E-ISSN: 2476-9703
Terbit sejak 2015

MUALLIMUNA: JURNAL MADRASAH IBTIDAIYAH
Alamat web jurnal: http://ojs.uniska-bjm.ac.id/index.php/jurnalmuallimuna
DOI: 10.31602/muallimuna.v12i1.24959

Vol. 12, No. 1, Oktober 2026
Halaman: 40-54

ROLE PLAY EFFECTS ON FOURTH-GRADE STUDENTS' NATURAL AND SOCIAL SCIENCES LEARNING OUTCOMES

Muhammad Fauzi[1], Siti Ramdiah[2*]
Universitas PGRI Kalimantan, Indonesia
[1]fauzimuhammad617@gmail.com, [2]sitiramdiah@upk.ac.id

Article history

Received:	*Revised:*	*Accepted:*	*Published:*
01 Agustus 2026	02 Oktober 2026	19 Oktober 2026	20 Oktober 2026

This is an open-access article under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0). Copyright © 2026 by authors.

Abstract: This study examined whether fourth-grade students receiving Role Play instruction demonstrated different Natural and Social Sciences learning outcomes from students receiving conventional instruction, while also describing classroom implementation. A quasi-experimental nonequivalent control-group pretest-posttest design involved 43 students, with 21 in the experimental group and 22 in the control group. Learning outcomes were measured using multiple-choice and essay tests, supported by observation sheets and a student-response questionnaire. One-Way ANOVA was the primary posttest comparison, and ANCOVA was used as a sensitivity analysis with pretest scores as a covariate. The experimental group obtained a higher posttest mean than the control group (89.00 vs. 80.86), $F(1, 41) = 32.17$, $p < 0.001$, eta squared = 0.440. ANCOVA confirmed the difference after baseline adjustment, $F(1, 40) = 31.08$, $p < 0.001$, partial eta squared = 0.437. Role Play was therefore associated with higher posttest achievement in this classroom context.

Keywords: Role Play; Learning Outcomes; Natural and Social Sciences

Abstrak: Penelitian ini menguji apakah siswa kelas IV yang memperoleh pembelajaran Role Play menunjukkan hasil belajar Ilmu Pengetahuan Alam dan Sosial yang berbeda dibandingkan siswa yang memperoleh pembelajaran konvensional, sekaligus mendeskripsikan keterlaksanaan pembelajarannya. Penelitian menggunakan desain kuasi-eksperimen nonequivalent control group pretest-posttest dengan 43 siswa, terdiri atas 21 siswa kelompok eksperimen dan 22 siswa kelompok kontrol. Hasil belajar diukur melalui tes pilihan ganda dan esai, sedangkan proses pembelajaran didukung data observasi dan angket respons siswa. One-Way ANOVA digunakan sebagai analisis utama untuk membandingkan posttest, sedangkan ANCOVA digunakan sebagai analisis sensitivitas dengan pretest sebagai kovariat. Rata-rata posttest kelompok eksperimen lebih tinggi daripada kontrol (89,00 berbanding 80,86), $F(1, 41) = 32,17$, $p < 0,001$, eta squared = 0,440. ANCOVA mengonfirmasi perbedaan setelah penyesuaian nilai awal, $F(1, 40) = 31,08$, $p < 0,001$, partial eta squared = 0,437. Dengan demikian, Role Play berkaitan dengan hasil posttest yang lebih tinggi dalam konteks pembelajaran yang diteliti.

Kata Kunci: Role Play; Hasil Belajar; Pembelajaran IPAS

INTRODUCTION

Primary education is expected to develop not only students' factual knowledge but also their capacity to use knowledge in situations that are meaningful to everyday life. In Natural and Social Sciences, learning outcomes therefore depend on more than exposure to definitions or teacher explanations; students must be able to connect concepts with observable events, social interaction, and decisions that they recognize from their own environment. When instruction remains dominated by explanation, note taking, and short question-and-answer exchanges, students may reproduce information without having sufficient opportunities to test how a concept operates in practice. This problem is especially relevant to topics that combine social roles, economic interaction, and decisions about needs, resources, and transactions because the meaning of the concepts is embedded in actions and relationships rather than in isolated terms alone.

At SDN 2 Barabai Darat, the instructional problem identified among fourth-grade students was not simply that the teacher used conventional methods, but that the learning process offered limited opportunities for students to enact the economic situations represented in the topic "How Do We Get All Our Needs?" Students were expected to understand needs and wants, limited resources, buying and selling, and the roles of people in economic activities, yet these ideas were largely encountered through verbal explanation. Their learning outcomes had not reached the expected level, and classroom participation remained primarily responsive to teacher direction. This created a mismatch between the nature of the content and the learning experience: the topic concerns interaction, choice, exchange, and role relationships, while the dominant classroom activity required students mainly to listen and answer questions. A method that could convert those abstract relationships into structured classroom experience was therefore pedagogically relevant.

Role Play was selected because its instructional structure corresponds directly with the representational demands of this IPAS topic. Through prepared scenarios, students can take the roles of buyers, sellers, producers, consumers, or other actors involved in economic activity. They must interpret the situation, communicate with other participants, make choices, follow the logic of their assigned roles, and respond to consequences within the scenario. In this way, concepts such as needs, priorities, exchange, and economic interdependence are not presented only as information but are represented through observable action. The method also changes the classroom position of students from recipients of explanation to participants who must coordinate their understanding with peers. This alignment between content and activity is the principal reason Role Play was considered more appropriate than merely increasing the amount of conventional explanation.

