



The Effectiveness of a Deep Learning Approach in Enhancing Elementary School Students' Physical Fitness during Physical Education

Angga Fitri Maharini^{1*}, Puspodari², Atrup³

^{1,2,3}Master's Program in Sports Education, Graduate School, Universitas Nusantara PGRI Kediri, Kediri, Indonesia

e-mail: af.maharini@gmail.com*, puspodari@unpkediri.ac.id, salam.atrup@unpkediri.ac.id

Abstract

This classroom action research documented implementation of pedagogical deep learning and within-class changes in physical fitness and participation among 84 fifth-grade students at SDN 2 Pakunden, Blitar City, during 2025/2026. The intervention comprised three iterative cycles with two Physical Education meetings per cycle. Archived records reported a composite fitness score and mastery status at baseline and after each cycle, and observed learning activity during Cycles I-III. Mean fitness scores increased descriptively from 68.12 at baseline to 75.13, 86.16, and 91.45 across the three cycles. The corresponding numbers meeting the school mastery criterion increased from 21/84 (25.00%) to 57/84 (67.86%), 78/84 (92.86%), and 84/84 (100%). Observed activity percentages were 75.08%, 84.67%, and 92.43% in Cycles I-III; no pre-intervention activity measure was available. These temporal changes describe one class cohort and do not establish causal effectiveness because there was no concurrent comparison group, inferential analysis, validated implementation measure, or control for maturation, repeated testing, teacher attention, and cumulative activity. The implementation is best regarded as a locally useful improvement cycle that requires evaluation with standardized measures and a controlled design.

Keywords: pedagogical deep learning; physical education; physical fitness; elementary students; classroom action research

corresponding author: af.maharini@gmail.com

Artikel Info:

Submitted: 04/08/2026

Revised: 01/09/2026

Accepted: 30/09/2026

Published: 04/10/2026

How to Cite: Maharini, A, F., Puspodarini., Atrup. (2026). The Effectiveness of a Deep Learning Approach in Enhancing Elementary School Students' Physical Fitness during Physical Education. *Journal Coaching Education Sports*, 7(2), 608-618.

Author's Contribution: a) Desain Penelitian; b) Pengumpulan Data; c) Analisis Statistik; d) Penyusunan Naskah; e) Pengumpulan Dana



Journal Coaching Education Sports is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

A. Introduction

Physical education supports children's movement competence, physical activity, health, and psychosocial development. At elementary level, instruction should provide sufficient movement practice while helping students understand why activities matter and how skills can be used beyond a lesson (Bailey et al., 2009; World Health Organization, 2020). In Indonesia, Pembelajaran Mendalam is a pedagogical policy framework organized around mindful, meaningful, and joyful learning; it is conceptually different from artificial-intelligence deep learning (Kementerian Pendidikan Dasar dan Menengah, 2025).

The present study did not collect systematic evidence that conventional Physical Education at the school was uniformly repetitive or teacher-directed. That characterization has therefore been removed. The practical problem was defined more narrowly from the archived baseline assessment: a large proportion of the participating class did not meet the school's fitness mastery criterion. This local condition motivated an iterative attempt to improve lesson design and student participation.

The baseline fitness assessment was administered to the 84 participating Grade V students before Cycle I during the 2025/2026 academic year. The archived

manuscript reports a mean composite score of 68.12 and 21 students (25.00%) meeting mastery, but it does not identify the exact date, assessor, standardized testing conditions, numerical mastery cut score, or institutional document authorizing that criterion. Accordingly, the baseline is presented as a school-defined classroom measure rather than a norm-referenced fitness classification.

In this article, deep learning always means pedagogy, not neural networks or artificial intelligence. The Indonesian framework emphasizes learning that is mindful, meaningful, and joyful and describes learning experiences that progress from understanding to application and reflection (Kementerian Pendidikan Dasar dan Menengah, 2025). The labels Surface Learning, Deep Learning, and Transfer Learning are used here as the teacher's instructional sequence: initial knowledge and guided performance, elaboration through reflection and collaboration, and application in a new or authentic movement task.

