Pedagogical Research
2026, 11(3), em0277
e-ISSN: 2468-4929
https://www.pedagogicalresearch.com

MODESTUM

Research Article

OPEN ACCESS

The effect of STEAM-based learning on values education: An experimental study with primary school students

Derya Kayıran [1]*, Hatice Zeynep İnan [2]

[1] Kahramanmaraş Sütçü İmam University, TÜRKİYE
[2] School of Teaching and Learning, Sam Houston State University, Texas, UNITED STATES
***Corresponding Author:** deryakeskinpalta@gmail.com

Citation: Kayıran, D., & İnan, H. Z. (2026). The effect of STEAM-based learning on values education: An experimental study with primary school students. *Pedagogical Research, 11*(3), em0277. https://doi.org/10.29333/pr/18942

ARTICLE INFO

Received: 08 Mar. 2026

Accepted: 16 Jun. 2026

ABSTRACT

This study explored the association between participation in STEAM-based learning activities and first-grade students' values development. Using a quasi-experimental single-group pretest–posttest design, the study was conducted with 30 students attending a public primary school. Data were collected through the Pictorial Values Scale for Children, which measures fifteen core values. Over six weeks, students engaged in twelve interdisciplinary STEAM activities designed to promote collaboration, problem-solving, and values-oriented learning. Benjamini–Hochberg adjusted significance levels were additionally considered to control inflated Type I error associated with multiple comparisons. The findings revealed differentiated patterns of change across value domains. Patriotism demonstrated the strongest and most consistent improvement, while friendship, empathy, diligence, and kindness showed moderate gains. By contrast, self-control, compassion, responsibility, and respect reflected only small changes, and honesty, patience, and helpfulness showed no significant overall improvement. Overall, STEAM-based instruction appeared to be more strongly associated with socially enacted values than with deeply internalized moral dispositions in short-term interventions.

Keywords: curriculum and instruction, pictorial values scale for children, primary school students, quasi-experimental design, STEAM, values education

INTRODUCTION

Values education in the primary school years constitutes a critical foundation for later social, ethical, and civic development, and recent curriculum reforms increasingly seek to integrate cognitive and socio-emotional goals within schooling (İnan, 2011; Kharozim, 2024). Early primary education represents a particularly sensitive period in which dispositions such as empathy, responsibility, and cooperation begin to take shape, making this stage crucial for structured values development (Tanabashi, 2021). Concurrently, STEAM (Science, Technology, Engineering, Arts, and Mathematics) education has gained prominence as an interdisciplinary approach promoting problem-solving, creativity, and conceptual understanding across domains (Spyropoulou et al., 2020). Just as cartoons on television can contribute to values development in young children (İnan, 2016), STEAM has also more recently been viewed as a pedagogical framework capable of supporting broader developmental outcomes, including socio-emotional skills and values formation, particularly in early grades where curriculum integration is more flexible (Tanabashi, 2021). Research indicates that arts-integrated, inquiry-based STEAM activities enhance engagement and creative expression while fostering collaboration, persistence, and responsibility processes closely associated with values development (Huang, 2024). Similarly, recent research suggests that STEAM-based science education, grounded in transformative and epistemically inclusive perspectives, can extend beyond simply improving task performance and contribute to the development of students' identities (Thapaliya & Luitel, 2026). Because human beings are a whole composed of mind, body, and soul, it is clear that children's education should support cognitive, social, and emotional growth in an integrated manner (Inan & Inan, 2015).

However, implementation studies also reveal persistent constraints: Many teachers lack sufficient training and practical experience to effectively apply STEAM pedagogy, limiting its potential impact on non-cognitive outcomes such as values formation (Chuong et al., 2023; Lyu et al., 2024). Thus, the translation of interdisciplinary learning into developmental gains depends heavily on instructional capacity and contextual support. Conceptually, STEAM pedagogy aligns well with mechanisms underlying values learning. Recent evidence also indicates that teachers tend to view innovative and interdisciplinary educational approaches positively, yet frequently report concerns regarding implementation readiness and the need for professional development support (Bilen et al., 2016; Kayıran et al., 2026). This suggests that the educational potential of approaches such as STEAM may

depend not only on instructional design but also on teachers' preparedness to integrate emerging pedagogical practices into classroom contexts.

Arts integration and experiential, object-based learning encourage perspective-taking, expression, and collaborative responsibility, supporting prosocial orientations central to values education (Kayıran, 2026; Kayıran & Bağçeci, 2025; Li, 2024). Comparative studies show that arts inclusion enhances motivation and conceptual understanding while providing alternative representational pathways for young learners, reinforcing respect for diverse perspectives and shared responsibility in group work (Corrigan et al., 2025). Problem-centered STEAM projects further create authentic contexts in which students negotiate roles, share responsibility, and collaboratively solve problems, enabling values such as fairness, cooperation, and helpfulness to be enacted rather than merely discussed (Spyropoulou et al., 2020; Tanabashi, 2021). Practitioner accounts likewise suggest that when art and design tasks address social or civic themes, they can support moral reasoning and value enactment. Yet such potential depends on curriculum designs that treat values as explicit learning goals rather than incidental outcomes (Inan et al., 2024; Kharozim, 2024).

Despite these theoretical and practice-based insights, rigorous empirical evidence examining STEAM's potential contribution to values development in early primary education remains limited. Current literature is characterized primarily by localized interventions, conceptual discussions, and calls to strengthen teacher preparation, rather than by systematic syntheses demonstrating consistent value-related outcomes (Adnan et al., 2023; Chuong et al., 2023; Huang, 2024; Lyu et al., 2024). Moreover, researchs linking STEAM to specific value domains among young learners remain scarce, and methodological heterogeneity complicates conclusions about STEAM-specific effects. Recent reviews therefore tend to highlight research needs rather than provide definitive estimates of STEAM's impact, emphasizing the necessity of stronger research designs, standardized outcome measures, and developmentally appropriate tools capable of capturing value-related changes among young children (Hsiao & Su, 2021; Huang, 2024; Lyu et al., 2024; Neldarisasmita et al., 2025). Program quality, implementation fidelity, and teacher capacity consistently emerge as moderators influencing outcomes, further underscoring the need for carefully designed and evaluated interventions. Addressing these gaps, the current quasi-experimental pretest-posttest study investigates how STEAM-based instructional activities influence fifteen core values among first-grade students: justice, family unity, diligence, friendship, honesty, empathy, kindness, self-control, patience, respect, love, responsibility, compassion, patriotism, and helpfulness. The study responds to three interrelated needs identified in recent scholarship: Empirically measuring values as explicit STEAM outcomes (Kharozim, 2024); employing developmentally appropriate pictorial assessment tools suitable for young learners (Huang, 2024); and generating classroom-based quasi-experimental evidence capable of informing curriculum design and teacher preparation aimed at maximizing STEAM's socio-emotional and moral potential (Lyu et al., 2024). Recent systematic reviews also emphasize that successful STEAM implementation depends heavily on instructional quality, interdisciplinary integration, and teacher preparedness (Amanova et al., 2025).

Recent studies also show that STEAM education is increasingly being discussed through broader cultural, artistic, and interdisciplinary perspectives. Plageras et al. (2020), for example, demonstrated how STEAM can be used to connect historical and cultural content with active learning experiences. Similarly, Lillis (2024) emphasized the importance of artistic pathways in moving from STEM to STEAM, highlighting the role of creativity and aesthetic engagement in learning. El Bedewy et al. (2024) further showed that transdisciplinary and technology-assisted STEAM practices can be shaped by cultural contexts. These studies support the view that STEAM is not limited to science and technology learning alone, but can also provide a flexible pedagogical space for social, cultural, and values-oriented learning.

