Student Creativity in Mechanical and Non-Mechanical Word Problems

Azobit Charles Awindago[a] and Clement Ayarebilla Ali[b]

[a]Department of Science and Mathematics Education, C. K. Tedam University of Technology and Applied Science, Ghana
[b]Department of Basic Education, School of Education and Lifelong Learning, University of Education, Winneba, Ghana

Article History: Received: 25 October 2024; Accepted: 7 December 2025; Published online: 30 December 2025

Abstract: The study distinguishes between mechanical word problems, which involve straightforward manipulation of mathematical symbols and procedures, and non-mechanical problems, which require interpreting real-life situations and translating them into mathematical models. To examine senior high school students' cognitive processes, the researchers employed Wallas's four-stage problem-solving model. Using a correlational research design, a sample of 350 students was randomly selected from a population of 638. Data were collected through questionnaires—chosen for scalability, anonymity, and accuracy—and observation checklists for their efficiency, objectivity, and systematic documentation. Findings for Research Question 1 indicated that Wallas's information-processing model provided progressive scaffolding that supported students within their Zone of Proximal Development when working on mechanical word problems. This scaffolding gradually faded, enabling a transfer of responsibility as students moved from mechanical to non-mechanical problems. For Research Question 2, students' knowledge levels were analyzed using mean, standard deviation, and frequency distributions. Research Question 3 employed Chi-square tests and Pearson correlations. The results showed that students possessed limited knowledge of non-mechanical word problems compared to mechanical ones. Additionally, no statistically significant relationships were found between problem-solving performance and gender, age, or program of study. The study recommended that stakeholders design interventions tailored to the specific learning needs of students.

Keywords: Mechanical, Non-mechanical, Students' creativity, Wallas model, Word problems

1. Introduction

Word problems provide opportunities for students to acquire the ability to solve problems in contextualized environments without resorting to memorization (Taley & Ndamenenu, 2022). These word problems are divided into mechanical and non-mechanical problems. In both mechanical and non-mechanical problems, students are expected to translate word problems, the way they understand them, into equations and apply their mathematical knowledge to solve them.

Mechanical problems involve straightforward manipulation of mathematical symbols and operations without the necessity of contextual understanding. Students can rely on procedural knowledge and algorithmic steps to arrive at their solutions (Qetrani et al., 2021). In contrast, non-mechanical problems require students to interpret a real-world scenario, extract relevant information, and then translate it into a mathematical model. Real-world problems require mathematical solutions that do not come in equations ready to be solved, but as words that need to be understood symbolically before being solved algebraically (Vale & Barbosa, 2023). This process necessitates a higher level of cognitive engagement, including reading comprehension, contextual reasoning, and the ability to identify and apply appropriate mathematical concepts (Qetrani et al., 2021).

1.1. The Problem

Hirashima (2015) contends that students often fail to solve problems when given problem-solving exercises. This practice permeates every high school mathematics class and tends to derail tedious efforts teachers make during mathematics instruction.

The second contention is the influences of gender, age, and programme differences in addition to attitudes, anxiety, failure, insecurity, self-confidence, and low perseverance (Khasawneh et al., 2019; Rembert et al., 2019; Wakhata et al., 2024). For instance, Bertoletti et al. (2023) discovered the myths of men being quantitative, while women are qualitative affect students' attitudes toward word problems. These problems were traced to inappropriate interventions and instructional strategies.

Lastly, Haavold and Sriraman (2021) opine that even though Polya's four-step model has become the most popular approach to teaching and learning problem-solving, it is principally focused almost entirely on heuristics or rules of thumb for making progress on difficult problems. This seemingly denies students of creativity and inventiveness to the discourse. Wallas's (1926) four-step problem-solving model has been tied to Vygotsky's (1978) Theory of Scaffolding to ensure learners themselves find solutions to the gaps between the teacher and the learner in a mathematics problem.

1.2. The Theoretical Framework of Scaffolding

Corresponding Author: Clement Ayarebilla Ali **email:** aliclement83@gmail.com
Citation Information: Ali, C. A. & Awindago, A. C. (2025). Student creativity in mechanical and non-mechanical word problems. *Turkish Journal of Mathematics Education*, *6*(3), 1-14.

In Figure 1, the study was guided by Malik and Wiseman's (2017) conception of Vygotsky's scaffolding theory of learning, whose main principle is to break down information or parts of a new skill into digestible pieces for the learner. The first stage is that the learner has no knowledge and CANNOT do any work at all. In the second stage, the learner CAN do activities WITH the support of teachers or experienced adults. In the third stage, the learner CAN do activities WITHOUT the support of the teacher. The implication is that students should be educated where they are capable of learning with peer support, instructional strategies, and regular assessment (Malik & Wiseman, 2017).

Figure 1. A theoretical framework for scaffolding (Adopted from Malik & Wiseman, 2017)

Also, Malik and Wiseman (2017) opine that Vygotsky's scaffolding is used when a teacher or capable student helps a learner within their ZPD. When the learner and teacher begin working together, the teacher models most of the work and explains how and why they do things to help the learner comprehend the content. As the learner becomes more comfortable with the material, the educator's assistance lessens, and the learner does more of the work on their own. The scaffolding continues to decrease until the student has mastered the content and no longer needs any scaffolding (Ali & Anderson, 2021).

Again, scaffolding started with the underlying cognitive system where learning takes place in memory of the mechanical word problems. The function was to overcome the limitations of mechanical word problems and zoom into non-mechanical word problems (van Nooijen et al., 2024). By so doing, students would bridge the gap between school and home mathematics as they apply real-life situations.

