Formatif: Jurnal Ilmiah Pendidikan MIPA, March 2026, 16 (1), 1-16
https://doi.org/10.30998/xdpy9m98
p-ISSN: 2088-351X e-ISSN: 2502-5457
Accredited (S1) by the Ministry of Higher Education, Science, and Technology, No. 295/C/C3/KPT/2026
Available online at https://newjournal.lppmunindra.ac.id/index.php/formatif

Tool Development and Validation for the Evaluation of Video Demonstration Presentation for a General Biology Laboratory Activity

Kevin Henson Gomez[1*] & Gemima Abayan Estrabillo[2]
[1]Holy Angel University, Angeles City, Pampanga, Philippines
[2]Department of Education – Division of Angeles City, Pampanga, Philippines

Abstract

This study aimed to develop and validate an evaluation tool for a video demonstration–based General Biology laboratory activity on cellular respiration using feedback from Grade 12 Science, Technology, Engineering, and Mathematics (STEM) students at Holy Angel University during School Year 2021–2022. Students' perceptions of facility of use, satisfaction, and self-reported laboratory competencies were examined to support instrument construction. A total of 260 Grade 12 STEM students enrolled in General Biology, all with prior exposure to laboratory-based science instruction, participated in the study. A quantitative evaluative research design was employed. The instrument consisted of two sections: (1) three open-ended questions capturing students' learning experiences, and (2) 26 Likert-type items measuring science process skills adapted from Safaah et al. (2017). A pilot test was conducted among Grade 12 General Academic Strand (GAS) students before administration to the main sample. Descriptive statistics, including means, standard deviations, and percentages, were used to summarize responses. Construct validity was assessed using the Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity, while internal consistency reliability was examined using Cronbach's alpha. Results showed acceptable levels of facility, satisfaction, and competency based on predefined scale descriptors. Factor analysis yielded two components with a KMO value of 0.91, indicating excellent sampling adequacy. Items 1, 3, 4, and 15 were removed due to low communalities and weak factor loadings, improving construct clarity. The final instrument demonstrated high reliability ($\alpha = 0.92$, pilot; $\alpha = 0.97$, post-test), supporting its validity for evaluating video-based biology laboratory activities.

Keywords: Video demonstration presentation method, cellular respiration, grade 12 STEM, tool development, tool validation

(*) Corresponding Author: kebingomez20@gmail.com

How to Cite: Gomez, K. H., & Estrabillo, G. A. (2026). Tool Development and Validation for the Evaluation of Video Demonstration Presentation for a General Biology Laboratory Activity. *Formatif : Jurnal Ilmiah Pendidikan MIPA*, *16*(1), 1-16. https://doi.org/10.30998/xdpy9m98

INTRODUCTION

The rapid development of educational technologies has transformed teaching, learning, and assessment practices, particularly in science education. In parallel with these developments, the imperative for rigorously developed research and evaluation instruments has grown increasingly pronounced. Effective assessment tools must be aligned with instructional design, learning objectives, and student competencies; however, mismatches between instructional approaches and available evaluation instruments remain a persistent issue in educational research (Kovacs et al., 2021). Inadequate evaluation tools may fail to capture authentic learning outcomes, resulting in incomplete or misleading conclusions about instructional effectiveness.

This challenge became more pronounced during the COVID-19 pandemic, when educational systems worldwide abruptly shifted from face-to-face instruction to online learning environments. While e-learning enabled instructional continuity, its effectiveness varied significantly across disciplines. STEM education, particularly subjects with laboratory components such as biology, chemistry, and physics, faced unique constraints due to the experiential and skill-based nature of laboratory learning (Gamage et al., 2020). Empirical studies have shown that although online modalities can support conceptual understanding, they often struggle to replicate hands-on experimentation, collaborative inquiry, and procedural skill development when not carefully designed and evaluated (Babincakova & Bernard, 2020; Kay et al., 2018).

