WebLog Journal of Community Medicine

Research Article
Published: 18 May, 2026

Instructor Presence Stabilizes Performance Distributions in Digitally Scaffolded CPR Training

*Felix Hoepfl**

Rosenheim Technical University of Applied Sciences, Faculty of Applied Health and Social Sciences, Rosenheim, Germany

Abstract

Background: Digital feedback systems are increasingly integrated into school-based cardiopulmonary resuscitation (CPR) training. Prior research has largely emphasized mean performance outcomes, whereas assessment theory suggests that distributional characteristics—particularly lower-tail reliability—may be equally relevant in safety-critical competence contexts.

Purpose: This study examined whether different instructional configurations under digitally scaffolded CPR training—peer-led and adult-led instruction—differ in distributional properties of performance, and how these formats compare to a minimal-scaffolding digital-only configuration.

Methods: A quasi-experimental comparative design was conducted across three conditions: peer-led instruction (n = 342), adult-led instruction (n = 69; individual-level subsample), and digital-only training without instructor guidance during assessment (n = 23). Mechanical CPR performance was assessed using QCPR-enabled manikins generating composite percentage scores. Non-parametric tests were used to compare distributions and frequencies of catastrophic low-performance cases (<50%), operationalized as lower-tail competence failures.

Results: Mean performance was comparable across groups (Peer: 84.45%; Adult: 83.54%; Digital-only: 82.52%). However, distributional differences emerged. Catastrophic lower-tail failures occurred exclusively in the digital-only condition (3/23; 13.0%) and were absent in instructor-present formats. Overall group differences were statistically significant (H = 11.38, p = .003), with pairwise contrasts indicating differences between instructor-present and digital-only conditions, but not between peer-led and adult-led instruction.

Conclusions: Although mean performance did not differ, instructor-present formats were associated with reduced lower-tail vulnerability. From an assessment perspective, distributional integrity may therefore complement mean-based interpretations when evaluating instructional effects in safety-critical skills training.

Keywords: CPR Training; Peer Instruction; Distributional Performance; Instructional Scaffolding; Health Professions Education; Competence Assessment; Simulation-Based Learning

OPEN ACCESS

Correspondence:
Felix Hoepfl, Rosenheim Technical University of Applied Sciences, Faculty of Applied Health and Social Sciences, Rosenheim, Germany, E-mail: Felix.Hoepfl@TH-Rosenheim.de/ ORCID: https://orcid.org/0000-0002-1499-4945

Received Date: *24 Apr 2026*
Accepted Date: *15 May 2026*
Published Date: *18 May 2026*

Citation:
Felix Hoepfl. Instructor Presence Stabilizes Performance Distributions in Digitally Scaffolded CPR Training. WebLog J Community Med. wjcm.2026.e1801. https://doi.org/10.5281/zenodo.20329043

Introduction

Out-of-hospital cardiac arrest remains a major public health challenge worldwide, with survival rates strongly dependent on early bystander cardiopulmonary resuscitation (CPR). International resuscitation bodies advocate the systematic implementation of basic life support (BLS) education in schools as a scalable population-level strategy to improve bystander response rates [1]. Klicken oder tippen Sie hier, um Text einzugeben. School-based CPR education is considered feasible across age groups and has been promoted as an essential component of community preparedness. However, while policy endorsement is strong, questions remain regarding optimal instructional formats and the mechanisms through which training quality is ensured.

Systematic reviews indicate that CPR training in schoolchildren can lead to meaningful knowledge acquisition and psychomotor skill development across various instructional models [2]. Both teacher-led and peer-led formats have been described as viable approaches, yet the literature remains heterogeneous with respect to instructor qualification, pedagogical structure, training duration, and assessment methods. Beyond differences in instructional format, professional perspectives regarding responsibility for BLS instruction in secondary education appear

heterogeneous. Empirical evidence suggests that neither teachers nor health professionals uniformly conceptualize CPR and AED training as an unequivocal curricular responsibility [3]. However, while role attribution remains debated, comparatively little evidence addresses how instructional configurations shape performance distributions in digitally scaffolded environments.

In particular, most studies have focused on mean performance outcomes and short-term retention, often relying on varied or non-standardized measurement instruments. Less attention has been paid to distributional characteristics of performance or to how digital feedback systems interact with different levels of instructor presence.

