN 2655-9153

difference being between 20.85 and 24.59, with a standardized effect of dz = 5.01. The lower limit of the confidence interval is 20.85. The reduction in standard deviation from 2.94 to 2.62 indicates that participants converged and improved, and that the effect was consistent with the standardizing function resulting from a shared visual reference. Given the context of a single-group design and a short interval, however, the improvement of such a large magnitude should be treated with caution. Though this design is susceptible to the testing, expectancy and maturation effects, the value represents the degree of change participants experienced under such effects. The value is not intended to represent the direct effect of the intervention. Dimension-level analysis addresses RO3 and reveals that the aggregate conceals differential movement. Gains were largest in airflow control ($\Delta = \Delta$, $p = p$) and smallest in embouchure control ($\Delta = \Delta$, $p = p$), with respiratory support intermediate ($\Delta = \Delta$, $p = p$). The ordering is mechanistically interpretable: air-stream direction is the component most directly visualisable by annotation and most amenable to an external attentional focus, whereas embouchure configuration is partly occluded by the mouthpiece itself, so that the learner must infer the target state from its consequences rather than observe it [22]. This is the first point at which the disaggregated design yields information that a global outcome measure could not, and it locates a boundary condition on the visual micro-skills construct: representational resolution declines as physical occlusion increases.

3.4. Qualitative themes

Reflexive thematic analysis of observation protocols, practice diaries, and reflective interviews produced four themes. The first three describe the mechanism of change; the fourth is a counter-theme that constrains it.

- Theme 1: From kinaesthetic guessing to visual referencing. Participants described their pre-intervention correction strategy as a search conducted by feel, in which adjustments were retained if the resulting sound improved and abandoned otherwise, with no representation of what had been changed. Following the intervention, correction relied more on an external referent, as participants identified checking a specific visual element, being aperture symmetry and abdominal displacement, instead of appraising an indistinct sensation. The change is not from lack of feedback to more feedback, but rather feedback that is specific compared to confounded feedback (sound as the only indicator to every component at once). This change directly addresses one of the conditions of deliberate practice accounts, where these conditions remain mostly unseen, and gives rationale as to why specific and normative auditory feedback is unproductive for novice performers.
- Theme 2: Decomposition of a fused action. Before the intervention, participants considered “playing a note” a singular act. Before the intervention, participants considered “playing a note” to be a single action; correction attempts were recorded that modified multiple parameters at once, thus the recorded improvements could not be traced to a specific source. After seeing dilated and annotated materials, participants then were able to systematically modify individual parameters, first adjusting inhalation without changing embouchure and then changing embouchure without altering inhalation. This is the segmentation principle at work on motor content. This is the segmentation principle at work on motor content [23]. Decomposition, on these grounds, is not a by-product of the medium, but the specific achievement of the representational operations defined in Section 3.2.
- Theme 3: Relocation of corrective authority. Participants shifted from waiting for external assessments to doing their own self-assessments on the model. Records show that instead of asking the moderator "is this correct?", the participants started asking moderators to confirm their decisions. Ultimately self-evaluation was possible thanks to the stable, replayable annotated reference, along with the observations that the self-regulation of instrumental behavior depends more on what is available to the learner than the disposition of the self-regulating learner, and it qualifies the standard framing of self-regulation as a characteristic of the learner: in this case it was a characteristic of the representation.

- Theme 4: Persistence of compensatory tension under performance pressure (counter-theme). In the excerpt task and during observed practice at higher dynamics and in the upper register, n participants reverted to elevated labial pressure and shortened inhalation, including several whose sustained-tone indicators had improved. Participants described knowing the correct configuration while being unable to maintain it when demand increased. Visual micro-skills therefore altered technical *understanding* and its application under controlled conditions more readily than they altered *automaticity* under load, consistent with the expectation that the transition to autonomous control requires extended practice beyond any representational intervention [20], [21]. This theme is the qualitative counterpart of the dimension-level asymmetry in Section 3.3 and is the principal internal check on over-reading the effect size.

3.5. Integration of quantitative and qualitative findings

The display shows that no pairing was fully discordant, but that the single complementary pairing carries the greater analytic weight. The magnitude of the composite gain, taken alone, would license a strong claim about the effectiveness of visual representation; the qualitative strand shows that this gain is unevenly distributed and unstable under load. Integration therefore functions here as a constraint on inference rather than as mutual corroboration, which is the substantive argument for having chosen a convergent design [27]. Table 4 presents the joint display merging the two strands at dimension level.