The scaffolded learning was a two-fold concept. One concept represented a new approach to intelligence testing, analyzing, and tailoring learners' specific learning needs and current talents. The other was a model for determining how social engagement with more experienced teachers affected learners' intellectual development. This was the specific case that established the connections between the minds of boys with girls, younger with older ones, and mathematics programmes with other programmes of study. This helped bridge the gaps between these variables (Raslan, 2024).

1.3. Conceptual Framework

Several cognitivists (Haavold & Sriraman, 2021; Matzin & Mundia, 2020; Sadler-Smith, 2015; Savic, 2016; Voskoglou, 2021) have proposed Wallas (1926) information-processing model of the problem-solving model.

The Wallas (1926) information-processing model of problem-solving is closely linked to the works of Vygotsky's scaffolding, especially ZPD (van Nooijen et al., 2024), and provides access to the ZPD in three components namely, contingency, fading, and transfer of responsibility. In this context, contingency scaffolding was employed by expert teachers to be dependent on and responsive to the learner's needs and responses. For example, whenever teachers provide techniques to solve problems, the learners innovate skills in performing the tasks. The fading was the gradual reduction in scaffolding techniques as the learners became more capable. The transfer of responsibility occurred when the learners successfully mastered the task at hand of moving from mechanical to non-mechanical word problems. For example, the learners now set up their tasks and solve them (van Nooijen et al., 2024).

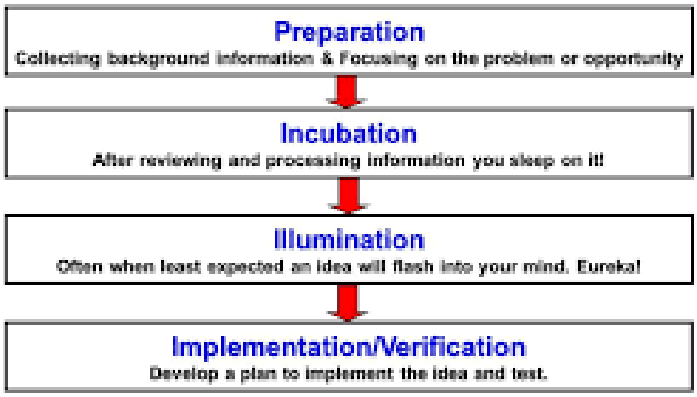

Figure 2. A conceptual framework for scaffolding (Adopted from Savic, 2016)

In Wallas's (1926) information-processing model of problem-solving, the mechanical milieu begins where the students initiate creativity about the problem at the preparation stage of the teachers' contingency. They then transition from the mechanical to the non-mechanical strategies by finding alternative creative strategies to the problem at the incubation stage of the teachers' fading. Finally, they deploy their creativity at the non-mechanical stage to solve the problem at the illumination stage, after which they check the reality and plausibility of the solution at the verification stage of the transfer stage (Savic, 2016; Voskoglou, 2021).

Savic (2016) makes the connection between problem-solving and mathematical creativity and contends that the stages of problem-solving are a subset of mathematical creativity. By exhausting the four steps and embedding them with Vygotsky's ZPD, creativity is more intrinsically enshrined in the process of problem-solving and less in the end product of mechanical and non-mechanical answers (Savic, 2016). Even though Sadler-Smith (2015) argued that Wallas's (1926) four-stage model of the creative process did not reflect accurately Wallas' full account of the creative process, the Intimation stage is a more authentic representation of creativity. Nazzal and Kaufman (2020) used the model to assess the quality of each stage and the overall creativity of students' work and found that creativity was influenced by the quality of the ideas generated and the relationships among the different interpretations of the ideas.

1.4. Research Questions

This study was guided by the following research questions:

1. What are the performances in mechanical and non-mechanical problems?
2. What relationships exist between gender, age, and programmes of study?

i. H_{01}: There are NO relationships between students' ages
ii. H_{02}: There is NO relationship between male and female students
iii. H_{03}: There are NO relationships among the programmes of study

2. Method

2.1. Research Design

In Figure 3, the correlational design compared two groups of students to try and identify their relationships. These variables were gender, age, and programme of study. The purpose of correlational research was to identify that these variables have connections in ways that changes in one affected change in the others. The researchers examined to demonstrate if there is a statistically considerable relationship between them without manipulating them (Ghanad, 2023).

Also, the design established positive or negative direction relationships. A positive correlation means that both variables move in the same direction. In contrast, a negative correlation means that the variables move in opposite directions or zero/no directions. However, the design was not used to establish causality the variables did not provide definitive proof that they lead to positive or negative results in the mechanical and non-mechanical problems (Ghanad, 2023).

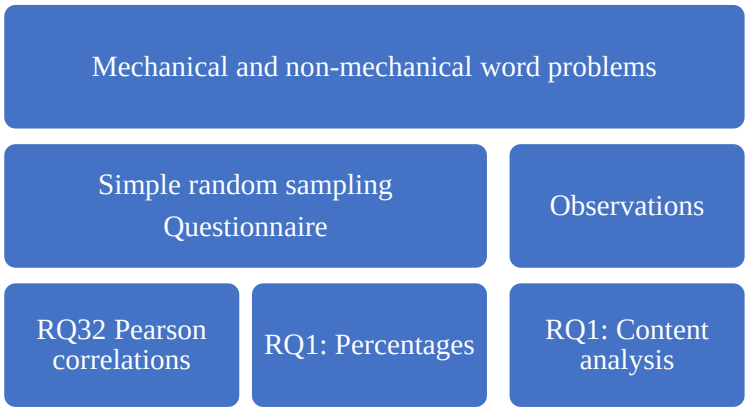

Figure 3. Correlational research design

2.2. Population, sample, and sampling techniques

A total of 638 students in two senior high schools were targeted for the study. This comprised 400 and 238 from schools A' and 'B' respectively, and a sample of 350 students that comprised 200 and 150 from schools A' and 'B' respectively. The sampling encompassed gender, age, programme of study, and the ten-word problems (Shukla, 2020).