Peer teaching has been widely discussed within health professions education, with theoretical grounding in cognitive and social congruence [4, 5]. While peer-led CPR instruction has shown promise in educational and public health contexts, less is known about how instructor presence—whether peer or professionally qualified adult—interacts with digitally scaffolded feedback systems. Emerging technologies such as QCPR-enabled manikins provide objective, parameter-based performance metrics, enabling standardized comparison across instructional formats. Yet, the integration of digital feedback with varying levels of human instructional presence has not been systematically examined.

Digitally scaffolded motor skill acquisition relies not only on the provision of performance metrics but also on learners' ability to interpret and regulate their actions in response to feedback. Classic scaffolding theory conceptualizes instructional support as a temporary regulatory structure that assists learners in identifying critical task features, reducing cognitive load, and maintaining goal-directed activity [6]. In contemporary educational research, scaffolding is understood as structured support that facilitates error detection, calibration, and self-regulation, particularly in technology-mediated learning environments [7]. Although feedback-enabled manikins provide quantitative task-level information, effective feedback uptake depends on learners' capacity to interpret signals and translate them into corrective motor adjustments [8]. Digital systems primarily operate at the level of task feedback, whereas instructors may additionally provide process-level and self-regulation guidance [9]. Although recent analyses describe technology-enhanced feedback as increasingly ecosystem-based [10]. The QCPR system in this study represents a structured task-feedback device. Its signals require interpretive mediation, a function that instructor presence may support.

We therefore conceptualize instructor presence—whether peer or professionally qualified adult—not primarily as a source of additional content delivery, but as a regulatory layer within a digitally scaffolded motor learning system. From this perspective, instructor presence may function to stabilize performance distributions by supporting feedback interpretation, error correction, and sustained task engagement.

Assessment theory emphasizes that performance interpretation depends on the intended use of scores and their consequences [11, 12]. In safety-critical domains such as CPR, competence may be less about maximizing average performance and more about ensuring reliable attainment of minimum acceptable standards. Under such a framework, distributional characteristics—including lower-tail performance—may provide additional insight beyond mean scores alone.

At the same time, advances in simulation technology have transformed procedural skills training. Feedback-enabled manikins, such as QCPR systems, provide real-time, directive feedback on compression depth, rate, and recoil, allowing learners to self-correct during task execution. Empirical work suggests that such systems offer reliable and standardized measurement of mechanical CPR parameters [13] and that structured digital feedback can meaningfully influence skill acquisition in simulation-based contexts [14]. When automated systems continuously signal performance deviations, the traditional instructional role of detecting and correcting technical errors may be partially redistributed to technology.

This technological shift raises a theoretically important question: if digital scaffolding provides immediate corrective input, what is the functional contribution of instructor presence? From an assessment perspective, instructor effects may not primarily manifest as differences in mean performance but as stabilization of performance distributions and safeguarding of minimum competence. In competency-based education, competence is conceptualized in terms of entrustability—the capacity to perform safely without direct supervision [4]. Within this framework, elimination of catastrophic low-performance outcomes represents a meaningful educational effect.

Despite extensive research on peer teaching and digital feedback systems separately, few studies have examined how instructor presence interacts with digitally scaffolded environments to shape distributional performance outcomes. In particular, the role of peer instructors in stabilizing performance under real-time feedback conditions remains insufficiently understood.

Stabilization is employed here as an analytic lens for interpreting distributional performance outcomes rather than as a formally developed theoretical construct; it serves to organize the interpretive framework through which instructor effects on score distributions are examined. Specifically, distributional stabilization refers to (a) absence of catastrophic lower-tail failures below a predefined performance threshold (<50%) as the primary indicator, and (b) attenuation of variance relative to comparison conditions as secondary support. This operationalization reflects a competence-reliability perspective in which instructional effectiveness is evaluated not solely by mean performance, but by the consistency and safety margin of performance distributions in skill-critical contexts.

Guided by assessment and competence theory, we conceptualize instructional effectiveness not solely in terms of mean performance but in relation to performance distribution and safeguarding of minimum standards. In safety-critical educational contexts, elimination of catastrophic underperformance may constitute a validity-relevant outcome.

Against this background, the study examines instructional configurations — specifically, instructor presence within structured training contexts — in digitally scaffolded CPR training in secondary school settings. Using objective QCPR performance data, three instructional formats are compared: peer-led instruction, adult-led instruction (individual-level subsample), and a digital-only minimal-scaffolding condition without instructor guidance during the assessment phase. It should be noted that instructor-present conditions were embedded within structured training sessions, whereas the digital-only condition represented a minimal-scaffolding performance format; the observed contrasts thus reflect differences in

instructional configuration broadly, rather than instructor presence in isolation. This quasi-experimental design permits comparison of instructor-present and minimal-scaffolding conditions while acknowledging differences in structured training exposure.