In sum, the convergence of growing interest in STEAM (e.g., Spyropoulou et al., 2020), theoretical links between integrative pedagogy and experiential values learning, and persistent evidence gaps regarding values outcomes in early education justify empirical investigations such as the present study. By providing quasi-experimental evidence across fifteen core values using a child-centered pictorial assessment approach, this research contributes to addressing existing gaps while offering practical insights for educators seeking to employ STEAM pedagogy to support both cognitive and moral development in primary education.

METHOD

Research Design

This study employed a quasi-experimental single-group pretest–posttest design to examine changes in first-grade students' values development following participation in STEAM-based learning activities. In this research, measurements obtained from the same group before and after intervention enable evaluation of instructional impact through comparative analysis.

Although a control group was not included, the single-group pretest-posttest design was preferred due to contextual and ethical considerations. The intervention was implemented within a naturally existing classroom setting, and all students were provided access to the STEAM-based instructional program to ensure educational equity. Therefore, the study focuses on within-group change over time rather than making strong causal comparisons between experimental and control conditions. While this design allows examination of instructional impact, potential internal validity threats such as maturation or testing effects should be considered when interpreting the findings. Accordingly, the findings should be interpreted as exploratory and associational rather than as definitive causal evidence regarding the effects of STEAM-based instruction on values development.

Participants

The study group consisted of 30 first-grade primary school students enrolled in a public primary school during the 2025–2026 academic year. Participants were selected using convenience sampling based on accessibility and voluntary participation. Institutional approval and parental consent were obtained prior to data collection, and ethical procedures concerning child participation were followed throughout the study.

The participating children were between 6 and 7 years of age, which is developmentally consistent with first-grade primary education in Türkiye. The school was located in a middle socioeconomic urban context. Because the study was conducted within a naturally existing classroom setting, all students in the classroom participated in the intervention process in order to preserve instructional continuity and educational equity.

Although demographic information such as age and grade level was collected, gender information was not systematically recorded during the data collection process. Therefore, gender-based analyses could not be conducted, which should be considered a limitation when interpreting the findings.

Data Collection Instrument

Values awareness was measured using the *Pictorial Values Scale for Children* (PVSC) developed by Kayıran (2025), which demonstrates satisfactory validity and reliability for preschool and primary school children. The PVSC was originally developed through comprehensive scale development procedures including expert review, content validation, exploratory analyses, and reliability testing appropriate for early childhood assessment. Construct validity evidence reported in the original study supports the multidimensional structure representing fifteen distinct values. Given that the present study employed an already validated instrument, additional item-level validity testing was not conducted. Instead, internal consistency reliability was re-examined within the current sample to confirm measurement stability. Although the original scale development study established satisfactory construct validity evidence for the instrument, construct validity was not re-examined within the present sample because the relatively small sample size was not suitable for stable factor-analytic procedures. Accordingly, the current study focused on internal consistency reliability and instructional application rather than psychometric revalidation. The scale measures fifteen values: justice, family unity, diligence, friendship, honesty, empathy, kindness, self-control, patience, respect, love, responsibility, compassion, patriotism, and helpfulness. Each value is represented through five visual scenarios appropriate to children's developmental levels. In each set, three visuals represent the target value, while two functions as distractors depicting inconsistent behaviors.

Scoring Procedure

Responses were collected using a three-point pictorial Likert-type scale suitable for young learners (smiling, neutral, and sad faces). For visuals representing the correct value, responses were scored from 1 to 3 points in ascending order, whereas scoring was reversed for distractor visuals. Scores obtained from correct visuals were summed and distractor scores subtracted, producing value scores ranging from 1 to 15. Interpretation intervals proposed by Kayıran (2025) classify mean scores as low, moderate, or high value recognition levels.

Implementation Process

The study was conducted in three stages: Pretest, intervention, and posttest. Initially, the PVSC was administered individually to participants. Subsequently, students participated in a six-week STEAM-based instructional program comprising twelve sessions conducted twice weekly. At the end of the intervention, the same instrument was re-administered to evaluate changes in values awareness.

Structure of STEAM-Based Instruction

The instructional program was designed using an interdisciplinary STEAM framework integrating science, technology, engineering, arts, and mathematics within values-oriented learning activities adapted to first-grade developmental characteristics. Activities promoted collaboration, problem-solving, creative production, and experiential learning while embedding values such as justice, cooperation, responsibility, empathy, and respect within authentic learning contexts.

Activities incorporated observation, collaborative design, artistic expression, dramatization, storytelling, music, and game-based learning to support both cognitive and affective development. Instruction emphasized active participation, cooperation, experiential engagement, and connections between classroom learning and daily life. STEAM tasks were intentionally structured to embed targeted value domains within authentic and collaborative learning experiences appropriate to first-grade developmental characteristics. Limited family involvement was also encouraged through home-based reflection and discussion activities.

To enhance methodological transparency and support replicability, the general structure of the six-week STEAM-based instructional program is summarized in **Table 1**. The table presents the thematic focus, interdisciplinary STEAM components, learning activities, targeted values, instructional roles, materials, and implementation contexts associated with each weekly session.

Table 1. Structure of the six-week STEAM-based values education program

Week	Session Theme	STEAM Components	Learning Activities	Targeted Values	Teacher Role	Student Role	Materials	Context
1	Helping Our Environment	Science + Art	Recycling design activity and environmental poster creation	Responsibility, helpfulness	Facilitator and guide	Collaborative designer	Recyclable materials, colored paper	Group work
1	Building Together	Engineering + Mathematics	Constructing simple structures collaboratively	Cooperation, friendship	Organizer	Team-based problem solver	Blocks, rulers, cardboard	Small-group activity
2	Fair Solutions	Engineering + Arts	Designing fair-sharing scenarios through role-play	Justice, empathy	Discussion moderator	Active participant	Scenario cards, drawings	Cooperative learning
2	Our Community Values	Science + Art	Community-themed visual storytelling	Patriotism, respect	Facilitator	Story creator	Art supplies, visuals	Whole-class discussion
3	Feelings and Communication	Arts + Technology	Digital emotion cards and dramatization	Empathy, kindness	Emotional support facilitator	Performer and observer	Tablets, emotion visuals	Pair work
3	Solving Problems Together	Engineering + Mathematics	Collaborative puzzle-solving tasks	Self-control, diligence	Process guide	Collaborative learner	Logic games, worksheets	Small-group activity
4	Designing Helpful Tools	Engineering + Science	Creating simple tools for helping others	Helpfulness, responsibility	Facilitator	Creative designer	Craft materials	Group-based production
4	Respect in Daily Life	Arts + Drama	Role-playing respectful behaviors	Respect, patience	Classroom moderator	Role-play participant	Drama cards	Interactive classroom
5	Stories of Honesty	Arts + Literacy	Story completion and illustration tasks	Honesty, justice	Story facilitator	Reflective participant	Story cards, markers	Individual and group work
5	Caring for Others	Science + Arts	Health and care-themed collaborative projects	Compassion, kindness	Guide and observer	Cooperative learner	Posters, visual materials	Cooperative learning
6	Working with Persistence	Mathematics + Engineering	Completing challenging design tasks	Diligence, self-control	Motivator	Problem solver	STEM kits, worksheets	Team-based learning
6	Celebrating Shared Values	Integrated STEAM	Group exhibition and reflective presentations	Patriotism, friendship, cooperation	Facilitator and evaluator	Presenter and collaborator	Student products, presentation materials	Whole-class presentation

Throughout the intervention process, activities were implemented by the classroom teacher in collaboration with the researchers using standardized instructional plans prepared prior to implementation. The instructional process emphasized active participation, collaborative learning, creative production, and value-oriented reflection appropriate to early primary education settings.

