

Numbered Heads Together and Fifth-Grade Pancasila Education Achievement: A Classroom Action Study

Abdul Mannan Bamu

Department of Primary School Teacher Education, Faculty of Teacher Training and Education, Universitas Pohnuato, Indonesia

Corresponding author: Abdul Mannan Bamu, Universitas Pohnuato, Indonesia

Email: AbdulMannanBamu@gmail.com

Article History	Received: 11/04/2026	Accepted: 10/05/2026	Published: 20/05/2026
-----------------	----------------------	----------------------	-----------------------

ABSTRACT

Pancasila Education requires students to understand civic norms and apply them in social situations, yet teacher-centred instruction can limit participation and achievement. This classroom action study examined changes in the learning achievement of 22 fifth-grade students at SDN 02 Tilamuta, Indonesia, after the implementation of Numbered Heads Together (NHT). The study included a baseline assessment and two cycles of planning, action, observation, and reflection. During each cycle, students worked in numbered groups, discussed teacher questions, prepared a shared response, and answered when their number was selected randomly. Achievement was measured with teacher-developed written tests scored from 0 to 100, with mastery defined as at least 75. Student-level scores were reanalysed using descriptive statistics, a Friedman test, Wilcoxon signed-rank comparisons with Holm adjustment, Kendall effect size, Cochran Q, and exact McNemar tests. Mean achievement increased from 60.00 at baseline to 75.00 in Cycle I and 90.68 in Cycle II. The repeated-measures change was statistically significant and large, $\chi^2(2) = 40.54$, $p < .001$, Kendall $W = .921$. Mastery increased from 31.8% to 72.7% and then 100%, Cochran $Q(2) = 22.80$, $p < .001$. The findings indicate strong local improvement associated with iterative NHT implementation. However, the absence of a comparison group and the lack of reported test validity and reliability prevent definitive causal conclusions.

Keywords: classroom action research; cooperative learning; elementary education; learning achievement; Numbered Heads Together; Pancasila Education

1. INTRODUCTION

Pancasila Education in Indonesian elementary schools is intended to develop knowledge of civic principles and the capacity to enact values such as responsibility, mutual respect, cooperation, and compliance with social norms. Achievement in this subject cannot be reduced to factual recall. Students need opportunities to explain concepts, examine everyday examples, compare rights and obligations, and justify appropriate conduct. Instruction that remains dominated by teacher explanation may provide information but can leave many students cognitively and socially passive.

The classroom problem addressed in this study was observed in Grade V at SDN 02 Tilamuta. Students showed limited focus, some continued to play or distract peers during



lessons, and participation was concentrated among a small number of students. The source thesis also reported that many students perceived Pancasila Education as uninteresting. Student-level baseline data confirmed low achievement: only 7 of 22 students reached the school mastery criterion of 75, although a separate narrative passage reported a conflicting figure. The individual score table provides the more auditable evidence and was therefore used in the present analysis.

Cooperative learning offers a structured response to passive participation. Its effectiveness depends on more than seating students in groups. Positive interdependence must be combined with individual accountability, promotive interaction, and explicit processing of group performance (Johnson & Johnson, 2009; Slavin, 2014). When these conditions are present, students are motivated to explain ideas, check one another's understanding, and contribute to a shared goal. Teachers nevertheless need a structure that prevents one confident learner from completing the task for the entire group.

Numbered Heads Together (NHT) is designed to strengthen both cooperation and accountability. Students are assigned numbers within a small group. The teacher poses a question, group members put their heads together to construct and verify an answer, and the teacher calls a number at random. Because any numbered member may have to respond, every student has a reason to understand the group answer. This structure can distribute participation, increase preparation, and expose misconceptions that remain hidden during whole-class questioning. Recent elementary studies have reported higher outcomes under NHT, although most evidence comes from specific subjects and short interventions (Al Kibtiyah & Iba, 2022; Saptono, 2024; Wati & Suarni, 2020).