Tabel 4. Joint display integrating dimension-level quantitative results with qualitative themes

Dimension	Quantitative result	Related qualitative evidence	Integration
Airflow control	Largest gain (Δ)	Themes 1 and 2: annotated air-stream trajectory gave a directly checkable referent; component isolated from embouchure adjustment	Convergent — visual annotation and measured gain point the same way
Respiratory support	Intermediate gain (Δ)	Themes 2 and 3: abdominal displacement externally observable, enabling independent self-check	Convergent
Embouchure control	Smallest gain (Δ)	Theme 4: reversion to labial pressure under load; target state occluded by the mouthpiece	Complementary — qualitative data explain the attenuated gain as an occlusion limit
Composite score	d_z = 5.01; SD narrowed 2.94 → 2.62	Theme 3: shared model reduced interpretive variance across participants	Convergent, with the discordance in Theme 4 retained as a scope condition

3.6. Digital pedagogy affordances underlying the mechanism

The findings above are attributable not to video as a technology but to specific affordances that the design mobilised. Table 5 maps each affordance to its pedagogical function and to its expression in trumpet micro-technique. Asynchronous access and replay are the affordances most extensively documented in the existing literature, where they are associated with flexibility, engagement, retention, and self-regulated study [43]–[46].

The present findings are consistent with that literature but reposition it: in this study neither affordance produced change on its own. Their function was to make the remaining four affordances repeatedly available. Dilation, signalling, contrast, and a shared model are the operations that carry the pedagogical work, and they are properties of how the material was constructed rather than of the delivery channel. This distinction matters for the field, because it implies that studies reporting null or weak effects of video in instrumental learning may be reporting on the delivery channel while leaving the representational construction unexamined [7], [47].

 ISSN 2655-9153

Tabel 5. Digital pedagogy affordances and their expression in trumpet micro-technique learning

Affordance	Pedagogical function	Expression in trumpet micro-technique
Asynchronous access	Decouples observation from instructional time	Learners revisited the target configuration at the point of their own error rather than at the moment it was taught
Replay	Enables iterative observation–imitation–correction cycles	Repeated comparison allowed incremental calibration of aperture and air-stream direction
Temporal dilation and freeze	Resolves events below the threshold of real-time perception	Made the onset sequence of inhalation, embouchure set, and release separately observable
Annotation and signalling	Directs attention to the controlling feature	Shifted attention from the body to the air-stream effect, inducing an external focus
Contrastive juxtaposition	Supplies a discriminative rather than normative reference	Made the learner's own deviation, not merely the target, perceptible
Shared visual model	Standardises the technical referent across learners	Reduced interpretive variance; consistent with the narrowing of score dispersion

3.7. Theoretical contribution

The first claim is that visual micro-skills describe a process that absent theories mention individually. Observational learning demands a model in which the relevant elements are observable [18]. In schema theory, an identifiable parameter must be allowed to vary [21]. From the attentional-focus perspective, a learner must be able to focus on a movement effect [22]. Deliberate practice relies on a precise task along with constructive feedback [24]. When dealing with a skill that is partially occluded, none of the requirements is satisfied by default. This construct specifies the representational operations of dilation, signaling, contrast, and shared reference, which make the requirements designable in the context of practice, rather than remaining an assumption. Secondly, the study shifts visualization in digital pedagogy from being a medium to being a mechanism. In much of the literature, the focus has been on whether digital tools have an impact on student outcomes [11], [12], [14]. The evidence presented here, in contrast, aims to describe what a representation must accomplish in order to impact a specific type of coordination. This shift in thinking enables the formation of testable hypotheses. For example, it is expected that procedures which maintain asynchronous access and replay, but that eliminate dilation, signaling, or contrast, will have weak effects on micro-coordination that will not impact the overall measures of global engagement. Third, the findings reveal a boundary condition that elaborates the construct as opposed to expanding it. Effects were weakest in conditions wherein the target configuration was physically occluded and was least stable to performance load, suggesting that representational resolution has a physical limit and that visual intervention impacts the cognitive-associative phases of motor learning (as opposed to the phase of automatisation) [20]. This is a scope condition, as opposed to a limit of the analysis, and indicates where additional modalities (*e.g.,* haptic, aerodynamic, or acoustic feedback) would be necessary [48]–[50].