Intervention Fidelity

To support implementation consistency, all activities were conducted by the same classroom teacher using pre-prepared instructional plans developed collaboratively with the researchers prior to implementation. Weekly activity guidelines, instructional sequences, materials, and targeted value components were standardized across sessions in order to maintain procedural consistency throughout the six-week intervention process.

In addition, the researchers monitored the overall implementation process through regular communication and review of instructional progress. However, no formal fidelity observation checklist or independent implementation monitoring procedure was employed. Therefore, consistency in instructional delivery could not be systematically verified, which constitutes a methodological limitation of the study.

Data Analysis

Prior to statistical analyses, the assumption of normality was examined for the difference scores (posttest–pretest) using skewness and kurtosis coefficients. The values were found to fall within the acceptable range of −1.5 to +1.5, indicating approximate normal distribution and supporting the use of parametric procedures (Tabachnick & Fidell, 2013).

To examine pretest–posttest changes across value domains, paired samples t-tests were conducted. Since the same participants were measured at two time points, the analyses were based on dependent observations. Degrees of freedom were calculated as $n - 1$ (df = 29). Statistical significance was evaluated at the .05 level.

In addition to total value scores, item-level analyses were conducted to examine students' differentiation between value-consistent and value-inconsistent visual scenarios, with illustrative examples for the Justice (Fairness) and Family Unity values presented in **Tables 2** and **3**, respectively. Because each value domain includes both target and distractor visuals, analyzing item-level change provides a more nuanced understanding of recognition accuracy and conceptual discrimination. This approach

Table 2. Paired samples t-test results for the justice (Fairness) Value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	1.87 (.63)	2.43 (.63)	-4.96	29	<.001	0.91	Large
Image 2	1.63 (.67)	1.07 (.25)	4.96	29	<.001	0.91	Large
Image 3	2.20 (.61)	2.93 (.25)	-6.28	29	<.001	1.15	Very Large
Image 4	2.20 (.76)	1.80 (.61)	3.03	29	.005**	0.55	Medium
Image 5	2.20 (.66)	2.53 (.63)	-2.76	29	.010*	0.50	Medium
Total	10.10 (1.16)	10.77 (.82)	-2.94	29	.006**	0.54	Medium

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 3. Paired samples t-test results for the family unity value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.00 (.37)	2.73 (.45)	-7.71	29	<.001	1.41	Very Large
Image 2	1.83 (.38)	1.00 (.00)	12.04	29	<.001	2.20	Very Large
Image 3	2.30 (.47)	2.90 (.31)	-5.84	29	<.001	1.07	Very Large
Image 4	2.13 (.35)	1.23 (.43)	16.16	29	<.001	2.95	Very Large
Image 5	2.27 (.45)	3.00 (.00)	-8.93	29	<.001	1.63	Very Large
Total	10.53 (.86)	10.87 (.35)	-1.84	29	.077	0.34	Small

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

allows detection of contextual shifts that may not be fully captured by aggregated total scores, particularly in short-term interventions. In addition to statistical significance testing, effect sizes were calculated using Cohen's *d* for dependent samples in order to determine the magnitude of observed changes. This approach enabled a more comprehensive interpretation of intervention effects across value domains.

Because multiple paired-samples t-tests were conducted across fifteen value domains, the Benjamini–Hochberg false discovery rate (FDR) correction procedure was additionally applied in order to reduce the risk of inflated Type I error while maintaining statistical power. Adjusted significance levels were considered when interpreting the statistical findings across value domains.

Given the sample size (N = 30), the study was considered sufficiently powered to detect medium-to-large within-group effects; however, smaller effects may not have been detected reliably. Therefore, findings associated with small effect sizes should be interpreted cautiously, particularly in relation to statistical stability and generalizability.

Reliability

Internal consistency of the PVSC was evaluated using Cronbach's alpha. The original scale reported an alpha value of .91 (Kayıran, 2025), while analysis within the present sample yielded an alpha coefficient of .89, indicating high reliability. Accordingly, the instrument was considered reliable for assessing values awareness in this study.

FINDINGS

The following section presents the findings regarding the examination of fifteen values within the scope of the STEAM-based intervention.

Pretest and posttest scores related to the value of justice (fairness) were compared using a paired samples t-test. The results indicated statistically significant changes in students' awareness of the justice value following the intervention. Among the images representing the correct value, posttest scores for Image 1 (M = 2.43, SD = .63) were significantly higher than pretest scores (M = 1.87, SD = .63), t(29) = -4.96, p < .001. Similarly, for Image 3, the posttest mean (M = 2.93, SD = .25) was significantly higher than the pretest mean (M = 2.20, SD = .61), t(29) = -6.28, p < .001. A significant increase was also observed for Image 5, t(29) = -2.76, p = .010. For the distractor images (Image 2 and Image 4), posttest scores significantly decreased (Image 2: t(29) = 4.96, p < .001; Image 4: t(29) = 3.03, p = .005). This finding suggests that students were better able to distinguish unfair behaviors following the intervention. When total scores were examined, the posttest mean (M = 10.77, SD = .82) was significantly higher than the pretest mean (M = 10.10, SD = 1.16), t(29) = -2.94, p = .006. These results demonstrate that the STEAM-based intervention significantly improved students' overall awareness of the justice value. Effect size analyses revealed large effects for Image 1, Image 2, and Image 3 (d ≥ .80), indicating a strong impact of the intervention on students' ability to recognize justice-related behaviors. The medium effect size obtained for the total score (d = .54) suggests that the intervention produced a statistically significant and pedagogically meaningful improvement in overall value awareness.

The analysis results for the family unity value indicate statistically significant differences between pretest and posttest scores across all five images (p < .001). For the images representing the correct value (Image 1, Image 3, and Image 5), posttest scores increased significantly. For the distractor images (Image 2 and Image 4), posttest scores significantly decreased, suggesting that students were better able to distinguish behaviors that do not represent family unity after the intervention. Effect size analyses revealed very large effects across all images (d > 1.00). In particular, the effect size obtained for Image 4 (d = 2.95) indicates a very strong intervention effect. However, the increase observed in total scores was not statistically significant, t(29) = -1.84, p = .077.