Research on NHT frequently reports only mean gains or percentages of students who meet a criterion. Such reporting is useful for classroom decisions but does not fully exploit repeated observations of the same students. It also risks overstating causality when there is no comparison group. The present article converts the original classroom action report into a transparent repeated-measures analysis. It distinguishes improvement over time from proof that NHT alone caused the improvement, reports practical as well as statistical change, and identifies inconsistencies in the source record rather than concealing them.

The study aimed to determine whether fifth-grade students' Pancasila Education achievement improved across baseline, Cycle I, and Cycle II during the implementation of NHT. Two questions guided the analysis: (1) Did achievement scores change significantly across the three occasions? (2) Did the proportion of students meeting the school mastery criterion increase across the three occasions?

2. LITERATURE REVIEW

2.1 Cooperative Learning and Individual Accountability

Social interdependence theory explains that the way goals are structured shapes interaction and learning. Positive interdependence exists when learners perceive that they can reach their goals only when their peers also reach theirs. This condition can promote explanation, assistance, and mutual regulation. Cooperative learning becomes academically productive when group goals are linked to individual learning rather than to the mere completion of a shared product (Johnson & Johnson, 2009; Slavin, 2014).

Implementation quality is therefore critical. Teachers must form manageable groups, give intellectually meaningful tasks, monitor participation, require individual preparation, and



provide feedback. Abramczyk and Jurkowski (2020) showed that teachers may value cooperative learning but do not always implement its evidence-based elements consistently. A classroom action design is useful in this context because the teacher can identify weaknesses in the first cycle and refine group composition, questioning, timing, or feedback in the next cycle.

2.2 The Numbered Heads Together Structure

NHT operationalises individual accountability through random selection. The instructional sequence normally includes numbering students, presenting a question, allowing group discussion, calling a number, and inviting the selected students to respond. Every learner must be ready because the respondent is not known during discussion. This uncertainty is pedagogically productive when the classroom climate is supportive and the questions require explanation rather than guessing.

Empirical studies in elementary contexts generally associate NHT with stronger learning outcomes, engagement, or motivation. Wati and Suarni (2020) found positive outcomes in social studies, while Al Kibtiyah and Iba (2022) reported an effect on elementary learning outcomes when NHT was supported by audiovisual media. Saptono (2024) also documented improved achievement after NHT implementation. These findings are consistent with the proposed mechanisms of peer explanation, rehearsal, and accountability. However, subject matter, teacher skill, instruments, and research design differ substantially across studies, so local classroom evidence must be interpreted within its own boundaries.

2.3 Achievement in Pancasila Education

Pancasila Education connects conceptual knowledge with civic judgement and conduct. In the present curriculum unit, students studied norms, rights, and obligations in everyday life. Cooperative discussion is relevant because such content often involves examples, reasons, and alternative responses rather than a single memorised definition. Students can compare cases, challenge an incomplete answer, and connect principles with school and community situations.

The theoretical expectation was that the NHT sequence would increase achievement by requiring all group members to process the material before an individual response could be requested. The expected benefit was not simply greater talk. Productive discussion had to be aligned with the learning objectives, followed by teacher feedback, and assessed individually. The novelty of the present analysis lies in testing the full student-level trajectory and mastery transitions rather than relying only on aggregate classroom percentages.

3. RESEARCH METHOD

3.1 Research Design

The study used classroom action research based on the recurring stages of planning, action, observation, and reflection. It comprised a baseline assessment followed by two instructional cycles. Cycle I was implemented on 1 December 2025 and Cycle II on 5 December 2025. The design was appropriate for improving an identified classroom

practice and evaluating whether the local success criterion was reached. It was not designed to estimate a causal treatment effect against a counterfactual condition.

3.2 Setting and Participants

Participants were all 22 students in one fifth-grade class at SDN 02 Tilamuta, Boalemo Regency, Gorontalo Province, Indonesia. The class comprised 13 boys and 9 girls. Total-class participation was appropriate for classroom action research because the action addressed an existing instructional problem and was implemented within the normal class. No sampling-based claim to represent other schools was made.