The simple random sampling technique was used to ensure that every student had an equal opportunity of being selected as inadequate logistics could not take all the over 600 students. The students' identification numbers were folded and put in a box. With a table of random numbers, the sample was generated starting from a predefined 'seeded' number at the beginning of the programme. This ensured the gender, age, and programme were adequately represented (Shukla, 2020). The gender comprised 229 male and 121 female students (see Table 6), the age group ranged between 15 and 23 years (see Table 4), and the programmes were grouped into General Science, General Arts, Home Economics, Agricultural Sciences, and Technical and Business (see Table 8).

2.3. Research Instruments

The research instruments were closed-ended and open-ended questionnaires and observation checklists. These two instruments of data collection complemented each other and enhanced the validity and reliability of the data. While the questionnaire catered for quick reach, scalability, anonymity, and accuracy, the observation checklist took care of efficiency, objectivity, standardization, and documentation (Ghanad, 2023).

2.4.1. Closed-ended questionnaire

The closed-ended questionnaire was structured into Sections A and B. Section A gathered data on age, gender, and Programmes of study. Section 'B' gathered data on the knowledge. The knowledge "Mechanical and Non-Mechanical Problem" scale consisted of eight items on a five-point Likert scale, ranging from Strongly Agree (5) to Strongly Disagree (1). Scores exceeding 3 indicated satisfactory knowledge, while scores below 3 indicated unsatisfactory knowledge, with scores closer to 4 indicating greater satisfaction and those closer to 2 indicating more dissatisfaction. Negative statements were interpreted by reversing responses to ensure uniform comparison. The essence of the closed-ended questionnaire was developed by (Ghanad, 2023) to establish statistical significance and show possible trends over some time.

2.4.2. Open-ended questionnaire

The open-ended questionnaire had two parts---A and B. The first part of Section B comprised ten (10) test items covering both mechanical and non-mechanical word problems. The first five items focused on mechanical, while the second five presented non-mechanical word problems. Each question in the mechanical problem category corresponded to an equivalent question in the non-mechanical word problems category (e.g., Q1 ↔ Q6, Q2 ↔ Q7, etc.). This means every mechanical item had a corresponding equivalent non-mechanical item to avoid different discrimination and difficulty levels. Students were expected to apply basic mathematical concepts to solve these problems, with the non-mechanical word problems requiring more critical thinking and problem-solving abilities.

The second part of Section B focused on the content of word problem-solving. Students were encouraged to utilize heuristic strategies to approach these problems. They required students to proceed in the order of Wallas's (1926) model. The problems were arranged in increasing levels of difficulty of preparation, incubation, illumination, and verification (Raslan, 2024). The essence of the open-ended questionnaire was to provide

quality information that helped to understand and interpret the qualitative behaviour of the model (Ghanad, 2023).

2.5. Reliability and validity

The validity of the instruments was established by reviewing two mathematics education experts to assess content, construct, and face components. These experts have had more than 10 years of teaching mathematics, published research papers in mathematics problem-solving, and reviewed journal articles on problem-solving strategies. They made revisions to the content of the mathematics tasks, the strategies to deploy to the content, and the participants to consider in the research. Following necessary revisions, the items were pilot-tested on a separate group of 50 students to evaluate reliability. Notably, the participants in the pilot group were not included in the actual study sample. Using the schools' nominal roll, 50 participants representing thirteen percent (13%) sample were randomly selected.

After obtaining approval from the school authorities, the questionnaires were distributed with the assistance of two teachers from the school. Participants were encouraged to complete the questionnaire honestly and sincerely. All distributed questionnaires were completed and returned within the specified timeframe.

In this study, rigorous research protocols were followed to minimize biases and errors, thereby enhancing the reliability of the findings.

The research design and execution were meticulously planned to ensure that if the study were replicated with different participants, respondents, or conditions, the results would remain consistent. This increased the reliability of the research outcomes. Even the pilot study showed a high level of internal consistency of Cronbach's alpha coefficient of 0.801 (Mohajan, 2017).

2.6. Method of Data Analysis

Research Question 1 used two tasks to show how the percentages of students who used mechanical and non-mechanical strategies. The characteristics of mechanical and non-mechanical problems were analysed by the way students responded to the questions. In each question, the researchers issued the problem, students provided their solutions, and the number of corrected responses was tabulated in percentages.

Research Question 2 used the Wallis-Kruskal Chi-square and Pearson's correlation coefficients to explore the students' relationships on age, gender, and programmes of study. This provided insights into the student's overall interactions with the word problems.

2.7. Ethical Consideration

The researchers obtained an introductory letter from the University and a consent form. The headmasters willingly agreed to the request and issued acceptance letters, thereby granting permission for data collection. These acceptance letters facilitated the commencement of data collection.

With consent from the Headmasters, the Heads of the Mathematics Departments of the participating schools received appropriate notification about the study. Data collection took place over one week at the three schools on different days. The research questionnaires were personally administered to the students, and responses were collected immediately upon completion. This approach ensured a 100% response rate.

The participants were not compelled to take part against their will, as neither the researcher nor the school administration enforced participation. While the teachers aiding the researcher did not receive direct benefits, the insights gleaned from the study could potentially inform their teaching approaches. No incentives, monetary or otherwise, were offered to participants to prevent any inadvertent influence on their decision to join.

To ensure anonymity, participants refrained from including their names on the questionnaires, which were assigned unique codes for easy tracking of mathematics performance. The data collected was exclusively used for the study's purposes and would be disposed of afterward. Participants' concerns were also prioritized, and they were given the option to halt questionnaire completion if they experienced any discomfort.

