Taushiah: Jurnal Hukum, Pendidikan dan Kemasyarakatan ISSN: 2599-1353
Vol. 15 No.2 (2025)

COMPARATIVE COGNITIVE LEARNING OUTCOMES FOLLOWING GAME-BASED LEARNING IN ISLAMIC CREED AND MORALS: A Quasi-Experimental Study

Mhd Rafli Halim, Ramlan Padang, Abu Bakar Khazali
Universitas Islam Sumatera Utara
Email. rafli.2019we@gmail.com, ramlan@fai.uisu.ac, abu.bakar@fai.uisu.ac.id

Abstrak

Penelitian ini membandingkan perubahan hasil belajar kognitif Akidah Akhlak setelah penerapan Game-Based Learning (GBL) dan pembelajaran konvensional di Madrasah Aliyah Negeri 3 Medan. Penelitian menggunakan desain eksperimen semu dengan pretest–posttest pada dua kelas yang dipilih secara purposif, masing-masing 30 siswa kelas XI MIPA. Data diperoleh melalui tes pilihan ganda sebanyak 15 butir, observasi pembelajaran, dan dokumentasi. Rata-rata kelompok GBL berubah dari 61,47 menjadi 84,50, sedangkan kelompok kontrol dari 60,47 menjadi 70,20. Selisih kenaikan rata-rata kedua kelompok adalah 13,30 poin. Statistik uji beda posttest yang dilaporkan, t(58) = 19,049 dengan p < 0,001, secara pendekatan sesuai dengan rerata dan simpangan baku yang dicantumkan. Namun, N-Gain yang dilaporkan sebesar 70,40% untuk kelompok eksperimen dan ukuran efek 1,09 belum dapat direkonsiliasi tanpa skor individual serta penjelasan rumus perhitungan. Temuan deskriptif menunjukkan kenaikan hasil belajar kognitif yang lebih besar pada kelompok GBL, sementara besarnya efek dan kesimpulan inferensial tetap memerlukan audit data asli. Penelitian ini tidak mengukur internalisasi nilai akhlak atau perubahan perilaku jangka panjang.

Kata kunci: Game-Based Learning, Hasil Belajar Kognitif, Eksperimen Semu

Abstract

This quasi-experimental study compared changes in cognitive achievement in Islamic Creed and Ethics following game-based learning (GBL) and conventional instruction at Madrasah Aliyah Negeri 3 Medan. Two purposively selected eleventh-grade science classes included 30 students each. Data were drawn from a 15-item multiple-choice achievement test, classroom observations, and instructional documentation. The GBL group's mean increased from 61.47 to 84.50, while the control group's mean increased from 60.47 to 70.20, yielding a between-group difference in mean gains of 13.30 points. The reported posttest statistic, t(58) = 19.049, p < .001, is approximately consistent with the published group means and standard deviations. However, the reported experimental N-gain of 70.40% and effect-size value of 1.09 cannot be reconciled without individual scores and explicit calculation procedures. The descriptive results indicate greater cognitive achievement gains in the GBL group, but effect magnitude and definitive inferential conclusions require verification against the original dataset. The study did not assess sustained moral behavior or the internalization of ethical values.

Keywords: Game-Based Learning, Cognitive Outcomes, Quasi-Experiment

INTRODUCTION

Islamic education plays a fundamental role in developing students' intellectual abilities, spiritual awareness, moral character, and social responsibility. Within the Indonesian Islamic educational system, Islamic Creed and morals constitutes an essential subject that integrates religious knowledge with the development of ethical behavior and personal integrity. Unlike subjects primarily emphasizing cognitive achievement, Islamic Creed and Ethics education requires students to understand religious principles, internalize moral values, and demonstrate appropriate behavior in everyday situations. Consequently, effective instruction should facilitate meaningful learning experiences that connect theoretical knowledge with practical moral decision-making. However, achieving these educational objectives remains challenging when instructional practices predominantly emphasize memorization, teacher explanations, and passive knowledge acquisition.