The aim of this study was to examine whether different instructor formats in digitally scaffolded CPR training—peer-led versus adult-led instruction—are associated with differences in mechanical CPR performance, with a particular focus on distributional characteristics relevant to minimum competence. A small digital-only condition with automated feedback but without human instructor guidance during the assessment phase was included to provide an exploratory minimal-scaffolding contrast.

The overarching research question guiding this study was:

How do peer-led and adult-led instructional formats in digitally scaffolded school-based CPR training relate to the distributional characteristics of QCPR performance, and how do these instructor-present formats compare descriptively to a minimal-scaffolding digital-only configuration?

To address this question, the following hypotheses were tested:

H1 (Peer–Adult Mean Equivalence). Mean QCPR performance will not differ substantially between peer-led and adult-led conditions.

H2 (Stabilization – Variability). Instructor-present conditions (peer-led and adult-led) will show reduced performance variability relative to the digital-only condition.

H3 (Lower-Tail Safety Margin). Catastrophic lower-tail failures (QCPR < 50%) will occur more frequently in the digital-only condition than in instructor-present conditions.

H4 (Peer–Adult Stabilization Equivalence). Peer-led and adult-led conditions will not differ substantially in performance variability or lower-tail outcomes.

Methods

Study Design

This study employed a quasi-experimental comparative design to examine the effect of instructor presence under digitally scaffolded CPR training. Three instructional conditions were analyzed at the individual level:

1. Peer-led instruction with real-time digital feedback
2. Adult-led instruction with real-time digital feedback (individual-level subsample)
3. Digital-only training without instructor guidance

In addition, aggregated course-level data from 20 adult-led courses were available and are reported descriptively. The author served as the instructor for the adult-led instruction. To reduce analytic bias, statistical procedures were predefined and applied consistently across groups.

Participants were not randomly assigned to conditions; instructional formats were implemented within routine educational settings.

Ethical aspects are detailed in the Declarations section.

Participants and Setting

The study was conducted at a public secondary school (Gymnasium) in Munich, Bavaria, Germany in two separate cohorts in 2022 and 2024. The Gymnasium represents the academic track of the German secondary education system. Participating students were enrolled in grades 7, 9, and 11, corresponding approximately to early and middle adolescence.

CPR training was implemented during regular school activities as part of routine educational programming. Peer-led training was delivered by students who had previously completed structured first-aid instruction. Adult-led sessions were conducted by certified first-aid instructors. The digital-only condition consisted of a standardized 2-minute compression-only performance assessment using feedback-enabled manikins without prior structured instruction during the testing session.

No medical history, health data, or sensitive personal information were collected. All outcome measures were based exclusively on anonymized performance metrics generated by QCPR-enabled manikins.

Peer-Led Condition

The peer-led cohort comprised 342 secondary school students drawn from two training waves conducted in 2022 (n = 245) and 2024 (n = 97). Instructional protocols, training materials, and QCPR measurement systems were identical across waves. Cohort-level mean performance was comparable across years (84.8% *vs.* 83.5%), supporting aggregation for analysis.

Peer instructors were trained student first responders who had completed formal CPR certification and instructional preparation prior to teaching.

Adult-Led Condition

Two adult-led datasets were available:

1. Aggregated dataset: 20 certified first-aid courses comprising 256 attending participants. Only course-level summary statistics were retained.
2. Individual-level subsample: Four adult-led courses (n = 69 participants) for which individual QCPR scores were available.

Only the individual-level subsample (n = 69) was included in inferential statistical analyses. The aggregated dataset is reported descriptively to contextualize adult-led performance across a broader course sample.

All adult-led courses were conducted by certified first-aid instructors using standardized curricula and the same QCPR measurement system.

Digital-Only Condition

The digital-only group consisted of 23 secondary school students from a comparable educational context. Participants performed a standardized two-minute compression-only CPR task using QCPR-enabled manikins without instructor guidance. Real-time automated feedback was available, but no human correction or structured practice beyond the assessment phase was provided.

Intervention Characteristics

All participants trained using Laerdal Little Anne QCPR manikins. Peer-led and adult-led sessions embedded structured instruction and supervised practice, followed by a standardized two-minute compression-only assessment phase. The digital-only group completed the same standardized two-minute compression-only task

Table 1: Descriptive statistics of QCPR performance across instructional conditions.