This emphasis on enactment is supported by broader evidence on experiential and participatory learning in primary education. Thomas and Rajdeep (2021) reported that activity-based modules combining games, role plays, drama, and reflection provided children with first-hand encounters with everyday-life situations and were followed by improvements in communication, participation, perceptions, and values. From a developmental perspective, Melik-Pashaev and Novlyanskaya (2022) likewise described plot-role-playing and dramatization as important forms of children's play through which creative potential and active meaning making can develop. These findings strengthen the rationale for using Role Play when the learning content is inherently relational, because

the instructional activity allows students to experience a structured situation rather than merely imagine it.

Previous studies provide evidence that Role Playing can support elementary learning outcomes, but the evidence is distributed across different subjects, outcomes, and research contexts. Yusnarti and Suryaningsih (2021) reported a positive relationship between Role Playing and elementary school learning outcomes, while Rauf (2024) also identified improved achievement following implementation of the model. More directly related to the present content, Darmawan et al. (2024) used Role Playing in fourth-grade IPAS on the topic of fulfilling needs and reported progressive improvement in learning completeness across classroom-action cycles. These findings indicate that role enactment can be productive when the subject matter can be translated into situations that students can perform. At the same time, the designs and aims of these studies differ, so their findings do not by themselves establish how two intact classes compare after different instructional approaches are applied to the same topic.

International studies further show that role-based learning can support achievement when the roles are tied closely to the target knowledge. Alghamdy (2022), in a study of elementary EFL learners, found statistically significant advantages for students learning through Role Play over those receiving traditional instruction, particularly in several communication-related skills. Kajiwara et al. (2023) showed that a machine-learning role-playing game could help learners across school-age levels understand selected processes by physically and conceptually enacting the stages of machine learning. Similarly, Tsuei and Lu (2025) reported significantly better insurance-related learning achievement among elementary students using a digital role-playing game than among students receiving traditional instruction. Although these studies addressed different subject domains, together they indicate that the educational value of Role Play may become stronger when abstract relationships are converted into explicit roles, decisions, and sequences of action.

The potential value of Role Play also extends beyond test performance. Role-playing activities require students to understand an assigned function, communicate with peers, coordinate actions, and respond to the perspectives of other participants. Saesari et al. (2023) linked role-playing activities with communication skills in elementary Natural and Social Sciences learning, and Nindialisma et al. (2025) reported improvement in cooperation and communication after Role Playing was implemented in IPAS. This emphasis on communication and collaboration is consistent with the broader orientation toward active classroom practices described by Realitawati et al. (2024) in elementary education. These process characteristics are relevant because learning through a scenario depends on students' participation in the social structure of the task. Nevertheless, communication and cooperation should not automatically be interpreted as mechanisms that caused higher achievement unless those relationships are directly tested; they are better treated as plausible learning processes that accompany the intervention.

The literature also suggests that affective and attentional processes can accompany this form of active participation, although gains in motivation do not automatically translate into higher test scores. Chen et al. (2024) found that role-play gamification in elementary classrooms increased the value and emotional dimensions of learning motivation, while Chinese and mathematics scores did not differ between groups. By contrast, Yu and Tsuei (2023) reported that digital game-based learning improved fourth-grade Chinese-language achievement and that different game mechanisms also influenced attention and self-efficacy. More recently, Shang et al. (2026) found that

gamified project-based learning, which included role-playing and other interactive elements, improved both mathematics achievement and learning motivation relative to traditional project-based learning. Taken together, these findings suggest that active participation may support several learning-related processes, but the relationship between engagement, motivation, and achievement is conditional on the instructional design and learning task.

Evidence from other subjects also suggests that Role Playing is adaptable when content can be embodied through roles or interaction. Susanto et al. (2024) reported differences in Indonesian-language learning outcomes for drama material, Nasution et al. (2024) applied Role Playing to food-chain content in elementary science, and Sihaloho et al. (2025) connected the method with critical-thinking development. The variety of these applications supports the idea that the instructional value of Role Play is not restricted to one subject. However, it also shows why a specific knowledge gap remains. Studies that demonstrate benefits in drama, food chains, communication, cooperation, or critical thinking do not necessarily show how Role Play functions in an IPAS economic topic where students must enact transactions and social roles, nor do they necessarily interpret outcome differences together with evidence about implementation fidelity and student response.

Recent technology-supported studies reinforce the importance of structured guidance within role-based or game-based activities. Hsieh and Hwang (2025) found that embedding concept maps into a digital role-playing environment improved elementary students' science and English achievement and encouraged more reflective and strategic learning patterns. Lin et al. (2023) similarly reported that an augmented-reality-supported contextualized dilemma approach with immersive role-playing scenarios improved primary students' achievement in AI ethical awareness, reasoning, and higher-order thinking. Koong et al. (2026) further showed that incorporating dynamic assessment into a role-playing game benefited English grammar achievement, particularly among students with lower prior achievement. Even outside Role Play specifically, Charamba (2022) demonstrated in fourth-grade Science and Technology that an instructional approach deliberately aligned with learners' classroom needs could produce significantly different posttest achievement from a traditional approach. Collectively, this evidence indicates that the effectiveness of active learning depends not only on participation itself but also on the alignment among content, guidance, learner characteristics, and the structure of the learning task.