3. Results

The Results have been divided into Research Question 1 and Research Question 2. In Research Question 1, the researchers analysed five questions in Table 1, and five questions in Table 2 on students' performance in mechanical and non-mechanical word problems respectively. In Research Question 2, the researchers used the Kruskal-Wallis Correlation Coefficients to analyse the relationships between the types of problems and the three demographic variables.

3.1. Research Question 1: Performance in Mechanical Word Problems

In Table 1, scaffolded learning analysed intelligence, analysis, and tailoring learners' specific learning solutions and percentages of performance. With a minimum score of 80%, many students nearly scored items 1 and 3. The teacher employed contingency scaffolding responsive to the learner's needs and responses. The fading was just gradual until the transfer of responsibility occurred.

Table 1. Characteristics of mechanical word problems

Problem	Students' Solution	Percentage Right
1. What is the value of x?	The value of x is -7	315 (90%)
2. $Numerator\frac{k-1}{2}+3k=10$, what is the value of k?	The value of k is 21/4	280(80%)
3. $3(a-2)=a+8$, what is the value of a?	The value of 'a' is 14	333(95%)
4. $6(y-2)=5y$, what is the value of y?	The value of y is 12	288(85%)
5. $5(m+4)=35$, what is the value of m?	The value of m is 3	280(80%)

Table 2 provides a scaffolded learning model for determining how social engagement with more experienced teachers affected learners' intellectual development. Question 1 was very simple for the majority of the students to solve. However, Questions 2 to 5 were progressively made more complex. As the complexity increased, the teachers transferred the tasks to the students to successfully master the tasks. With a sample of 350 students, as low as 5% of the students got item 4, and the highest was item 5 (40%).

Table 2. Performance in Non-Mechanical Word Problems

Problem	Students' Solution	Performance
1. Four times two less than a number is the same as thrice one more than the number. Find the number.	4x-2 = x+3	105(30%)
2. A certain number is added to 12, and the result is divided by 2. If the answer is the same as subtracting the number from 4, what is the number?	(x+12)/2 = 4-x	53(15%)
3. When thrice a certain number is added to 3 and the sum is doubled, the result is 13 more than 5 times the original number. Find the original number.	2(3x+3) = 5(x+13)	70(20%)
4. In a particular community, there are 250 men more than women. The number of children is twice the number of women and the number of men is twice the number of women and children combined. How many are there?	(W+250)+2w+2(w+25)	18(5%)
5. Five times a certain number is added to 3 and the result is multiplied by 4. The final result is 72. What is the number?	4(5x+3) = 72	140(40%)

3.2. RQ2: Relationships between Mechanical and Non-Mechanical Problems

In this section, the relationship between students' solutions in mechanical and non-mechanical word problems was compared. The results can be found in Tables 1-14.

Null hypothesis (H_0): There is no significant relationship between student's mechanical and word problem-solving.

Table 3. Spearman's rank correlation between mechanical and non-mechanical problems

		Mechanical	Non-Mechanical
Mechanical	Correlation Coefficient	1.000	.590**
	Sig. (2-tailed)	.	.000
	N	350	348
Non-Mechanical	Correlation Coefficient	.590**	1.000
	Sig. (2-tailed)	.000	.
	N	348	348

Table 3 presents Spearman's rank correlation coefficient between mechanical and non-mechanical word problems. The coefficient of 0.590 reflects a moderate positive correlation, suggesting that students who excel in one type of problem-solving tend to perform well in the other as well. Furthermore, the low p-value of 0.000 indicates statistical significance at the 0.05 level (2-tailed), indicating a meaningful relationship between mechanical and word problem-solving abilities.

Consequently, we rejected the null hypothesis that *There is no significant relationship between student's mechanical and word problem-solving*. The correlation coefficient's magnitude and the p-value both affirm a statistically significant relationship between students' performances on mechanical and word problems. This provides compelling evidence for the existence of a significant and moderate positive relationship between mechanical and word problems. This implies that proficiency in one type of problem-solving tends to correspond with proficiency in the other.

H_{01}: There are NO relationships between students' ages

Table 4 presents an insightful breakdown of students' age groups. In the mechanical problems, it was evident that older students in the 22 and 23 age groups, exhibited higher mean ranks. Age 23 had the highest mean rank of 206.23, closely followed by age 23 with a mean rank of 206.38. Conversely, younger students (e.g. 20-year age group) demonstrated lower mean ranks, with age 20 recording the lowest mean rank of 141.00. This disparity suggested a positive correlation between age and mechanical problem-solving skills.

Table 4. Mean ranks of students' ages

Problems	Age	N	Mean Rank
Mechanical Problems	15	41	159.85
	16	53	185.76
	17	99	186.95
	18	49	168.22
	19	37	167.28
	20	23	141.00
	21	18	154.78
	22	18	198.86
	23	12	206.38
Non-Mechanical Problems	15	40	181.12
	16	53	178.47
	17	98	185.46
	18	49	172.35
	19	37	162.53
	20	23	133.78
	21	18	164.56
	22	18	153.22
	23	12	215.92

On the non-mechanical word problems, the 23-year age group showcased superior performance with the highest mean rank of 215.92, and age 17 obtained strong performance with a mean rank of 185.46. Again, the 20-year age group exhibited the lowest mean rank of 133.78. This pattern reaffirmed that age correlated positively with non-mechanical word problem-solving skills, with older students displaying a notable advantage over their younger peers.