3.8. Limitations

Four limitations bound these claims. The single-group pre-test–post-test design contains no control condition. Therefore, the maturation, repeated testing, and expectancy effects cannot be separated from the intervention. Consequently, the effect size documents change within the context of the study, and an isolated causal effect is not determined. The sample consisted of 25 non-randomly selected participants from one city, all of whom were in early stages of non-conservatory learning. Therefore, the findings are likely to be analytically generalizable (as opposed to statistically generalizable) to other similar semi-formal learning contexts, but not to conservatory or professional contexts. The post-test was administered directly after the practice phase, and therefore, retention beyond that period is unknown. However, Theme 4 provides an evidence-based rationale to expect that some of the gains are, in fact, unstable. The final limitation is that several representational operations were combined in the intervention, and therefore, their individual effects cannot be determined. Isolating the contribution of each of the operations would require a factorial or a dismantling design.

4. Conclusion

This research investigates the efficacy of video-based visual micro-skills for the correction of technical errors in novice trumpet players. Assessments resulted in the measured trumpet playing technique demonstrating a statistically significant improvement ($t(24) = 25.03$, $p < 0.001$, $dz = 5.01$, 95% CI [20.85, 24.59]) with a mean improvement from 62 to 85 among 25 participants and a narrowing of dispersion. Changes in the qualitative segment showed that participants shifted from relying on kinesthetic feedback to visual feedback, decomposed the fused actions to control the separate actions and moved the control of the actions to the participants. Integration showed the limits of the intervention. The control of embouchure showed improvement but a lack of control of the body resulted in nervous tension, with embouchure control showing the least improvement. The intervention focused on changing technical understanding and control of movements, but automatic control of the movements was not achieved. The study suggests that visual micro-skills include particular representational processes such as temporal dilation and attentional signaling, as well as processes involving contrastive juxtaposition and a shared reference model. These processes 'manufacture' the perceptual conditions for the presuppositions of observational learning, schema theory, attentional focus research, and deliberate practice as they pertain to partially obfuscated skills. This would reposition visualization within digital pedagogy from being a mode of delivery to being an active mechanism. In a more practical sense, the study offers a validated, three-dimensional, 120-minute, five-stage protocol that can be both implemented and appraised by instrumental teachers, as well as the evidence that asynchronous access and replay are inadequate in and of themselves. It is how the material is constructed that does the pedagogical work. The limitations of this study, which include working with two fundamental domains of error, early-stage learners, one semi-structured context, and a single demographic group with no control, restrict the generalizability of the findings. Future studies should test the construct in a controlled comparative design, disaggregate representational operations to assess the contribution of each, expand the retention interval, and examine the effects of adding haptic and aerodynamic feedback to the proposed model for overcoming the occlusion effect in embouchure control.

Acknowledgment

The authors thank the wind-instrument students of Universitas Negeri Surabaya and the trumpet players of the Surabaya drum-band communities for their participation, and colleagues who provided constructive comments on earlier drafts. The authors also thank the expert judges who reviewed the assessment rubric.

Declarations

Author contribution	:	All authors contributed equally to the conceptualization, methodology, intervention design, data collection, formal analysis, validation, and writing of the original draft, review, and editing of the manuscript. All authors read and approved the final version of the manuscript.
Funding statement	:	This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
Conflict of interest	:	The authors declare no conflict of interest.
Additional information	:	No additional information is available for this paper.

References

[1] J. López-Pineda, M. C. Rodríguez-Martínez, R. Gómez-Rodríguez, L. García-Casares, and N. García-Casares, "Biomechanical Assessments in Woodwind Musicians: A Systematic Review," *Healthcare*, vol. 11, no. 11, p. 1621, Jun. 2023, doi: 10.3390/healthcare11111621.

[2] V. Fréour and G. P. Scavone, "Acoustical interaction between vibrating lips, downstream air column, and upstream airways in trombone performance," *J. Acoust. Soc. Am.*, vol. 134, no. 5, pp. 3887–3898, Nov. 2013, doi: 10.1121/1.4823847.

[3] R. Nelkenstock, P. W. Iltis, D. Voit, J. Frahm, E. Passarotto, and E. Altenmüller, "Movement patterns in tuba playing: comparison of an embouchure dystonia case with healthy professional tuba players using real-time MRI imaging," *Front. Neurol.*, vol. 14, May 2023, doi: 10.3389/fneur.2023.1106217.