Table 4. Paired samples t-test results for the diligence value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.00 (.53)	2.57 (.57)	-5.46	29	<.001	1.00	Large
Image 2	1.57 (.63)	1.17 (.46)	4.40	29	<.001	0.80	Large
Image 3	2.60 (.50)	3.00 (.00)	-4.40	29	<.001	0.80	Large
Image 4	2.17 (.59)	1.73 (.78)	4.71	29	<.001	0.86	Large
Image 5	2.43 (.57)	2.80 (.41)	-3.61	29	.001**	0.66	Medium–Large
Total	10.77 (1.38)	11.27 (1.11)	-2.48	29	.019*	0.45	Medium

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 5. Paired samples t-test results for the friendship value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.87 (.35)	3.00 (.00)	-2.11	29	.043*	0.39	Small–Medium
Image 2	2.33 (.71)	2.27 (.74)	0.39	29	.702	0.07	Negligible
Image 3	2.80 (.41)	3.00 (.00)	-2.69	29	.012*	0.49	Medium
Image 4	2.03 (.89)	2.40 (.81)	-1.65	29	.110	0.30	Small
Image 5	2.37 (.61)	2.83 (.38)	-3.75	29	.001**	0.68	Medium–Large
Total	12.40 (1.35)	13.50 (1.25)	-3.37	29	.002**	0.62	Medium–Large

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 6. Paired samples t-test results for the honesty value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.07 (.78)	2.23 (.90)	-1.22	29	.231	0.22	Small
Image 2	1.23 (.57)	1.27 (.52)	-0.30	29	.769	0.05	Negligible
Image 3	2.60 (.62)	2.67 (.66)	-0.40	29	.690	0.07	Negligible
Image 4	1.63 (.85)	1.67 (.80)	-0.20	29	.845	0.04	Negligible
Image 5	2.70 (.60)	2.63 (.67)	0.49	29	.625	0.09	Negligible
Total	10.23 (1.79)	10.47 (2.30)	-0.55	29	.590	0.10	Small

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

The effect size for the total score was small (d = .34). This suggests that while the intervention had a strong effect on individual behavior discrimination, it did not produce a statistically significant overall change in the total scale score.

Pretest and posttest scores related to the diligence value were compared using a paired samples t-test, and the results are presented in **Table 4**. Statistically significant differences were found across all images (p < .01). For the correct images (Image 1, Image 3, and Image 5), posttest means increased significantly. Notably, a large effect size was observed for Image 1 (d = 1.00). In Image 3, all students obtained the highest score in the posttest (SD = .00), indicating a strong intervention effect. For the distractor images (Image 2 and Image 4), posttest scores decreased significantly, suggesting improved discrimination ability. When total scores were examined, the posttest mean (M = 11.27) was significantly higher than the pretest mean (M = 10.77), t(29) = -2.48, p = .019. The effect size for the total score was medium (d = .45), indicating a statistically significant and pedagogically meaningful improvement in overall diligence awareness.

Pretest and posttest scores for the friendship value were compared using a paired samples t-test, and the results are presented in **Table 5**. Significant increases were observed in Image 1, Image 3, and Image 5. In particular, a medium-to-large effect size was obtained for Image 5 (d = .68). No statistically significant differences were found for Image 2 and Image 4 (p > .05), suggesting that the intervention did not produce significant changes in discrimination ability for certain behavioral examples. When total scores were examined, the posttest mean (M = 13.50) was significantly higher than the pretest mean (M = 12.40), t(29) = -3.37, p = .002. The effect size for the total score was medium-to-large (d = .62), indicating a statistically significant and pedagogically meaningful improvement in overall friendship awareness.

Pretest and posttest scores related to the honesty value were compared using a paired samples t-test, and the results are presented in **Table 6**. The results indicated no statistically significant differences across any image or in the total score (p > .05). Effect size analyses revealed that all effect sizes were small or negligible (d < .20). These findings suggest that the intervention did not produce a meaningful change in students' awareness of the honesty value. Regarding the total score, the posttest mean (M = 10.47) did not significantly differ from the pretest mean (M = 10.23), t(29) = -0.55, p = .590. The effect size was small (d = .10).

Pretest and posttest scores related to the empathy value were compared using a paired samples t-test, and the results are presented in **Table 7**. The results indicated no statistically significant differences at the individual image level (p > .05). For Image 1, the pretest and posttest means were identical, indicating no observable change. However, when total scores were examined, the posttest mean (M = 14.27) was significantly higher than the pretest mean (M = 13.60), t(29) = -2.57, p = .016. The effect size for the total score was medium (d = .47). These findings suggest that while the intervention did not produce marked changes in individual behavior discrimination, it resulted in an overall increase in general empathy awareness.

Table 7. Paired samples t-test results for the empathy value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.97 (.18)	2.97 (.18)	0.00	29	1.000	0.00	None
Image 2	2.57 (.68)	2.80 (.61)	-1.65	29	.109	0.30	Small
Image 3	2.87 (.35)	2.93 (.25)	-1.00	29	.326	0.18	Small
Image 4	2.63 (.56)	2.77 (.57)	-1.28	29	.211	0.23	Small
Image 5	2.57 (.57)	2.80 (.41)	-1.76	29	.090	0.32	Small
Total	13.60 (1.35)	14.27 (1.17)	-2.57	29	.016*	0.47	Medium

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 8. Paired samples t-test results for the kindness value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.90 (.40)	3.00 (.00)	-1.36	29	.184	0.25	Small
Image 2	2.10 (.84)	2.17 (.99)	-0.70	29	.489	0.13	Negligible
Image 3	2.90 (.40)	2.97 (.18)	-1.00	29	.326	0.18	Small
Image 4	2.13 (.94)	2.20 (.96)	-0.81	29	.423	0.15	Negligible
Image 5	2.77 (.50)	2.83 (.46)	-1.44	29	.161	0.26	Small
Total	12.80 (1.81)	13.17 (1.86)	-2.16	29	.039*	0.40	Medium (Borderline)

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 9. Paired samples t-test results for the self-control value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.17 (.65)	2.73 (.45)	-4.26	29	<.001	0.78	Large
Image 2	2.07 (.69)	1.13 (.35)	6.91	29	<.001	1.26	Very Large
Image 3	1.83 (.79)	2.73 (.52)	-4.79	29	<.001	0.88	Large
Image 4	1.93 (.64)	1.20 (.48)	4.10	29	<.001	0.75	Large
Image 5	2.13 (.73)	2.70 (.54)	-2.98	29	.006**	0.54	Medium
Total	10.13 (1.53)	10.50 (.73)	-1.32	29	.197	0.24	Small

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 10. Paired samples t-test results for the patience value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	1.57 (.77)	1.67 (.80)	-0.50	29	.620	0.09	Negligible
Image 2	1.20 (.48)	1.23 (.57)	-0.23	29	.823	0.04	Negligible
Image 3	2.27 (.78)	2.30 (.84)	-0.15	29	.879	0.03	Negligible
Image 4	2.00 (.79)	1.87 (.73)	0.81	29	.423	0.15	Small
Image 5	2.77 (.50)	2.73 (.58)	0.30	29	.769	0.05	Negligible
Total	9.80 (2.12)	9.80 (2.23)	0.00	29	1.000	0.00	None

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Pretest and posttest scores related to the kindness value were compared using a paired samples t-test, and the results are presented in **Table 8**. No statistically significant differences were found at the individual image level (p > .05). Effect sizes were small or negligible (d < .30). However, when total scores were examined, the posttest mean (M = 13.17) was significantly higher than the pretest mean (M = 12.80), t(29) = -2.16, p = .039. The effect size for the total score was small-to-medium (d = .40). These findings indicate that while the intervention did not produce strong changes in individual behavior discrimination, it resulted in a statistically significant increase in overall kindness awareness.

Pretest and posttest scores for the self-control value were compared using a paired samples t-test, and the results are presented in **Table 9**. Statistically significant differences were found across all images (p < .01). For the correct images (Image 1, Image 3, and Image 5), posttest scores increased significantly. A very large effect was observed for Image 2 (d = 1.26). For the distractor images (Image 2 and Image 4), posttest scores significantly decreased, indicating improved discrimination of behaviors not representing self-control. However, the increase observed in the total score was not statistically significant, t(29) = -1.32, p = .197. The effect size for the total score was small (d = .24), suggesting that while the intervention strongly affected individual behavior discrimination, it did not produce a significant overall change in total self-control awareness.