Traditional teacher-centered instruction continues to influence classroom practices in various educational contexts. Although lectures, discussions, and question-and-answer activities provide structured opportunities for knowledge transmission, their effectiveness may be limited when students have insufficient opportunities for active participation and experiential learning. Freeman et al. (2014) demonstrated that active learning approaches can improve academic performance compared with conventional lecturing. Similarly, Prince (2004) emphasized that meaningful student engagement contributes to deeper conceptual understanding. These findings highlight the importance of instructional strategies that encourage learners to construct knowledge actively rather than merely receive information. In Islamic Creed and Ethics education, such strategies are particularly relevant because students must understand abstract religious concepts and apply ethical principles to real-life situations.

The increasing integration of educational technology and interactive pedagogical approaches has encouraged educators to explore game-based learning (GBL) as an alternative instructional strategy. Game-based learning incorporates structured gameplay into educational activities to achieve specific learning objectives through challenges, problem-solving, feedback, and meaningful interaction. Unlike conventional instruction, GBL creates opportunities for students to participate actively while developing knowledge through purposeful learning experiences. Plass et al. (2015) explained that game-based learning integrates cognitive, motivational, affective, and sociocultural dimensions of learning. Likewise, Garris et al. (2002) proposed that instructional games facilitate learning through repeated cycles of engagement, feedback, and reflection. These characteristics suggest that appropriately designed educational games may enhance students' understanding of complex academic concepts.

Empirical research provides substantial evidence supporting the educational potential of game-based learning. Wouters et al. (2013) found that serious games can improve learning and retention, particularly when supported by appropriate instructional guidance. Clark et al. (2016) similarly reported positive effects of digital games on learning outcomes across educational settings. Sailer and Homner (2020) identified beneficial effects of gamification on cognitive learning outcomes, although effectiveness varied according to instructional design. Furthermore,

Hamari et al. (2016) demonstrated that educational games can influence learning through increased engagement and perceived challenge. These findings indicate that learning effectiveness depends not merely on incorporating games but also on aligning game mechanics with educational objectives, student characteristics, and assessment requirements.

From a theoretical perspective, game-based learning is closely associated with constructivist learning principles, experiential learning, and motivational theories. Students develop understanding through active exploration, interaction, and reflection, while immediate feedback enables them to identify misconceptions and improve performance. Deci and Ryan (2000) emphasized that autonomy, competence, and relatedness contribute to intrinsic motivation. These psychological needs can be supported through educational games involving meaningful choices, achievable challenges, and collaborative participation. Additionally, Ryan et al. (2006) demonstrated the motivational significance of psychological need satisfaction in gameplay. Nevertheless, instructional effectiveness requires careful consideration of cognitive demands because excessive game complexity may distract learners from essential educational content (Mayer & Moreno, 2003).

In Islamic Creed and Ethics instruction, game-based learning offers opportunities to transform abstract religious concepts into contextualized educational experiences. Interactive quizzes, collaborative challenges, ethical decision-making simulations, and scenario-based activities may help students distinguish commendable behavior from prohibited conduct while strengthening their conceptual understanding. Research concerning educational games indicates that well-designed gameplay can support knowledge acquisition, motivation, and engagement (Connolly et al., 2012; Boyle et al., 2016). Moreover, learning activities incorporating meaningful feedback and structured challenges may improve students' persistence and academic involvement (Domínguez et al., 2013; Su & Cheng, 2015). However, improvements in cognitive achievement should not automatically be interpreted as evidence of successful moral internalization, which requires additional behavioral assessment.

Preliminary observations at Madrasah Aliyah Negeri 3 Medan described reliance on teacher explanations, lectures, and question-and-answer activities during Islamic Creed and Ethics lessons. These accounts provide a local rationale for examining more participatory instruction, but the source material does not report dated observation logs, baseline participation counts, or specific misconceptions identified before intervention. Accordingly, this context should be interpreted as a preliminary pedagogical concern rather than a quantified diagnosis of low engagement. The study compares two existing eleventh-grade science classes to examine whether educational games, collaborative tasks, and immediate feedback are accompanied by greater improvement in cognitive test scores. Its outcomes concern knowledge and understanding; they do not directly assess religious commitment or moral conduct.

