INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 9

Assessment of the Senior High School Students' Performance in Science Using the Multimodal Text

DOI: 10.5281/zenodo.24579446

REBA B. LORCA
rebalorca@gmail.com
Cebu Technological University, Main Campus, R. Palma St., Cebu City
Patao National High School – Deped Cebu Province

Abstract:

This study assessed the effectiveness of multimodal text in improving the science performance of Grade 11 students in Earth and Life Science, focusing on geological processes, namely earthquakes, volcanic eruptions, and landslides, at Patao National High School, Bantayan, Cebu, during School Year 2024–2025. A quasi-experimental pre-test–post-test control group design was employed with two intact classes: the control group (Grade 11–Mars A, n = 40) taught through conventional chalk-and-talk instruction, and the experimental group (Grade 11–Jupiter B, n = 38) taught through an Enhanced Learning Packet using multimodal text strategies that integrated diagrams, photographs, hazard maps, infographics, videos, simulations, written explanations, and interactive activities. Both groups answered the same researcher-made 30-item multiple-choice test, with 10 items per competency, before and after instruction, and the experimental group answered a perception questionnaire on difficulty and agreement. At pre-test, both groups performed at the Did Not Meet Expectation level, with overall means of 6.13 for the control group and 6.57 for the experimental group, and no significant difference was found across the three competencies ($p > .05$). At post-test, the control group improved modestly to 7.27, still at Did Not Meet Expectation, whereas the experimental group rose to 8.75, at the Very Satisfactory level. Independent-samples t-tests showed highly significant differences in favour of the experimental group for all three competencies ($p < .001$). The experimental group also rated all three topics as easy to understand (overall weighted mean = 1.89 on a 2-point scale) and agreed that they enjoyed learning, understood geological processes, and that their teacher used visual, written, and combined multimodal texts (overall weighted means of 3.06 and 3.14 on a 4-point scale). The findings support multimodal text as an effective, inclusive, and engaging approach for teaching abstract geological concepts and provide an evidence base for the Enhanced Learning Packet in Geological Processes.

Keywords: multimodal text; science performance; Grade 11 students; geological processes; Earth and Life Science; Enhanced Learning Packet

Introduction:

Science learning becomes meaningful only when learners can connect abstract ideas to concrete, observable phenomena. In Earth and Life Science, that connection is especially difficult because geological processes unfold over scales of time, space, and force that cannot be directly observed in a classroom. An earthquake lasts seconds but is driven by stresses accumulated over decades; a volcanic eruption is preceded by subtle ground deformation and gas emissions; a landslide may be triggered by a single night of rain acting on a slope weakened over years by deforestation, quarrying, or poor land use. When these processes are taught only through words on a board or in a module, learners must imagine what they have never seen, and many are left with fragmented recall rather than usable understanding. This study starts from that practical classroom problem: how can a teacher make invisible, rapid, large-scale, and hazardous processes visible, intelligible, and memorable for Grade 11 learners?

At Patao National High School in Guinlakgan, Patao, Bantayan, Cebu, the problem was concrete and urgent. The school serves a coastal island community exposed to earthquakes, typhoons, and related hazards, so geological literacy is not merely an academic requirement; it is a life skill. Yet teachers observed that learners struggled to describe hazards, to read hazard maps, and to explain how human activities speed up or trigger landslides. These are the three competencies prescribed by the Most Essential Learning Competencies (MELCs) for the first quarter of Earth and

Life Science, and they demand more than memorisation. Describing hazards requires vocabulary tied to images of ground rupture, liquefaction, tsunami, lahar, tephra fall, and slope failure. Using hazard maps requires spatial reading: colour, legend, scale, proximity, and exposure. Explaining human triggers requires causal reasoning that links deforestation, mining, construction, and waste disposal to slope stability. A single mode of presentation, spoken explanation or plain text, is poorly matched to all three demands at once.

Multimodal text offers a principled response. Following Jancsary, Höllerer, and Meyer (2016), multimodal text is meaning-making through the combination of verbal language and other semiotic resources such as images, layout, colour, typography, symbols, gesture, and sound. In science education, Alvermann and Wilson (2011) argued that learners must interpret multiple sign systems to construct meaning about a phenomenon.

A hazard map beside a short explanatory paragraph, a photograph of a lahar beside a labelled diagram of a volcano, or a short simulation of ground shaking beside a definition of liquefaction does not merely decorate the lesson; it distributes meaning across modes so that each mode does work the others cannot do. The visual shows form and spatial relations; the written text names, sequences, and explains cause; the video or simulation shows process over time. The New London Group (2000) and Cope and Kalantzis (2009) called this a pedagogy of multiliteracies, in which learners design meaning by combining available resources rather than passively receiving a single authoritative text.

Three theories explain why this combination should help. First, the theory of multimodality, articulated by Kress and van Leeuwen (2001) and Bezemer and Kress (2008), holds that a mode is a socially and culturally shaped resource for meaning-making, and that learning is more effective when learners engage with multiple forms of representation rather than relying solely on written language. Second, Paivio's (1986) dual coding theory holds that verbal and nonverbal systems are distinct but interconnected, so information encoded in both words and images is more likely to be remembered and retrieved than information encoded in words alone. Third, Gilbert's (2008) work on visualisation distinguishes internal representations, the mental images and frameworks learners construct, from external representations, the diagrams, models, and simulations teachers provide for shared inspection. Well-designed external representations scaffold the internal ones. Together, these theories predict that multimodal instruction should improve comprehension, engagement, and retention, particularly for abstract science topics, provided the modes are carefully selected, sequenced, and explicitly linked to learning objectives rather than added as decoration.

Empirical work supports that prediction but also warns against assuming that any use of visuals is enough. Cabual (2021) and Annamalai and Sazan (2025) reported that most learners prefer visual and hands-on experiences that sustain attention and improve comprehension. Freeman et al. (2014) showed, in a large meta-analysis, that active learning increases performance in science, engineering, and mathematics. Hansen and Richland (2020) emphasised that learners must attend to relevant information and connect ideas across representations, which requires teacher guidance, not merely exposure. Klein et al. (2020) found that mobile devices, video analysis, and multiple representations could reduce unnecessary cognitive load and improve conceptual understanding in physics when used in inquiry.

Tang (2023) showed that written language and diagrams in textbooks complement each other when they are integrated, and Danielsson et al. (2023) showed that primary science learners use speech, writing, images, and actions together to represent content. The gap this study addresses is local and practical: evidence was still needed on whether a deliberately designed multimodal packet, aligned to the MELCs and to the realities of a Philippine public school classroom, would raise Grade 11 performance on geological processes in Bantayan, Cebu, and whether learners themselves would experience the topics as easier and more engaging.

Accordingly, this study assessed the effectiveness of integrating multimodal text in teaching geological processes to Grade 11 students at Patao National High School during School Year 2024–2025. It compared a control group taught by conventional instruction with an experimental group taught through an Enhanced Learning Packet using multimodal text strategies. It asked whether the groups differed at pre-test and post-test, whether the difference was statistically significant, how difficult learners perceived the three topics to be, and how far they agreed that multimodal strategies were used and helped them learn.

The ultimate output is an Enhanced Learning Packet in Geological Processes that other Earth and Life Science teachers can adapt. The study is significant for learners, who gain multiple entry points into difficult content; for teachers, who gain a tested packet and a clearer rationale for selecting modes; for school leaders, who gain evidence for resourcing visual and digital materials; and for future researchers, who gain a baseline, an instrument, and a design that can be replicated, extended, or tested with delayed post-tests, larger samples, and other topics.

Literature Review:

Multimodality as a theory of meaning-making. The starting point for this study is the claim that communication and learning are never carried by words alone. Kress and van Leeuwen (2001) described multimodal discourse as the interplay of modes and media in contemporary communication, while Bezemer and Kress (2008) defined a mode as a socially and culturally shaped resource for meaning-making. In that view, a hazard map, a photograph of a volcanic eruption, a colour-coded legend, a gesture tracing a fault, and a written definition are not redundant illustrations of the same message; each contributes a distinct kind of meaning.

Siegel (2006) and Hassett and Curwood (2009) added that learners, especially children and adolescents, are inherently multimodal: they make meaning through talk, gesture, drama, drawing, and the integration of linguistic signs. Campbell and Olteanu (2023) extended this argument to a postdigital age, arguing for a literacy framework that integrates digital and non-digital experiences rather than treating print as the default. For science teaching, this matters because scientific meaning itself is multimodal: a geologist reads a map, interprets a seismograph, narrates a process, and draws a cross-section in the same act of explanation. Teaching that withholds all but the verbal mode therefore withholds part of the science itself.