This operational sequence is related to, but not identical with, the classic distinction between surface and deep approaches to learning. In that tradition, a deep approach involves seeking meaning and connecting ideas, whereas a surface approach is oriented toward limited

reproduction (Marton & Saljo, 1976; Biggs & Tang, 2011). Constructivist and meaningful-learning theories provide additional interpretive support for activating prior knowledge, collaborative sense-making, and application, but the present study did not measure students' learning approaches or cognitive structures directly.

Evidence that school-based physical activity can improve fitness depends strongly on program duration, intensity, measurement, and comparison conditions (Kriemler et al., 2011). The specific contribution of this study is therefore local and methodological rather than causal or internationally novel: it documents three iterative classroom cycles using the Indonesian pedagogical framework alongside repeated school fitness and participation measures. No transparent review located evidence sufficient to claim that this precise sequence is uniquely effective in elementary Physical Education.

The study aimed to (1) describe how pedagogical deep learning was organized across three classroom-action cycles, (2) report within-class changes in the archived composite fitness score and mastery counts from baseline through Cycle III, and (3) describe learning-activity observations across Cycles I-III. Psychomotor skill

scores were removed from the aims and abstract because the source manuscript did not provide a definition, rubric, assessor information, reliability evidence, or a complete result table.

B. Methods

A classroom action research design followed the planning-action-observation-reflection cycle (Kemmis et al., 2014). Three consecutive cycles were conducted, each with two instructional meetings. The design supported iterative classroom improvement but did not include a concurrent comparison group. Consequently, observed temporal change cannot be attributed uniquely to the pedagogy; maturation, repeated testing, increased practice, teacher attention, and cycle-specific modifications remain plausible explanations.

The study involved one Grade V cohort of 84 students enrolled in regular Physical Education at SDN 2 Pakunden, Blitar City, during 2025/2026. All enrolled students were described as participating. The retained manuscript did not report age distribution, sex/gender composition, inclusion or exclusion criteria, recruitment, attendance by meeting, attrition, missing data, or procedures for students with disabilities, illness, or exercise limitations. These details should be reconstructed from

class and attendance records before publication.

Implementation used three instructional functions within each cycle: Surface Learning introduced movement concepts and guided practice; Deep Learning used questioning, reflection, collaboration, problem solving, peer feedback, and self-assessment; and Transfer Learning required application in contextual movement tasks. Across cycles, the teacher used observation and reflection to revise subsequent instruction. However, the session calendar, minutes per meeting, exercise intensity, detailed activities, equipment, workload, attendance, and specific adaptations following each reflection were not preserved in the manuscript. Therefore, the intervention cannot yet be replicated as a fixed protocol. Peer and self-assessment were formative only; the archived text does not support treating them as outcome scores.

Fitness was reported as a 0-100 composite covering muscular strength, speed, flexibility, and cardiorespiratory endurance. The source manuscript did not name the component tests, protocols, units, equipment, number of trials, age/sex norms, normalization formula, component weights, assessors, or safety procedures. Component-level scores and dispersion statistics were also unavailable. The

measure is therefore labeled the school's composite fitness score; it should not be interpreted as a validated standardized fitness battery. A complete test manual and deidentified component data are required for independent replication.

Learning activity was recorded during each intervention cycle using a classroom observation sheet. The archived report supplies only cycle percentages and does not retain observation items, response scale, denominator, observers, training, inter-rater agreement, or procedures separating implementation fidelity from student activity. No pre-intervention activity observation was conducted. Field notes were used for practical cycle reflection, but the qualitative corpus, coding, coders, and credibility procedures were not documented; therefore, no formal qualitative themes are claimed.

C. Result and Discussion

Result

Baseline Physical Fitness

Analysis was descriptive at the class level. For each assessment stage, the available mean composite fitness score, exact mastery count, and mastery percentage were reported. Activity percentages were reported for Cycles I-III only. Standard deviations, individual change scores, component-test

distributions, score ranges, and missing-data details were unavailable. No inferential test, effect size, confidence interval, or repeated-measures model was conducted; terms such as significant, effect, and effectiveness are therefore avoided. Repeated exposure to the same or similar tests may have produced practice effects, but test equivalence and administration conditions were not documented.

The archived classroom dataset included 84 students. At baseline, the mean school composite fitness score was 68.12, and 21 students (25.00%) met the school-defined mastery criterion while 63 (75.00%) did not. The numerical cut score and its curricular basis were not documented, so mastery percentages should be read as local administrative categories rather than external fitness norms.