[4] A. López Requena, A. Baño Alcaraz, P. Escolar Reina, E. Ferrández Gómez, and G. Cánovas Ambit, "Effectiveness of a cervical treatment in wind-instrument musicians with temporomandibular dysfunction: A randomized clinical trial," *J. Bodyw. Mov. Ther.*, vol. 40, pp. 1867–1873, Oct. 2024, doi: 10.1016/j.jbmt.2024.10.060.

[5] G. Macovei *et al.*, "Changes in Dento-Facial Morphology Induced by Wind Instruments, in Professional Musicians and Physical Exercises That Can Prevent or Improve Them A Systematic Review," *Life*, vol. 13, no. 7, p. 1528, Jul. 2023, doi: 10.3390/life13071528.

[6] A. Acquilino and G. Scavone, "Current State and Future Directions of Technologies for Music Instrument Pedagogy," *Front. Psychol.*, vol. 13, Mar. 2022, doi: 10.3389/fpsyg.2022.835609.

[7] B. Behzadaval, L. Serra Marin, and L. Nijs, "Application of Visual Information in Music Education Digital Technologies: A Scoping Review," *Educ. Sci.*, vol. 16, no. 2, p. 309, Feb. 2026, doi: 10.3390/educsci16020309.

[8] A. Michałko, A. Campo, L. Nijs, M. Leman, and E. Van Dyck, "Toward a meaningful technology for instrumental music education: Teachers' voice," *Front. Educ.*, vol. 7, Oct. 2022, doi: 10.3389/feduc.2022.1027042.

[9] K. Fredriksson, O. Zandén, and C. Wallerstedt, "Teaching and learning in music education a meta-synthesis," *Music Educ. Res.*, vol. 26, no. 2, pp. 193–204, Mar. 2024, doi: 10.1080/14613808.2024.2319579.

[10] N. J. Hodges and I. M. Franks, "Modelling coaching practice: the role of instruction and demonstration," *J. Sports Sci.*, vol. 20, no. 10, pp. 793–811, Jan. 2002, doi: 10.1080/026404102320675648.

[11] L. Yihan, T. Van Cuong, B. Kiss, T. Z. Oo, N. Szabó, and K. Józsa, "The Use and Effectiveness of Digital Tools in Elementary Music Education: A Systematic Review," *Music Sci.*, vol. 8, Aug. 2025, doi: 10.1177/20592043251363338.

[12] K. I. C. H. Weatherly, B. Wright, and E. Y. Lee, "Digital game–based learning in music education: A systematic review from 2011 to 2023," *Res. Stud. Music Educ.*, vol. 48, no. 1, pp. 6–24, Apr. 2026, doi: 10.1177/1321103X241270819.

[13] R. Fauzy Ananda, R. Milyartini, and Y. Sukmayadi, "Mapping the Use of Digital Game-Based Learning in Music Education: A Scoping Review," *J. Pendidik. Indones.*, vol. 6, no. 11, pp. 4982–4989, Nov. 2025, doi: 10.59141/japendi.v6i11.8772.

[14] Y. Zhang, W. F. Beh, and C. Zhang, "A Systematic Review of Blended Learning Strategies and Outcomes in Music Education," *Harmon. J. Arts Res. Educ.*, vol. 24, no. 2, pp. 247–262, Dec. 2024, doi: 10.15294/harmonia.v24i2.18523.

[15] G. Rahmawati, B. Abduljabar, and R. W. A. Rozak, "Project-Based Learning in Music Education: A Scopus-Based Bibliometric Review of Research Trends and Intellectual Structure," *J. Paedagogy*, vol. 13, no. 1, pp. 266–277, Jan. 2026, doi: 10.33394/jp.v13i1.19068.

[16] M. Mödinger, A. Woll, and I. Wagner, "Video-based visual feedback to enhance motor learning in physical education a systematic review," *Ger. J. Exerc. Sport Res.*, vol. 52, no. 3, pp. 447–460, Sep. 2022, doi: 10.1007/s12662-021-00782-y.

[17] S. C. Youssef, A. Aydin, A. Canning, N. Khan, K. Ahmed, and P. Dasgupta, "Learning Surgical Skills Through Video-Based Education: A Systematic Review," *Surg. Innov.*, vol. 30, no. 2, pp. 220–238, Apr. 2023, doi: 10.1177/15533506221120146.