Dual coding, cognitive load, and multiple representations. Paivio (1986) proposed that human cognition comprises verbal and nonverbal systems that can operate independently but are most powerful when referentially connected. Mayer (2009, 2014) developed this into a cognitive theory of multimedia learning, with principles that are directly relevant to packet design: combine words and pictures, place labels near the graphics they describe, remove extraneous material, and segment complex processes into learner-paced steps. Gilbert (2008) distinguished internal from external representations and argued that visualisation is a metacognitive skill: learners must be taught not only to look at diagrams but to interrogate them, to translate between a diagram, a photograph, and a written explanation, and to check whether their mental image matches the evidence.

Hansen and Richland (2020) showed that multiple representations help only when learners attend to relevant information and connect ideas across forms, which implies that a packet must do more than present many modes; it must cue attention, for example by asking learners to locate a hazard on a map, label a diagram, or explain a video still in their own words. Klein et al. (2020) found that mobile video analysis and multiple representations could reduce extraneous cognitive load in physics inquiry, a finding that supports short, focused simulations of earthquakes, landslides, and eruptions rather than long, undifferentiated videos.

Multimodal texts in science classrooms. The New London Group (2000) and Cope and Kalantzis (2009) argued that a pedagogy of multiliteracies should help learners move between available designs, designing, and redesigned meanings. In science, that movement is visible when a learner reads a short text about lahar, watches a video of a lahar after rainfall, traces the lahar path on a hazard map, and then writes a safety advisory for a community. Alvermann and Wilson (2011) emphasised that comprehension of multimodal science texts requires strategies for interpreting multiple sign systems, not merely decoding words. Unsworth et al. (2019) showed that multimodal literacy in science education narrows the gap between learners' diverse prior knowledge and the demands of scientific discourse, because photographs, diagrams, and written supports provide alternative routes into the same concept. Tang (2023) analysed junior high school textbooks and found that diagrams and written language function together in scientific explanations, with diagrams often carrying spatial and causal information that prose carries only with difficulty. Danielsson et al. (2023) combined long and short timescales to show how primary learners used speech, writing, images, and actions to represent science content over time, reinforcing the idea that learning is distributed across modes and moments, not contained in a single lesson or text.

Visual representations, diagrams, and hazard maps. Evagorou, Erduran, and Mäntylä (2015) argued that visual representations engage learners not only cognitively but also affectively, helping them appreciate the nature of science. Tomas, Mills, and Lewthwaite (2018) found that student-constructed animation about plate tectonics improved middle school learning, suggesting that production, not only reception, deepens understanding, which is why the present packet includes drawing, poster-making, and map-reading tasks. Sudaria and Sanchez (2023) reported that visual aids and formative assessment improved comprehension of earthquakes among Philippine learners, a finding close to the present context.

Hazard maps deserve special attention because they are a distinctive multimodal genre: colour, symbol, legend, scale, and place names must be read together. A learner who can decode a hazard map can answer a practical question, 'Is our barangay exposed, and where should we evacuate?' That is a different competence from recalling a definition, and it is one that conventional chalk-and-talk instruction, which rarely puts an authentic map in learners' hands, often fails to develop. Dewi, Nugroho, and Prasetyo (2023) argued that integrating risk perception into science education enhances preparedness for geological hazards, linking classroom learning to community resilience, a link that is salient in an island municipality such as Bantayan.

Learner preferences, engagement, and active learning. A consistent finding in the Philippine and international literature is that learners prefer visual and hands-on experiences. Cabual (2021) reported that students prefer visual and hands-on learning to sustain attention and improve comprehension, and Annamalai and Sazan (2025) reported similar preferences for multimodal learning in science classrooms. Yeo and Nielsen (2020) framed multimodal science teaching and learning as a field in its own right, not a supplement to 'real' science teaching. Freeman et al. (2014) meta-analysed 225 studies and found that active learning increased performance in science, engineering, and mathematics and reduced failure rates compared with lecturing alone. That finding cautions against a design in which learners merely watch videos; the present packet therefore pairs each multimodal input with an activity, eliciting prior knowledge, engaging with photographs, explaining a process, elaborating through a case, and evaluating through a test and a creative task. Jaimes, Libarkin, and Conrad (2020) and Libarkin (2008) reminded researchers that learners hold persistent conceptions about Earth and life over time, so a pre-test is not a formality; it reveals the conceptions that instruction must confront, for example the idea that all earthquakes produce tsunamis, or that volcanic eruption is the only trigger of lahar.

Synthesis and research gap. Taken together, the literature supports three propositions that guided this study. First, meaning in science is multimodal, and instruction that uses only one mode impoverishes the concept. Second, multiple representations help when they are selected, sequenced, and explicitly linked to objectives and activities, and when cognitive load is managed. Third, active, multimodal learning is preferred by learners and is associated with higher performance and engagement.

What remained uncertain was whether a low-cost, teacher-made Enhanced Learning Packet, using locally available photographs, maps, videos, and print materials and aligned to the MELCs for geological processes, would produce measurable gains for Grade 11 learners in a Philippine public school, and whether learners would perceive the three hazard topics as easier after experiencing it. That is the gap this study fills. It does not claim that multimodal text is a panacea, nor that technology alone improves learning. It tests a specific, documented intervention, on specific competencies, with a specific population, and reports both achievement and perception so that the packet can be judged, adapted, and improved.

Methodology:

Research design

This study employed a quasi-experimental pre-test–post-test control group design. The design was chosen because the study used two intact Grade 11 classes rather than randomly assigning individual learners to groups, preserving normal class organisation and the regular school setting while still permitting a comparison of outcomes. The control group received conventional instruction; the experimental group received instruction through the Enhanced Learning Packet using multimodal text strategies. Both groups were taught the same three MELC competencies over the same first-quarter period, took the same pre-test before instruction and the same post-test after instruction, and were compared

within and between groups. The perception questionnaire was administered to the experimental group after the intervention. The design is appropriate for a school-based effectiveness question, but it carries the usual quasi-experimental limitation that intact classes may differ in unmeasured ways despite similar pre-test performance, and that the teacher-researcher effect cannot be fully separated from the packet effect.

Research environment and respondents

The study was conducted at Patao National High School, a public secondary school in Sitio Guinlakgan, Barangay Patao, Municipality of Bantayan, Cebu, established in 2005 to bring accessible secondary education closer to Patao and nearby communities. At the time of the study the school enrolled 1,486 learners, 1,051 in Junior High School and 435 in Senior High School, with 256 in Grade 11 and 179 in Grade 12. The respondents were two heterogeneous Grade 11 sections in the first semester of School Year 2024–2025: Grade 11–Mars A (morning session, n = 40) as the control group and Grade 11–Jupiter B (afternoon session, n = 38) as the experimental group, for a total of 78 learners. The distribution by sex was balanced: the control group had 20 males and 20 females, and the experimental group had 20 males and 18 females, with 40 males and 38 females overall. No sampling was applied beyond the selection of the two intact sections; all enrolled learners in those sections participated as part of regular instruction, with permission from school authorities.

Research instruments

Two instruments were used. First, a researcher-made 30-item multiple-choice achievement test, with 10 items per competency, was adapted from the first-edition (2020) module used by the researcher and aligned to a Table of Specifications: 60% knowledge, 30% process or skills, and 10% understanding, covering 10 teaching hours on geological processes. The three competencies were: (1) describing the various hazards that may happen in the event of earthquakes, volcanic eruptions, and landslides; (2) using hazard maps, identifying areas prone to hazards brought about by earthquakes, volcanic eruptions, and landslides; and (3) identifying human activities that speed up or trigger landslides. The test was dry-run with Grade 12 learners in the same school who had previously studied the topic under the spiral curriculum, and items were reviewed for content validity, clarity, and alignment before use. Scoring followed DepEd performance-rating bands: 90–100% (9.0–10.0) Outstanding; 85–89% (8.5–8.9) Very Satisfactory; 80–84% (8.0–8.4) Satisfactory; 75–79% (7.5–7.9) Fairly Satisfactory; and below 75% (below 7.5) Did Not Meet Expectation, applied per 10-item competency subtest. Second, a perception questionnaire used a 4-point Likert scale for agreement (4 = Strongly Agree, 3 = Agree, 2 = Disagree, 1 = Strongly Disagree; with 3.26–4.00 Strongly Agree, 2.51–3.25 Agree, 1.76–2.50 Disagree, 1.00–1.75 Strongly Disagree) and a 2-point scale for perceived difficulty (2 = Easy, 1 = Difficult; 1.50–2.00 Easy to Understand, 1.00–1.49 Difficult to Understand). The questionnaire asked learners to rate difficulty for each of the three topics, to rate their own enjoyment, understanding, and independent reading or research, and to rate their teacher's use of visual texts, written texts alongside visuals, combined text types, and multimodal strategies. The packet itself included an 'I Never Forget' reflection and a poster task with a rubric on content, design and layout, use of graphics, and creativity.