Laboratory activities play a central role in science education by fostering scientific reasoning, observational skills, data interpretation, and experimental competencies (Bretz, 2019). In General Biology, laboratory instruction supports students' ability to connect abstract biological processes to observable phenomena and scientific procedures. During the pandemic, educators employed alternative laboratory approaches, including simulations, home-based tasks, and video demonstrations, to compensate for restricted physical laboratory access. Among these, video demonstration–based laboratories emerged as a widely adopted pedagogical strategy because they allowed instructors to model experimental procedures while maintaining safety and accessibility.

Despite their widespread use, previous studies on online laboratory implementations have primarily focused on student perceptions, engagement, or instructional feasibility, often overlooking the quality of evaluation tools used to assess laboratory outcomes. Many existing instruments lack strong validation evidence, clear measurement dimensions, or alignment with science process competencies, limiting their ability to accurately assess learning quality and skill attainment. Poorly designed evaluation tools may obscure gaps in student competencies, inflate perceived effectiveness, or fail to inform instructional improvement.

This study addresses this gap by focusing on the systematic development and validation of an evaluation tool for video demonstration–based laboratory activities. The tool is designed to measure three dimensions: facility of use, student satisfaction, and laboratory competencies grounded in science process skills. The laboratory topic of Cellular Respiration was selected because it integrates conceptual understanding with procedural explanation and was delivered entirely through video demonstration in the General Biology course at Holy Angel University during School Year 2021–2022.

The study does not claim immediate universal generalizability; rather, it provides an empirically validated instrument that may serve as a reference framework for evaluating similar video-based laboratory activities in biology and other STEM disciplines. By establishing the validity and reliability of the tool, this research contributes to improving the quality of laboratory assessment and supports evidence-based pedagogical decision-making in online and blended science education.

METHODS

Research Design

This study employed a quantitative evaluative research design to examine the suitability of a video demonstration–based approach for conducting online laboratory activities in General Biology. Evaluative research is appropriate when the purpose is to collect systematic evidence to inform judgments about educational programs, instructional strategies, or assessment tools (Cristobal & Cristobal, 2017). In this study, the focus was not on measuring instructional effectiveness through achievement scores, but on developing and validating an evaluation instrument that captures students' perceptions of an online laboratory method.

The video demonstration presentation method was selected because it allows teachers to model laboratory procedures, explain concepts step-by-step, and maintain consistency in instruction when access to physical laboratories is limited. This approach also reduces safety risks in home-based settings. However, it may limit hands-on manipulation, collaborative inquiry, and immediate feedback, which underscores the need for a validated tool to assess how students perceive the facility of use, satisfaction, and competency development in such laboratory formats.

The evaluated laboratory activity focused on cellular respiration, specifically anaerobic respiration, through the production of wine and yoghurt using readily available household materials. Students viewed a teacher-prepared video demonstration and then performed the activity independently at home before completing the assessment questionnaire.

Participants of the Study

The participants were Grade 12 senior high school students enrolled in General Biology at Holy Angel University during School Year 2021–2022. Although both STEM and General Academic Strand (GAS) students took General Biology, only STEM students were included due to their stronger alignment with laboratory-intensive science instruction. From a population of 733 STEM students, those who met the inclusion criteria were invited to participate.

Eligibility criteria required that participants: (1) were bona fide Grade 12 STEM students; (2) were enrolled during S.Y. 2021–2022; (3) had completed General Biology 1 and 2; and (4) had participated in the video demonstration–based laboratory activity on cellular respiration. Participants were 18 years old or above, of mixed sexes, and had basic access to digital devices and stable internet connectivity. While individual technological proficiency was not formally measured, all participants regularly engaged in online learning activities during the pandemic, ensuring baseline familiarity with video-based instruction and online surveys.