Condition	n	Mean (%)	SD	Median (%)	IQR	Range (%)
Peer-led	342	84.45	8.93	84	14	69–100
Adult-led (individual-level subsample)	69	83.54	10.28	86	16	63–100
Digital-only	23	82.52	30.02	94	7	3–99

Note: Mean = arithmetic mean; SD = standard deviation; IQR = interquartile range (Q3 − Q1). QCPR scores represent composite percentage scores generated by the Laerdal Little Anne QCPR system based on compression parameters (depth, rate, recoil, and hand position). Catastrophic low-performance cases (<50%) occurred exclusively in the digital-only condition (3/23; 13.0%).

without instructor presence.

Thus, performance measurement duration was comparable across conditions, although only instructor-present groups received structured training prior to assessment.

Instrumentation and Performance Scoring

CPR performance was measured using the Laerdal Little Anne QCPR system, which records compression depth, rate, chest recoil, and hand position accuracy in real time and integrates these parameters into a composite percentage score according to manufacturer specifications.

Feedback-enabled manikins have demonstrated reliable measurement of mechanical CPR parameters aligned with guideline recommendations [13]. Structured digital feedback systems have further been shown to influence skill acquisition in simulation-based training contexts [14].

The composite QCPR score therefore represents a standardized index of mechanical CPR quality. The proprietary weighting algorithm underlying the composite score is not publicly disclosed. Conclusions of the present study pertain specifically to mechanically defined CPR performance.

Outcome Measures

The primary outcome was the composite QCPR percentage score.

Performance stabilization was operationalized through:

- Measures of dispersion (standard deviation and interquartile range)
- Variance comparisons across groups
- Frequency of catastrophic low-performance cases

Catastrophic low performance was defined a priori as a composite QCPR score <50%, representing substantial deviation from guideline-aligned mechanical CPR performance.

Statistical Analysis

All analyses were conducted at the individual level for peer-led (n = 342), adult-led subsample (n = 69), and digital-only (n = 23) conditions.

Because score distributions were non-normal and extreme lower-tail values were observed in the digital-only condition, non-parametric testing was applied.

- Group differences across three conditions were examined using the Kruskal–Wallis test.
- Pairwise comparisons were conducted using Mann–Whitney U tests with Holm correction.
- Effect sizes were calculated using rank-biserial correlations.
- Variance differences were assessed using Levene's test (centered on the median).
- Differences in the proportion of catastrophic low-performance cases (<50%) were tested using Fisher's exact test.

The significance threshold was set at $\alpha = .05$ (two-tailed).

Results

Descriptive Statistics

Peer-Led Condition (n = 342):

The peer-led cohort demonstrated a mean QCPR score of 84.45% (SD = 8.93), with a median of 84% and a range from 69% to 100%. The interquartile range (IQR) was 14 percentage points. No participant scored below 50%.

Adult-Led Condition (Individual-Level Subsample; n = 69):

The adult-led subsample achieved a mean score of 83.54% (SD = 10.28), with a median of 86% and a range from 63% to 100%. The IQR was 16 percentage points. No participant scored below 50%.

Digital-Only Condition (n = 23):

The digital-only group achieved a mean score of 82.52% (SD = 30.02), with a median of 94% and a range from 3% to 99%. The IQR was 7 percentage points. Three of 23 participants (13.0%) scored below 50%, including two cases at 3% and one at 29%. The low IQR reflects the concentration of 20 of 23 participants above 85%, while the three extreme low-performance cases fell below 50%, producing a strongly polarized score distribution with a heavily loaded lower tail. The large discrepancy between IQR (7 pp) and SD (30.02 pp) is a direct consequence of this distributional structure.

Descriptive statistics across instructional conditions are summarized in Table 1.

The distributional characteristics of performance across instructional conditions are displayed in Figure 1, highlighting differences in variability and central tendency that are not visible from mean scores alone.

H1: Peer–Adult Mean Equivalence

Mean performance did not differ significantly between peer-led and adult-led conditions (Mann–Whitney U, $p = .669$; $r \approx .03$), supporting H1. Mean scores were comparable in magnitude, and effect size estimates indicated negligible differences.

H2: Stabilization Hypothesis (Variance)

A Kruskal–Wallis test across all three conditions indicated significant differences in overall score distributions ($H = 11.38$, $p = .003$).