Intervention and procedure

The research flow moved from input through process to output. Input comprised topic selection, literature review, and adviser confirmation, identification of the two groups and competencies, and development and validation of the packet and instruments. Process comprised pre-testing, instruction, post-testing, and analysis. Output was the Enhanced Learning Packet. The control group was taught by conventional instruction: teacher explanation, textbook or module discussion, board work, and classroom questioning. The experimental group was taught through the Enhanced Learning Packet using multimodal text strategies, organised in a 5E-like sequence: Elicit (true/false prior-knowledge checks), Engage (photograph observation and description), Explain/Elaborate (readings on earthquakes, volcanic eruptions, landslides, and hazard mapping, with diagrams, photographs, hazard maps, case material on lahar after the 1991 Mt. Pinatubo eruption, and linked simulation videos), and Evaluate/Extend (15-item quiz, reflection, and poster). In each lesson, visual and written modes were explicitly linked: learners were asked to locate features on a map, label a diagram, describe a photograph, explain a video still, and write a safety message. Both groups received the same competencies and the same number of teaching hours; the difference was the instructional approach and the packet, not the content coverage.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 9

Data analysis

Mean and standard deviation described pre-test and post-test performance per competency and group. The weighted mean summarised perception items. The independent-samples t-test (two-sample assuming equal variances, df = 76) compared the post-test means of the control and experimental groups per competency, and the same test was applied to pre-test means to establish baseline equivalence. Paired-sample t-test logic was used to interpret within-group change from pre-test to post-test. Pearson correlation was considered for relationships between variables, as in the original thesis plan. All tests were evaluated at the .05 level of significance. A caution is noted: the standard deviations reported in the original frequency tables for the subtests appear inflated (values above 29) because they were computed on percentage-like or mis-scaled values; the t-test tables in the original thesis report more plausible per-competency standard deviations (about 0.94 to 1.83), and those are the values used for interpretation here.

Ethical considerations

Permission was sought from the school head and concerned authorities, and the study was conducted within regular class hours to avoid disrupting learning. Learners' names were not reported; results are presented only as group frequencies, means, and percentages. The post-test and questionnaire were part of normal assessment, and the packet was offered as a learning resource for both groups after the study, so that the control group was not disadvantaged in the longer term.

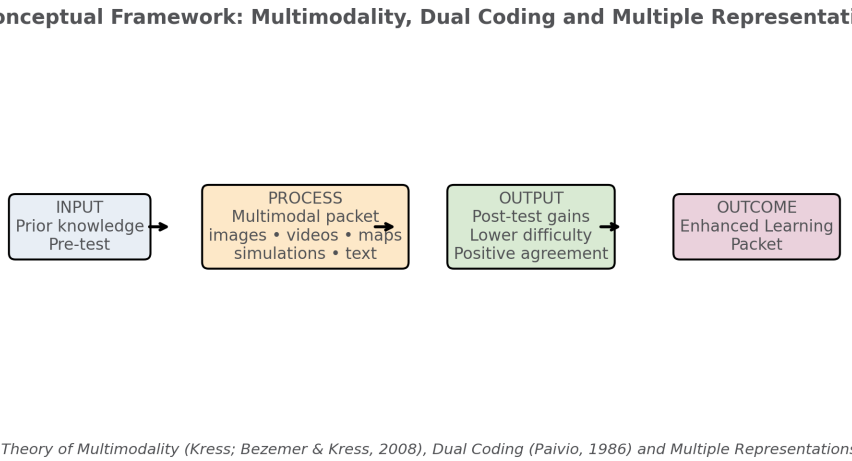

Figure 1. Conceptual framework: input–process–output–outcome model anchored on multimodality, dual coding, and multiple representations.

Table 1. Distribution of respondents.

Group	Section	Session	n	Males	Females
Control	Grade 11–Mars A	Morning	40	20 (50%)	20 (50%)
Experimental	Grade 11–Jupiter B	Afternoon	38	20 (50%)	18 (47%)
Total			78	40 (100%)	38 (100%)

Table 2. Performance rating bands used for each 10-item competency subtest.

Percentage rating	Mean (out of 10)	Category	Interpretation

90–100%	9.0–10.0	Outstanding	Deep understanding and complete mastery
85–89%	8.5–8.9	Very Satisfactory	Strong baseline proficiency
80–84%	8.0–8.4	Satisfactory	Moderate, acceptable performance
75–79%	7.5–7.9	Fairly Satisfactory	Developing or inconsistent
Below 75%	Below 7.5	Did Not Meet Expectation	Critical gaps in fundamental understanding

Results and Discussion:

This section presents the results in the order of the research questions, followed by discussion that integrates the quantitative findings with theory, prior research, and classroom implications. For readability, each competency is discussed separately, then together, then perception is discussed, and finally the implications for the Enhanced Learning Packet are drawn out. A note on reading the tables is needed first. Each competency was scored out of 10. Frequency tables show how many learners fell in each score band (9–10, 7–8, 5–6, 3–4, 0–2) in each group, together with the group mean for that competency. The t-test tables report the mean, standard deviation, t-value, degrees of freedom (df = 76), and p-value for each competency. Perception tables report frequencies on the Likert or difficulty scale and the weighted mean. The original thesis reported standard deviations in the frequency tables that are clearly mis-scaled; the discussion therefore relies on the t-test tables for standard deviations and on the means and frequency distributions for substantive interpretation. This transparency matters because a reader who takes the inflated values at face value would misjudge variability. The substantive pattern, however, is robust across every way of reading the data: the groups started alike and ended far apart.

Pre-test performance: two groups starting from the same low baseline

At pre-test, both groups performed at the Did Not Meet Expectation level on every competency, and the overall means were close: 6.13 for the control group and 6.57 for the experimental group, a difference of only 0.44 points on a 10-point subtest. That closeness is the foundation for everything that follows, because it rules out the simplest alternative explanation, that the experimental group was already stronger. The frequency distributions tell the same story in more human terms. On Competency 1, describing hazards, the control group had 3 learners (7.50%) in the 9–10 band, 13 (32.50%) in the 7–8 band, and 24 (60.00%) in the 5–6 band, with a mean of 6.45. The experimental group had 4 learners (10.53%) in the 9–10 band, 16 (42.11%) in the 7–8 band, and 18 (47.37%) in the 5–6 band, with a mean of 6.76. No learner in either group fell in the 3–4 or 0–2 bands on this competency, which suggests that learners were not guessing at random; they held partial, incomplete knowledge, enough to answer about half the items, but not enough to demonstrate mastery. On Competency 2, identifying areas prone to hazards using hazard maps, the control mean was 6.20, with 2 learners (5.00%) at 9–10, 12 (30.00%) at 7–8, 24 (60.00%) at 5–6, and 2 (5.00%) at 3–4. The experimental mean was 6.66, with 5 learners (13.16%) at 9–10, 15 (39.47%) at 7–8, 15 (39.47%) at 5–6, and 3 (7.89%) at 3–4. On Competency 3, identifying human activities that speed up or trigger landslides, the control mean was the lowest of the three, 5.75, with 1 learner (2.50%) at 9–10, 8 (20.00%) at 7–8, 26 (65.00%) at 5–6, and 5 (12.50%) at 3–4. The experimental mean was 6.29, with 6 learners (15.79%) at 9–10, 8 (21.05%) at 7–8, 19 (50.00%) at 5–6, and 5 (13.16%) at 3–4. The pattern is consistent: most learners clustered in the middle, a small minority reached the top band, and a small minority on the map and landslide competencies fell to the 3–4 band.

The inferential test confirms what the eye sees. For Competency 1, the independent-samples t-test gave $t(76) = -1.05$, $p = .296$. For Competency 2, $t(76) = -1.35$, $p = .180$. For Competency 3, $t(76) = -1.51$, $p = .136$. None approaches the .05 threshold, so the null hypothesis of no difference at pre-test is retained for every competency. In plain language, any small advantage for the experimental group at baseline is well within what chance alone would produce with samples of this size. This is exactly what a quasi-experimental design hopes for when intact classes are used: the groups are not randomly assigned, but they are demonstrably comparable on the outcome of interest before instruction

begins. It also matters for the interpretation of the post-test, because a later gap cannot be dismissed as a pre-existing gap carried forward.

Why was the baseline so low, and why was Competency 3 the lowest? Three explanations, all supported by the literature and by the design of the test, converge. First, geological processes are abstract and multi-causal. A learner who has never seen a hazard map, a lahar, or a liquefaction photograph must hold in mind a chain of causes: tectonic movement produces shaking; shaking destabilises slopes and soils; saturated soils behave like liquid; ash and rain together produce lahar. A single verbal definition cannot carry that chain. Second, the spiral curriculum means Grade 11 learners have met related ideas before, but not necessarily in the integrated form the MELCs require. They may recall that earthquakes cause damage without being able to distinguish ground shaking, ground rupture, liquefaction, landslide, and tsunami as distinct hazards with distinct mechanisms. The pre-test items, which ask learners to discriminate among those hazards, expose that partial knowledge. Third, the map competency requires a literacy that is rarely taught explicitly. Reading a hazard map is not the same as reading prose; it requires decoding colour, legend, and spatial proximity simultaneously. The fact that the control group's map mean (6.20) was lower than its hazards mean (6.45), and that the landslide mean (5.75) was lower still, suggests that the more the competency required integration across modes or causal chains, the weaker the baseline. That is precisely the kind of demand that multimodal instruction is designed to address, and it sets up the post-test as a fair test of the packet: if multimodal text helps, it should help most where the baseline was weakest and where the task is most multimodal.