Previous studies identify potential benefits of instructional games across different educational settings (Vogel et al., 2006; Sitzmann, 2011; Girard et al., 2013), but many address digital games, simulations, or other school subjects. Gamification, which introduces game elements

into non-game tasks, is not identical to game-based learning, in which playing structured instructional games is central to the learning process. The present study concerns group challenges, question-based activities, points, and feedback in Islamic Creed and Ethics lessons; further documentation is needed to determine the exact boundaries of this intervention. Its contribution is a contextual comparison of cognitive outcomes between existing madrasah classes rather than proof that the method is generally superior across religious-education contexts. Because a comprehensive review of comparable Indonesian madrasah studies is not documented, claims of novelty are necessarily limited.

Accordingly, the study has three objectives: to describe the instructional activities labeled as GBL, to compare pretest and posttest cognitive scores in the experimental and comparison classes, and to examine the reported statistical evidence for a posttest difference. The dependent variable is performance on an Islamic Creed and Ethics cognitive achievement test, not moral behavior or character internalization. Both groups studied equivalent subject matter over four meetings, but the available record does not establish whether all other conditions, including teaching personnel and classroom circumstances, were identical. The interpretation therefore distinguishes observed score differences from an unqualified causal effect and highlights the need for transparent intervention records and independently verifiable calculations.

METHOD

This study employed a quantitative, nonrandomized quasi-experimental design with pretests and posttests for two existing classes. The experimental group received instruction using structured educational games; the control group received conventional teaching. The outcome was cognitive achievement on a 15-item test in Islamic Creed and Ethics. Because students were not randomly assigned, baseline characteristics and classroom conditions could confound between-group comparisons. The study focuses on comparisons of cognitive scores, not changes in students' moral behavior. Observational accounts are supplementary descriptions of classroom implementation rather than standardized measures of motivation.

The study was conducted at Madrasah Aliyah Negeri 3 Medan during the second semester of the 2025/2026 academic year. Two eleventh-grade science classes were purposively selected: XI Science 3 (30 students, experimental group) and XI Science 4 (30 students, control group). The manuscript does not document a formal class-selection rationale, sample-size calculation, allocation procedure, or comparison of teacher characteristics. It also does not provide precise dates for each intervention meeting. Consequently, the two groups should be described as existing classes rather than randomly equivalent experimental units, and generalization beyond the two classes should be made cautiously.

Both groups completed a cognitive pretest. Over four instructional meetings, the experimental class participated in group challenges, questions, point accumulation, and immediate instructional feedback; the control class received lectures, discussions, and teacher-directed questioning on equivalent materials. A posttest followed the instructional period. The source does not provide complete game rules, time allocation by activity, scoring rubrics, examples of game

questions, or a monitoring checklist for implementation fidelity. Nor does it establish that the same teacher delivered both conditions or that pretest and posttest used identical or parallel test forms. These missing procedural details should be supplied from authentic lesson plans and field notes before strong claims of replicability or controlled implementation are made.

Data came from a 15-item multiple-choice cognitive test, classroom observations, and teaching documents. The manuscript reports item–total correlations of 0.339–0.876 against a stated critical r value of 0.334 and Cronbach's alpha of 0.687. This alpha indicates that measurement consistency warrants caution and does not by itself establish strong test quality. The number and identity of students involved in instrument testing, item blueprint, response scoring, treatment of invalid items, and procedures for converting fifteen item scores onto a 0–100 scale were not provided. These details must be checked against the original instrument and testing output rather than inferred.

Statistical summaries were produced using SPSS version 29 and included pretest and posttest means, standard deviations, Shapiro–Wilk tests, Levene's tests, independent-samples t-tests, normalized gain, and an effect-size estimate. The manuscript reports a posttest t statistic; using its rounded posttest means, standard deviations, and group sizes gives approximately $t = 19.05$, consistent with the reported $t(58) = 19.049$. Nevertheless, statistical significance in a posttest-only comparison does not adjust for all baseline and class-level differences. A stronger analysis would prespecify a between-group gain comparison or ANCOVA adjusting for pretest, subject to appropriate assumptions. N-gain must be defined as the mean of individual gains or the gain computed from group means, since these are not interchangeable. The original individual records and SPSS output were not available for verification; reported N-gain and effect size are therefore retained only as unverified values, not as validated effect estimates.