Variance comparisons using Levene's test (centered on the median) revealed significant differences across groups ($F = 9.43$, $p <$

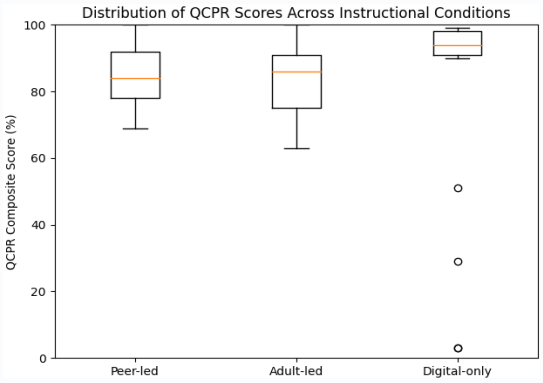

Figure 1: Distributional characteristics of QCPR performance across instructional configurations.
Boxplots depict median, interquartile range (IQR), and full score range (whiskers). The digital-only condition includes extreme lower-tail values (<50%) not observed in instructor-present formats.

.001). Pairwise comparisons showed significantly greater variance in the digital-only condition compared to the peer-led condition ($F = 17.26$, $p < .001$). The adult-led versus digital-only comparison showed a non-significant trend toward greater dispersion in the digital-only group ($F = 3.22$, $p = .076$). No significant variance difference was observed between peer-led and adult-led conditions ($F = 2.27$, $p = .132$).

Taken together, H2 was partially supported: variance attenuation was statistically evident in the peer-led comparison and directionally consistent, though not statistically significant, in the adult-led comparison. The partial support for variance attenuation reflects the modest statistical power of Levene's test in small-n comparisons. Importantly, lower-tail elimination (H3) and variance attenuation represent conceptually distinct distributional phenomena; the absence of catastrophic failures does not depend on variance attenuation and constitutes the stronger operational indicator of distributional stabilization in the present framework.

H3: Catastrophic Low Performance

Catastrophic lower-tail failures (<50%) occurred exclusively in the digital-only condition (3/23; 13.0%) and were absent in both peer-led (0/342) and adult-led (0/69) conditions.

Fisher's exact tests confirmed significantly higher lower-tail frequency in the digital-only condition compared to:

- Peer-led instruction ($p < .001$)
- Adult-led instruction ($p = .014$)

These findings support H3.

H4: Peer–Adult Stabilization Equivalence

Peer-led and adult-led conditions did not differ significantly in performance variability or lower-tail outcomes. Both groups demonstrated absence of catastrophic low-performance cases and comparable dispersion measures. These findings support H4.

A summary of inferential statistical results by hypothesis is provided in Table 2.

Discussion

This study examined instructor presence under digitally scaffolded CPR training using a distributional perspective. In addressing the research question of how different instructor formats relate to performance distribution under digital scaffolding, the findings indicate that peer-led and adult-led instruction yielded comparable performance profiles, whereas the minimal-scaffolding digital-only configuration was associated with distinct lower-tail vulnerability. Although mean performance levels were broadly similar across conditions, distributional properties differed. Catastrophic low-performance cases (<50%) occurred exclusively in the digital-only condition and were statistically more frequent than in both instructor-present formats. No statistically significant differences were observed between peer-led and adult-led conditions. Importantly, these findings do not suggest that instructor presence enhances mean performance. Rather, the absence of instructor guidance was associated with presence of extreme lower-tail variability and a higher likelihood of extreme underperformance. From a conventional mean-centered perspective, instructional effects might appear limited. However, when interpreted through an assessment lens emphasizing reliability and minimum safety thresholds, distributional integrity may be educationally meaningful.

Validity theory emphasizes that score interpretation depends on intended use and consequences [11, 12]. In safety-critical domains such as CPR, competence may be understood in terms of consistent attainment of minimally safe performance rather than maximal achievement. Within an entrustability-oriented framework [4], rare but substantial performance failures carry particular interpretive weight. Under this perspective, the absence of catastrophic low-performance cases in instructor-present conditions warrants careful

Table 2: Inferential Statistical Results by Hypothesis.