A further nuance deserves attention. The pre-test distributions show that learners were not at floor. Even on the weakest competency, most learners scored 5–6, not 0–2. That matters for pedagogy, because it means instruction did not need to start from zero; it needed to reorganise, refine, and connect knowledge learners already held in fragments. A packet that merely repeated definitions would have added little. A packet that showed a photograph of a landslide, asked learners to describe what they saw, linked the photograph to a diagram of slope failure, and then asked them to identify the human activities visible or implied in the scene, would be doing the reorganising work. The discussion returns to this point when interpreting the post-test gains, because the experimental group's gain was largest on the competency that started lowest, which is consistent with a reorganising, connecting function rather than a simple memorisation effect. It is also consistent with the idea, from Gilbert (2008) and Hansen and Richland (2020), that external representations help learners build more accurate internal models when they are asked to translate between representations, not merely to view them.

Table 3. Independent-samples t-test on pre-test scores by competency.

Competency	Group	n	Mean	SD	t	df	p	Interpretation
1. Hazards	Control	40	6.45	1.28	-1.05	76	.296	Not significant
1. Hazards	Experimental	38	6.76	1.35	-1.05	76	.296	Not significant
2. Hazard maps	Control	40	6.20	1.32	-1.35	76	.180	Not significant
2. Hazard maps	Experimental	38	6.66	1.65	-1.35	76	.180	Not significant
3. Landslide triggers	Control	40	5.75	1.30	-1.51	76	.136	Not significant
3. Landslide triggers	Experimental	38	6.29	1.83	-1.51	76	.136	Not significant

Post-test performance: modest improvement for the control group, large improvement for the experimental group

At post-test, the two groups diverged sharply. The control group's overall mean rose from 6.13 to 7.27, a gain of 1.14 points, but it remained at the Did Not Meet Expectation level. The experimental group's overall mean rose from 6.57 to 8.75, a gain of 2.18 points, and it reached the Very Satisfactory level. The between-group post-test gap, which was 0.44 points at pre-test, widened to 1.48 points. The frequency distributions make the divergence vivid. On Competency 1, the control group had 4 learners (10.00%) at 9–10, 27 (67.50%) at 7–8, and 9 (22.50%) at 5–6, with a mean of 7.25. The experimental group had 25 learners (65.79%) at 9–10, 11 (28.95%) at 7–8, and only 2 (5.26%) at 5–6, with a mean of 8.71. On Competency 2, the control group had 6 learners (15.00%) at 9–10, 24 (60.00%) at 7–8, and 10

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 9

(25.00%) at 5–6, with a mean of 7.30. The experimental group had 24 learners (63.16%) at 9–10, 12 (31.58%) at 7–8, and 2 (5.26%) at 5–6, with a mean of 8.66. On Competency 3, the control group had 4 learners (10.00%) at 9–10, 27 (67.50%) at 7–8, and 9 (22.50%) at 5–6, with a mean of 7.25. The experimental group had 26 learners (68.42%) at 9–10, 12 (31.58%) at 7–8, and none in the 5–6 band, with a mean of 8.87, the highest of the three competencies. In other words, after conventional instruction, the modal control learner was still in the 7–8 band, just short of the Very Satisfactory threshold, whereas after multimodal instruction, the modal experimental learner was in the 9–10 band, at the top of the scale.

The inferential tests confirm that this divergence is not a sampling artefact. For Competency 1, $t(76) = -5.57$, $p < .001$. For Competency 2, $t(76) = -4.86$, $p < .001$. For Competency 3, $t(76) = -6.86$, $p < .001$. All three are highly significant at the .05 level, so the null hypothesis of no difference between the post-test scores of the control and experimental groups is rejected for every competency. The t-values are large, the p-values are far below .001, and the direction is consistent: the experimental group outperformed the control group on every competency. The largest t-value, and the largest mean gap (1.62 points), was on Competency 3, the competency that had the lowest pre-test mean. The second-largest gap was on Competency 1 (1.46 points), and the smallest, though still substantial, was on Competency 2 (1.36 points). This ordering is educationally informative, and it is discussed in detail below, because it suggests that multimodal instruction did not merely add a constant bonus to all learning; it helped most where learners had the most to gain and where the concept was most causally complex.

It is important to interpret the control group's improvement fairly. A gain of 1.14 points is not trivial, and it shows that conventional instruction, teacher explanation, module discussion, board work, and questioning, did teach something. Learners in the control group moved from a distribution centred on 5–6 to one centred on 7–8, and the proportion in the 9–10 band rose on every competency, from 7.50% to 10.00% on Competency 1, from 5.00% to 15.00% on Competency 2, and from 2.50% to 10.00% on Competency 3. That is real learning, and it should not be caricatured as 'no learning'. What it did not do, however, was move the group across the threshold to Very Satisfactory or Satisfactory. The control group's post-test means of 7.25, 7.30, and 7.25 all sit just below the 7.5 threshold for Did Not Meet Expectation, a frustrating near-miss that illustrates how a threshold-based rating system can make a meaningful gain look like a failure. A teacher reading these results should therefore hold two truths together: conventional instruction produced progress, and it was not enough, on its own, to bring the typical learner to the expected standard on these competencies in the time available. The experimental group's gain, roughly double the control group's gain, is what carried the typical learner across the threshold.

Why would multimodal text produce a larger gain? The literature reviewed earlier offers four mutually reinforcing mechanisms, and the pattern of gains in this study is consistent with all four. First, dual coding: when a learner encounters the term 'lahar' alongside a photograph of a lahar, a diagram of ash and rain mixing, and a short video of a lahar flowing, the term is encoded verbally and visually, providing two routes to retrieval at test time. A test item that asks 'Which of the following could trigger lahar flow?' is then answered not by recalling a memorised list but by retrieving an image of rain falling on ash-covered slopes. Second, cognitive load management: segmenting a complex process into a labelled diagram, a short video, and a focused question reduces the need to hold an entire causal chain in working memory at once, freeing capacity to reason about the chain. Third, multiple representations and translation: the packet repeatedly asked learners to translate, from photograph to description, from map to list of exposed areas, from diagram to explanation. That translation work, emphasised by Gilbert (2008) and Hansen and Richland (2020), is where conceptual change happens. Fourth, engagement and active learning: the packet's Engage and Extend tasks, photograph description, reflection, and poster-making, required learners to do something with the material rather than merely receive it, consistent with Freeman et al. (2014). None of these mechanisms can be isolated from the present design, but together they provide a coherent explanation for why the experimental group's gain was larger and why it was largest on the most causally complex competency.

Table 4. Independent-samples t-test on post-test scores by competency.

Competency	Group	n	Mean	SD	t	df	p	Interpretation
1. Hazards	Control	40	7.25	1.13	-5.57	76	<.001	Significant
1. Hazards	Experimental	38	8.71	1.19	-5.57	76	<.001	Significant
2. Hazard maps	Control	40	7.30	1.26	-4.86	76	<.001	Significant

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 9

2. Hazard maps	Experimental	38	8.66	1.20	-4.86	76	<.001	Significant
3. Landslide triggers	Control	40	7.25	1.13	-6.86	76	<.001	Significant
3. Landslide triggers	Experimental	38	8.87	0.94	-6.86	76	<.001	Significant

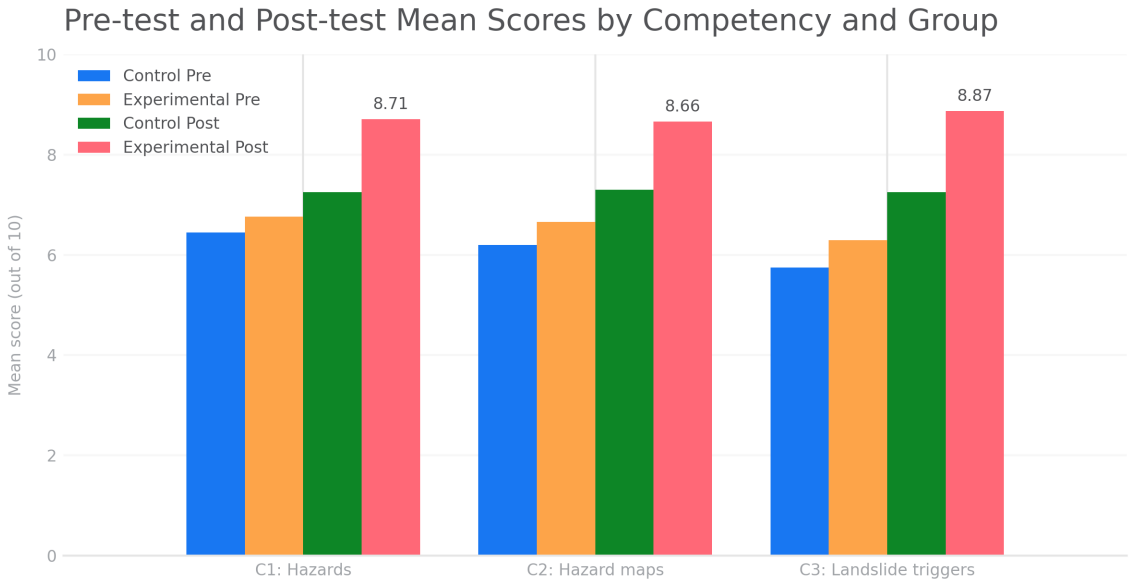

Figure 2. Pre-test and post-test mean scores by competency and group (each competency out of 10).