RESULTS AND DISCUSSIONS

Implementation of Game-Based Learning in Islamic Creed and Ethics Instruction at Madrasah Aliyah Negeri 3 Medan

The intervention involved two existing classes of 30 students each, studying equivalent Islamic Creed and Ethics topics over four meetings. The experimental class used educational challenges, collaborative responses, point-based participation, and immediate feedback, whereas the comparison class used conventional instruction. The text does not contain a full instructional protocol with game rules, precise duration of each activity, or a classroom-fidelity instrument. Thus, the evidence supports a general description of the treatment but not complete replication of its specific mechanics. This limitation also makes it difficult to separate the effects of gameplay from those of increased interaction, additional practice, and immediate feedback.

The GBL class began with lesson objectives and explanations of the activity, followed by group questioning, response challenges, feedback, and discussion. The teacher guided students and clarified misunderstandings. The manuscript does not provide authentic examples of the questions, answer keys, or scoring procedures used during games. Without these records, the alignment

between individual game tasks and the cognitive competencies tested cannot be independently evaluated. A future revision should insert examples from the original teaching materials and an activity schedule if those records exist, rather than reconstructing an intervention retrospectively.

Students in the experimental class were organized into small groups to answer questions, discuss alternatives, and receive immediate feedback. Classroom observations described participation and enthusiasm during these activities. However, no structured observation scale, frequency table, or independently recorded comparison of participation across classes was provided. These descriptions therefore indicate perceived classroom engagement and cannot establish measured improvement in motivation or cognitive competence. Although active participation is pedagogically relevant, the principal study outcome remains the independently administered cognitive test, and enthusiasm during gameplay should not be treated as equivalent to higher academic achievement.

The implementation process corresponds with the theoretical framework developed by Plass et al. (2015), who emphasized the integration of cognitive, motivational, affective, and sociocultural processes within game-based learning environments. Similarly, Garris et al. (2002) explained that educational games facilitate learning through engagement, feedback, and repeated participation. In the present study, these instructional mechanisms were reflected in students' opportunities to answer questions, receive corrective explanations, and reconsider their understanding of Islamic religious concepts. Such activities may strengthen conceptual learning by encouraging students to process information actively rather than depend exclusively on memorization.

The collaborative characteristics of GBL also provide a relevant explanation for the observed classroom interactions. According to Deci and Ryan (2000), intrinsic motivation is influenced by autonomy, competence, and relatedness. Educational games incorporating achievable challenges and group cooperation may support these psychological needs. Ryan et al. (2006) similarly demonstrated that competence and autonomy are important motivational components of gameplay. Within the experimental classroom, students' participation in collaborative challenges provided opportunities to demonstrate understanding, negotiate answers, and receive immediate recognition for successful performance.

Furthermore, Hamari et al. (2016) identified relationships among challenge, engagement, and learning in game-based environments. Domínguez et al. (2013) also demonstrated that game-related instructional elements can influence educational participation, although their effectiveness depends on implementation conditions. These findings support the interpretation that structured educational games may facilitate active classroom involvement. Nevertheless, the distinction between game-based learning and gamification must be recognized because the present intervention incorporated educational gameplay rather than merely adding points or rewards to conventional activities.

In contrast, the control group received instruction primarily through lectures, textbook explanations, note-taking, and teacher-directed questioning. Although these activities supported structured content delivery, they provided fewer opportunities for interactive challenges and

collaborative problem-solving. Freeman et al. (2014) demonstrated the academic benefits of active learning, while Prince (2004) emphasized the importance of meaningful student participation. These studies provide theoretical support for interpreting differences in classroom engagement between the two instructional conditions.

Overall, the implementation findings indicate that GBL was integrated into Islamic Creed and Ethics instruction through organized educational activities, collaborative participation, instructional feedback, and reflective learning. The classroom observations support its feasibility as an interactive instructional approach. However, determining its effectiveness in improving cognitive achievement requires examination of the pretest and posttest results, which constitute the principal quantitative evidence of this investigation.