Accordingly, the research gap in the present study is not simply the absence of a Role Play study at one particular school. A new school location alone would constitute contextual replication rather than a strong scholarly contribution. The more relevant gap concerns the integration of three elements within one classroom comparison: first, an IPAS topic whose concepts can be represented through economic roles and transactions; second, a comparison of post-instruction learning outcomes between Role Play and conventional instruction in two intact fourth-grade classes; and third, process evidence showing whether the planned Role Play stages were implemented and how students participated and responded. Examining these elements together makes it possible to interpret the outcome difference in relation to the instructional process without claiming that the process indicators themselves prove causation.

The contribution of this study is therefore framed analytically rather than geographically. It examines Role Play as a content-aligned instructional structure for elementary IPAS in which conceptual knowledge is enacted through social and economic

interaction, and it combines outcome measurement with classroom-process indicators. This approach adds to existing evidence by showing whether a between-group posttest difference remains when baseline scores are also considered in a sensitivity analysis, while documenting how completely the planned stages were carried out. The contribution is bounded by the study design: the participants came from two intact classes in one school, the intervention addressed one IPAS topic, and the findings concern immediate learning outcomes rather than long-term retention or transfer. These boundaries are important for interpreting the results without extending them beyond the conditions that were actually studied.

Based on this rationale, the study addressed two questions. First, did fourth-grade students who received Role Play instruction demonstrate different posttest learning outcomes from students who received conventional instruction? Second, how was the Role Play learning process implemented in terms of teacher activity, student activity, and student responses? The study therefore aimed to compare learning outcomes between the two instructional groups and to describe the implementation of Role Play during the topic "How Do We Get All Our Needs?" The statistical interpretation focused on observed between-group differences, while the classroom-process data were used to characterize how the intervention was delivered and experienced.

METHODS

This study used a quantitative quasi-experimental nonequivalent control-group pretest-posttest design. The design was appropriate because the research was conducted in existing school classes rather than by randomly assigning individual students to newly created groups. Both groups completed a pretest before instruction and a posttest after the instructional period. The experimental group received instruction using the Role Play method, whereas the control group received conventional instruction based on teacher explanation and teacher-directed question-and-answer activities. The pretest was used to describe baseline learning outcomes before either instructional condition was implemented. The principal inferential comparison concerned posttest performance, while an additional baseline-adjusted analysis was used to examine whether the substantive conclusion remained when pretest scores were statistically controlled.

The study was conducted at SDN 2 Barabai Darat during the second semester of the 2025/2026 academic year. The population comprised 65 fourth-grade students distributed across Classes IV A, IV B, and IV C. The study sample consisted of 43 students selected through purposive sampling from two intact classes. Class IV C, consisting of 21 students, served as the experimental group, and Class IV A, consisting of 22 students, served as the control group. Because the classes existed before the study, group membership was not determined by individual randomization. This feature is consistent with the nonequivalent control-group design and is taken into account in the interpretation of the findings. The design notation is presented in Tabel 1.

Table 1. Research Design

Group	Pretest	Treatment	Posttest
Experimental	O_1	Role Play	O_2
Control	O_3	Conventional instruction	O_4

The independent variable was the Role Play method and the dependent variable was students' learning outcomes in Natural and Social Sciences. Both groups studied the topic "How Do We Get All Our Needs?" The experimental-group scenarios represented situations involving needs and wants, the use of limited resources, buying and selling, and

interaction among people involved in economic activity. The Role Play procedure consisted of explaining learning objectives, preparing and introducing scenarios, distributing roles, clarifying activity rules, conducting the simulation, discussing what occurred, and completing a reflection that connected the scenario with the intended concepts. Students performed the assigned economic roles, while the teacher organized the sequence, facilitated interaction, and guided conceptual consolidation after the performance.

The control group studied the same general learning content through conventional instruction without the Role Play method. Teacher explanation and directed question-and-answer activities were used to present and review the concepts. The distinction between the groups therefore concerned the instructional approach rather than exposure to different subject matter. The Role Play condition was observed across four meetings. Observation focused on whether the planned stages were implemented and whether students participated in role preparation, simulation, cooperation, discussion, and reflection. These observations were treated as implementation indicators rather than as independent evidence of treatment effectiveness.

The research instruments comprised a learning-outcome test, teacher-activity observation sheets, student-activity observation sheets, a student-response questionnaire, and documentation. The administered learning-outcome test contained eight multiple-choice items and seven essay items and was used for both pretest and posttest measurement. Each correct multiple-choice response received a score of 1, while an incorrect or unanswered response received 0. Essay responses were scored from 0 to 4 according to accuracy and completeness. The multiple-choice and essay components were converted to percentage scores and combined with equal weighting to obtain the final learning-outcome score. The raw scoring records therefore preserve both item-level performance and the combined score used in the descriptive and inferential analyses.

Before use, the learning-outcome instruments had undergone validity and reliability testing. The final administered test structure, item-level scoring, and equal weighting of the multiple-choice and essay components were applied consistently at pretest and posttest. Teacher-activity observation sheets assessed the implementation of the Role Play stages, including preparation, role distribution, simulation, discussion, and reflection. Student-activity sheets assessed participation during role preparation, performance, cooperation, discussion, and reflection. The student-response questionnaire contained ten Yes/No statements addressing students' interest, comfort, and perceived ease of understanding the material. These process instruments were used descriptively to characterize implementation and acceptance rather than to serve as additional learning-outcome measures.