3.3 NHT Intervention and Cycle Refinement

Students were organised into five groups containing approximately four or five members. Each member received a number. The teacher introduced the topic, posed questions or distributed a worksheet, and allowed group members to discuss and agree on the strongest answer. The teacher then called a number at random, and the student with that number represented the group. Responses were followed by feedback and additional questions. Individual written work was administered at the end of the lesson.

Cycle I focused on rules and norms in everyday life. Reflection indicated that several students still required stronger engagement and support, and six remained below the mastery threshold. Cycle II retained the core NHT structure while improving monitoring, reinforcement, and the opportunity for every member to prepare an answer. The content remained within the same Pancasila Education unit concerning norms, rights, and obligations. Because the source thesis does not provide a complete fidelity checklist, implementation changes are described at the level documented in the classroom record.

3.4 Instruments and Outcome Definition

The primary outcome was the student's written achievement score on a 0-100 scale at baseline, Cycle I, and Cycle II. The school mastery criterion was 75. The thesis describes the assessments as pretest, posttest, worksheets, or written tests, but it does not report the number of items, item format, scoring blueprint, evidence of content validation, reliability, or equivalence of difficulty across occasions. The present article therefore treats the scores as teacher-developed classroom assessments and avoids claims of standardised measurement.

Observation sheets were also used to describe cognitive, affective, and psychomotor classroom activity. Their scoring structure and rater reliability were not documented sufficiently for defensible inferential analysis. These observations informed reflection between cycles but were not combined with written-test scores or used as a second outcome in the statistical reanalysis.

3.5 Data Analysis

Means, standard deviations, medians, interquartile ranges, score ranges, and mastery percentages were calculated for each occasion. Because the same 22 students were measured three times and the scores were discrete with evident ceiling effects in Cycle II, an omnibus Friedman test was used. Pairwise changes were examined with two-sided Wilcoxon signed-rank tests and Holm adjustment. Kendall W quantified the magnitude of the overall repeated-measures change, and matched-pairs rank-biserial correlations quantified pairwise effects. Mastery status was analysed with Cochran Q and exact



McNemar tests with Holm adjustment. The significance level was .05. Statistical results were recalculated directly from the student-level score tables in the thesis.

3.6 Ethical Considerations

The source thesis describes school-based coordination and classroom implementation but does not report an institutional ethics approval number, written parental consent, or child assent procedures. Student names were removed from the present manuscript and only de-identified scores were analysed. Before journal submission, the author should retain any available school permission and consent documentation and respond transparently to editorial ethics queries. No ethics approval information was invented.

4. RESULTS

4.1 Descriptive Change in Achievement

Achievement increased at each measurement occasion. The mean rose by 15.00 points from baseline to Cycle I and by a further 15.68 points from Cycle I to Cycle II. The baseline-to-final mean gain was 30.68 points. The median shifted from 60 to 80 and then 90. Dispersion remained similar during the first cycle but decreased markedly in Cycle II, indicating that the class became both higher-performing and more homogeneous. Eight students obtained the maximum score of 100 in Cycle II, which signals a possible ceiling effect.

Table 1

Descriptive Statistics and Mastery Across Measurement Occasions

Occasion	M	SD	Median	IQR	Range	Mastery, n (%)
Baseline	60.00	16.33	60	40-80	40-80	7 (31.8)
Cycle I	75.00	16.26	80	65-90	50-90	16 (72.7)
Cycle II	90.68	8.63	90	82.5-100	75-100	22 (100.0)

Note. Scores range from 0 to 100. Mastery was defined by the school criterion of a score of at least 75. IQR = interquartile range. The baseline table in the thesis contained 7 mastered students, or 31.8%; an earlier narrative statement of 45.45% was inconsistent with the student-level data and was not used.