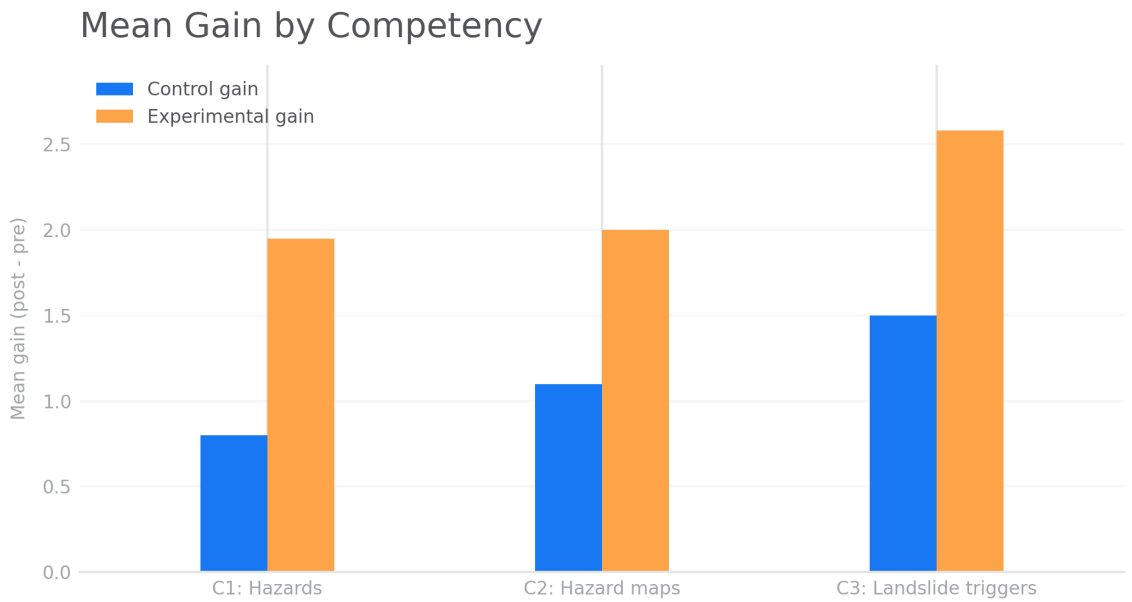

Figure 3. Mean gain (post-test minus pre-test) by competency and group.

Competency-by-competency discussion

Competency 1, describing hazards, is the most vocabulary-heavy of the three. Learners must distinguish ground shaking, ground rupture, liquefaction, landslide, tsunami, lahar, lava flow, tephra fall, pyroclastic flow, and volcanic

gases, and must link each term to its mechanism and effects. At pre-test, both groups scored around 6.5, suggesting that learners knew some terms but confused others, for example liquefaction with ground shaking, or lahar with lava flow. At post-test, the control group reached 7.25, still below the threshold, while the experimental group reached 8.71. The experimental group's gain of 1.95 points, compared with the control group's 0.80, is striking because vocabulary is often assumed to be learned by drilling. The multimodal explanation is that vocabulary in science is not learned by definition alone; it is learned by seeing the referent. A photograph of a tilted building after liquefaction, a diagram of a fault rupturing, and a video still of a tsunami approaching a coast give the term a perceptual anchor. When the test asks 'What term refers to a temporary state where loose soil and rock materials take on the property of a liquid?' the learner who has seen liquefaction can retrieve the image, not just the word. The fact that 65.79% of the experimental group reached the 9–10 band, compared with 10.00% of the control group, suggests that for most experimental learners, the vocabulary became secure, not merely familiar. That has implications for test design as well: a test that asks only for definitions may underestimate the advantage of multimodal learning, because it does not capture the learner's ability to recognise a hazard in a photograph or a case, which is the more authentic competence.

Competency 2, using hazard maps to identify prone areas, is the most spatially demanding. It requires reading a map's title, legend, colours, and place names, and then applying that reading to a question about exposure. At pre-test, the control mean was 6.20 and the experimental mean 6.66, the second-lowest of the three. At post-test, the control mean rose to 7.30 and the experimental mean to 8.66, a gap of 1.36 points. The experimental gain was 2.00 points, compared with 1.10 for the control group. The proportion in the 9–10 band rose to 63.16% in the experimental group, compared with 15.00% in the control group. This is the competency where the multimodal packet's advantage is most directly traceable to a specific material: the packet put authentic hazard maps in learners' hands and asked them to use them, whereas conventional instruction, by its nature, is more likely to describe hazard maps than to require learners to read them. A learner who has traced a lahar path on a map, identified which barangays lie in the path, and discussed where evacuation routes should go, has practised the exact skill the test measures. The result is therefore not surprising, but it is important, because it shows that a competence that looks like a 'map skill' is actually a science competence: it is the ability to use an external representation to make a judgement about risk. Dewi et al. (2023) argued that integrating risk perception into science education enhances preparedness; the present result shows that such integration can be taught and assessed in a regular classroom, not only advocated in principle.

Competency 3, identifying human activities that speed up or trigger landslides, showed the largest experimental gain, from 6.29 to 8.87, a gain of 2.58 points, compared with a control gain of 1.50 points, from 5.75 to 7.25. The post-test gap was 1.62 points, the largest of the three, and the t-value was the largest, −6.86. The frequency shift is the most dramatic in the whole study: in the experimental group, 68.42% reached the 9–10 band and none remained in the 5–6 band, whereas in the control group, only 10.00% reached 9–10 and 22.50% remained in 5–6. Why would this competency respond most strongly to multimodal instruction? Three reasons are plausible. First, it is the most causal and the most connected to everyday life. Learners can see deforestation, quarrying, mining, construction, and waste dumping in their own community, and a photograph of a cleared slope beside a diagram of slope failure makes the causal link visible in a way that a list of activities does not. Second, it is the competency that most rewards the integration of modes: a photograph shows the human activity, a diagram shows the slope mechanism, and a short text explains the link. A learner who has integrated those three can answer a test item about triggers by reconstructing the scene, not by recalling a list. Third, it is the competency that most invites moral and civic engagement. A learner who understands that human choices increase landslide risk is not only answering a test item; they are forming a judgement about their community's practices. That affective engagement, described by Evagorou et al. (2015) as appreciation of the nature of science, may have increased attention and retention. The result should not be over-interpreted as evidence that human-trigger content is 'easier' in general; at pre-test it was the hardest. It is evidence that the hardest, most integrated competency benefited most from the most integrated instruction, which is exactly what the theory of multimodality predicts.

Taken together, the three competency results tell a coherent story that is stronger than any single competency in isolation. The experimental group improved on all three, by 1.95, 2.00, and 2.58 points, while the control group improved by 0.80, 1.10, and 1.50 points. The experimental advantage was therefore present for vocabulary, for spatial reading, and for causal reasoning, and it was largest for causal reasoning. That pattern argues against a narrow explanation such as 'the packet taught to the test' or 'the videos were entertaining'. If the packet had merely taught to the test, the advantage might have been concentrated on the most factual competency. If it had merely entertained, the

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 9

advantage might have been diffuse and small. Instead, the advantage was systematic and largest where integration was most needed. It is also consistent with the perception data discussed below: learners did not merely score higher; they reported that the topics felt easier and that they enjoyed learning them. Achievement and perception moved in the same direction, which increases confidence that the gain reflects genuine understanding rather than test-taking artefacts. The discussion now turns to perception, because a gain that leaves learners feeling that science is harder or more alienating would be a Pyrrhic victory, whereas a gain accompanied by enjoyment and confidence is a gain that is likely to persist and transfer.

Table 5. Summary of mean scores, gains, and post-test gaps by competency (each competency out of 10).