Data were summarized using the mean, median, standard deviation, minimum, and maximum. Distributional assumptions were examined using the Shapiro-Wilk test for each pretest and posttest dataset. Homogeneity of posttest variance between groups was examined using Levene's test, with values above .05 interpreted as not providing evidence against the homogeneity assumption. A baseline comparison of pretest scores was also examined to describe whether the two intact classes displayed a statistically detectable difference before the instructional period. This baseline comparison was descriptive in purpose and was not treated as evidence that the classes were identical.

One-Way Analysis of Variance was used as the primary hypothesis test for the posttest comparison because the principal research question concerned whether final learning outcomes differed between the Role Play and conventional-instruction groups.

The analysis was evaluated at an alpha level of .05 and was reported with the F statistic, degrees of freedom, exact probability value, and eta squared as an effect-size estimate. Because the design also contained pretest information and the classes were nonequivalent intact groups, an Analysis of Covariance was performed as a sensitivity analysis, with posttest score as the dependent variable, instructional group as the fixed factor, and pretest score as the covariate. This additional analysis did not replace the primary ANOVA; it evaluated whether the observed posttest difference remained after statistical adjustment for baseline performance.

For the ANCOVA sensitivity analysis, the homogeneity-of-regression-slopes assumption was examined by testing the interaction between instructional group and pretest score. The interaction was not statistically significant, $p = 0.169$, supporting use of a common pretest slope across groups. Model residuals did not show evidence of departure from normality in the Shapiro-Wilk assessment, $p = 0.500$. Adjusted group means and the 95% confidence interval for the adjusted mean difference were calculated to aid interpretation. Effect size for the group term was expressed as partial eta squared. The use of both the primary posttest ANOVA and the baseline-adjusted sensitivity analysis allowed the robustness of the between-group result to baseline variation to be evaluated directly.

RESULTS AND DISCUSSION

Result

The pretest was administered before the instructional period to describe the students' initial learning outcomes. As shown in Table 2, the control group had a pretest mean of 43.95 (SD = 15.395), while the experimental group had a mean of 46.19 (SD = 14.204). The mean difference was 2.24 points, both groups had the same minimum score of 20, and their score dispersion was similar. A baseline One-Way ANOVA did not identify a statistically significant difference between the two pretest means, $F(1, 41) = 0.24$, $p = 0.624$, eta squared = 0.006. This result indicates that no statistically detectable baseline difference was observed in the measured learning outcome; however, it is not interpreted as proof that the intact classes were equivalent on every relevant characteristic.

Table 2. Pretest Descriptive Statistics

Group	N	Mean	Median	Standard Deviation	Minimum	Maximum
Control	22	43.95	41.00	15.395	20	71
Experimental	21	46.19	46.00	14.204	20	67

After instruction, the pattern of scores differed more clearly. Table 3 shows that the experimental group obtained a posttest mean of 89.00 (SD = 4.593), compared with 80.86 (SD = 4.804) in the control group, producing an unadjusted mean difference of 8.14 points. The experimental group also had a higher minimum score (80 vs. 72) and a higher maximum score (96 vs. 90). These descriptive results indicate that the distribution of final scores for the Role Play group was shifted upward relative to the conventional-instruction group. Because the outcome of interest was the post-instruction difference rather than the amount of within-group change alone, the inferential analysis focused primarily on the posttest comparison and then examined whether that comparison remained after baseline adjustment.

Table 3. Posttest Descriptive Statistics

Group	N	Mean	Median	Standard Deviation	Minimum	Maximum
Control	22	80.86	81.00	4.804	72	90
Experimental	21	89.00	91.00	4.593	80	96

Assumption testing and primary hypothesis analysis

The Shapiro-Wilk results are presented in Table 4. All four pretest and posttest datasets had significance values above 0.05: control pretest p = 0.319, control posttest p = 0.912, experimental pretest p = 0.328, and experimental posttest p = 0.146. Levene's test for the posttest comparison produced p = 0.970, so the data did not provide evidence against the equal-variance assumption. These diagnostic results supported the planned parametric comparison. The primary One-Way ANOVA showed a statistically significant between-group difference in posttest learning outcomes, F(1, 41) = 32.17, exact p = 1.27 x 10^-6, eta squared = 0.440. Thus, the students who received Role Play instruction demonstrated higher posttest scores than students who received conventional instruction under the conditions of this study.

Table 4. Shapiro–Wilk Normality Test

Dataset	N	Significance	Decision
Control pretest	22	0.319	Normal
Control posttest	22	0.912	Normal
Experimental pretest	21	0.328	Normal
Experimental posttest	21	0.146	Normal

Baseline-adjusted sensitivity analysis

The ANCOVA sensitivity analysis yielded the same substantive interpretation. After pretest scores were included as a covariate, the instructional-group term remained statistically significant, F(1, 40) = 31.08, exact p = 1.87 x 10^-6, partial eta squared = 0.437. The adjusted posttest mean was 88.95 for the Role Play group and 80.92 for the control group. The adjusted mean difference was 8.03 points, with a 95% confidence interval from 5.12 to 10.94. The pretest covariate itself was not statistically significant within this model, F(1, 40) = 0.92, p = 0.344, but the covariate was retained because the purpose of the sensitivity analysis was to determine whether the group comparison was robust to baseline adjustment rather than to test pretest as an intervention outcome. Table 5 summarizes the primary and sensitivity analyses. The adjusted difference was only slightly smaller than the unadjusted difference of 8.14 points, which shows that baseline adjustment changed the estimated group contrast very little. The 95% confidence interval remained entirely above zero, providing an interval estimate that is consistent with the hypothesis-test result. Taken together, the primary and sensitivity analyses indicate that the observed separation between groups was not dependent on whether the pretest score was omitted from or included in the statistical model.