4.2 Repeated-Measures and Pairwise Tests

The Friedman test demonstrated a strong ordered change across the three occasions. Kendall W of .921 indicates that the measurement occasion accounted for a very large and consistent ordering of student performance. Pairwise analysis showed that 21 of 22 students improved from baseline to Cycle I, while one student declined by 10 points. From Cycle I to Cycle II, 19 improved and 3 retained the same score. Every student scored higher in Cycle II than at baseline. The pairwise effect sizes were correspondingly large.

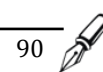


Table 2
Repeated-Measures Analysis and Pairwise Score Comparisons

Analysis	Statistic	p	Holm-adjusted p	Effect size	Direction of change
Friedman test	chi-square(2) = 40.54	< .001	Not applicable	Kendall W = .921	Overall increase
Baseline vs Cycle I	Wilcoxon W = 7.00	< .001	< .001	r _{rb} = .945	21 higher, 1 lower
Cycle I vs Cycle II	Wilcoxon W = 0.00	< .001	< .001	r _{rb} = 1.000	19 higher, 3 equal
Baseline vs Cycle II	Wilcoxon W = 0.00	< .001	< .001	r _{rb} = 1.000	22 higher

Note. $n = 22$. Pairwise tests were two-sided Wilcoxon signed-rank tests using the available nonzero differences; p values were adjusted across the three contrasts with the Holm procedure. r_{rb} = matched-pairs rank-biserial correlation. Kendall W was calculated as $\text{chi-square}/[n(k - 1)]$.

4.3 Mastery Transitions

The classroom success criterion was not reached in Cycle I because 72.7% achieved mastery, slightly below the thesis target of at least 75%. The action therefore appropriately continued. By Cycle II, all 22 students reached at least 75. Cochran Q confirmed that mastery status changed across occasions. Exact paired tests showed significant gains at each transition, including the six students who moved from non-mastery in Cycle I to mastery in Cycle II. No student lost mastery between occasions.

Table 3
Mastery-Status Changes Across Measurement Occasions

Comparison	Newly mastered	Lost mastery	Unchanged	Exact McNemar p	Holm-adjusted p
Baseline to Cycle I	9	0	13	.004	.008
Cycle I to Cycle II	6	0	16	.031	.031
Baseline to Cycle II	15	0	7	< .001	< .001

Note. Cochran Q across all three occasions was 22.80, $df = 2$, $p < .001$. Newly mastered indicates a change from below 75 to at least 75. No student moved from mastery to non-mastery in any comparison.

5. DISCUSSION

5.1 Interpretation of the Achievement Trajectory

The results show a pronounced classroom-level improvement during the NHT action. The pattern was not driven only by a small group of already successful students. Nearly the entire class improved in Cycle I, and all students improved between baseline and Cycle II. The reduction in the Cycle II standard deviation is educationally important because it suggests that initially lower-performing students moved closer to their peers rather than the class mean rising while gaps remained unchanged.



The NHT structure provides a plausible mechanism. Numbering and random response create individual accountability within a cooperative task. Students cannot rely completely on the most confident group member because any number may be called. The heads-together phase also allows students to rehearse language, compare examples, and correct incomplete interpretations before public response. These features align with social interdependence theory and with the requirement that cooperative learning connect group success to individual understanding (Johnson & Johnson, 2009; Slavin, 2014).

The improvement from Cycle I to Cycle II reinforces the action-research logic. NHT is not a fixed technique that produces an automatic effect after one use. The first cycle exposed continuing difficulties and created evidence for refinement. Stronger teacher monitoring, reinforcement, and repeated experience with the response routine likely reduced uncertainty and expanded participation. Research on teachers' use of cooperative learning similarly indicates that outcomes depend on faithful implementation of accountability and interaction structures (Abramczyk & Jurkowski, 2020).

5.2 Relationship to Previous Research

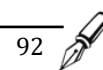
The direction of the findings is consistent with elementary studies reporting higher achievement under NHT. Saptono (2024) found improved learning outcomes after NHT, and Al Kibtiyah and Iba (2022) reported an effect in an elementary quasi-experiment. Wati and Suarni (2020) linked NHT-supported instruction with social studies outcomes and motivation. Although those studies used different subjects and designs, they support the proposition that structured peer discussion and random individual responding can improve academic engagement and performance.