Table 5. Kruskal-Wallis test for age groups

	Mechanical Problems	Non-Mechanical Problems
Chi-Square	9.37	8.86
df	8	8
Asymp. Sig.	.31	.35

The Kruskal-Wallis test was conducted to examine the relationships between mechanical and non-mechanical word problems across the different age groups. For the mechanical problems, the Chi-Square test statistic was 9.37 with 8 degrees of freedom. The associated *p*-value (Asymptotic. Sig.) is .312, which was greater than .05.

This indicated that there was no significant difference in mechanical problems across different age groups ($p > .05$).

Similarly, for the non-mechanical word problem, the Chi-Square test statistic was 8.863 with 8 degrees of freedom, and the associated *p*-value was .35. Thus, there were no significant relationships in non-mechanical word problem scores across different age groups ($p > .05$). Based on the Kruskal-Wallis test results, we failed to reject the null hypothesis for both mechanical and non-mechanical word problems.

H_{02}: There is NO relationship between male and female students

Table 6 displays the mean ranks of students' mechanical and word problem scores by gender. It was shown that the male students exhibited slightly higher mean ranks in mechanical problems (178.59) as compared to females (169.65). This suggests a modest performance advantage for male students in mechanical problem-solving tasks. In the non-mechanical word problems, male students also demonstrated had a slight edge (175.17,) as compared to female students (173.24), giving male students a marginally better performance than female students.

Table 6. Mean ranks of students' gender

	Gender	N	Mean Rank
Mechanical Problems	male	229	178.59
	female	121	169.65
	Total	350	
Non-Mechanical Problems	male	229	175.17
	female	121	173.24
	Total	350	

This notwithstanding, the discrepancies were relatively small. This implies a less pronounced gender gap in problem-solving abilities, particularly in non-mechanical word problems compared to mechanical problems.

Table 7. Kruskal-Wallis test for gender

	Mechanical Problems	Non-Mechanical Problems
Chi-Square	.658	.030
df	1	1
Asymp. Sig.	.417	.863
a. Kruskal Wallis Test		b. Grouping Variable: Gender

Table 7 presents the results of the Kruskal-Wallis test on gender. In the mechanical problem, the Chi-Square test statistic is 0.658 with 1 degree of freedom, resulting in an asymptotic significance level (Asymp. Sig.) of .417. Since .417 is greater than the typical significance level of .05, we failed to reject the null hypothesis and concluded that there was no significant difference in mechanical problems between male and female students.

Similarly, for the non-mechanical word problems, the Chi-Square test statistic is 0.030 with 1 degree of freedom, resulting in an asymptotic significance level of .863. Once again, .863 is greater than .05, indicating that there is no significant difference between male and female students. Based on the Kruskal-Wallis test results, we do not find evidence to reject the null hypothesis for either mechanical or non-mechanical word problems. This suggests that there were no significant relationships in problem-solving abilities between male and female students.

H_{03}: There are NO relationships among the programmes of study

Table 8 provides insights into students' problem-solving performance across different programmes of study. For mechanical problems, students in the General Science programme obtained the highest mean rank of 267.12, indicating the best performance among all programmes. This was followed by students in the General Agriculture programme with a mean rank of 191.20. Conversely, students in the Home Economics programme demonstrated the lowest mean rank of 138.37, suggesting comparatively weaker performance in mechanical problem-solving.

Table 8. Mean ranks of students' programmes of study

Mechanical	Gen. Arts	125	181.87
	Gen. Science	16	267.12
	Technical	26	143.94
	Gen. Agriculture	64	191.20
	Home Economics	59	138.37
	Business	60	171.23
	Total	350	
Non-Mechanical	Gen. Arts	123	171.01
	Gen. Science	18	231.72
	Technical	26	148.98
	Gen. Agriculture	64	184.31
	Home Economics	59	141.70
	Business	60	199.24
	Total	350	

In contrast, for non-mechanical word problem scores, students in the General Science programme still exhibited the highest mean rank of 231.72, maintaining its superiority over the other programs' performance. Students in the Business programme had the second-highest mean rank of 199.24, while the Home Economics programme stayed in the lowest mean rank of 141.70.

The analysis reveals varying levels of performance across different programmes of study. While students in the General Science programme tended to perform well in both mechanical and non-mechanical, students in the Home Economics programme showed relatively weaker performance in both groups. These findings underscore the importance of considering programme-specific factors in designing targeted interventions to enhance students' problem-solving abilities.

Table 9. Kruskal-Wallis test for mechanical and word problem scores by programme

	Mechanical Problems	Non-Mechanical Problems
Chi-Square	27.411	17.733
Df	5	5
Asymp. Sig.	.000	.003

Table 9 presents the results of the Kruskal-Wallis test conducted to assess whether there are significant differences in mechanical and non-mechanical word problems across different programmes of study. For mechanical problems, the Chi-Square test statistic was 27.411 with 5 degrees of freedom, resulting in an asymptotic significance level (Asymp. Sig.) of .000. We, therefore, rejected the null hypothesis, and concluded that there was a significant difference across different programmes of study.

Similarly, for non-mechanical word problem scores, the Chi-Square test statistic was 17.733 with 5 degrees of freedom, resulting in an asymptotic significance level of .003. The .003 indicated that there was a significant difference in non-mechanical word problems across different programmes.

In addition, the Kruskal-Wallis test results showed that we rejected the null hypothesis for both mechanical and non-mechanical problem scores. This suggested that there were significant differences in problem-solving abilities across different programmes of study. Some post-hoc tests were carried out to identify specific differences between specific programmes.

Table 10. Comparison of General Arts and Home Economics programmes

	Mechanical	Non-Mechanical
Mann-Whitney U	2809.000	3018.500
Wilcoxon W	4579.000	4788.500
Z	-2.675	-1.850
Asymp. Sig. (2-tailed)	.007	.064

Table 10 compares test statistics for mechanical and non-mechanical problems of a test using Mann-Whitney U tests. The higher values for non-mechanical problems indicated superior rankings compared to the mechanical problems. This suggested potential performance differences.