Pretest and posttest scores for the patience value were compared using a paired samples t-test, and the results are presented in **Table 10**. No statistically significant differences were found across any image or in the total score (p > .05). Effect sizes were negligible across all measures (d < .20). The identical pretest and posttest total means (t(29) = 0.00, p = 1.000) indicate that the intervention did not produce any observable change in overall patience awareness. This finding may suggest that patience is more resistant to short-term interventions or that baseline levels were already relatively stable.

Table 11. Paired samples t-test results for the respect value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.87 (.43)	2.90 (.40)	-0.57	29	.573	0.10	Negligible
Image 2	1.90 (.80)	1.87 (.73)	0.21	29	.839	0.04	Negligible
Image 3	2.50 (.78)	2.73 (.58)	-1.32	29	.199	0.24	Small
Image 4	1.33 (.48)	1.23 (.43)	0.83	29	.415	0.15	Small
Image 5	2.90 (.31)	3.00 (.00)	-1.80	29	.083	0.33	Small–Medium
Total	11.50 (1.04)	11.73 (.74)	-0.96	29	.345	0.18	Small

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 12. Paired samples t-test results for the love value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.03 (.76)	2.77 (.43)	-4.63	29	<.001	0.84	Large
Image 2	2.37 (.67)	1.47 (.51)	8.12	29	<.001	1.48	Very Large
Image 3	2.20 (.76)	2.60 (.56)	-2.69	29	.012	0.49	Medium
Image 4	1.93 (.78)	1.37 (.49)	4.26	29	<.001	0.78	Large
Image 5	2.03 (.85)	2.53 (.57)	-2.92	29	.007	0.53	Medium
Total	10.57 (1.87)	10.73 (1.48)	-0.48	29	.634	0.09	Negligible

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 13. Paired samples t-test results for the responsibility value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	1.93 (.87)	1.63 (.76)	1.56	29	.130	0.28	Small
Image 2	1.23 (.50)	1.30 (.65)	-0.44	29	.662	0.08	Negligible
Image 3	2.50 (.68)	2.43 (.73)	0.34	29	.738	0.06	Negligible
Image 4	2.00 (.74)	1.87 (.78)	0.81	29	.423	0.15	Small
Image 5	2.77 (.43)	2.70 (.65)	0.47	29	.645	0.09	Negligible
Total	10.43 (2.01)	9.93 (2.45)	0.96	29	.346	0.17	Small

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 14. Paired samples t-test results for the compassion value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	1.70 (.79)	1.77 (.86)	-0.42	29	.677	0.08	Negligible
Image 2	1.23 (.57)	1.20 (.48)	0.33	29	.745	0.06	Negligible
Image 3	2.63 (.61)	2.33 (.76)	1.80	29	.083	0.33	Small–Medium
Image 4	1.63 (.85)	1.33 (.66)	1.66	29	.107	0.30	Small
Image 5	2.70 (.60)	2.53 (.78)	0.90	29	.378	0.16	Small
Total	9.90 (1.45)	9.17 (2.12)	1.60	29	.122	0.29	Small

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Pretest and posttest scores for the respect value were compared using a paired samples t-test, and the results are presented in **Table 11**. No statistically significant differences were found across individual images or in the total score ($p > .05$). Effect sizes were small or negligible ($d < .35$). Although an increase was observed in Image 5, this difference was not statistically significant ($p = .083$). Similarly, total respect scores did not differ significantly between pretest and posttest ($t(29) = -0.96$, $p = .345$), indicating no meaningful intervention effect on respect awareness.

Pretest and posttest scores for the love value were analyzed using a paired samples t-test, and the results are presented in **Table 12**. Significant differences were observed across all five images ($p < .05$). Image 1 showed a significant increase in posttest scores ($d = .84$, large effect). Image 2 demonstrated a significant decrease with a very large effect ($d = 1.48$). Images 3 and 5 showed medium-level effects ($d \approx .50$), while Image 4 showed a large effect ($d = .78$). However, no statistically significant difference was found in the total love score ($t(29) = -0.48$, $p = .634$, $d = .09$). This suggests that despite directional changes in individual items, the overall total score may have balanced out.

Pretest and posttest scores for the responsibility value were analyzed using a paired samples t-test, and the results are presented in **Table 13**. No statistically significant differences were found in any sub-item or in the total score ($p > .05$). Effect sizes were small or negligible across all measures. Although the pretest mean for Image 1 was higher than the posttest mean, this difference was not significant ($t(29) = 1.56$, $p = .130$, $d = .28$). Overall, the intervention did not produce a meaningful cognitive or attitudinal change in responsibility.

The paired-samples t-test results for the compassion value are presented in **Table 14**. No statistically significant differences were found for compassion at the item or total level ($p > .05$). Image 3 approached significance ($p = .083$) with a small-to-medium effect size ($d = .33$), yet this difference did not reach statistical significance. Overall, the intervention did not produce a statistically meaningful change in compassion awareness.

Table 15. Paired samples t-test results for the patriotism value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.20 (.66)	2.87 (.43)	-4.55	29	<.001	0.83	Large
Image 2	1.67 (.76)	1.17 (.46)	2.92	29	.007	0.53	Medium
Image 3	2.03 (.67)	2.83 (.46)	-4.74	29	<.001	0.87	Large
Image 4	1.90 (.88)	1.10 (.31)	4.56	29	<.001	0.83	Large
Image 5	2.03 (.76)	3.00 (.00)	-6.92	29	<.001	1.26	Very Large
Total	9.83 (1.12)	10.97 (.56)	-5.19	29	<.001	0.95	Large

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 16. Paired samples t-test results for the helpfulness value (N = 30)

Measure	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Cohen's d	Effect Size
Image 1	2.57 (.73)	2.90 (.31)	-2.57	29	.016	0.47	Medium
Image 2	2.17 (.79)	1.50 (.78)	3.25	29	.003	0.59	Medium
Image 3	2.40 (.81)	2.83 (.46)	-3.26	29	.003	0.60	Medium
Image 4	1.77 (.73)	1.40 (.77)	1.88	29	.070	0.34	Small–Medium
Image 5	2.70 (.60)	2.97 (.18)	-2.50	29	.018	0.46	Medium
Total	11.60 (2.25)	11.60 (1.40)	0.00	29	1.000	0.00	None

*p < .05, **p < .01, ***p < .001; Note. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

Table 17. Comparison of total scores across all values with Benjamini–Hochberg adjusted significance levels (N = 30)

Rank	Value	Pretest Mean (SD)	Posttest Mean (SD)	t	df	p	Adjusted p (BH-FDR)	Cohen's d	Effect Size
1	Patriotism	9.83 (1.12)	10.97 (.56)	-5.19	29	<.001	<.001	0.95	Large
2	Friendship	12.40 (1.35)	13.50 (1.25)	-3.37	29	.002	.015	0.62	Medium–Large
3	Justice	10.10 (1.16)	10.77 (.82)	-2.94	29	.006	.030	0.54	Medium
4	Empathy	13.60 (1.35)	14.27 (1.17)	-2.57	29	.016	.060	0.47	Medium
5	Diligence	10.77 (1.38)	11.27 (1.11)	-2.48	29	.019	.057	0.45	Medium
6	Kindness	12.80 (1.81)	13.17 (1.86)	-2.16	29	.039	.098	0.40	Small–Medium
7	Family Unity	10.53 (.86)	10.87 (.35)	-1.84	29	.077	.165	0.34	Small
8	Compassion	9.90 (1.45)	9.17 (2.12)	1.60	29	.122	.229	0.29	Small
9	Self-Control	10.13 (1.53)	10.50 (.73)	-1.32	29	.197	.329	0.24	Small
10	Respect	11.50 (1.04)	11.73 (.74)	-0.96	29	.345	.518	0.18	Small
11	Responsibility	10.43 (2.01)	9.93 (2.45)	0.96	29	.346	.472	0.17	Small
12	Honesty	10.23 (1.79)	10.47 (2.30)	-0.55	29	.590	.737	0.10	Very Small
13	Love	10.57 (1.87)	10.73 (1.48)	-0.48	29	.634	.731	0.09	Very Small
14	Patience	9.80 (2.12)	9.80 (2.23)	0.00	29	1.000	1.000	0.00	None
15	Helpfulness	11.60 (2.25)	11.60 (1.40)	0.00	29	1.000	1.000	0.00	None