Student Learning Outcomes in Islamic Creed and Ethics Following the Implementation of Game-Based Learning

Cognitive achievement was assessed through fifteen multiple-choice items intended to cover Islamic Creed and Ethics knowledge. The manuscript reports item–total correlations of 0.339–0.876 and Cronbach's alpha of 0.687; the item-testing sample, scoring conversion, content blueprint, and comparability of pretest and posttest instruments were not documented. These gaps limit independent appraisal of validity and reliability. The reported statistics may be presented as provisional properties of the existing instrument, but a transparent item blueprint and anonymized assessment records are needed before the measurement can be considered fully auditable.

According to the manuscript's descriptive summaries, the experimental class (n = 30) increased from M = 61.47 (SD = 3.093) to M = 84.50 (SD = 3.037); the control class (n = 30) increased from M = 60.47 (SD = 3.093) to M = 70.20 (SD = 2.772). Both groups improved, with a larger change in the GBL class. The summary values alone do not establish whether any individual scores were missing, how fifteen responses were converted to a 100-point scale, or whether all scores correspond to the same tested cohort. The original records must be reconciled with these summaries.

Table 1. Descriptive Statistics of Student Learning Outcomes

Group	Assessment	Mean	SD	Minimum	Maximum
Experimental	Pretest	61.47	3.093	56	67
Experimental	Posttest	84.50	3.037	79	90
Control	Pretest	60.47	3.093	55	66
Control	Posttest	70.20	2.772	65	75

Source: Summary values reproduced from the research manuscript. Group-level statistics and the score-conversion procedure require verification against anonymized student records before publication.

The experimental group's mean increased by 23.03 points, compared with 9.73 points in the control group, a descriptive difference in gains of 13.30 points. Its posttest mean also exceeded the control group's by 14.30 points. These are arithmetic differences between the reported group

means, not yet pretest-adjusted estimates of the intervention's causal effect. Any inferential comparison of gain scores requires the standard deviation of individual changes, while ANCOVA would require the individual pretest and posttest data. Consequently, confidence intervals and gain-score statistical tests cannot be reconstructed reliably from the published summaries alone.

The findings are consistent with previous research identifying positive relationships between educational games and academic achievement. Wouters et al. (2013) reported that serious games can improve learning outcomes and knowledge retention, particularly when integrated with appropriate instructional support. Clark et al. (2016) similarly demonstrated that digital game-based learning can produce favorable educational outcomes across different learning environments. Furthermore, Vogel et al. (2006) identified positive effects of computer games and interactive simulations on cognitive learning. Although these investigations addressed broader educational contexts, their findings provide relevant comparative evidence supporting the potential effectiveness of interactive instructional approaches.

The observed improvement may also be interpreted through the relationship between active participation and cognitive processing. During GBL activities, students were required to retrieve information, discuss alternative responses, evaluate questions, and apply previously acquired knowledge. These activities potentially strengthened conceptual understanding through repeated exposure and meaningful interaction. Connolly et al. (2012) identified knowledge acquisition and engagement among the frequently reported outcomes of educational games. Similarly, Boyle et al. (2016) emphasized the importance of instructional design in determining whether game-based activities produce meaningful learning benefits.

The experimental group's performance also aligns with findings from studies conducted in specific educational disciplines. Papastergiou (2009) demonstrated that digital game-based learning improved educational effectiveness and motivation in secondary-level computer science education. Kebritchi et al. (2010) identified favorable effects of mathematics computer games on academic achievement. Although these studies investigated different subjects, they suggest that interactive educational environments may support knowledge acquisition when gameplay is appropriately aligned with instructional objectives.

The achievement results should nevertheless be interpreted within the boundaries of the research design. The assessments primarily measured cognitive understanding rather than behavioral transformation or sustained moral development. Therefore, higher examination scores cannot independently establish that students consistently demonstrate honesty, responsibility, humility, or other Islamic moral characteristics outside the classroom.

Overall, the descriptive findings indicate that students receiving GBL instruction demonstrated greater improvements in Islamic Creed and Ethics achievement than students receiving conventional instruction. The results support the potential educational value of interactive learning activities while emphasizing the importance of rigorous statistical testing, reliable measurement, and careful interpretation of cognitive achievement.