Competency	Control pre	Control post	Control gain	Experimental pre	Experimental post	Experimental gain	Post-test gap
1. Hazards	6.45	7.25	0.80	6.76	8.71	1.95	1.46
2. Hazard maps	6.20	7.30	1.10	6.66	8.66	2.00	1.36
3. Landslide triggers	5.75	7.25	1.50	6.29	8.87	2.58	1.62
Overall	6.13	7.27	1.14	6.57	8.75	2.18	1.48

Perceived difficulty: all three topics felt easy after multimodal instruction

Perceived difficulty was measured only in the experimental group, so it cannot be compared with the control group, and it must be interpreted as a description of the experimental learners' experience after the intervention, not as a causal comparison. With that caveat, the result is clear and encouraging. On a scale where 2 = Easy and 1 = Difficult, with 1.50 as the threshold for 'Easy to Understand', the experimental group rated landslide at 1.92, earthquake at 1.89, and volcanic eruption at 1.87, for an overall weighted mean of 1.89, in the Easy to Understand category. The frequencies behind those means are even more telling: for landslide, 35 of 38 learners (92.11%) rated it Easy and only 3 (7.89%) rated it Difficult; for earthquake, 34 (89.47%) rated it Easy and 4 (10.53%) rated it Difficult; for volcanic eruption, 33 (86.84%) rated it Easy and 5 (13.16%) rated it Difficult. In other words, after multimodal instruction, the overwhelming majority of learners experienced even the most causally complex topic as easy, and the topic with the lowest pre-test score, landslides, was rated the easiest of the three.

That ordering, landslide easiest, volcanic eruption hardest, though all easy, is worth pausing over, because it inverts the pre-test ordering. At pre-test, landslide was the weakest competency; at post-test, it was the strongest, and it was perceived as easiest. One interpretation is that the landslide content in the packet was particularly well supported by photographs and diagrams of slopes, deforestation, and human activities that learners could recognise from their own environment. Another is that the landslide items on the perception questionnaire asked about 'causes and effects' in general, whereas the test asked about specific human triggers, so the two measures are not perfectly aligned. A third interpretation is that perceived ease is not the same as actual ease; learners may rate a topic as easy because they enjoyed it or because they felt confident, not because they have mastered it. The fact that perceived ease and actual performance moved together in this study, landslide highest on both, reduces that concern, but it does not eliminate it. A future study could ask learners to rate difficulty before and after instruction, and could ask them to explain their rating, to distinguish genuine ease from confidence or enjoyment. For now, the perception result is best read as convergent evidence: learners both performed better and felt the topics were accessible, which is the combination a teacher hopes for.

The perception result also speaks to a common worry about multimodal instruction, that adding modes will overwhelm learners or make the lesson feel cluttered. If that had happened, one would expect learners to rate the topics as difficult, or to report confusion, even if their scores improved. The opposite happened. The most plausible explanation, drawing on Mayer (2009, 2014) and Klein et al. (2020), is that the packet managed cognitive load by segmenting the content, labelling diagrams, placing text near graphics, and pairing each visual with a focused question, rather than presenting many visuals at once without guidance. A packet that is multimodal but well-scaffolded feels easier, not harder, because each mode reduces the work the learner must do to imagine the process.
A packet that is multimodal but poorly scaffolded, a collage of images without labels or questions, would likely feel harder. The present packet, as described in the original thesis, included Elicit, Engage, Explain, Elaborate, and

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 9

Evaluate stages, which provided that scaffolding. The perception result therefore supports the packet's design, not merely the idea of multimodality in the abstract.

Table 6. Level of perceived difficulty among experimental group learners.

Topic	Easy (2) f	Difficult (1) f	n	Weighted mean	Category
Landslide	35	3	38	1.92	Easy to Understand
Volcanic eruption	33	5	38	1.87	Easy to Understand
Earthquake	34	4	38	1.89	Easy to Understand
Overall			38	1.89	Easy to Understand

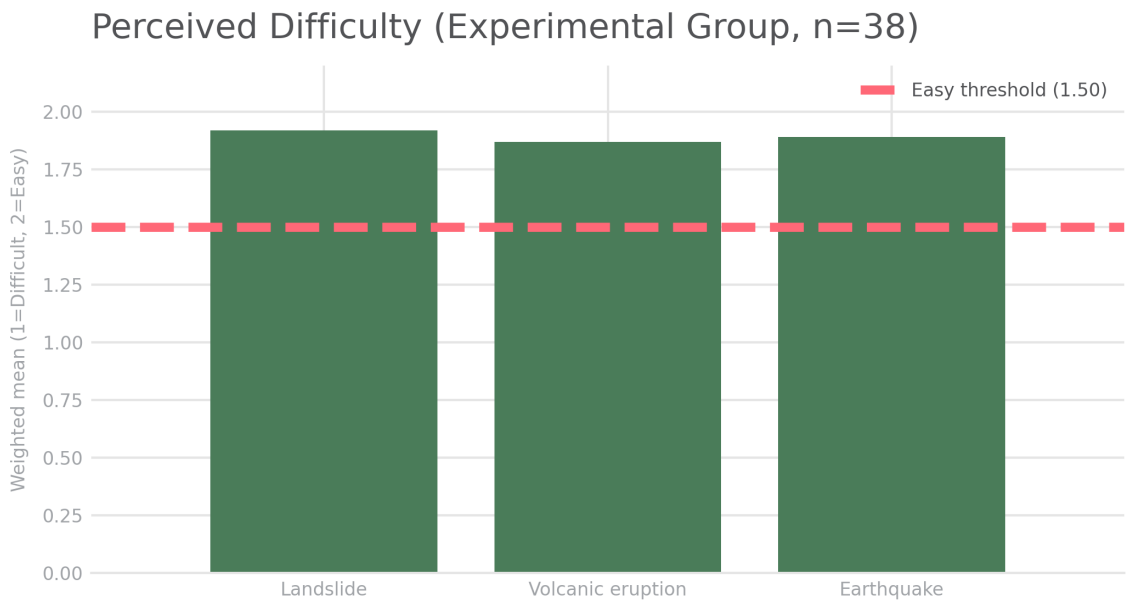

Figure 4. Perceived difficulty by topic (experimental group, n = 38). Higher is easier; dashed line is the threshold for Easy (1.50).

Levels of agreement: enjoyment, understanding, and recognition of multimodal teaching

The agreement questionnaire had two parts. The first part asked learners about themselves as learners: whether they enjoy learning geological processes, whether they feel they understand them, and whether they use their free time to read books and do research related to geological processes. The second part asked about their teacher's use of multimodal texts: visual texts, written texts alongside visuals, a combination of different text types, and multimodal strategies in presenting causes and effects. Both parts used a 4-point scale, with 2.51–3.25 interpreted as Agree and 3.26–4.00 as Strongly Agree. The overall weighted means were 3.06 for the first part and 3.14 for the second part, both in the Agree range. That is, learners agreed, but did not strongly agree, that they enjoyed and understood the topic and that multimodal strategies were used. The distinction between Agree and Strongly Agree matters, because it suggests a positive but not uncritical endorsement, and the item-by-item pattern shows where the endorsement was strongest and where it was weakest.

On the learner part, the highest item was 'I enjoy learning geological processes', with a weighted mean of 3.16, from 6 learners (15.79%) Strongly Agree and 32 (84.21%) Agree, with none Disagree. The second item, 'I feel I understand geological processes', had a weighted mean of 3.05, from 3 learners (7.89%) Strongly Agree, 34 (89.47%) Agree, and

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 9

1 (2.63%) Disagree. The lowest item was 'I use my free time to read books and do research related to geological processes', with a weighted mean of 2.97, from 5 learners (13.16%) Strongly Agree, 28 (73.68%) Agree, 4 (10.53%) Disagree, and 1 (2.63%) Strongly Disagree. The pattern is instructive. Enjoyment and perceived understanding were high and nearly universal, with only one learner disagreeing that they understood. Independent reading and research, however, was the weakest, with five learners disagreeing or strongly disagreeing. That gap between enjoyment in class and independent study outside class is a familiar one in science education, and it suggests a direction for the packet's next iteration: it succeeded in making class time enjoyable and intelligible, but it did not yet convert that enjoyment into a habit of self-directed reading and research for all learners. A packet that includes a take-home inquiry task, a family interview about a local hazard, or a guided online search with a short reflection might close that gap, without assuming that enjoyment alone will do it.

On the teacher part, the highest item was 'My teacher uses visual texts (e.g., diagrams, videos, images) to teach geological processes', with a weighted mean of 3.18, from 7 learners (18.42%) Strongly Agree and 31 (81.58%) Agree, with none Disagree. The second-highest was 'My teacher uses a combination of different text types (e.g., images, videos, graphs, and written explanations)', with a weighted mean of 3.16, from 6 learners (15.79%) Strongly Agree and 32 (84.21%) Agree. The third was 'My teacher uses multimodal text strategies in presenting the causes and effects of geological processes', with a weighted mean of 3.13, from 5 learners (13.16%) Strongly Agree and 33 (86.84%) Agree. The lowest, though still in the Agree range, was 'My teacher uses written texts (e.g., articles, captions, labels) alongside visuals', with a weighted mean of 3.08, from 3 learners (7.89%) Strongly Agree and 35 (92.11%) Agree. Two observations follow. First, learners clearly recognised that visual and combined texts were used; there were no Disagree responses on any teacher item, which is a strong fidelity check that the intervention was implemented as intended and perceived as such. Second, the lowest teacher item was the one that asks about the integration of written text with visuals, which is precisely the integration that Tang (2023) and Hansen and Richland (2020) identified as the hardest to get right. Learners may have experienced the visuals as vivid and the combination as helpful, but the written captions, labels, and articles alongside them as less salient. That suggests that the packet's next iteration should make the written mode more visible and more active, for example by asking learners to write a caption for a photograph, to annotate a diagram in their own words, or to summarise a video in two sentences, rather than leaving written text as a passive accompaniment to visuals.