Table 5. Primary and Baseline-Adjusted Inferential Analyses

Analysis	**Statistic**	**df**	**Exact p**	**Effect size**
Baseline pretest comparison	F = 0.24	1, 41	0.624	$\eta^2 = 0.006$
Primary posttest ANOVA	F = 32.17	1, 41	1.27 × 10^-6	$\eta^2 = 0.440$
ANCOVA group effect	F = 31.08	1, 40	1.87 × 10^-6	partial η^2 = 0.437

Implementation and student responses

Classroom-process data were collected across four Role Play meetings. All planned teacher-implementation indicators were recorded as completed in each meeting, including scenario preparation, role distribution, simulation, discussion, and reflection,

resulting in a teacher-activity score of 100% on the observation checklist. Student-activity indicators also reached 100% across the observed meetings: students participated in role preparation, performed the scenarios, communicated with group members, took part in discussion, and completed reflection. The ten-item response questionnaire was completed by all 21 students in the experimental group, and every student selected "Yes" for every statement, producing 100% positive responses. These percentages indicate complete endorsement of the checklist indicators used in this study. Because the observation and questionnaire measures produced ceiling-level values, they are interpreted as descriptive evidence of implementation and acceptance rather than as sensitive measures of differences in the quality or intensity of participation.

Discussion

The principal finding is that students in the Role Play class demonstrated higher posttest learning outcomes than those in the conventional-instruction class. The primary ANOVA identified an unadjusted difference of 8.14 points. After pretest scores were controlled, ANCOVA produced an adjusted difference of approximately eight points and the group term remained statistically significant. The close agreement between the analyses indicates that the observed posttest separation was not driven by the small baseline mean difference. However, the design remains quasi-experimental: the two conditions were represented by two pre-existing classes rather than individually randomized students. The statistical association therefore supports a robust between-group difference under the study conditions but does not establish that every part of that difference was caused exclusively by Role Play.

The effect-size estimates complement the probability tests by describing the magnitude of separation between the two groups within the observed data. The primary ANOVA produced eta squared = 0.440, and the baseline-adjusted model produced partial eta squared = 0.437 for the instructional-group term. These values indicate that group membership was strongly associated with variation in posttest performance in this sample, and the close similarity between the unadjusted and adjusted estimates is consistent with the stability of the result after baseline control. The effect size should not be interpreted as the percentage of learning caused by Role Play, because eta-squared measures describe variance within a statistical model rather than a causal proportion attributable to an intervention. Reporting the effect size alongside the adjusted means and confidence interval therefore provides a more informative description of the observed difference than relying on statistical significance alone.

The direction of the finding is consistent with earlier evidence that Role Playing can support elementary learning outcomes. Yusnarti and Suryaningsih (2021) reported higher learning outcomes associated with the method, and Rauf (2024) similarly described improved achievement following its implementation. The present study extends that general pattern by examining the method in an IPAS topic centered on needs and economic activity and by comparing two instructional groups rather than interpreting one group in isolation. The result is also closely aligned with Darmawan et al. (2024), who applied Role Playing to the fourth-grade IPAS topic of fulfilling needs and reported progressive improvement in classroom mastery. Whereas that study used a classroom-action approach to document improvement across cycles, the present study contributes a between-group comparison and a baseline-adjusted sensitivity analysis. The studies therefore address related but not identical questions, and their convergence is more informative than treating one design as a replication of the other.

The present result also converges with evidence from international elementary-school contexts. Alghamdy (2022) observed stronger achievement and several communication outcomes under Role Play than under traditional instruction, while Tsuei and Lu (2025) reported higher achievement in an elementary insurance-education intervention using a digital role-playing game. These parallels are noteworthy because the substantive domains differ, yet each intervention required students to act within a defined situation rather than receive information only through conventional explanation. The consistency across domains therefore supports a cautious interpretation that structured enactment can be academically productive when the enacted roles correspond meaningfully with the concepts being learned.

The content of the learning topic offers a plausible explanation for why Role Play was associated with higher final scores. Concepts such as needs, wants, limited resources, buying and selling, producers, and consumers describe relationships among people and choices made in specific situations. When these ideas are presented only through teacher explanation, students must mentally construct the social context while also learning the terminology. In the Role Play condition, the context was externalized: students enacted the roles, interacted with other actors, made choices, and could observe the sequence of an exchange. This does not demonstrate a tested mediating mechanism, but it provides a reasonable interpretation of how the instructional format may have reduced the distance between abstract concepts and familiar daily-life situations. The reflective discussion after the simulation further allowed the teacher to reconnect the performed activity with the formal concepts, which is important because activity alone does not guarantee conceptual learning.

This interpretation is compatible with the experiential-learning pattern reported by Thomas and Rajdeep (2021), whose primary-school intervention combined role play, games, and reflection so that children could encounter and discuss situations directly. It is also consistent with Kajiwara et al. (2023), where learners used role enactment to represent otherwise abstract machine-learning processes. In the present study, the economic topic similarly provided a natural mapping between concept and action: students could embody producers, consumers, buyers, or sellers and observe how choices and transactions unfolded. Such alignment does not prove a specific cognitive mechanism, but it strengthens the pedagogical interpretation that Role Play may make relational concepts more concrete by organizing them into observable sequences of interaction.