Table 1. Baseline Physical Fitness of Students

Indicator	Result
Number of students	84
Mean physical fitness score	68.12
Students achieving mastery	21
Students not achieving mastery	63
Mastery learning (%)	25.00
Non-mastery (%)	75.00

These values establish the starting point for the same cohort observed in Cycles I-III. They do not show how students were distributed across individual strength, speed, flexibility, or endurance components because component scores and measures of dispersion were unavailable.

Improvement in Physical Fitness Across the Classroom Action Research

Table 2. Improvement in Students' Physical Fitness Across the Research Cycles

Assessment Stage	Mean Score	Mastery Learning (%)
Baseline	68.12	25.00
Cycle I	75.13	67.86
Cycle II	86.16	92.86
Cycle III	91.45	100.00

The mean composite score increased from 68.12 at baseline to 75.13 in Cycle I,

Cycles

The class mean increased at each recorded stage. Exact mastery counts were 21/84 at baseline, 57/84 in Cycle I, 78/84 in Cycle II, and 84/84 in Cycle III. The same or similar classroom assessment was reportedly repeated, but archived records did not document standardized conditions or procedures to reduce practice effects.

86.16 in Cycle II, and 91.45 in Cycle III. This is an observed temporal pattern within

one cohort, not an estimate of causal impact. Iterative instructional revision, cumulative physical practice, familiarity with testing, maturation, attendance, and assessor expectations may all have contributed.

Observed Learning Activity During the Intervention

Learning-activity percentages were

Table 3. Students' Learning Activity During the Intervention

Research Cycle	Learning Activity (%)
Cycle I	75.08
Cycle II	84.67
Cycle III	92.43

The cycle values show progressively higher observer-recorded activity. Interpretation remains limited because the construct, denominator, observer training, and inter-rater agreement were not preserved.

Summary of Recorded Changes

Table 4. Overall Improvement Following the Implementation of the Deep Learning Approach

Indicator	Baseline	Cycle III	Improvement
Mean physical fitness score	68.12	91.45	+23.33
Mastery learning (%)	25.00	100.00	+75.00
Learning activity (%)	75.08*	92.43	+17.35

By Cycle III, the class mean fitness score was 23.33 points above baseline and all 84 students met the school-defined mastery criterion. Observer-recorded learning activity was 17.35 percentage points higher than in Cycle I. Without dispersion statistics, individual-level trajectories, standardized measures, or a control group, the magnitude and generalizability of these changes cannot be determined.

available only after implementation began: 75.08% in Cycle I, 84.67% in Cycle II, and 92.43% in Cycle III. Because no pre-intervention observation was recorded, the study cannot estimate change from baseline. The percentages describe differences across cycles under an incompletely documented observation instrument.

Table 4 presents the first and final available observations. For fitness, the first observation was baseline; for learning activity, the first observation was Cycle I. These different starting points are labeled explicitly and should not be treated as equivalent baselines

Discussion

This classroom action research observed progressively higher class-level fitness scores, mastery counts, and activity percentages during three iterative cycles. The pattern indicates that the teacher's improvement process coincided with favorable changes. It does not demonstrate that pedagogical deep learning caused those changes, because intervention components

were not isolated and competing explanations were not controlled.

The instructional rationale was to combine movement practice with attention, meaning, reflection, collaboration, and contextual application. These features are compatible with the Indonesian Pembelajaran Mendalam framework and with established accounts of deep approaches to learning (Biggs & Tang, 2011; Kementerian Pendidikan Dasar dan Menengah, 2025). Nevertheless, the study did not measure mindful, meaningful, or joyful experience, learning approach, mediation, or implementation fidelity, so the proposed mechanism remains a hypothesis.

The observed fitness change must also be interpreted physiologically. School-based physical activity interventions can improve fitness, but effects depend on exposure, intensity, duration, adherence, and assessment quality (Kriemler et al., 2011). Here, six meetings were reported without minutes, workload, intensity, or attendance. It is therefore unclear whether changes reflect physiological adaptation, test familiarity, scoring changes, increased effort, or a combination of processes.

Constructivist and meaningful-learning theories can inform lesson design by encouraging learners to connect prior experience, feedback, and new movement

challenges. Peer discussion and self-reflection may help students understand technique and purpose. In this study, however, those activities were not linked analytically to fitness change, and peer/self-assessment procedures were not validated. They should be described as instructional components rather than demonstrated causal mechanisms.