[18] A. Bandura, "Social foundations of thought and action," *Englewood Cliffs, NJ*, vol. 1986, no. 23–28, p. 2, 1986.

[19] Y. Han, S. K. Bin Syed Ali, and L. Ji, "Use of Observational Learning to Promote Motor Skill Learning in Physical Education: A Systematic Review," *Int. J. Environ. Res. Public Health*, vol. 19, no. 16, p. 10109, Aug. 2022, doi: 10.3390/ijerph191610109.

[20] and M. I. P. Paul Morris Fitts, *Human Performance*. Belmont, CA: Brooks/Cole, 1967.

[21] R. A. Schmidt, T. D. Lee, C. J. Winstein, G. Wulf, and H. N. Zelaznik, *Motor control and learning: A behavioral emphasis*, vol. 4. Human kinetics Champaign, IL, 2005.

[22] G. Wulf, "Attentional focus and motor learning: a review of 15 years," *Int. Rev. Sport Exerc. Psychol.*, vol. 6, no. 1, pp. 77–104, Sep. 2013, doi: 10.1080/1750984X.2012.723728.

[23] R. E. Mayer, "Cognitive Theory of Multimedia Learning," in *The Cambridge Handbook of Multimedia Learning*, Cambridge University Press, 2021, pp. 57–72, doi: 10.1017/9781108894333.008.

[24] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, "The role of deliberate practice in the acquisition of expert performance.," *Psychol. Rev.*, vol. 100, no. 3, pp. 363–406, Jul. 1993, doi: 10.1037/0033-295X.100.3.363.

[25] J. W. Creswell and V. L. Plano Clark, *Designing and Conducting Mixed Methods Design*, Third Edit. Los Angeles: SAGE Publication Inc, 2018.

[26] T. C. Guetterman, M. D. Fetters, and J. W. Creswell, "Integrating Quantitative and Qualitative Results in Health Science Mixed Methods Research Through Joint Displays," *Ann. Fam. Med.*, vol. 13, no. 6, pp. 554–561, Nov. 2015, doi: 10.1370/afm.1865.

[27] M. D. Fetters, L. A. Curry, and J. W. Creswell, "Achieving Integration in Mixed Methods Designs Principles and Practices," *Health Serv. Res.*, vol. 48, no. 6pt2, pp. 2134–2156, Dec. 2013, doi: 10.1111/1475-6773.12117.

[28] A. L. Brown, "Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings," *J. Learn. Sci.*, vol. 2, no. 2, pp. 141–178, Apr. 1992, doi: 10.1207/s15327809jls0202_2.

[29] L. A. Palinkas, S. M. Horwitz, C. A. Green, J. P. Wisdom, N. Duan, and K. Hoagwood, "Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research," *Adm. Policy Ment. Heal. Ment. Heal. Serv. Res.*, vol. 42, no. 5, pp. 533–544, Sep. 2015, doi: 10.1007/s10488-013-0528-y.

[30] L. R. Aiken, "Three Coefficients for Analyzing the Reliability and Validity of Ratings," *Educ. Psychol. Meas.*, vol. 45, no. 1, pp. 131–142, Mar. 1985, doi: 10.1177/0013164485451012.

[31] T. K. Koo and M. Y. Li, "A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research," *J. Chiropr. Med.*, vol. 15, no. 2, pp. 155–163, Jun. 2016, doi: 10.1016/j.jcm.2016.02.012.

[32] D. Lakens, "Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs," *Front. Psychol.*, vol. 4, 2013, doi: 10.3389/fpsyg.2013.00863.

[33] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, Jan. 2006, doi: 10.1191/1478088706qp063oa.

[34] A. Schiavio and L. Nijs, "Implementation of a Remote Instrumental Music Course Focused on Creativity, Interaction, and Bodily Movement. Preliminary Insights and Thematic Analysis," *Front. Psychol.*, vol. 13, May 2022, doi: 10.3389/fpsyg.2022.899381.

[35] A. F. Hayes and K. Krippendorff, "Answering the Call for a Standard Reliability Measure for Coding Data," *Commun. Methods Meas.*, vol. 1, no. 1, pp. 77–89, Apr. 2007, doi: 10.1080/19312450709336664.

[36] L. Birt, S. Scott, D. Cavers, C. Campbell, and F. Walter, "Member Checking," *Qual. Health Res.*, vol. 26, no. 13, pp. 1802–1811, Nov. 2016, doi: 10.1177/1049732316654870.

