APPENDIX B

Pre-Intervention Analysis

To ensure the comparability between groups before implementing the Realistic Mathematics Education (RME) approach, a pre-test was administered to both the experimental class (12A2) and the control class (12A3). The objective was to determine whether students in both classes demonstrated equivalent academic proficiency prior to the intervention.

Descriptive Statistics

Descriptives

	Class			Statistic	Std. Error
Grade	12A2	Mean		8.650	.1991
		95% Confidence Interval for Mean	Lower Bound	8.248	
			Upper Bound	9.052	
		5% Trimmed Mean		8.767	
		Median		9.150	
		Variance		1.665	
		Std. Deviation		1.2902	
		Minimum		4.8	
		Maximum		10.0	
		Range		5.2	
		Interquartile Range		1.5	
		Skewness		-1.376	.365
		Kurtosis		1.308	.717
	12A3	Mean		8.593	.1735
		95% Confidence Interval for Mean	Lower Bound	8.242	
			Upper Bound	8.943	
		5% Trimmed Mean		8.658	
		Median		8.900	
		Variance		1.204	
		Std. Deviation		1.0974	
		Minimum		5.5	
		Maximum		10.0	
		Range		4.5	
		Interquartile Range		1.5	
		Skewness		-.818	.374
		Kurtosis		.355	.733

Indicator	Class 12A2 (Experimental)	Class 12A3 (Control)
Mean	8.650	8.593
Median	9.150	8.900
Standard Deviation	1.2902	1.0974
Score Range	4.8 – 10.0	5.5 – 10.0
Skewness	–1.376	–0.818
Kurtosis	1.308	0.355

The descriptive statistics indicate that the mean, median, and variability between the two classes were relatively similar. Both distributions were slightly left-skewed, and the skewness and kurtosis values were within acceptable ranges.

Test of Normality

Normality was assessed using the Kolmogorov–Smirnov and Shapiro–Wilk tests:

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Class	Statistic	df	Sig.	Statistic	df	Sig.
Grade	12A2	.250	42	<.001	.842	42	<.001
	12A3	.145	40	.034	.927	40	.013

a. Lilliefors Significance Correction

Class	K–S (Sig.)	Shapiro–Wilk (Sig.)
12A2	0.034	0.013
12A3	< 0.001	< 0.001

All p-values were below 0.05, indicating that the data in both groups did not follow a normal distribution. Thus, a non-parametric test was required for further comparison.

Mann–Whitney U Test

Ranks

	Class	N	Mean Rank	Sum of Ranks
Grade	12A2	42	43.18	1813.50
	12A3	40	39.74	1589.50
	Total	82		

Test Statistics[a]

	Grade
Mann-Whitney U	769.500
Wilcoxon W	1589.500
Z	-.657
Asymp. Sig. (2-tailed)	.511

a. Grouping Variable: Class

Class	N	Mean Rank	Sum of Ranks
12A2	42	43.18	1813.50
12A3	40	39.74	1589.50

- Mann–Whitney U = 769.500
- Z = –0.657
- Sig. (2-tailed) = 0.511

The p-value greater than 0.05 indicates no statistically significant difference between the pre-intervention scores of the two classes.

Interpretation

The results suggest that students in both the experimental (12A2) and control (12A3) classes had comparable academic performance prior to the RME intervention. This baseline equivalence enhances the reliability and objectivity of subsequent intervention analyses.

Post-Intervention Analysis

After Class 12A2 received instruction based on the Realistic Mathematics Education approach, a post-test was administered to both groups to assess the impact of the intervention.

Descriptive Statistics

Descriptives

	Class			Statistic	Std. Error
Grade	12A2	Mean		8.693	.1693
		95% Confidence Interval for Mean	Lower Bound	8.351	
			Upper Bound	9.035	
		5% Trimmed Mean		8.769	
		Median		9.000	
		Variance		1.204	
		Std. Deviation		1.0971	
		Minimum		5.5	
		Maximum		10.0	
		Range		4.5	
		Interquartile Range		1.5	
		Skewness		-.980	.365
		Kurtosis		.604	.717
	12A3	Mean		7.625	.2198
		95% Confidence Interval for Mean	Lower Bound	7.180	
			Upper Bound	8.070	
		5% Trimmed Mean		7.639	
		Median		8.000	
		Variance		1.933	
		Std. Deviation		1.3902	
		Minimum		5.0	
		Maximum		10.0	
		Range		5.0	
		Interquartile Range		2.8	
		Skewness		-.001	.374
		Kurtosis		-.807	.733

Indicator	12A2 (Experimental)	12A3 (Control)
Mean	8.693	7.625
95% Confidence Interval	[8.351 – 9.035]	[7.180 – 8.070]
Median	9.000	8.000
Standard Deviation	1.0971	1.3902

http://doi.org/10.22342/jme.v16i2.pp603-632

Skewness	–0.980	–0.001
Kurtosis	0.604	–0.807
Min – Max	5.5 – 10.0	5.0 – 10.0
Interquartile Range	1.5	2.2

The results show that Class 12A2 had a higher mean, median, and more consistent performance than Class 12A3. The experimental class distribution was more concentrated and slightly left-skewed.