This study extends that evidence to a Pancasila Education lesson on norms, rights, and obligations. The subject is particularly compatible with accountable discussion because students must relate principles to lived examples. However, the final score distribution was compressed near the maximum. The finding may reflect strong learning, easier Cycle II assessment, or both. Without documented test equivalence, the magnitude should be interpreted as the observed classroom trajectory rather than a precise estimate of an NHT effect.

5.3 Practical Implications

Teachers applying NHT should retain its defining sequence. Groups need to be small enough for every student to speak. Questions should require reasoning or application, not only recall. The teacher should allow sufficient heads-together time, circulate to detect misconceptions, and call numbers unpredictably. Feedback should address both the accuracy of the answer and the quality of the group explanation.

Classroom monitoring should examine individual score trajectories and mastery transitions rather than only a class percentage. A high mastery rate can hide students who regress, while a mean gain can hide persistent non-mastery. In this study, paired records showed that no student lost mastery and that all six remaining non-mastered students crossed the criterion in Cycle II. This information provides a stronger basis for deciding whether the action should stop or continue.



5.4 Limitations and Threats to Validity

The study has important limitations. It involved one intact class and no control or comparison condition. Score gains may partly reflect repeated practice, maturation, increased teacher attention, familiarity with classroom routines, or differences in assessment difficulty. The teacher and researcher roles were closely connected, which may also introduce expectancy effects.

The achievement tests were not described in enough detail to establish content equivalence, construct validity, internal consistency, or scoring reliability. The Cycle II ceiling effect limits discrimination among higher-performing students. The intervention lasted only two closely spaced cycles, so retention and transfer to later Pancasila Education topics are unknown. Observational ratings could not be analysed rigorously because the source did not provide sufficient scoring and reliability information. Future studies should use parallel validated tests, independent or blinded scoring where feasible, fidelity checks, delayed follow-up, and a comparison class when the research aim shifts from local improvement to causal evaluation.

6. CONCLUSION

The implementation of Numbered Heads Together was associated with a strong and consistent improvement in fifth-grade Pancasila Education achievement at SDN 02 Tilamuta. Mean scores increased from 60.00 at baseline to 75.00 in Cycle I and 90.68 in Cycle II. Mastery increased from 31.8% to 72.7% and then 100%. Repeated-measures and mastery-status analyses confirmed statistically significant changes with large effect sizes. The instructional value of NHT appears to lie in its combination of peer explanation, shared preparation, random individual accountability, and iterative teacher feedback. The results justify NHT as a practical classroom-improvement option in this setting. They do not, however, prove that NHT alone caused the gains because the study lacked a comparison group and did not establish the psychometric quality or equivalence of its tests.

ACKNOWLEDGEMENTS

The author thanks the principal, Grade V teacher, students, and school community of SDN 02 Tilamuta for supporting the classroom implementation and data collection.

AUTHOR CONTRIBUTIONS

Abdul Mannan Bamu: Conceptualisation, methodology, classroom implementation, observation, data curation, formal analysis, writing of the original draft, review, and editing. The author approved the final manuscript and accepts responsibility for its integrity.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

DATA AVAILABILITY

The de-identified achievement scores supporting this article may be requested from the corresponding author, subject to school permission and applicable confidentiality requirements.



CONFLICTS OF INTEREST

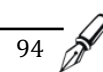
The author declares no conflicts of interest related to this research, authorship, or publication.

GENERATIVE ARTIFICIAL INTELLIGENCE DISCLOSURE

OpenAI ChatGPT was used to assist English-language editing, manuscript restructuring, and statistical reanalysis of the original de-identified score matrix. It was not used to create participants, raw scores, references, findings, or ethical approval information. The author reviewed the revised manuscript and remains responsible for its accuracy and integrity.