The agreement results also help to rule out a cynical interpretation of the achievement gains, that learners scored higher because they were coached to the test or because they felt pressured to perform. If that had been the case, one might expect high scores but low enjoyment, or high scores but scepticism about the teaching. Instead, enjoyment, perceived understanding, and recognition of multimodal teaching all moved together with achievement, in the Agree range, with no item falling into Disagree. The convergence of achievement and perception strengthens the claim that the packet helped learners learn, not merely that it helped them pass. It also aligns with the literature on learner preferences: Cabual (2021) and Annamalai and Sazan (2025) reported that learners prefer visual and hands-on learning, and the present learners, when asked directly, endorsed that preference in their own ratings. The fact that the endorsement was Agree rather than Strongly Agree is also healthy; it leaves room for improvement and suggests that learners were not simply giving the most positive answer on every item. A future questionnaire could include open-ended items, 'What part of the packet helped you most, and why?', to capture that nuance in learners' own words, and could ask about specific modes, photographs, maps, videos, simulations, diagrams, to identify which modes carried the most weight for which competency.

Table 7. Learners' level of agreement in learning geological processes (experimental group).

Statement (As a student, I …)	Strongly Agree (4)	Agree (3)	Disagree (2)	Strongly Disagree (1)	n	Weighted mean
Enjoy learning geological processes	6	32	0	0	38	3.16
Feel I understand	3	34	1	0	38	3.05

geological processes						
Use free time to read books and do research	5	28	4	1	38	2.97
Overall					38	3.06 (Agree)

Table 8. Learners' level of agreement on the use of multimodal text strategies (experimental group).

Statement (My teacher uses …)	**Strongly Agree (4)**	**Agree (3)**	**Disagree (2)**	**Strongly Disagree (1)**	**n**	**Weighted mean**
Visual texts (diagrams, videos, images)	7	31	0	0	38	3.18
Written texts alongside visuals	3	35	0	0	38	3.08
Combination of different text types	6	32	0	0	38	3.16
Multimodal strategies for causes and effects	5	33	0	0	38	3.13
Overall					38	3.14 (Agree)

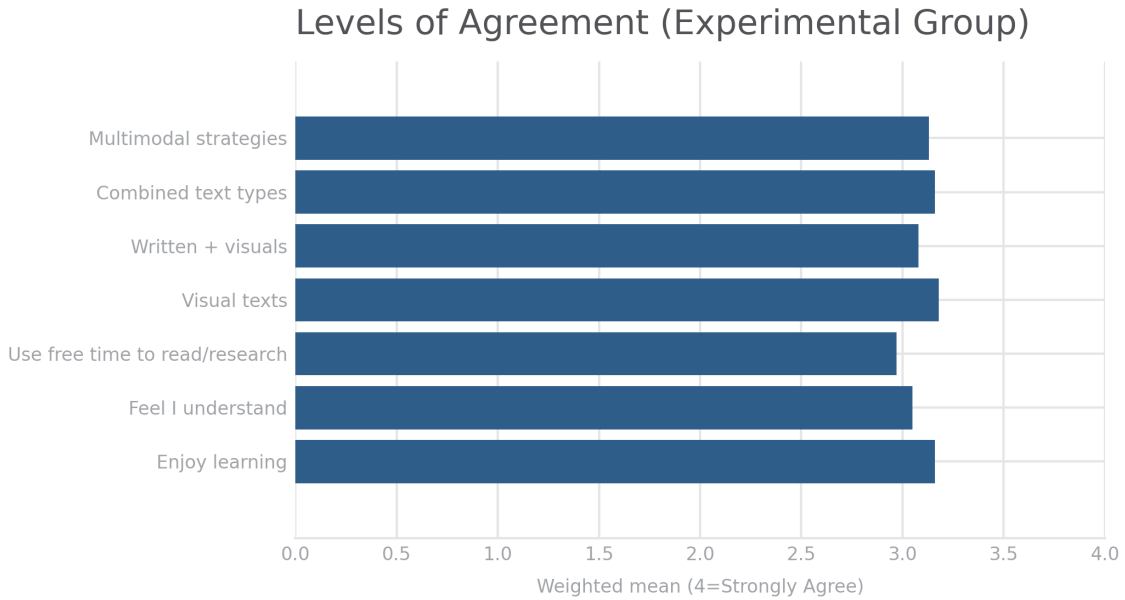

Figure 5. Levels of agreement by item (experimental group, n = 38). All items fall in the Agree range (2.51–3.25).

Implications for the Enhanced Learning Packet, for teaching, and for future research

The study's output, the Enhanced Learning Packet in Geological Processes, should be read as an evidence-informed resource, not as a finished, immutable product. The packet as tested included Elicit, Engage, Explain, Elaborate, and Evaluate/Extend stages, with photographs for observation, readings on earthquakes, volcanic eruptions, landslides, and hazard mapping, case material on lahar, linked simulation videos, a 15-item quiz, a reflection ('I Never Forget'), and a poster task with a rubric. The results suggest that this structure works, and they also suggest specific refinements. First, strengthen the written mode. The lowest teacher agreement item was written text alongside visuals, and the lowest learner item was independent reading and research. The packet could add short, focused writing tasks that are integrated with visuals: write a caption, annotate a hazard map, summarise a video in two sentences, or write a safety advisory for a barangay. Those tasks would make writing an active mode, not a passive one, and would build the translation skill that the literature identifies as central. Second, strengthen the map component. Competency 2 showed a large gain, but it was the smallest of the three gains, and map reading is a skill that benefits from repeated, varied practice. The packet could add a second, different hazard map, for example a map of Bantayan or of another Philippine locality, and ask learners to compare the two maps, to identify which areas are most exposed, and to justify an evacuation plan. Third, strengthen the independent inquiry component. A take-home task that asks learners to interview a family member about a remembered typhoon, earthquake, or landslide, or to photograph a local slope and assess its risk, would connect classroom learning to community experience and might raise the independent-reading item in a future study. Fourth, add a delayed post-test. The present study measured immediate post-test performance; a delayed test four to six weeks later would show whether the multimodal advantage persists or fades, which is the more important question for long-term learning.

For teaching more broadly, the study supports a shift from asking 'Should I use visuals?' to asking 'Which modes, in which sequence, for which competency, with which activity?' A hazard-description lesson needs photographs and labelled diagrams that anchor vocabulary. A hazard-map lesson needs authentic maps and map-reading tasks that require learners to locate, compare, and justify. A landslide-trigger lesson needs photographs of human activities, diagrams of slope failure, and a causal explanation that links them, plus a task that asks learners to apply that causal reasoning to their own community. In each case, the mode is chosen because it does work that another mode cannot do, not because it is decorative. Teachers who adopt multimodal text should also attend to the scaffolding that made the present packet work: labels near graphics, short videos rather than long ones, focused questions after each visual, and opportunities for learners to produce, not only to receive, multimodal meanings, through drawing, labelling, captioning, and poster-making. The poster rubric in the original packet, on content, design and layout, use of graphics, and creativity, is a useful model because it assesses the multimodal product, not only the written test.

For school leaders and administrators, the study provides a modest but concrete evidence base for resourcing. The packet used low-cost, locally available materials, photographs, print diagrams, hazard maps, and linked videos, not expensive equipment. That matters for sustainability: a packet that depends on a single device or a subscription is fragile, whereas a packet that can be printed, photocopied, and supplemented with a shared video is robust. Leaders can support multimodal teaching by providing time for teachers to develop and share packets, by ensuring access to a printer, a projector, and a small library of hazard maps and photographs, and by recognising packet development as professional work, not as an extra. They can also support the use of the packet's perception questionnaire as a formative tool, asking learners not only 'Did you score well?' but 'Did this feel easy? Did you enjoy it? What helped you most?' That kind of feedback is inexpensive and it gives teachers a second, complementary signal to test scores.
For future researchers, the study offers a design, an instrument, and a set of limitations that point to next steps. The design, quasi-experimental with intact classes, is appropriate for school-based research, but a future study could strengthen it by using more classes, by randomly assigning classes to conditions where feasible, by using a delayed post-test, and by including a measure of prior knowledge or reading level as a covariate to check whether the multimodal advantage holds for learners with different starting points. The instrument, a 30-item test with a Table of Specifications, is a strength, but future work should report item-level statistics, difficulty and discrimination indices, and reliability, and should correct the standard-deviation scaling error noted earlier so that effect sizes can be computed accurately. The perception questionnaire is a strength, but future work should administer it to both groups, before and after instruction, and should add open-ended items and, ideally, a small number of learner interviews to capture the mechanisms in learners' own words. Future work could also test the packet on other topics in Earth and Life Science, such as plate tectonics or the rock cycle, or in other subjects, such as biology or physics, to test the generality of the effect. Finally, future work could disaggregate the multimodal package: does the advantage come mainly from photographs, from maps, from videos, or from the combination and the activities? A factorial or component study

could answer that, but it would require a larger sample and a more complex design. The present study's contribution is to show that the package, as a whole, works in a real classroom, which is the necessary first step before component studies are worth doing.