Hypothesis	Comparison	Statistical Test	Statistic	p-value	Effect Size (r)
H1 (Peer–Adult Mean Equivalence)	Peer-led vs Adult-led	Mann–Whitney U (Holm-corrected)	—	.669	.03
H2 (Stabilization – Variance)	All three groups	Kruskal–Wallis	H = 11.38	.003	—
	Peer-led vs Digital-only	Levene's test (median-centered)	F = 17.26	< .001	—
	Adult-led vs Digital-only	Levene's test (median-centered)	F = 3.22	.076	—
	Peer-led vs Adult-led	Levene's test (median-centered)	F = 2.27	.132	—
H3 (Catastrophic Low Performance)	Peer-led vs Digital-only	Fisher's exact test	—	< .001	—
	Adult-led vs Digital-only	Fisher's exact test	—	.014	—
H4 (Peer–Adult Stabilization Equivalence)	Peer-led vs Adult-led	Variance & lower-tail comparison	—	ns	—

Note: MWU = Mann–Whitney U test with Holm correction for multiple comparisons.
Effect sizes for MWU comparisons are reported as rank-biserial correlations (r).
Levene's test was centered on the median.
Catastrophic low performance was defined a priori as QCPR < 50%.
ns = not statistically significant ($p \geq .05$).

consideration.

The pattern was consistent across peer-led and adult-led formats. The absence of statistically significant differences between these two conditions is consistent with, though does not establish, the possibility that instructor presence rather than qualification level may be the relevant distributional factor within digitally scaffolded contexts. This finding should be interpreted as indicating no detectable difference within the present sample rather than functional equivalence; replication in larger and randomly allocated samples is required before stronger conclusions can be drawn.

Several mechanisms may plausibly, though speculatively, account for the observed distributional differences between instructor-present and digital-only conditions; the following are offered as theoretically motivated hypotheses rather than empirical conclusions. First, automated feedback requires accurate perception, interpretation, and motor translation by the learner. Real-time performance metrics may signal deviations (e.g., insufficient compression depth or rate), yet learners may misinterpret, overlook, or fail to prioritize these cues. Feedback effectiveness depends not only on the availability of information but also on learners' capacity to interpret and apply it appropriately [15]. Instructor presence may therefore support translation of metric feedback into targeted motor adjustments.

The concept of feedback literacy — the learner's capacity to understand, evaluate, and act upon performance information — is particularly relevant in motor skill contexts. Novice learners engaging with automated parameter feedback may lack the metacognitive resources to translate quantitative deviation signals (e.g., "compression depth insufficient") into appropriate biomechanical adjustments. Carless and Boud (2018) distinguish between feedback availability and feedback uptake, arguing that information alone does not generate learning unless the recipient possesses the interpretive capacity to respond. In the context of CPR training, this distinction maps onto the difference between digital feedback as a signal source and instructor presence as a signal interpreter. Peer or adult instructors may provide the cognitive scaffolding that bridges metric output and corrective action — a function that automated systems, operating solely at the task-information level [9], cannot fulfill.

Second, certain biomechanical errors—such as suboptimal hand placement, elbow positioning, or body weight transfer—may not be fully resolved through parameter-based feedback alone. Motor learning research suggests that augmented feedback can guide performance but does not automatically correct underlying movement patterns without additional instructional input [16].

Schmidt and Lee (2020) distinguish between augmented feedback — externally provided information about performance outcomes or kinematics — and intrinsic feedback, which arises from the learner's own sensory systems. Augmented feedback from QCPR systems operates at the outcome level: it signals whether compression depth or rate met guideline thresholds. However, the underlying kinematic errors that produce these deviations — improper elbow lock, insufficient hip hinge, or inadequate body weight transfer — are not directly reported by the system. Correcting these foundational patterns requires either sophisticated self-monitoring capacity or external observational input. For adolescent novice learners with limited proprioceptive awareness, the latter may be essential. Instructor presence thus provides a form of kinematic scaffolding that the digital system structurally cannot supply: real-time observation, interpretation of movement quality, and targeted verbal or physical correction. This distinction helps explain why digital feedback alone may be insufficient to prevent catastrophic underperformance, even when task-level outcome data are available.

Third, instructor presence may exert a regulatory function. Supervised performance may reduce disengagement, incomplete effort, or performance variability during short assessment tasks. Such regulatory support may attenuate variance and reduce the likelihood of extreme underperformance.

This regulatory dimension of instructor presence may operate through at least two pathways. First, social evaluation effects — well documented in performance research — suggest that observed performance tends to elicit greater effort and sustained engagement than unobserved performance, particularly in adolescent populations. The awareness of being monitored by a peer or adult instructor may reduce task-disengagement, which could manifest as the extreme underperformance cases observed in the digital-only condition. Second, instructors may provide real-time motivational prompts and behavioral regulation cues — verbal encouragement, reminders to maintain compression rate, or pacing support — that attenuate effort dropout during the two-minute assessment window. Peer instructors may be particularly effective in this regard, as cognitive and social congruence [4, 5] may facilitate the kind of informal, proximal regulation that reduces performance disengagement without the evaluative pressure associated with adult-led formats. Together, these pathways suggest that instructor presence may function as a regulatory stabilizer independently of technical skill transmission.