Note. Benjamini–Hochberg false discovery rate correction was applied across the fifteen total-score comparisons in order to control inflated Type I error associated with multiple statistical testing. *p < .05, **p < .01, ***p < .001. A paired samples t-test was conducted. Cohen's d values were calculated using the formula d = t/√n. Effect sizes were interpreted as approximately .20 = small, .50 = medium, and .80 = large.

As shown in **Table 15**, significant differences were observed across all items and in the total score for patriotism (p < .01). Large effect sizes were found for most items, with a very large effect for Image 5 (d = 1.26). The total score also increased significantly (t(29) = -5.19, p < .001, d = .95), indicating a strong and consistent intervention effect on patriotism.

As shown in **Table 16**, significant differences were found in Images 1, 2, 3, and 5 (p < .05), with medium-level effect sizes (d ≈ .46–.60). Image 4 did not reach statistical significance. However, total helpfulness scores remained identical between pretest and posttest (t(29) = 0.00, p = 1.000), suggesting that directional item-level changes balanced out at the aggregate level.

Table 17 presents the comparison of total scores across all values following the Benjamini–Hochberg false discovery rate correction procedure. Friendship and justice also retained comparatively stronger patterns following adjustment for multiple comparisons.

Empathy, diligence, and kindness demonstrated comparatively more tentative patterns of change. Although these domains showed statistically significant differences in the unadjusted analyses, they did not remain statistically significant after the Benjamini–Hochberg correction procedure. Therefore, these changes should not be interpreted as robust intervention outcomes, but rather as tentative improvements that require cautious interpretation and further confirmation in future studies.

Values such as self-control, family unity, compassion, responsibility, respect, honesty, patience, and helpfulness demonstrated limited or non-significant changes at the total-score level following adjustment procedures.

Smaller effect sizes were identified for self-control, family unity, compassion, responsibility, and respect. These dimensions are more closely associated with self-regulation, moral judgment, and internalized control mechanisms. The relatively modest magnitude of change suggests that such values may be less susceptible to short-term instructional interventions and may instead require sustained, experiential, and behavior-oriented learning processes to achieve more pronounced development.

Values that remained statistically non-significant or negligible at the total score level included helpfulness, honesty, and patience. Notably, although certain item-level changes were detected for helpfulness and love, the absence of change in aggregate scores suggests that increases in some behavioral representations may have been offset by more critical evaluations in others. This pattern may reflect a contextual recalibration of value interpretation rather than a uniform directional shift.

Taken together, the findings indicate that the intervention exerted stronger influence on emotionally salient and collectively oriented values, while producing comparatively limited change in character-based and deeply internalized moral dispositions. From a theoretical standpoint, this distinction is noteworthy: short-term or scenario-based instructional designs may be particularly effective in mobilizing social and collective value domains, whereas enduring character traits likely require longer-term, practice-based, and experience-driven pedagogical approaches.

Overall, the integrated pattern demonstrates that the intervention exhibited a selective effect profile, with the magnitude and direction of change varying systematically according to the structural nature of the value domain.

DISCUSSION

Previous research has shown that STEAM-based instructional activities can integrate values education effectively (Adnan et al., 2023). However, most studies do not examine pretest–posttest changes across distinct value domains. The present findings provide a more differentiated picture of how specific values respond to short-term STEAM-based interventions. Importantly, because multiple statistical comparisons were conducted across value domains, Benjamini–Hochberg adjusted significance levels were also considered. Accordingly, some findings should be interpreted as exploratory and tentative rather than definitive.

The results reveal a selective pattern of change rather than a uniform transformation across value domains. The strongest and most consistent pattern of change emerged in patriotism, with a large effect size (d = .95). Moderate effects were observed for friendship, which remained statistically significant after the Benjamini–Hochberg adjustment. Empathy, diligence, and kindness showed small-to-moderate changes in the unadjusted analyses; however, these improvements did not remain statistically significant after correction for multiple comparisons and should therefore be interpreted as tentative. These findings suggest that participation in STEAM-based learning environments may be associated with increased sensitivity toward collective identity and socially interactive values, while having more limited short-term influence on deeply internalized character dispositions.

The strong effect on patriotism may be explained by the collaborative and product-oriented nature of STEAM activities. Project-based learning environments emphasize shared goals, group accountability, and public presentation of outcomes. Such contexts may support the development of collective belonging and civic-oriented attitudes, especially when tasks are framed around community or societal themes (Sochacka et al., 2016; Wu & Chang, 2023). Research indicates that collaborative design and maker-centered pedagogies increase students' sense of group participation and shared responsibility (Girdzijauskienė & Šmitienė, 2020). When learners jointly construct tangible outputs, values related to collective identity may be reinforced more readily than individually internalized moral dispositions. Within the present study, patriotism was operationalized primarily in terms of collective participation, social belonging, and community-oriented cooperation rather than political or ideological orientation.

Similarly, the moderate improvements observed in friendship, empathy, diligence, and kindness align with literature highlighting the social-interactional dimension of STEAM pedagogy. Arts integration and collaborative problem-solving provide structured opportunities for peer negotiation, role-taking, and mutual support (Leskinen et al., 2023). Such repeated interaction patterns have been associated with increased prosocial engagement in classroom contexts (Thornberg & Oğuz, 2013). Importantly, these values are relational in nature; therefore, they can become visible through short-term behavioral engagement.

In contrast, minimal or small changes in honesty, patience, responsibility, compassion, and respect are consistent with moral development theories emphasizing the gradual internalization of ethical dispositions. Deeply rooted character traits typically develop through long-term socialization processes involving family, community, and sustained reflection (Emre et al., 2024; Juujärvi & Myyry, 2022; Thornberg & Oğuz, 2013). Short-term project cycles, even when collaborative, may not provide sufficient duration to support observable change in deeply internalized moral orientations.

Notably, helpfulness showed significant change at the item level but not at the total score level, indicating a possible contextual reinterpretation rather than consistent value internalization. This pattern suggests that students may differentiate between situational helping behaviors and broader moral commitments, an issue that warrants further qualitative investigation.

From an ecological perspective, value transformation requires multi-context reinforcement beyond isolated classroom experiences (Avendaño-Uribe et al., 2022). While STEAM's hands-on and collaborative structure may facilitate observable social engagement, sustained curricular continuity and intentional ethical framing are necessary for deeper moral internalization. Longitudinal and mixed-method approaches are therefore recommended to capture gradual value consolidation over time (Hughes et al., 2023).