The Effect of Game-Based Learning on Improving Student Learning Outcomes in Islamic Creed and Ethics

The article reports distributional tests, variance-homogeneity tests, independent-samples t-tests, normalized gains, and an effect-size estimate for the experimental and control classes. The reported posttest independent-samples t-test is arithmetically plausible given group means, standard deviations, and n = 30 per group; this is separate from the unresolved N-gain and effect-size reporting. The distinction is important: a correctly computed posttest difference does not by itself establish that gameplay caused the difference, nor does it validate every additional statistic reported in the manuscript.

The manuscript reports Shapiro–Wilk p values of 0.679 and 0.549, a pretest Levene p value of 0.472, and a pretest comparison of $t(58) = 1.252$, $p = .216$. The pretest t statistic is consistent with the displayed means (61.47 versus 60.47) and SDs (both 3.093), up to rounding. A nonsignificant pretest comparison does not demonstrate that the two purposively selected classes were equivalent in all respects. Other differences in prior knowledge, learning conditions, or classroom composition could remain. The analysis therefore requires a careful statement of its assumptions and limits.

The reported posttest means are 84.50 and 70.20 (difference = 14.30), with standard deviations of 3.037 and 2.772 and 30 students per group. An independent-samples t statistic reconstructed from these rounded values is approximately 19.05, close to the manuscript's reported $t(58) = 19.049$, $p < .001$. This consistency supports the arithmetic plausibility of that particular test but does not substitute for verifying the underlying scores or the homogeneity assumption. Moreover, comparing only posttest means is not identical to estimating a treatment effect that accounts for the baseline measurements.

Table 2. Summary of Reported Inferential Statistical Results

Statistical analysis	Reported result	Interpretation
Pretest independent-samples t-test	$t(58) = 1.252$, $p = .216$	Not significant
Pretest Levene's test	$p = .472$	Equal variances not rejected
Posttest Levene's test	$p = .457$	Equal variances not rejected
Posttest independent-samples t-test	$t(58) = 19.049$, $p < .001$	Significant difference
Experimental N-gain	70.40%	Reported value; method unverified
Control N-gain	24.70%	Reported value; method unverified
Unverified reported effect size	1.09	Metric and computation not documented

Source: Statistical values as reported in the submitted article.

Table 2 reports N-gain values of 70.40% for the experimental group and 24.70% for the control group. Using the published group means and a maximum score of 100, the gain calculated from group means is approximately 59.8% and 24.6%, respectively. This discrepancy does not alone

prove the experimental N-gain was miscalculated: averaging student-level normalized gains can produce a different value. However, the manuscript does not specify which approach was used or provide individual values, so the reported 70.40% cannot be verified. The reported effect size of 1.09 similarly lacks a named statistic, formula, and supporting output. It should not be interpreted as a validated Cohen's d without recalculation. Authors should supply the original dataset and analysis output before these quantities are used to support a claim of effect magnitude.

The statistical findings are broadly consistent with the educational effectiveness literature. Sitzmann (2011) demonstrated that computer-based simulation games can support learning when appropriately integrated into instructional environments. Girard et al. (2013) identified educational benefits associated with serious games, while Sailer and Homner (2020) reported positive effects of gamification on cognitive learning outcomes. Together, these studies suggest that interactive instructional elements can contribute to academic achievement, although effectiveness depends on learning objectives, instructional quality, and implementation conditions.

The findings may also be interpreted through motivational and cognitive learning mechanisms. Su and Cheng (2015) demonstrated that mobile gamification can improve learning motivation and achievement. Mayer and Moreno (2003) emphasized that instructional design must manage cognitive demands to facilitate meaningful learning. Accordingly, GBL activities may be most effective when educational challenges reinforce relevant concepts without introducing unnecessary distractions. In the present study, collaborative questioning, instructional feedback, and repeated practice potentially contributed to students' stronger posttest performance.