REFERENCES

- Abramczyk, A., & Jurkowski, S. (2020). Cooperative learning as an evidence-based teaching strategy: What teachers know, believe, and how they use it. *Journal of Education for Teaching*, 46(3), 296-308.
<https://doi.org/10.1080/02607476.2020.1733402>
- Al Kibtiyah, M., & Iba, K. (2022). The effect of Numbered Heads Together (NHT) cooperative learning model assisted by audio visual media on learning outcomes of elementary school students. *Jurnal Paedagogy*, 9(4).
<https://doi.org/10.33394/jp.v9i4.5494>
- Astutik, P., & Wulandari, S. S. (2021). Analisis model pembelajaran Numbered Head Together dalam meningkatkan keaktifan siswa. *Jurnal Pendidikan Administrasi Perkantoran*, 9(1), 154-168. <https://doi.org/10.26740/jpap.v9n1.p154-168>
- Cahya, L. P. (2024). NHT cooperative model for learning independence of Pancasila Education in elementary school: Why teacher-centered must change? *Cakrawala: Journal of Citizenship Teaching and Learning*, 2(2), 161-168.
<https://doi.org/10.70489/kpb3nk82>
- Haliadi, Satibi, O., & Muji Utami, N. C. (2024). Comparison of social science learning results using cooperative learning models of Think Pair Share and Numbered Heads Together type. *ELSE: Elementary School Education Journal*, 8(2).
<https://doi.org/10.30651/else.v8i2.21585>
- Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, 38(5), 365-379. <https://doi.org/10.3102/0013189X09339057>
- Palupi, D. I., Rahmani, E., Yusnita, E., Pertiwi, H., Gustina, H., & Priyanti, N. (2023). Mengenal model kooperatif Numbered Head Together untuk pembelajaran anak usia dini. *EDUKASIA: Jurnal Pendidikan dan Pembelajaran*, 4(1), 21-28.
<https://doi.org/10.62775/edukasia.v4i1.89>
- Ristiana, B., Ni'mah, N. U., & Sutriyani, W. (2022). Peranan pembelajaran model Number Head Together terhadap hasil belajar materi aproksima di sekolah dasar. *Jurnal PGSD Indonesia*, 8(2), 1-9.
- Saptono, B. (2024). The effectiveness of Numbered Head Together in learning. *Jurnal Prima Edukasia*, 12(2), 293-299. <https://doi.org/10.21831/jpe.v12i2.74975>



- Segundo Marcos, R. I., López Fernández, V., Daza González, M. T., & Phillips-Silver, J. (2020). Promoting children's creative thinking through reading and writing in a cooperative learning classroom. *Thinking Skills and Creativity*, 36, 100663. <https://doi.org/10.1016/j.tsc.2020.100663>
- Slavin, R. E. (2014). Cooperative learning and academic achievement: Why does groupwork work? *Anales de Psicología*, 30(3), 785-791. <https://doi.org/10.6018/analesps.30.3.201201>
- Suradi, F. M., & Aliyyah, R. R. (2023). Peningkatan hasil belajar matematika melalui model pembelajaran Numbered Head Together di sekolah dasar. *Jurnal Pengajaran Sekolah Dasar*, 2(1), 113-124. <https://doi.org/10.56855/jpsd.v2i1.351>
- Wati, N. K. S., & Suarni, N. K. (2020). Social studies learning with Numbered Head Together model improves learning outcomes viewed from student learning motivation. *International Journal of Elementary Education*, 4(2), 244-254. <https://doi.org/10.23887/ijee.v4i2.25250>
- Ziliwu, M. H., Bawamenewi, A., Lase, B. P., & Harefa, H. O. N. (2024). Peranan Pendidikan Pancasila dalam membentuk karakter peserta didik di sekolah. *JIIIP: Jurnal Ilmiah Ilmu Pendidikan*, 7(9), 9956-9965. <https://doi.org/10.54371/jiip.v7i9.5839>