Test of Normality

Normality tests again indicated that neither class followed a normal distribution:

Class	Kolmogorov–Smirnov (p)	Shapiro–Wilk (p)
12A2	0.041	0.003
12A3	0.015	0.027

Therefore, the Mann–Whitney U test was used for comparison.

Mann–Whitney U Test

Ranks

	Class	N	Mean Rank	Sum of Ranks
Grade	12A2	42	50.14	2106.00
	12A3	40	32.42	1297.00
	Total	82		

Test Statistics[a]

	Grade
Mann-Whitney U	477.000
Wilcoxon W	1297.000
Z	-3.386
Asymp. Sig. (2-tailed)	<.001

a. Grouping Variable: Class

Class	N	Mean Rank	Sum of Ranks
12A2	42	50.14	2106.00
12A3	40	32.42	1297.00

- Mann–Whitney U = 477.000
- Z = –3.386
- Asymptotic Sig. (2-tailed) < 0.001

The p-value below 0.001 indicates a statistically significant difference in performance between the two groups, with the experimental group outperforming the control group.

Interpretation

The post-intervention analysis confirms that the RME method led to significantly better academic outcomes compared to traditional instruction. This provides strong empirical support for the effectiveness of RME in enhancing mathematics performance at the secondary level.

One-Month Follow-Up Analysis

To evaluate the sustainability of the RME approach, a follow-up test was administered to the experimental class (12A2) one month after the intervention. Scores immediately after the intervention were labeled T1, and scores one month later were labeled T2. The difference variable was defined as D = T1 – T2.

Descriptive Statistics

Descriptives

			Statistic	Std. Error
D	Mean		.0667	.04631
	95% Confidence Interval for Mean	Lower Bound	-.0269	
		Upper Bound	.1602	
	5% Trimmed Mean		.0368	
	Median		.0000	
	Variance		.090	
	Std. Deviation		.30014	
	Minimum		-.30	
	Maximum		1.50	
	Range		1.80	
	Interquartile Range		.32	
	Skewness		2.680	.365
	Kurtosis		11.920	.717

Indicator	Value
Mean (D)	0.0667
Median	0.0000
Standard Deviation	0.3001
95% Confidence Interval (Mean)	[–0.0269 ; 0.1602]
Skewness	2.680
Kurtosis	11.920
Minimum – Maximum	–0.30 – 1.50
Interquartile Range	0.36

Although the average D value was positive, indicating a slight decline in scores after one month, the median was zero, suggesting that most students did not experience a meaningful change. However, the high skewness and kurtosis indicated non-normality.

Test of Normality

Normality tests confirmed that the distribution of D was not normal:

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
D	.171	42	.003	.762	42	<.001

a. Lilliefors Significance Correction

Test Type	p-value
Kolmogorov–Smirnov	0.003
Shapiro–Wilk	< 0.001

Consequently, the Wilcoxon Signed-Rank Test was used.

Wilcoxon Signed-Rank Test

Ranks

		N	Mean Rank	Sum of Ranks
T2 - T1	Negative Ranks	20[a]	20.60	412.00
	Positive Ranks	16[b]	15.88	254.00
	Ties	6[c]		
	Total	42		

a. T2 < T1

b. T2 > T1

c. T2 = T1

Test Statistics[a]

	T2 - T1
Z	-1.260[b]
Asymp. Sig. (2-tailed)	.208

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Comparison	N	Mean Rank	Sum of Ranks
T2 < T1 (Score decrease)	20	20.60	412.00
T2 > T1 (Score increase)	16	15.88	254.00
T2 = T1 (No change)	6	–	–

- Z = –1.260
- Asymptotic Sig. (2-tailed) = 0.208

With a p-value greater than 0.05, the test indicates no statistically significant difference between the two time points.

Interpretation

Despite minor individual variations, there was no significant overall change in student scores one month after the intervention. These findings suggest that the effects of RME were sustained over time, contributing to long-term retention of learning.