Again, the negative Z-scores for both groups (-2.675 and -1.850) indicated lower mean ranks for the mechanical group. The p-value for the mechanical problems (.007) was statistically significant, while the non-mechanical problems (.064) were not. Thus, there was a significant performance gap in the mechanical problems, whereas differences in the non-mechanical ones were less pronounced.

Table 11. Comparison of General Arts and General Science Programmes

	Mechanical Problems	Non-Mechanical Problems
Mann-Whitney U	520.000	638.500
Wilcoxon W	8395.000	8264.500
Z	-3.267	-2.296
Asymp. Sig. (2-tailed)	.001	.022

Table 11 compares the performance of General Arts and General Science Programmes in mechanical and non-mechanical problems using Mann-Whitney U tests. General Arts students scored a Mann-Whitney U of 520.000 for mechanical problems, and 638.500 for non-mechanical problems. The Z-scores (-3.267 and -2.296) indicated lower mean ranks for mechanical scores.

In addition, the Asymp. Sig. (2-tailed) values (.001 and .022) signified significant major gaps between the two programmes, particularly in mechanical. This underscored varied competencies across General Arts and General Science programmes, emphasizing the need for tailored instructional strategies to address disparities effectively.

Table 12. Comparison of General Science and Business programmes

	Mechanical Problems	Non-Mechanical Problems
Mann-Whitney U	190.000	380.500
Wilcoxon W	2020.000	2210.500
Z	-3.866	-1.276
Asymp. Sig. (2-tailed)	.000	.202

As seen in Table 12, the Mann-Whitney U test revealed significant differences in mechanical problems (U = 190.000, Z = -3.866, p = 0.000) but not in non-mechanical problems (U = 380.500, Z = -1.276, p = 0.202). Therefore, the null hypothesis was rejected for mechanical problems, indicating a significant difference in scores between the two programmes, while it was retained for non-mechanical problems.

Table 13. Comparison of General Science and Home Economics programmes

	Mechanical Problems	Non-Mechanical Problems
Mann-Whitney U	144.000	237.000
Wilcoxon W	1914.000	2007.000
Z	-4.354	-3.069
Asymp. Sig. (2-tailed)	.000	.002

The Mann-Whitney U test results for mechanical problems (U = 144.000, Z = -4.354, p = 0.000) and non-mechanical problems (U = 237.000, Z = -3.069, p = 0.002) suggested significant differences between both groups. Therefore, the null hypothesis was rejected to indicate significant differences in scores in the two Programmes.

Table 14. Comparison of General Science and Agricultural Science programmes

	Mechanical Problems	Non-Mechanical Word Problems
Mann-Whitney U	282.000	371.000
Wilcoxon W	2362.000	2451.000
Z	-2.958	-1.705
Asymp. Sig. (2-tailed)	.003	.088

In Table 14, the Mann-Whitney U test results for Mechanical (U = 282.000, Z = -2.958, p = 0.003) indicated a significant difference between groups, suggesting rejection of the null hypothesis. However, for non-mechanical (U = 371.000, Z = -1.705, p = 0.088), there appeared to be a difference, which was not significant at 0.05, leading to the retention of the null hypothesis.

Table 15. Comparison of General Science and Technical programmes

	Mechanical Problems	Non-Mechanical Problems
Mann-Whitney U	70.000	113.500
Wilcoxon W	421.000	464.500
Z	-3.750	-2.465
Asymp. Sig. (2-tailed)	.000	.014

The Mann-Whitney U test results in Table 13 for mechanical problems (U = 70.000, Z = -3.750, p = 0.000) indicated a significant difference between groups, warranting rejection of the null hypothesis. Similarly, for non-mechanical problems (U = 113.500, Z = -2.465, p = 0.014), although the difference was less pronounced, it remained statistically significant at 0.05, leading to the rejection of the null hypothesis.

Table 16. Comparison of Technical and Business programmes

	Mechanical Problems	Non-Mechanical Problems
Mann-Whitney U	646.000	567.000
Wilcoxon W	997.000	918.000
Z	-1.300	-2.018
Asymp. Sig. (2-tailed)	.194	.044

The Mann-Whitney U test results for mechanical problems (U = 646.000, Z = -1.300, p = 0.194) suggested no significant difference between the two programmes (see Table 14), leading to the retention of the null hypothesis. However, for non-mechanical problems (U = 567.000, Z = -2.018, p = 0.044), the results indicated a statistically significant difference at 0.05, prompting rejection of the null hypothesis.

4. Discussion

In Research Question 1, the Wallas (1926) information-processing model progressively scaffolded and provided students with access to the ZPD's contingency scaffolding in mechanical word problems. This gradually faded into the transfer of responsibility from mechanical to non-mechanical word problems (van Nooijen et al., 2024). An overwhelming number of students (80%) gathered information about the problems at the preparation and incubation stages in the mechanical word problems. However, the solutions at the illumination and verification stages at the transfer stage are very minimal. As low as 4o% of the students got the highest score (Savic, 2016; Voskoglou, 2021).

In Research Question 2, Spearman's rank correlation coefficient of 0.590 indicates a moderate positive correlation between students' mean scores on mechanical word problems. This suggests that proficiency in solving mechanical types of problems tended to correspond with proficiency in non-mechanical types. This emphasizes the transferability of skills (Raslan, 2024) from information gathering at the preparation stage to the solution at the illumination stage and checking of solutions at the verification stage (Savic, 2016; Voskoglou, 2021).