Activity observations increased across cycles, but no true activity baseline was available. Greater observed activity may represent improved engagement, growing familiarity with classroom routines, observer expectations, or easier alignment with later-cycle tasks. Future studies should define activity behavior, use trained independent observers or objective movement measures, and collect comparable observations before implementation and throughout follow-up.

Relative to controlled school-based physical-activity research, this study has a shorter, less standardized exposure and a weaker causal design (Kriemler et al., 2011). Its value lies in showing how a teacher used iterative planning and reflection to organize a local improvement process, not in estimating an intervention effect. The local result can inform a more rigorous feasibility protocol with prespecified sessions, fidelity criteria, standardized fitness tests, and adverse-

event monitoring.

The three labels used by the teacher should also be interpreted cautiously. Surface Learning, Deep Learning, and Transfer Learning were sequential instructional functions in this classroom, whereas classic surface and deep approaches describe how learners engage with tasks (Marton & Saljo, 1976). Future research should establish the conceptual mapping and test whether students actually experience the intended shift in processing and transfer.

Practical implications are provisional. Teachers may consider testing lessons that make exercise purposes explicit, provide structured practice and feedback, invite reflection, and require application in varied movement contexts. Such recommendations should be evaluated rather than assumed effective. Schools should also ensure that workload, accessibility, health screening, hydration, and progression are appropriate for children.

The contribution is primarily local: an initial implementation record of Indonesia's pedagogical deep-learning framework in elementary Physical Education. Policy recency alone does not establish international novelty, and the current evidence cannot show superiority over cooperative learning, problem-based

learning, direct instruction, or usual practice. Comparative research is needed.

Methodological limitations are substantial. The study used one school, one cohort, no control group, and iterative changes between cycles. Measures were underspecified and not shown to be valid or reliable; assessors may have known the cycle; repeated testing may have improved scores; and participant characteristics, attendance, missingness, and potential confounders were not analyzed. No dispersion, individual change, inferential uncertainty, or follow-up was available.

Future studies should employ experimental or quasi-experimental designs; preregister outcomes and analysis; report age, sex, health status, disability accommodations, attendance, attrition, and adverse events; use named age-appropriate fitness protocols with component scores; validate observation and psychomotor rubrics; measure implementation fidelity; and report individual-level change with uncertainty. Longer follow-up is needed to distinguish short-term test familiarity from sustained fitness development.

D. Conclusion

During three classroom-action cycles involving 84 fifth-grade students at one elementary school, mean school composite fitness scores, mastery counts, and

observed activity percentages increased over time. The findings document within-class improvement during iterative implementation of pedagogical deep learning; they do not establish causal effectiveness. Interpretation is limited by the absence of a comparison group, short and incompletely specified exposure, unvalidated composite and observation measures, missing dispersion statistics, possible assessor and repeated-test effects, and incomplete participant reporting.

The implementation provides a basis for developing a replicable Physical Education protocol that combines guided movement practice, reflection, collaboration, and transfer tasks. Its effectiveness and safety should be tested using standardized fitness measures, independent or blinded assessment where feasible, implementation-fidelity evidence, and controlled longitudinal designs.

E. Acknowledgments

The authors would like to express their sincere gratitude to the Principal, Physical Education teachers, and fifth-grade students of SDN 2 Pakunden, Blitar City, Indonesia, for their cooperation and active participation throughout this study. The authors also acknowledge the support provided by the Master's Program in

Sports Education, Graduate School, Universitas Nusantara PGRI Kediri, which facilitated the completion of this research.

F. Conflict of Interest

The authors declare no conflicts of interest. Research involved children and physical exertion. The archived manuscript confirms school cooperation but does not contain an ethics-review body and approval number, written parental/guardian consent, child assent, confidentiality procedures, permission for record access, health screening, or adverse-event monitoring. These protections must be verified from original field documentation and reported before publication; they must not be inferred retrospectively.