Nevertheless, the quasi-experimental design introduces limitations concerning causal interpretation. Because participants were selected from existing classes rather than randomly assigned individually, unmeasured differences in classroom characteristics, instructional interactions, and learning conditions may have influenced the outcomes. Furthermore, the study examined achievement immediately following a relatively brief instructional intervention, limiting conclusions about long-term retention and sustained behavioral development.

The displayed posttest summary statistics indicate a substantial between-class score difference, and the reported t test is approximately consistent with those summaries. However, interpretations of practical impact should not rely on the unverified N-gain and effect-size entries. Because students came from intact, purposively selected classes, unmeasured class differences and instructional conditions remain alternative explanations for some of the observed gap. The defensible conclusion is that this cohort receiving GBL showed greater descriptive achievement gains than its comparison class. Claims about exact effect magnitude, general effectiveness, and long-term outcomes require a reconciled dataset, explicit formulas, and stronger baseline-adjusted analysis.

CONCLUSIONS

The study compared cognitive learning outcomes in Islamic Creed and Ethics between two existing eleventh-grade classes, each comprising 30 students. The reported experimental mean rose from 61.47 to 84.50, while the control mean increased from 60.47 to 70.20, a descriptive difference in gains of 13.30 points. A posttest independent-samples t statistic of approximately

19.05 can be reconstructed from the reported means and standard deviations, consistent with the manuscript's $t(58) = 19.049$, $p < .001$. However, the reported experimental N-gain and effect-size value still require verification from individual student scores and clearly stated calculation procedures. The groups were purposively selected rather than randomly assigned, and conditions beyond the instructional method were not fully documented. The evidence therefore supports a larger observed cognitive score increase in the GBL class, not an unrestricted causal claim or proof of improved moral behavior. Future studies should provide complete assessment records, treatment-fidelity documentation, baseline-adjusted analyses, and measures of longer-term knowledge retention and moral practice.

REFERENCES

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. Computers & Education, 94, 178–192. https://doi.org/10.1016/j.compedu.2015.11.003

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. Review of Educational Research, 86(1), 79–122. https://doi.org/10.3102/0034654315582065

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A systematic literature review of empirical evidence on computer games and serious games. Computers & Education, 59(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. J. (2013). Gamifying learning experiences: Practical implications and outcomes. Computers & Education, 63, 380–392. https://doi.org/10.1016/j.compedu.2012.12.020

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. Simulation & Gaming, 33(4), 441–467. https://doi.org/10.1177/1046878102238607

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. Journal of Computer Assisted Learning, 29(3), 207–219. https://doi.org/10.1111/j.1365-2729.2012.00489.x

Hamari, J., Shernoff, D. J., Rowe, E., Coller, B., Asbell-Clarke, J., & Edwards, T. (2016). Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. Computers in Human Behavior, 54, 170–179. https://doi.org/10.1016/j.chb.2015.07.045

Kebritchi, M., Hirumi, A., & Bai, H. (2010). The effects of modern mathematics computer games on mathematics achievement and class motivation. Computers & Education, 55(2), 427–443. https://doi.org/10.1016/j.compedu.2010.02.007

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. Educational Psychologist, 38(1), 43–52. https://doi.org/10.1207/S15326985EP3801_6

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. Computers & Education, 52(1), 1–12. https://doi.org/10.1016/j.compedu.2008.06.004

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. Educational Psychologist, 50(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Prince, M. (2004). Does active learning work? A review of the research. Journal of Engineering Education, 93(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. Motivation and Emotion, 30(4), 344–360. https://doi.org/10.1007/s11031-006-9051-8

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. Cognitive Research: Principles and Implications, 5, Article 6. https://doi.org/10.1186/s41235-019-0202-z

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. Personnel Psychology, 64(2), 489–528. https://doi.org/10.1111/j.1744-6570.2011.01190.x

Su, C.-H., & Cheng, C.-H. (2015). A mobile gamification learning system for improving the learning motivation and achievements. Journal of Computer Assisted Learning, 31(3), 268–286. https://doi.org/10.1111/jcal.12088

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. Journal of Educational Computing Research, 34(3), 229–243. https://doi.org/10.2190/FLHV-K4WA-WPVQ-H0YM

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. Journal of Educational Psychology, 105(2), 249–265. https://doi.org/10.1037/a0031311