Additionally, age appears to correlate positively with word problems across both mechanical and word problem-solving tasks. Older students, particularly those in the 22 and 23 age groups, consistently demonstrate higher mean ranks compared to younger age groups. This finding underscores the developmental aspect of problem-solving skills, where maturity and experience likely contributed to improved proficiency (Bertoletti et al., 2023; Kaiser & Zhu, 2022). Taley and Ndamenenu (2022) found that word problems provide opportunities to contextualize environments without resorting to memorization. This suggests factors such as increased cognitive development, accumulated experience, and improved critical thinking skills with age likely contribute to these differences. Sarı and Yüce (2020) emphasize the role of cultural and socioeconomic factors in shaping problem-solving abilities, suggesting that older students may benefit from more extensive educational opportunities and resources.

Also, the analysis reveals no significant relationships between male and female students. While male students show slightly higher mean ranks in both problems, the gender gap was minimal and not statistically significant according to the Kruskal-Wallis test results. This suggests that gender does not play a significant role in determining word problem skills (Bertoletti et al., 2023; Kaiser & Zhu, 2022). These findings aligned with

Ahmad and Duskri (2018), and Di Tommaso et al. (2024), who contended that while boys often report higher confidence in their mathematical abilities than girls, this confidence gap can be mitigated through supportive and inclusive teaching practices. Supporting these findings, Bertoletti et al. (2023) debunked the myths that women are qualitative and men are quantitative. Kaiser and Zhu (2022) equally argued that differences in mathematical success between males and females are more related to attitudes and expectations of success rather than inherent abilities.

Furthermore, there were significant differences across various programmes of study, as revealed by the Kruskal-Wallis test. Students in the General Science programme consistently demonstrated superior performance in both mechanical and non-mechanical word problem tasks as compared to their peers in other programmes. In contrast, students in the Home Economics programme exhibited comparatively lower performance. These findings underscore the importance of considering programme-specific factors when designing interventions to enhance problem-solving abilities. Tailoring educational strategies to the unique needs and strengths of different programmes can help improve overall student outcomes.

Similarly, this study found that students in General Science and General Agriculture scored higher than their counterparts in other programmes, while those in Home Economics consistently performed lower (Adu et al., 2015). This suggests that curriculum and instructional approaches should be customized to leverage the strengths and address the weaknesses of students in different fields of study, ensuring that all students receive the support necessary to enhance their problem-solving abilities.

These findings highlight the importance of considering age, gender, and programme-specific factors when designing educational interventions. By addressing the unique challenges and leveraging the strengths associated with each demographic and academic group, educators can create more effective and equitable learning environments that promote mathematical proficiency and confidence among all students.

5. Conclusions

The Wallas (1926) model progressively scaffolded and provided students with access to the ZPD's contingency in scaffolding in mechanical word problems. As the students moved from simple to complex tasks, they utilized creativity to solve the problems. This was evident in the relationships between mechanical and non-mechanical problems (0.590). This suggests that students who excelled in creativity were likely to perform well in the tasks, emphasizing the interconnected nature of the model. The skill development in the interconnection between the stages was the by-product of mathematical creativity. The cogent processes and the innovations enshrined in the transitions were ample evidence of creativity.

Also, older students generally outperformed younger peers across both mechanical and word problem-solving tasks. Gender-based differences in performance were minimal, highlighting the need for educational strategies that cater to individual learning needs rather than relying on gender stereotypes. There were also statistically significant differences between the programmes, and this was more pronounced in General Science.

Again, there were significant differences in mechanical problems but not in non-mechanical problems for some programmes. This underscores the importance of tailored educational approaches to meet the specific needs and competencies of students in each programme.

6. Recommendations

In Research Question 1, it was found that the students gradually transferred their responsibility from mechanical to non-mechanical word problems. It was recommended that teachers design more programme-specific interventions to be implemented using the Wallas model. This will consolidate creativity in solving mathematics word problems.

In Research Question 2, Spearman's rank correlation coefficients showed moderate positive correlations. It was recommended that stakeholders should design interventions tailored to the specific needs of students' ages, gender, and academic programmes. This will enhance problem-solving abilities within disciplinary contexts to optimize learning outcomes.

Funding: The study received no funding support.

Ethics Declaration: All participants into the current study were given a consent form and asked to opt in whether they wanted to participate. No additional ethical approval was ¬required to conduct the study.

Data Availability: The corresponding author has access to the data supporting the findings of this study upon request.

Conflict of Interest: The author affirms that there were no conflicts of interest associated with the study.

References

Adu, E., Assuah, C. K., & Asiedu-Addo, S. K. (2015). Students' errors in solving linear equation word problems: Case study of a Ghanaian senior high school. *African Journal of Educational Studies in Mathematics and Sciences*, 11, 17-30.

Ahmad, A., & Duskri, M. (2018). Gender differences of mathematical critical thinking skills of secondary school students. *The 6th South East Asia Design Research International Conference* (6th SEA-DR IC), 1–6. https://doi.org/10.1088/1742-6596/1088/1/012054

Ali, C. A., & Anderson, H. K. (2021). Pre-service teachers' pedagogical content knowledge in transferring from basic musical notations to basic fractions. *Journal of STEM Education: Innovations and Research (JSTEM), 22*(2), 46-50.