These mechanistic interpretations remain hypothetical, as the present study did not collect process-level or observational data. The design was quasi-experimental and does not permit causal attribution. Instructor-present conditions were embedded within structured training sessions, whereas the digital-only condition involved a standardized two-minute performance task without structured preparatory instruction during that session. The digital-only format therefore represents a minimal-scaffolding configuration rather than a pure absence-of-instructor contrast. Observed differences may thus reflect the presence of adaptive regulatory support within structured learning contexts rather than instructor presence per se. Experimental designs that disentangle instructional structure from instructor presence are needed to clarify underlying causal mechanisms.

These findings carry potential implications for the design of school-based CPR training programmes. Programmes relying exclusively on digital feedback systems without human instructor oversight may achieve comparable mean performance but remain vulnerable to a small proportion of learners who fall catastrophically below minimum competence thresholds. From a competency-based perspective, even a 13% rate of catastrophic underperformance in a minimally scaffolded condition represents a potentially meaningful signal in safety-critical skill contexts. Within the present sample, the introduction of either peer or adult instructor presence was associated with complete absence of catastrophic underperformance, suggesting this as a potentially meaningful differential. However, given the small digital-only sample (n = 23) and wide confidence intervals, this observation should be interpreted as exploratory rather than definitive. For programme planners operating under resource constraints, this finding suggests that the functional contribution of instructor presence may not require expert qualifications: trained peer instructors appear equally capable of providing the stabilizing

regulatory function described above. This has direct relevance to scalability, as peer-led formats are significantly less resource-intensive than professionally certified instructor provision.

Additional limitations warrant consideration. The digital-only sample was relatively small (n = 23), limiting precision in estimating lower-tail frequency. Although Fisher's exact tests indicated statistically significant differences in catastrophic performance rates, confidence intervals around such proportions in small samples remain wide. The 95% confidence interval for the observed 13.0% catastrophic failure rate (3/23) ranges approximately from 2.8% to 33.6%, indicating substantial uncertainty around the point estimate. The adult-led individual-level dataset was likewise limited in scope. Replication in larger and randomized samples would strengthen inference.

Furthermore, the QCPR composite score captures mechanical aspects of CPR performance only. While feedback-enabled manikins provide standardized and reliable measurement of compression parameters, broader competence dimensions—including situational awareness, communication, and adaptive decision-making—were not assessed. The proprietary weighting of score components constrain construct-level interpretability. Conclusions therefore pertain specifically to mechanically defined CPR performance under digitally scaffolded conditions.

Taken together, the findings suggest that peer-led and adult-led instruction under digitally scaffolded conditions produced comparable performance distributions, whereas a minimal-scaffolding digital-only format was associated with greater lower-tail vulnerability. Whether this association reflects instructor presence, structured practice exposure, or interaction effects between human guidance and automated feedback requires further investigation. Importantly, the results indicate that mean equivalence does not preclude meaningful differences in distributional safety margins. In skill-critical educational contexts, instructional effectiveness may therefore be evaluated not only by average performance but also by the integrity and reliability of performance distributions.

Future research should experimentally disentangle instructor presence from instructional structure, examine process-level mechanisms of feedback interpretation and regulation, and test whether distributional stabilization extends beyond mechanical CPR metrics to broader competence domains within health professions education.

Declarations

Ethics approval

This study was conducted in a secondary school educational setting in Munich, Bavaria, Germany. According to the institutional guidelines of Rosenheim Technical University of Applied Sciences and applicable Bavarian school regulations governing educational research in schools, formal review by an institutional ethics committee was not required for the secondary analysis of fully anonymized classroom performance data without biomedical intervention.

The study involved analysis of anonymized performance metrics generated during routine cardiopulmonary resuscitation (CPR) training activities and did not involve collection of personal health information or medical interventions.

Prior to implementation, the study procedures were approved by the school administration in accordance with Bavarian school regulations for educational research in schools (§24 BaySchO). The study was conducted in accordance with the ethical standards of Rosenheim Technical University of Applied Sciences and the principles of the Declaration of Helsinki as applicable to educational research.

Consent to participate

Participation in the performance assessment and accompanying questionnaire was voluntary. Students and their legal guardians received written information about the purpose of the study, the anonymous handling of the collected data, and the voluntary nature of participation.