References

- Alm, K., Beery, T. H., Eiblmeier, D., & Fahmy, T. (2022). Students' learning sustainability – implicit, explicit or non-existent: A case study approach on students' key competencies addressing the SDGs in HEI program. *International Journal of Sustainability in Higher Education*, 23(8), 60–84. <https://doi.org/10.1108/IJSHE-12-2020-0484>
- Jia, Y., Anida Abdullah, N., Eliza, H., Lu, Q., Si, D., Guo, H., & Wang, W. (2025). A narrative review of deep learning applications in sports performance analysis: Current practices, challenges, and future directions. *BMC Sports Science*,

- Medicine and Rehabilitation*, 17(1), 249.
<https://doi.org/10.1186/s13102-025-01294-0>
- Jufriadi, J., & Wahibah, W. (2025). Faith-Driven Innovation in Practice: Investigating MBKM within Islamic Higher Education in Indonesia. *Utamax: Journal of Ultimate Research and Trends in Education*, 7(1), 71–86.
<https://doi.org/10.31849/utamax.v7i1.24627>
- Li, F. (2025). AI-Enhanced Curriculum Design and Deep-Learning-Based Assessment in International Sports Communication Education: *International Journal of Information and Communication Technology Education*, 21(1), 1–18.
<https://doi.org/10.4018/IJICTE.391357>
- Liu, W. (2024). Design and user behavior analysis of an English learning social platform based on digital entertainment content recommendation algorithm. *Entertainment Computing*, 51, 100734–100734.
<https://doi.org/10.1016/j.entcom.2024.100734>
- Lozano-Lozano, M., Fernández-Lao, C., Cantarero-Villanueva, I., Noguerol, I., Álvarez-Salvago, F., Cruz-Fernández, M., Arroyo-Morales, M., & Galiano-Castillo, N. (2020). A Blended Learning System to Improve Motivation, Mood State, and Satisfaction in Undergraduate Students: Randomized Controlled Trial. *Journal of Medical Internet Research*, 22(5), e17101.
<https://doi.org/10.2196/17101>
- Pasaribu, A. M. N., Budidarma, A., & Nasution, S. D. K. (2025). Collaboration between the Ministry of Education and Culture and the Ministry of Youth and Sports in Successful National Sports Grand Design (DBON) Through Early Identification of Student Talents. *Journal Coaching Education Sports*, 6(1), 102–113.
<https://doi.org/10.31599/d3f7cx14>
- Pietraszewski, P., Terbalyan, A., Rocznik, R., Maszczyk, A., Ornowski, K., Manilewska, D., Kuliś, S., Zając, A., & Gołaś, A. (2025). The Role of Artificial Intelligence in Sports Analytics: A Systematic Review and Meta-Analysis of Performance Trends. *Applied Sciences*, 15(13), 7254.
<https://doi.org/10.3390/app15137254>
- Purnomo, E., Ma'mun, A., Winarno, M. E., Mardesia, P., & Jermaina, N. (2025). Integrating Values Education in Physical Education Learning: A Holistic Approach to Enhancing Student Well-Being. *Seminars in Medical Writing and Education*, 4, 769.
<https://doi.org/10.56294/mw2025769>
- Sumarsono, S., & Wibowo, M. P. (2025). Artificial Intelligence in Recordkeeping: A Systematic Review of Machine Learning Applications for Automated Records Classification. *Khizanah Al-Hikmah: Jurnal Ilmu Perpustakaan, Informasi, Dan Kearsipan*, 13(1), 148–159.
<https://doi.org/10.24252/v13i1a12>
- Syahrudin, Saleh, M. S., & Saleh, Muh. S. (2025). A Study on Physical Education Outcomes and Motor Competence Based on Students' Sports Interests in School Settings. *COMPETITOR: Jurnal Pendidikan Kepratihan Olahraga*, 17(1), 519–536.
<https://doi.org/10.26858/cpjok.v17i1.220>
- Uçar, M. H. B., Solak, S., Jimale, A. O., Ünal, H., & Eken, S. (2025). A comprehensive survey on pose

estimation and tracking in sports:
Methodologies, datasets,
challenges, and future directions.
Artificial Intelligence Review,
59(2), 54.
<https://doi.org/10.1007/s10462-025-11344-1>

Zhang, Y., Fu, X., & Wu, Z. (2025). A Case Study of Big Data and Artificial Intelligence in Physical Education: *Journal of Cases on Information Technology*, 27(1), 1–20.
<https://doi.org/10.4018/JCIT.386841>