The social structure of Role Play may also help explain why the learning process differed from teacher-centered instruction. Students had to understand their own roles, monitor what other participants were doing, communicate, coordinate, and respond to events in the scenario. Saesari et al. (2023) identified communication as an important outcome of role-playing activities in elementary IPAS, while Nindialisma et al. (2025) reported improvement in cooperation and communication following Role Playing. Realitawati et al. (2024) likewise emphasized communication and collaboration as central elements of active twenty-first-century learning in elementary classrooms. The present observation data are consistent with these process characteristics because students participated in preparation, simulation, discussion, and reflection. However, the study did not statistically test communication or cooperation as mediators of achievement. They should therefore be understood as features of the learning environment that may support knowledge construction, not as confirmed causal pathways.

The importance of structure is further illustrated by recent role-based learning research. Hsieh and Hwang (2025) found that concept-map scaffolding within a digital role-playing environment was associated with higher science and English achievement and more reflective learning behavior, while Lin et al. (2023) reported that contextualized role-playing supported by visual guidance and feedback improved achievement and higher-order thinking in AI ethics. These studies suggest that the learning value of role enactment is strengthened when students receive clear representational support and opportunities to organize what occurs during the activity. This is consistent with the present intervention, in which role distribution and simulation were followed by discussion and reflection rather than treated as isolated performance activities.

Comparisons with applications in other subjects further clarify the boundary of the method. Susanto et al. (2024) reported an effect of Role Playing in Indonesian-language drama learning, Nasution et al. (2024) applied it to food-chain material, and Sihaloho et al. (2025) linked the method with critical-thinking development. These studies suggest that Role Play can be adapted beyond social or economic topics, but they also imply that successful application depends on whether the content can be meaningfully represented through roles, sequences, or interaction. Drama naturally contains characters and dialogue, food-chain relationships can be embodied as interacting components, and economic transactions can be enacted as exchanges among actors. In contrast, topics that cannot be represented coherently through roles may require other forms of active learning. The practical implication is not that Role Play should replace conventional instruction universally, but that it is particularly suitable when conceptual relationships can be made visible through structured enactment.

Teacher implementation is another important contextual factor. The teacher-activity checklist reached 100% across the four observed meetings, indicating that all planned stages were recorded as completed. This matters because Role Play is not equivalent to asking students simply to perform a scene. The learning objective must be explicit, the scenario must represent the target concept, roles must be sufficiently clear for students to act productively, and the simulation must be followed by discussion and reflection. Without these stages, students may remember the performance but fail to connect it with academic content. The complete implementation observed in this study therefore helps establish that the comparison involved the intended instructional procedure rather than a partial or inconsistent version of it. Nevertheless, a 100% checklist score should not be interpreted as evidence that every stage was implemented with identical quality, because a completion checklist is less sensitive to variation in depth, timing, or pedagogical skill.

The same caution applies to the student-activity and response data. Student activity reached 100% on the observation indicators, and every experimental-group student gave positive Yes/No responses to all ten questionnaire statements. These results show that students participated in the behaviors included on the instruments and that they did not express negative responses through the available questionnaire options. They do not establish that all students were equally engaged, equally comfortable, or equally helped by the method. Binary response options can produce limited differentiation, and ceiling-level percentages reduce the capacity of the measure to distinguish stronger from weaker reactions. Future studies could supplement such measures with Likert-type response scales, qualitative student reflections, or structured observation ratings that capture intensity and quality as well as completion. This would allow the process

evidence to contribute more analytically to understanding how students experience Role Play.

The ceiling-level response data should also be considered in light of evidence showing that positive affective responses and academic gains do not always move together. Chen et al. (2024) found that role-play gamification enhanced learning motivation without producing differences in Chinese or mathematics scores, whereas Yu and Tsuei (2023) reported concurrent benefits for achievement and selected dimensions of attention and self-efficacy. Shang et al. (2026) similarly observed gains in both achievement and motivation in a gamified project-based learning condition that incorporated role-playing elements. These contrasting patterns reinforce the need to keep the present 100% positive responses conceptually separate from the achievement effect: the questionnaire documents favorable acceptance of the learning experience, but it does not establish that positive response was the mechanism producing the higher posttest scores.

The baseline results also require careful interpretation. The control and experimental groups had relatively similar pretest means, and the baseline comparison was not statistically significant. This supports the descriptive statement that there was no statistically detectable difference in measured prior performance. It does not prove that the classes were equivalent in all respects, nor does it make the pretest irrelevant. Pretest scores represent prior knowledge measured before treatment and can still be related to later performance even though they cannot be caused by the subsequent intervention. For this reason, retaining ANOVA as the primary posttest comparison while adding ANCOVA as a sensitivity analysis provides a balanced analytical strategy. The primary analysis remains consistent with the research question focused on final outcomes, and the sensitivity analysis shows that the main group difference persists when baseline scores are incorporated into the model.