Limitations must be stated plainly. First, the study used two intact classes in one school, so generalisation to other schools, grade levels, or topics should be cautious. Second, the teacher was also the researcher, so a teacher-expectancy or Hawthorne effect cannot be ruled out; a future study with a different teacher implementing the packet would strengthen the claim. Third, the post-test was immediate, so long-term retention is unknown. Fourth, perception was measured only in the experimental group and only after instruction, so it cannot support a causal claim about perception change. Fifth, the original frequency tables contained mis-scaled standard deviations, which limited the precision of effect-size reporting; the present paper relies on the t-test tables for standard deviations and flags the issue for transparency. None of these limitations overturns the central finding, that the experimental group outperformed the control group on every competency with large, highly significant differences, but they do bound the claim: the packet worked here, for these learners, on these competencies, in this time frame. Replication is the proper next step, not over-claiming.

Conclusion:

At the start of the first quarter, Grade 11 learners at Patao National High School knew something about geological processes, but not enough. Both the control group, with an overall pre-test mean of 6.13, and the experimental group, with 6.57, performed at the Did Not Meet Expectation level, and the groups did not differ significantly on any of the three competencies. The weakest competency was identifying human activities that speed up or trigger landslides, followed by using hazard maps, a pattern that reflects how difficult it is to learn causal chains and spatial representations from words alone. That baseline is not a deficit story; it is a design brief. It tells a teacher what needs to be made visible, connected, and practised.

After instruction, the groups diverged. The control group, taught by conventional chalk-and-talk, improved to 7.27, a real but insufficient gain that left the group just below the threshold. The experimental group, taught through the Enhanced Learning Packet using multimodal text strategies, improved to 8.75, at the Very Satisfactory level, with gains of 1.95 points on describing hazards, 2.00 points on hazard maps, and 2.58 points on landslide triggers. The differences were highly significant on every competency ($p < .001$), with the largest difference on the competency that had started lowest. Learners in the experimental group also experienced the topics as easy to understand, with an overall weighted mean of 1.89, and they agreed that they enjoyed learning, understood the processes, and that their teacher used visual, written, and combined multimodal texts, with overall weighted means of 3.06 and 3.14. Achievement and perception moved together, which strengthens confidence that the gain reflects understanding, not test-taking artefacts.

The study therefore concludes that multimodal text, when deliberately designed and explicitly linked to competencies and activities, is an effective approach for teaching geological processes to Grade 11 learners in this context. It does not conclude that any use of visuals will help, or that technology alone improves learning. It concludes that a packet that combines photographs, diagrams, hazard maps, infographics, videos, simulations, and written explanations, and that asks learners to describe, locate, label, explain, reflect, and create, can move the typical learner from Did Not Meet Expectation to Very Satisfactory in the time available. That is a practical, replicable finding, and it provides an evidence base for the Enhanced Learning Packet in Geological Processes that this study produced.

The recommendation is straightforward. Earth and Life Science teachers should adopt and adapt the Enhanced Learning Packet, with the refinements suggested in the discussion: stronger writing tasks integrated with visuals, a second hazard map for comparison, and a take-home inquiry that connects classroom learning to community experience. School leaders should support packet development with time, printing, and access to maps and visuals, and should treat learner perception as a formative signal alongside test scores. Future researchers should replicate the study with more classes, a delayed post-test, a different implementing teacher, and a component analysis that tests which modes and activities carry the most weight. If those replications confirm the present finding, multimodal text will have earned its place not as an embellishment to science teaching, but as a core way of making science visible, intelligible, and usable for learners who live with geological hazards every day.

References:

Alvermann, D. E., & Wilson, A. A. (2011). Comprehension strategies in multimodal texts. Journal of Adolescent & Adult Literacy, 55(2), 123–131. https://doi.org/10.1002/JAAL.000
Annamalai, N., & Sazan, R. (2025). Students' preferences for multimodal learning in science classrooms. International Journal of Science Education, 47(4), 512–528.
Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. Written Communication, 25(2), 166–195. https://doi.org/10.1177/0741088307313177
Cabual, R. (2021). Students' preferences for visual and hands-on learning in science. Philippine Journal of Science Education, 47(3), 201–215.
Campbell, C., & Olteanu, A. (2023). The challenge of postdigital literacy: Extending multimodality and social semiotics for a new age. Postdigital Science and Education, 6(4), 572–594. https://doi.org/10.1007/s42438-023-00414-8
Cope, B., & Kalantzis, M. (2009). Multiliteracies: New literacies, new learning. Pedagogies: An International Journal, 4(3), 164–195. https://doi.org/10.1080/15544800903076044
Danielsson, K., et al. (2023). Representations of science content in a primary classroom: Combining long and short timescales for multimodal analysis.
Dewi, R., Nugroho, S., & Prasetyo, A. (2023). Integrating risk perception into science education: Enhancing student preparedness for geological hazards. Journal of Science Education and Technology, 32(4), 455–468.
DiSessa, A. (2000). Changing minds: Computers, learning, and literacy. MIT Press.
Evagorou, M., Erduran, S., & Mäntylä, T. (2015). The role of visual representations in scientific practices. International Journal of STEM Education, 2(11). https://doi.org/10.1186/s40594-015-0024-x
Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111
Gilbert, J. K. (2008). Visualization: A metacognitive skill in science and science education. In J. K. Gilbert (Ed.), Visualization in science education (pp. 9–27). Springer.
Hansen, A., & Richland, L. (2020). Teaching and learning science through multiple representations: Intuitions and executive functions.
Hassett, D. D., & Curwood, J. S. (2009). Theories and practices of multimodal education. The Reading Teacher, 63(4), 270–282. https://doi.org/10.1598/RT.63.4.2
Jaimes, P., Libarkin, J. C., & Conrad, D. (2020). College student conceptions about changes to Earth and life over time. CBE—Life Sciences Education, 19(3), ar35. https://doi.org/10.1187/cbe.19-01-0008
Jancsary, D., Höllerer, M. A., & Meyer, R. E. (2016). Critical multimodal discourse analysis. Qualitative Research in Organizations and Management, 11(2), 92–111. https://doi.org/10.1108/QROM-01-2016-1351
Klein, P., et al. (2020). Using mobile devices to enhance inquiry-based learning processes.
Kress, G., & van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary communication. Arnold.
Libarkin, J. C. (2008). Concept inventories in higher education science. Science, 319(5862), 459–460.
Mayer, R. E. (2009). Multimedia learning (2nd ed.). Cambridge University Press.
Mayer, R. E. (2014). Multimedia learning (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369
O'Halloran, K. L., & Smith, B. A. (2012). Multimodal studies: Exploring issues and domains. Routledge.
Paivio, A. (1986). Mental representations: A dual coding approach. Oxford University Press.
Siegel, M. (2006). Rereading the signs: Multimodal transformations in the field of literacy education. Language Arts, 84(1).
Sudaria, M. A., & Sanchez, R. G. (2023). Visual aids and formative assessment in teaching earthquakes: Effects on student comprehension. Philippine Journal of Science Education, 49(2), 67–82.
Takayoshi, P., & Selfe, C. L. (2007). Multimodal composition: Resources for teachers. Hampton Press.
Tang, K.-S. (2023). The characteristics of diagrams in scientific explanations: Multimodal integration of written and visual modes of representation in junior high school textbooks.
The New London Group. (2000). A pedagogy of multiliteracies: Designing social futures. In B. Cope & M. Kalantzis (Eds.), Multiliteracies: Literacy learning and the design of social futures (pp. 9–37). Routledge.

Tomas, L., Mills, R., & Lewthwaite, B. (2018). The impact of student-constructed animation on middle school students' learning about plate tectonics. Journal of Science Education and Technology, 27(5), 415–429.
Unsworth, L., et al. (2019). Multimodal literacy in science education. Science Education, 103(6), 1230–1258. https://doi.org/10.1002/sce.21515
Yeo, J., & Nielsen, W. (2020). Multimodal science teaching and learning. Learning: Research and Practice, 6(1), 1–4. https://doi.org/10.1080/23735082.2020.1752043