Overall, the present findings suggest that STEAM-based instruction may be associated with differentiated patterns of change across value domains. The strongest short-term changes were observed in collective and socially enacted value, whereas deeply internalized moral dispositions remain comparatively resistant to rapid change. These findings highlight the importance of integrating long-term pedagogical planning, ethical reflection components, and cross-context collaboration in order to maximize STEAM's contribution to holistic values education.

CONCLUSION

This study investigated changes in first-grade students' values development following a STEAM-based instructional program. The findings revealed differentiated patterns of change across value domains rather than uniform improvement. The most substantial and consistent effect was observed in patriotism, which demonstrated a large effect size and statistically significant improvement at both the item and total score levels. The observed increase in patriotism scores may indicate heightened awareness of collective identity and civic-oriented values when learning tasks emphasize shared goals and group accountability during the STEAM activities. More robust patterns of change were observed in patriotism, friendship, and justice. Although empathy, diligence, and kindness showed promising improvements in the initial analyses, these changes did not remain statistically significant after adjustment for multiple comparisons and should therefore be considered tentative. These domains are closely linked to peer interaction, collaboration, and joint problem-solving core components of STEAM pedagogy. The results suggest that STEAM environments may support prosocial engagement and cooperative interaction among young learners. In contrast, honesty, patience, responsibility, compassion, and respect demonstrated small or non-significant changes. These values are more deeply internalized and regulated through long-term moral socialization processes. The limited short-term change in these domains suggests that while STEAM activities can enhance socially visible behaviors, they may not be sufficient on their own to rapidly transform stable moral dispositions. Notably, helpfulness and love displayed item-level fluctuations but no significant change at the total score level, indicating contextual variation rather than consistent value internalization. This pattern suggests that students may reinterpret helping behaviors differently across situational contexts. Importantly, the overall scale did not demonstrate uniform large gains across all values, but rather showed domain-specific responsiveness. These findings highlight that STEAM-based instruction supports values education in a differentiated manner, with stronger short-term effects on collectively enacted and interaction-based values than on internally regulated moral traits.

Overall, STEAM pedagogy appears to provide a potentially supportive context for socially observable and collaborative values in early primary education. However, deeper moral internalization likely requires sustained, multi-context reinforcement extending beyond short-term classroom interventions. In addition, although the Benjamini–Hochberg correction procedure was applied to reduce inflated Type I error associated with multiple comparisons, the relatively small sample size may still limit the stability of some statistical estimates. Additionally, the relatively small sample size, the use of convenience sampling, and the single-school context limit the broader generalizability of the findings. The study was conducted exclusively with first-grade students attending one public primary school in Türkiye; therefore, the results may reflect context-specific developmental, cultural, and educational characteristics. Although the study was capable of detecting medium-to-large effects, smaller changes across certain value domains may have remained statistically underpowered. Consequently, the findings should be interpreted cautiously when considering transferability to different grade levels, school settings, or cultural contexts. Furthermore, although standardized instructional plans were used throughout the intervention, implementation fidelity was not formally assessed through independent observational procedures. Additionally, teacher-related variables such as pedagogical content knowledge, prior STEAM implementation experience, instructional beliefs, and orientation toward values education were not systematically controlled within the present study. Variability in these instructional characteristics may have influenced implementation quality and students' responsiveness to the intervention. Because the measurement process relied on children's recognition and evaluation of socially desirable behaviors presented through pictorial scenarios, the potential influence of social desirability bias cannot be completely ruled out. Participants may have tended to select responses perceived as socially appropriate rather than reflecting fully internalized value orientations. Additionally, because gender information was not systematically recorded, potential gender-related differences in values responsiveness could not be examined.

RECOMMENDATIONS

Based on the differentiated impact observed across value domains, future STEAM implementations should prioritize longer-term and developmentally sustained interventions. Deeply internalized values such as honesty, responsibility, patience, and compassion appear to require extended exposure, repeated practice, and structured ethical reflection beyond short project cycles. Therefore, integrating explicit moral dialogue, guided reflection on dilemmas, and value-oriented questioning into STEAM activities may strengthen internalization processes. In addition, alignment between school practices and family or community contexts should be encouraged, as value development is reinforced through ecological consistency. Incorporating experiential strategies such as drama, service-learning, and real-life problem scenarios may further support affective engagement and transfer. Future research should employ longitudinal and mixed-method designs to capture gradual developmental change, while sustained professional development for teachers is essential to ensure that ethical framing and reflective components are intentionally embedded within STEAM pedagogy. Overall, maximizing STEAM's contribution to values education requires extended duration, contextual alignment, reflective depth, and pedagogical continuity.

Author contributions: DK & HZİ: study design the study, methodology, formal analyses, writing – original draft, writing – review & editing; **DK:** data curation. Both authors agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Acknowledgements: The authors would like to thank the participating students and school administration for their cooperation during the implementation of the study.

Ethical statement: The authors stated that ethical approval for this study was obtained from the Social Sciences Ethics Committee of Kahramanmaraş Sütçü İmam University on 26.12.2025. Written informed consents were obtained from the participants.

AI statement: The authors stated that, during the preparation of this manuscript, the author used an artificial intelligence–based language assistance tool to improve the clarity, grammar, and readability of the text. The AI tool was used solely for language editing and structuring purposes. All conceptualization, study design, data collection, statistical analyses, interpretation of results, and final decisions regarding the content of the manuscript were carried out entirely by the author. The author takes full responsibility for the integrity and accuracy of the work.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Adnan, M., Ahmad, C. N. C., Ibharim, L. F. M., Khalid, F. A. M., Prihaswati, M., Khasanah, U., & Hidayat, R. (2023). Development and usability of STEAM textbook integrated character education with local wisdom themes for primary school students. *Journal of Higher Education Theory and Practice*, *23*(5), 54-61. https://doi.org/10.33423/jhetp.v23i5.5925

Amanova, A. K., Butabayeva, L. A., Abayeva, G. A., Umirbekova, A. N., Abildina, S. K., & Makhmetova, A. A. (2025). Systematic review on the implementation of STEAM education in schools. *Eurasia Journal of Mathematics, Science and Technology Education, 21*(1), Article em2568. https://doi.org/10.29333/ejmste/15894

Avendaño-Uribe, B., Ojeda-Ramírez, S., & Perez-Baron, J. (2022). Resourcefulness, narratives, and identity in science, technology, engineering, arts and mathematics education: A perspective of makerspaces for rural communities in Colombia. *Frontiers in Education, 7,* 1-8. https://doi.org/10.3389/feduc.2022.1055722

Bilen, K., Ercan, O., & Doğruluk, S. (2016). Fen ve teknoloji öğretmenlerinin mesleki tükenmişlikleri ile öğretim programına ilişkin görüşlerinin incelenmesi [Examining the professional burnout of science and technology teachers and their views on the curriculum.]. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, *18*(1), 115-136. https://doi.org/10.17556/jef.00879

Chuong, H. L. A., Thien, N. H., Hanh, T. T. T., & Hanh, D. C. (2023). Assessing the competence of early childhood education students at teacher education universities in Vietnam in terms of implementing STEAM education. *European Journal of Contemporary Education, 12*(2), 385-398. https://doi.org/10.13187/ejced.2023.2.385

Corrigan, M., Wong, J., Grove, D., Andersen, S., & Hughes, B. (2025). Enhancing elementary students' conceptual understandings of scientific phenomena: The impact of STEAM-first and STEM-first approaches. *Science Education, 109*(5), 1336-1364. https://doi.org/10.1002/sce.21942