This analytical strategy also clarifies the language used to describe the findings. Statements such as "Role Play caused improvement" or "Role Play was proven to increase achievement" would imply a stronger causal conclusion than the design can support. The more defensible conclusion is that students receiving Role Play demonstrated significantly higher posttest learning outcomes than students receiving conventional instruction, and that this difference remained after baseline adjustment. This wording preserves the substantive result while acknowledging the design. It is particularly important because the treatment condition and class membership coincide: one intact class received Role Play and one intact class received conventional instruction. Unmeasured class-level characteristics could therefore contribute to the observed difference even though the pretest adjustment did not materially change the result.

The scholarly contribution of the study lies in the combination of content alignment, outcome comparison, and process description. The novelty is not the use of Role Play in a new school or the claim that the method has never been studied. Existing studies already show that Role Playing can be applied to learning outcomes, communication, cooperation, drama, science content, and critical thinking. The present contribution is narrower: it provides evidence from an IPAS economic topic in which the role structure of the method is closely matched to the conceptual structure of the content, and it interprets the posttest difference alongside evidence that the planned instructional stages were carried out and accepted by students. The baseline-adjusted sensitivity analysis further strengthens the transparency of the statistical interpretation while preserving the posttest comparison as the primary analysis.

The emerging literature also indicates that instructional effects can differ across learner profiles and implementation designs. Koong et al. (2026) reported particular benefits of a dynamically assessed role-playing game for students with lower prior achievement, and Tsuei and Lu (2025) found that learning benefits in digital role-playing insurance education varied across gender and prior-achievement groups. The present study was not designed or powered to test such subgroup differences, so no comparable claim is made here. However, these findings identify an important direction for extending the current evidence: future research could examine whether the advantage associated with Role Play is uniform across students or is stronger for learners who begin with different levels of prior knowledge, confidence, or engagement.

Several limitations define the appropriate scope of interpretation. First, the sample consisted of only 43 students from two intact classes at one school. The findings therefore should not be generalized directly to all elementary schools, grade levels, or student populations. Second, individual random assignment was not used, so residual differences between classes cannot be excluded even when pretest scores are controlled statistically. Third, the intervention concerned one topic and the outcome was measured immediately after instruction; the study does not show whether the observed difference would persist over time or transfer to topics that are less easily represented through roles. Fourth, the observation and response instruments produced ceiling values, limiting their ability to discriminate variation in implementation quality and student experience. Finally, the psychometric description available for the learning-outcome instrument is limited to the statement that validity and reliability testing was completed before use; this constrains independent evaluation of measurement quality and should be considered when interpreting the findings.

Within these boundaries, the findings have practical implications for elementary IPAS teaching. Role Play can be considered when teachers need students to understand concepts that involve people, roles, exchange, sequence, or decision making. A useful implementation sequence is to begin with a clearly defined conceptual objective, construct a scenario that requires students to use the target concepts, assign roles with understandable responsibilities, conduct the simulation, and then use structured discussion and reflection to translate experience back into academic language. Time management is also important so that performance does not displace conceptual consolidation. For teachers, the central value of the method is therefore not entertainment or performance itself but the opportunity to make relationships visible and discussable through action.

Future research could test the same instructional logic with a larger number of classes and schools so that class-level variation can be separated more clearly from treatment effects. Studies could also include delayed posttests to examine retention, use more discriminating measures of engagement and response, and assess whether communication or cooperation statistically mediates learning outcomes. Comparing Role Play with other active-learning approaches would also help determine whether the observed advantage is specific to role enactment or reflects broader benefits of student-centered participation. Such work would move the evidence beyond the question of whether Role Play can be used successfully toward a more precise understanding of when it is most useful, for which kinds of IPAS content, and through which classroom processes it supports learning.

CONCLUSION

Students who received Role Play instruction demonstrated higher posttest Natural and Social Sciences learning outcomes than students who received conventional instruction in the topic "How Do We Get All Our Needs?" The primary One-Way ANOVA showed a significant between-group difference, and the ANCOVA sensitivity analysis confirmed that the difference remained after pretest performance was statistically adjusted. The study's contribution is therefore not a claim that Role Play is novel in itself, but evidence that a role-based simulation aligned with an elementary IPAS economic topic can be associated with higher final achievement while being implemented through structured scenario preparation, role enactment, discussion, and reflection. Teacher and student activity checklists and student responses indicated complete endorsement of the observed implementation indicators. Interpretation should remain bounded by the use of two intact classes from one school, the absence of individual random assignment, the focus on one topic, and the use of immediate posttest outcomes. Role Play can consequently be considered a context-appropriate instructional alternative for IPAS content that can be represented meaningfully through social roles and interaction, provided that teachers connect the enacted experience explicitly with the intended concepts.

REFERENCES

Alghamdy, R. (2022). How Do EFL Learners Interact and Learn English Communication Skills during a Role-Play Strategy? World Journal of English Language, 12(1), 313–320. https://doi.org/10.5430/wjel.v12n1p313

Charamba, E. (2022). Pushing linguistic boundaries: Translanguaging in a bilingual Science and Technology classroom. Journal of Multilingual and Multicultural Development, 43(10), 951–965. https://doi.org/10.1080/01434632.2020.1783544

Chen, P.-Z., Chang, T.-C., & Wu, C.-L. (2024). Class of Oz: Role-play gamification integrated into classroom management motivates elementary students to learn. Educational Studies, 50(6), 1373–1388. https://doi.org/10.1080/03055698.2022.2081788

Darmawan, A. D., Purwati, & Andriani, D. N. (2024). Penerapan metode Role Playing untuk meningkatkan hasil belajar siswa kelas IV SDN 01 Klegen pada mata pelajaran IPAS topik pemenuhan kebutuhan. Pendas: Jurnal Ilmiah Pendidikan Dasar, 9(3), 747–758. https://doi.org/10.23969/jp.v9i3.15443