[37] A. H. Watson and K. Price, "Embouchure Muscle Activity in Student and Elite Trumpeters," *Med. Probl. Perform. Art.*, vol. 35, no. 1, pp. 42–53, Mar. 2020, doi: 10.21091/mppa.2020.1006.

 ISSN 2655-9153

[38] K. Kula, H. Z. Cilingir, G. Eckert, J. Dagg, and A. Ghoneima, "The association of malocclusion and trumpet performance," *Angle Orthod.*, vol. 86, no. 1, pp. 108–114, Jan. 2016, doi: 10.2319/111214-812.1.

[39] L. M. T. Jesus, "New Insights into Laryngeal Articulation and Breathing Control of Trumpeters: Biomedical Signals and Auditory Perception," *Appl. Sci.*, vol. 14, no. 19, p. 8957, Oct. 2024, doi: 10.3390/app14198957.

[40] A. Türk-Espitalier, M. Bertsch, and I. Cossette, "Effect of Expiratory Muscle Strength Training on the Performance of Professional Male Trumpet Players," *Med. Probl. Perform. Art.*, vol. 39, no. 1, pp. 18–26, Mar. 2024, doi: 10.21091/mppa.2024.1003.

[41] L. S. Acheson, P. G. Rogers, C. Thamrin, M. J. Peters, C. S. Farah, and L. M. Seccombe, "Acute cardiopulmonary responses during expert trumpet performance," *Respir. Physiol. Neurobiol.*, vol. 281, p. 103493, Oct. 2020, doi: 10.1016/j.resp.2020.103493.

[42] K. Sakai, T. Kawasaki, Y. Ikeda, J. Tanabe, A. Matsumoto, and K. Amimoto, "Differences in the early stages of motor learning between visual-motor illusion and action observation," *Sci. Rep.*, vol. 13, no. 1, p. 20054, Nov. 2023, doi: 10.1038/s41598-023-47435-8.

[43] Saluky and Nurul Bahiyah, "Revolutionizing Education: Unlocking the Potential of Asynchronous Video for Interactive Online Learning," *Int. J. Educ. Humanit.*, vol. 3, no. 2, pp. 187–196, Jul. 2023, doi: 10.58557/(ijeh).v3i2.149.

[44] Q. Wang, Y. Wen, and C. L. Quek, "Engaging learners in synchronous online learning," *Educ. Inf. Technol.*, vol. 28, no. 4, pp. 4429–4452, Apr. 2023, doi: 10.1007/s10639-022-11393-x.

[45] E. Navarrete, A. Nehring, S. Schanze, R. Ewerth, and A. Hoppe, "A Closer Look into Recent Video-based Learning Research: A Comprehensive Review of Video Characteristics, Tools, Technologies, and Learning Effectiveness," *Int. J. Artif. Intell. Educ.*, vol. 35, no. 4, pp. 1631–1694, Dec. 2025, doi: 10.1007/s40593-025-00481-x.

[46] Z. Qi, J. Wang, Y. Du, and X. Zhang, "The effect of blended learning approaches on vocational school students' practical skills performance and learning behaviors: virtual simulations or hands-on activities while accessing instructional videos," *Instr. Sci.*, vol. 54, no. 1, p. 39, Jun. 2026, doi: 10.1007/s11251-026-09793-4.

[47] J. Shao, "Enhancing learning outcomes in violin music education: a systematic review," *Front. Educ.*, vol. 10, Jan. 2026, doi: 10.3389/feduc.2025.1683580.

[48] M. G. Luciani, A. Cortelazzo, and A. M. Proverbio, "The role of auditory feedback in the motor learning of music in experienced and novice performers," *Sci. Rep.*, vol. 12, no. 1, p. 19822, Nov. 2022, doi: 10.1038/s41598-022-24262-x.

[49] W. Li, P. Manoharan, X. Cui, F. Liu, K. Liu, and L. Dai, "The effects of musical feedback training on metacognition and self-directed learning," *Front. Hum. Neurosci.*, vol. 17, p. 1304929, Dec. 2023, doi: 10.3389/fnhum.2023.1304929.

[50] M. Reybrouck and A. Schiavio, "Music performance as knowledge acquisition: a review and preliminary conceptual framework," *Front. Psychol.*, vol. 15, Feb. 2024, doi: 10.3389/fpsyg.2024.1331806.