El Bedewy, S., Lavicza, Z., Sabitzer, B., Houghton, T., & Nurhasanah, F. (2024). Exploring transdisciplinary, technology-assisted, and architectural modelling STEAM practices through a cultural lens. *European Journal of Science and Mathematics Education, 12*(2), 211-235. https://doi.org/10.30935/scimath/14304

Emre, O., Özyazıcı, K., Keskinkılıç, A., & Arslan, Z. (2024). The role of adolescents' mental health and well-being in predicting their smoking status. *Archives of Psychiatric Nursing*, *51*, 137-142. https://doi.org/10.1016/j.apnu.2024.06.005

Girdzijauskienė, R., & Šmitienė, G. (2020). Integration of arts in STEAM projects: Experience of primary school teachers. *Gamtamokslinis Ugdymas / Natural Science Education, 17*(2), 74-84. https://doi.org/10.48127/gu-nse/20.17.74

Hsiao, P., & Su, C. (2021). A study on the impact of STEAM education for sustainable development courses and its effects on student motivation and learning. *Sustainability, 13*(7), Article 3772. https://doi.org/10.3390/su13073772

Huang, W. (2024). Young students' aesthetic literacy and online usage during COVID-19 in Taiwan. *Journal of Media and Management, 6*(1) 1-18. https://doi.org/10.47363/jmm/2024(6)166

Hughes, J., Morrison, L., Robb, J., & Hagerman, M. (2023). "It feels like I have a camera in my eye": New methods for literacies research in maker-oriented classrooms. *Frontiers in Education, 8*, Article 1098410. https://doi.org/10.3389/feduc.2023.1098410

İnan, H. Z. (2011). Okul öncesi eğitim programında değerler eğitiminin kapsamı [The scope of values education in preschool education programs.]. In A. Arıkan (Ed.), *Okul öncesi dönemde değerler eğitimi* (pp. 95-122). Anadolu Üniversitesi Yayınları.

Inan, H. Z., & Inan, T. (2015). 3*H*s Education: Examining hands-on, heads-on and hearts-on early childhood science education. *International Journal of Science Education*, *37*(12), 1974-1991. https://doi.org/10.1080/09500693.2015.1060369

Inan, H. Z., Jozwiak, M., Inan, Y. S., & Sarac, S. (2024). Examining concept development classroom interaction quality and children's developmental progress in state pre-K/head start programs using pre-K CLASS and LAP-3. *Acta Psychologica*, *245*, Article 104201. https://doi.org/10.1016/j.actpsy.2024.104201

İnan, T. (2016). Çocuk medyasında evrensel ve yerel değerlerin aktarımı: TRT Çocuk kanalı örneği [The transmission of universal and local values in children's media: The example of TRT Çocuk channel.]. *Akademik Sosyal Araştırmalar Dergisi*, *4*(25), 200-212. https://doi.org/10.16992/ASOS.1135

Juujärvi, S., & Myyry, L. (2022). Online dilemma discussions as a method of enhancing moral reasoning among health and social care graduate students. *International Journal of Ethics Education, 7*(2), 271-287. https://doi.org/10.1007/s40889-022-00143-9

Kayıran, D. (2025). Development of a pictorial values scale for children: A validity and reliability study. *Değerler Eğitimi Dergisi, 23*(50), 527-556. https://doi.org/10.34234/ded.1803856

Kayıran, D. (2026). Education for sustainability in curriculum-based activities: A study on the attitudes and activity plans of in-service and preservice teachers towards environmental sustainability education. *Journal of Curriculum Studies Research*, *8*(1), 44-60. https://doi.org/10.46303/jcsr.2026.3

Kayıran, D., & Bağçeci, B. (2025). Effectiveness of a values education program for preschool students: a teacher-based evaluation study. *Journal of Culture and Values in Education*, *8*(3), 61-94. https://doi.org/10.46303/jcve.2025.27

Kayıran, D., Sönmez, M., Avcı, A., & Haji Mohamud, R. Y. (2026). Teachers' perceptions of artificial intelligence in curriculum integration: Opportunities, concerns, and professional development needs. *Frontiers in Artificial Intelligence*, *9*, Article 1806165. https://doi.org/10.3389/frai.2026.1806165

Kharozim, M. (2024). STEAM learning model in Madrasah Ibtidaiyah. *Elementary, 2*(2), 58-62. https://doi.org/10.55210/elementary.v2i2.445

Leskinen, J., Kajamaa, A., & Kumpulainen, K. (2023). Learning to innovate: Students and teachers constructing collective innovation practices in a primary school's makerspace. *Frontiers in Education, 7*, Article 936724. https://doi.org/10.3389/feduc.2022.936724

Li, Z. (2024). Reflections of STEAM education concept on art and design education in the context of digitalization. *Education Reform and Development, 6*(3), 176-184. https://doi.org/10.26689/erd.v6i3.6637

Lillis, M. (2024). From STEM to STE(A)M: STEM education through artistic pathways. *Hellenic Journal of STEM Education*, *3*(2), 21-27. https://doi.org/10.51724/hjstemed.v3i2.19

Lyu, S., Niu, S., Yuan, J., & Zhan, Z. (2024). Developing professional capital through technology-enabled university-school-enterprise collaboration: An innovative model for C-STEAM preservice teacher education in the Greater Bay area. *Asia Pacific Journal of Innovation and Entrepreneurship, 18*(3), 270-299. https://doi.org/10.1108/apjie-01-2024-0014

Neldarisasmita, Haryanto, & Wicaksana, E. J. (2025). Uncovering STEAM research patterns and gaps: Systematic literature review in education journal. *Jurnal Penelitian Pendidikan IPA, 11*(4), 1-10. https://doi.org/10.29303/jppipa.v11i4.10976

Plageras, A., Kourtis, S., Xenakis, A. C., Kalovrektis, K., Psycharis, S., & Vavougios, D. (2020). Understanding ancient Greek civilizations: A STEAM teaching perspective. *Hellenic Journal of STEM Education*, *1*(2), 45-57. https://doi.org/10.51724/hjstemed.v1i2.9

Sochacka, N., Guyotte, K., & Walther, J. (2016). Learning together: A collaborative autoethnographic exploration of STEAM (STEM + the arts) education. *Journal of Engineering Education, 105*(1), 15-42. https://doi.org/10.1002/jee.20112

Spyropoulou, C., Wallace, M., Vassilakis, C., & Poulopoulos, V. (2020). Examining the use of STEAM education in preschool education. *European Journal of Engineering and Technology Research*, (Special Issue: CIE), 1-6. https://doi.org/10.24018/ejeng.2020.0.cie.2309

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Pearson.

Tanabashi, S. (2021). STEAM education using sericulture ukiyo-e: Object-based learning through original artworks collected at a science university museum in Japan. *Interdisciplinary Journal of Environmental and Science Education, 17*(4), Article e2248. https://doi.org/10.21601/ijese/10962

Thapaliya, P., & Luitel, B. C. (2026). STEAM-based inclusive science education: Envisioning Brahma-Vishnu-Shiva epistemology. *Educational Philosophy and Theory*, (SPECIAL ISSUE: CIE 2020), 1-20. https://doi.org/10.1080/00131857.2026.2628276

Thornberg, R., & Oğuz, E. (2013). Teachers' views on values education: A qualitative study in Sweden and Turkey. *International Journal of Educational Research, 59*, 49-56. https://doi.org/10.1016/j.ijer.2013.03.005

Wu, S., & Chang, Y. (2023). Designing and implementing maker curriculum for promoting 7th graders' maker competence. *Frontiers in Education, 8*, Article 1201534. https://doi.org/10.3389/feduc.2023.1201534