Hsieh, S.-F., & Hwang, G.-J. (2025). Concept Maps as a Content and Language Integrated Learning Facilitator in Digital Gaming Contexts: Triangulated Evidence From Students' Learning Achievements, Perceptions and Behavioural Patterns. Journal of Computer Assisted Learning, 41(6). https://doi.org/10.1111/jcal.70136

Kajiwara, Y., Matsuoka, A., & Shinbo, F. (2023). Machine learning role playing game: Instructional design of AI education for age-appropriate in K-12 and beyond. Computers and Education: Artificial Intelligence, 5. https://doi.org/10.1016/j.caeai.2023.100162

Koong, C.-S., Tseng, H.-C., & Chen, C.-H. (2026). Effects of incorporating dynamic assessment into a role-playing game on students' English grammar achievements and intrinsic motivation. Innovation in Language Learning and Teaching. https://doi.org/10.1080/17501229.2026.2730677

Lin, X.-F., Wang, Z., Zhou, W., Luo, G., Hwang, G.-J., Zhou, Y., Wang, J., Hu, Q., Li, W., & Liang, Z.-M. (2023). Technological support to foster students' artificial

intelligence ethics: An augmented reality-based contextualized dilemma discussion approach. Computers and Education, 201. https://doi.org/10.1016/j.compedu.2023.104813

Melik-Pashaev, A. A., & Novlyanskaya, Z. N. (2022). Artist in a child. National Psychological Journal, 47(3), 26–34. https://doi.org/10.11621/npj.2022.0304

Nasution, N. F., Dalimawaty, Afriwani, I., & Sutarini. (2024). Penggunaan metode Role Playing untuk meningkatkan hasil belajar peserta didik pada materi rantai makanan kelas 5 SD Negeri 060924 Medan Amplas. Cendekia Pendidikan, 4(4), 50–54.

Nindialisma, T. C., Sadtyadi, H., & Ngadat, N. (2025). Pengaruh metode Role Playing terhadap keterampilan kerjasama dan komunikasi siswa dalam pembelajaran IPAS berbasis Kurikulum Merdeka di kelas vb sdn 7 wonogiri. SOCIAL: Jurnal Inovasi Pendidikan IPS, 5(2), 590–600. https://doi.org/10.51878/social.v5i2.5296

Rauf, A. (2024). Pengaruh implementasi model Role-Playing guna meningkatkan hasil belajar siswa. UNISAN Jurnal, 3(7), 253–266.

Realitawati, R., Ikrom, F. D., Herawan, E., & Kadarsah, D. (2024). Penerapan 4C skills dalam pembelajaran abad 21 di sekolah dasar. Muallimuna: Jurnal Madrasah Ibtidaiyah, 10(1), 22–32. https://doi.org/10.31602/muallimuna.v10i1.15533

Saesari, A. A. I., Untari, M. F. A., & Nuvitalia, D. (2023). Analisis metode bermain peran terhadap keterampilan berkomunikasi pada mata pelajaran Ilmu Pengetahuan Alam dan Sosial kelas IV sekolah dasar. Didaktik: Jurnal Ilmiah PGSD STKIP Subang, 9(2), 4561–4570. https://doi.org/10.36989/didaktik.v9i2.1149

Shang, J., Huang, W., & Li, X. (2026). Effects of Gamified Project-Based Learning on Mathematics Achievement and Motivation: A Comparative Study in Elementary Education. International Journal of Science and Mathematics Education, 24(2). https://doi.org/10.1007/s10763-025-10644-w

Sihaloho, M. S., Sihombing, D. J., & Ginting, F. Y. A. (2025). Meningkatkan kemampuan berpikir kritis siswa SD melalui penerapan metode pembelajaran Role Playing. Jurnal Pendidikan dan Keguruan, 2(11), 2248–2253.

Susanto, A. H., Widyawati, Y., Setyaningsih, N., & Sumardi, S. (2024). Pengaruh metode Role Playing terhadap hasil belajar Bahasa Indonesia materi drama pada siswa kelas V sekolah dasar. Pendas: Jurnal Ilmiah Pendidikan Dasar, 9(3), 200–207. https://doi.org/10.23969/jp.v9i03.16758

Thomas, B., & Rajdeep, C. (2021). Life Skill Education: Enhancing Empowerment among Rural Primary School Children in Gujarat, India. Space and Culture, India, 8(4), 22–32. https://doi.org/10.20896/saci.v8i4.898

Tsuei, M., & Lu, H.-C. (2025). Enhancing Children's Insurance-Related Knowledge and Learning Attitudes Through Digital Game-Based Learning. International Journal of Serious Games, 12(2), 77–90. https://doi.org/10.17083/ijsg.v12i2.879

Yu, Y.-T., & Tsuei, M. (2023). The effects of digital game-based learning on children's Chinese language learning, attention and self-efficacy. Interactive Learning Environments, 31(10), 6113–6132. https://doi.org/10.1080/10494820.2022.2028855

Yusnarti, M., & Suryaningsih, L. (2021). Pengaruh model pembelajaran Role Playing terhadap hasil belajar siswa sekolah dasar. Ainara Journal: Jurnal Penelitian dan PKM Bidang Ilmu Pendidikan, 2(3), 253–261