For students under 16 years of age, information about the study was provided to their legal guardians in advance in accordance with school procedures. Completion and submission of the questionnaire or the digital survey was considered to constitute informed consent to participate.

All data were collected and analyzed in anonymized form.

Consent for publication

Not applicable. The manuscript does not contain identifiable individual data.

Competing interests

The author declares that he has no competing interests.

Funding

The author received no specific funding for this research.

Author contributions

Felix Höpfl conceived the study, designed the research protocol, conducted the instructional interventions, collected and analyzed the data, and wrote the manuscript.

Acknowledgements

The author thanks Katharina Heß for her valuable support in the preparation of the questionnaire and her assistance during the implementation of the 2024 CPR training sessions.

Data availability

The datasets generated and analyzed during the current study are available from the author on reasonable request.

References

1. Schroeder DC, Semeraro F, Greif R, Bray J, Morley P, Parr M, et al. KIDS SAVE LIVES: Basic Life Support Education for Schoolchildren: A Narrative Review and Scientific Statement from the International Liaison Committee on Resuscitation. Circulation. 2023; 147: 1854–68. doi:10.1161/CIR.0000000000001128.

2. Plant N, Taylor K. How best to teach CPR to schoolchildren: a systematic review. Resuscitation. 2013; 84: 415–21. doi:10.1016/j.resuscitation.2012.12.008.

3. Chaves J, Lorca-Marín AA, Vázquez-Bernal B. Beliefs of health professionals and teachers about basic life support in secondary education. Discov Educ. 2025. doi:10.1007/s44217-025-00917-5.

4. Cate O ten. Entrustability of professional activities and competency-based training. Med Educ. 2005; 39: 1176–7. doi:10.1111/j.1365-2929.2005.02341.x.

5. Lockspeiser TM, O'Sullivan P, Teherani A, Muller J. Understanding the experience of being taught by peers: the value of social and cognitive congruence. Adv Health Sci Educ Theory Pract. 2008; 13: 361–72.

doi:10.1007/s10459-006-9049-8.

6. Wood D, Bruner JS, Ross G. The role of tutoring in problem solving. J Child Psychol Psychiatry. 1976; 17: 89–100. doi:10.1111/j.1469-7610.1976.tb00381.x.

7. Belland BR. Scaffolding: Definition, Current Debates, and Future Directions. In: Spector JM, Merrill MD, Elen J, Bishop MJ, editors. Handbook of Research on Educational Communications and Technology. New York, NY: Springer New York; 2014. p. 505–518. doi:10.1007/978-1-4614-3185-5_39.

8. Nicol DJ, Macfarlane-Dick D. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. Studies in Higher Education. 2006; 31: 199–218. doi:10.1080/03075070600572090.

9. Hattie J, Timperley H. The Power of Feedback. Review of Educational Research. 2007; 77: 81–112. doi:10.3102/003465430298487.

10. Masoumian Hosseini M, Qayumi K, Zolfaghari M, Hosseini STM, Haghighi E, Rajabzadhe S, Sabet B. Unleashing the potential of education: embracing a new era of learning through feedback evolution. Discov Educ 2026. doi:10.1007/s44217-026-01110-y.

11. Messick S. Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist. 1995; 50: 741–9. doi:10.1037/0003-066X.50.9.741.

12. van der Vleuten CP. The assessment of professional competence: Developments, research and practical implications. Adv Health Sci Educ Theory Pract. 1996; 1: 41–67. doi:10.1007/BF00596229.

13. Spartinou A, Karageorgos V, Sorokos K, Darivianaki P, Petrakis EC, Papapanagiotou M, et al. Effects of peer-education training on cardiopulmonary resuscitation knowledge and skill retention of secondary school students: a feasibility study. BMJ Open. 2024; 14: e075961. doi:10.1136/bmjopen-2023-075961.

14. Lin L, Ni S, Liu Y, Xue J, Ma B, Xiong D, et al. Effect of peer videorecording feedback CPR training on students' practical CPR skills: a randomized controlled manikin study. BMC Med Educ. 2022; 22: 484. doi:10.1186/s12909-022-03563-9.

15. Carless D, Boud D. The development of student feedback literacy: enabling uptake of feedback. Assessment & Evaluation in Higher Education. 2018; 43: 1315–25. doi:10.1080/02602938.2018.1463354.

16. Schmidt R, Lee T. Motor learning and performance: From principles to application. 6th ed. Champaign, Ill., Windsor, On, Stanningley, Leeds: Human Kinetics; 2020.