Bertoletti, A., Cannistrà, M., Lema, M.D., Chiara Masci, M., Mergoni, A., Rossi, L., & Soncin, M. (2023). The Determinants of Mathematics Achievement: A Gender Perspective Using Multilevel Random Forest. *Economies 11:* 32. https://doi.org/10.3390/economies11020032

Di Tommaso, M. L., Contini, D., De Rosa, D., Ferrara, F., Piazzalunga, D., & Robutti, O. (2024). Tackling the gender gap in mathematics with active learning methodologies. *Economics of Education Review, 100*, 102538. https://doi.org/10.1016/j.econedurev.2024.102538

Emanuel, E. P. L., Kirana, A., & Chamidah, A. (2021). Enhancing students' ability to solve word problems in Mathematics. *Journal of Physics. Conference Series, 1832*(1), 012056. https://doi.org/10.1088/1742-6596/1832/1/012056

Ghanad, A. (2023). An Overview of Quantitative Research Methods. *International Journal of Multidisciplinary Research and Analysis, 6*(8), 3794-3803. https://doi.org/10.47191/ijmra/v6-i8-52

Haavold, P. O., & Sriraman, B. (2021). Creativity in problem-solving: integrating two different views of insight. *ZDM-The International Journal of Mathematics Education, 54*(5), 1-14. https://doi.org/10.1007/s11858-021-01304-8

Hirashima, T. (2015). *Increasing Problem Simplification as Scaffolding in Exercises*. Hiroshima University. Retrieved from https://www.researchgate.net/publication/266413339

Kaiser, G & Zhu, Y. (2022). Gender differences in mathematics achievement: A secondary analysis of Programme for International Student Assessment data from Shanghai. *Asian Journal for Mathematics Education*, *1*(1), 115–130.

Khasawneh, E., Gosling, C. & Williams, B. (2021). What impact does maths anxiety have on university students?. *BMC Psychol 9*, 37. https://doi.org/10.1186/s40359-021-00537-2

Malik, S. A., & Wiseman, A. W. (2017). Revisiting and re-representing scaffolding: The two gradient model. *Cogent Education, 4*(1), 1-15 https://doi.org/10.1080/2331186X.2017.1331533

Matzin, E. S. & Mundia, L. (2020). Efficacy of the Bar Model Method of Teaching Mathematics to Year 7 Students: Case Study of Teachers in Brunei Darussalam. *Journal of Educational Issues, 6* (1), 402-421. https://doi.org/10.5296/jei.v6i1.17154

Nazzal, L. J., & Kaufman, J. C. (2020). The relationship of the quality of creative problem-solving stages to overall creativity in engineering students. *Thinking Skills and Creativity, 38* (1), 1-16. https://doi.org/10.1016/j.tsc.2020.100734

Qetrani, S., Ouailal, S., & Achtaich, N. (2021). Enhancing Students' Conceptual and Procedural Knowledge Using a New Teaching Approach of Linear Equations Based on the Equivalence Concept. *Eurasia Journal of Mathematics Science and Technology Education, 17*(7), em1978. https://doi.org/10.29333/ejmste/10938

Raslan, G. (2024). The Impact of the Zone of Proximal Development Concept (Scaffolding) on the Students Problem Solving Skills and Learning Outcomes. In: Al Marri, K., Mir, F.A., David, S.A., Al-Emran, M. (eds). *BUiD Doctoral Research Conference 2023. Lecture Notes in Civil Engineering*, vol 473. Springer, Cham. https://doi.org/10.1007/978-3-031-56121-4_6

Rembert, D. M., Mack, N. A., & Gilbert, J. E. (2019). Exploring the needs and interests of fifth graders for personalized math word problem generation. *Proceedings of the 18th ACM International Conference on Interaction Design and Children*, IDC 2019, 592–597. https://doi.org/10.1145/3311927.3325309

Sadler-Smith, E. (2015). Wallas' Four-Stage Model of the Creative Process: More Than Meets the Eye? *Creativity Research Journal 27*(4):342-352. https://doi.org/10.1080/10400419.2015.1087277

Sarı, M. H., & Yüce, E. (2020). Problems Experienced in Classrooms with Students from Different Cultures. *Journal on Efficiency and Responsibility in Education and Science, 13*(2), 90–100. https://doi.org/10.7160/eriesj.2020.130204

Savic, M. (2016). Mathematical Problem-Solving via Wallas' Four Stages of Creativity: Implications for the Undergraduate Classroom. *The Mathematics Enthusiast, 13* (3), 255-278. https://doi.org/10.54870/1551-3440.1377

Shukla, S. (2020). *Research Methodology and Statistics*. Ahmedabad: Rishit Publications. Retrieved from https://www.researchgate.net/publication/346426707_CONCEPT_OF_POPULATION_AND_SAMPLE

Taley, B. I. (2022). Teacher and student views of mathematics word problem-solving task at senior high school level. *FNAS Journal of Mathematics and Science Education, 3*(2), 39-49. https://doi.org/10.1016/0364-0213(88)90023-7

Vale, I., & Barbosa, A. (2023). Active learning strategies for an effective mathematics teaching and learning. *European Journal of Science and Mathematics Education, 11*(3), 573-588. https://doi.org/10.30935/scimath/13135

van Nooijen, C.C.A., de Koning, B.B., Bramer, W.M. et al. (2024). A Cognitive Load Theory Approach to Understanding Expert Scaffolding of Visual Problem-Solving Tasks: A Scoping Review. *Educational Psychology Review, 36* (12), 1-42. https://doi.org/10.1007/s10648-024-09848-3

Voskoglou, M. (2021). Problem Solving and Mathematical Modelling. *American Journal of Educational Research, 9 (1),* 85-90. https://doi.org/10.12691/education-9-2-6

Wakhata, R., Balimuttajjo, S., & Mutarutinya, V. (2024). Relationship between students' attitude towards, and performance in mathematics word problems. *PloS One, 19*(2), e0278593. https://doi.org/10.1371/journal.pone.027859

Wallas, G. (1926). *The Art of Thought*. New York: Harcourt, Brace & Company. Retrieved from https://search.worldcat.org/formats-editions/1114115