Instrumentation

The primary data collection instrument was a self-administered questionnaire developed specifically for this study. It consisted of two sections aligned with the study's evaluative purpose. The first section included three open-ended questions designed to capture students' qualitative feedback regarding the facility of use, satisfaction, and perceived learning experience associated with the video demonstration method. The second section consisted of 26 Likert-type statements, rated on a four-point scale (1 = strongly disagree to 4 = strongly agree) to avoid neutral responses. This section measured three dimensions: (1) facility of use, (2) student satisfaction, and (3) laboratory competencies grounded in science process skills. Item development was guided by a synthesis of literature on online laboratory learning (e.g., Corcoran et al., 2009) and established frameworks on science process skills. Core competencies such as observing, interpreting results, following procedures, and relating theory to practice were systematically mapped to questionnaire items to ensure conceptual alignment.

Validity and Reliability

Content validity was established through expert review by a panel of science educators with doctoral-level qualifications in education. Experts evaluated item clarity, relevance, and alignment with the intended measurement dimensions. A pilot test involving 12 Grade 12 GAS students was conducted to assess clarity, preliminary reliability, and feasibility. While modest in size, this pilot served as an initial screening step. Full-scale validity and reliability were subsequently examined using data from the main sample, ensuring more robust psychometric evaluation through factor analysis and internal consistency testing.

Data Collection

After institutional approval and informed consent, data were collected online via Google Forms. Participants were given three weeks to complete the questionnaire, requiring approximately 15–20 minutes. Responses were anonymous to protect confidentiality. After adequate responses were obtained, the form was deactivated, and data were securely downloaded, stored, and handled in accordance with Holy Angel University's research data management policies.

Data Analysis

Data were analyzed using SPSS version 18.0. Descriptive statistics (means, standard deviations, and percentages) were used to summarize participant responses. Cronbach's alpha assessed internal consistency reliability. Kaiser–Meyer–Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity evaluated the suitability of the data for factor analysis. Exploratory factor analysis was conducted to examine the underlying structure of the instrument and to identify items with weak psychometric properties. This analytical approach supports tool development by refining its construct clarity rather than determining learning gains.

Scope and Limitations

This study assessed students' perceptions and self-reported competencies, not direct learning outcomes or performance-based laboratory skills. Therefore, conclusions are limited to the validation of an evaluation tool, not the causal effectiveness of video demonstration laboratories on achievement.

RESULTS & DISCUSSION

Results

This chapter presents the analyzed data gathered from the research instrument. The results interpreted from this chapter were essential to address the specific problem statements, which are as follows:

Perceptions of Grade 12 STEM Students from S.Y. 2021-2022 on the Video Presentation Method used for the Conduct of the General Biology Laboratory Activity "Cellular Respiration."

Table 1. Perception of Grade 12 STEM Students from S.Y. 2021-2022 on the facility of the video presentation method used for the conduct of the General Biology laboratory activity, Cellular Respiration

Response	Percentage
Accessible	.5
Astonishing	.5
Clean	.5
Concise	.5
Convenient	.5
Easy	1
Educational	.5
Effective	1.5
Equipped	.5
Excellent	2
Fitting	.5
Fun	1.0

Good	62.5
Great	16.1
Informative	.5
Innovative	.5
Interactive	.5
Interesting	.5
Nice	1.0
Okay	6.3
Orderly	1
Practical	.5
Prudent	.5
Unique	.5
Useful	.5
Total	100.0

The respondents stated their perceptions of the facility of the method used for the laboratory activity on Cellular Respiration. The tallied responses above show that the item with the highest percentage is Good at 62.5% - a value greater than half of the sample size, considered a majority of the respondents. The remaining respondents were positive affirmations regarding the facility of the method used for the said laboratory activity.

Table 2. Perception of Grade 12 STEM Students from S.Y. 2021-2022 on the satisfaction of the video presentation method used for the conduct of the General Biology laboratory activity, Cellular Respiration

Response	Percentage
Adequate	1.0
Average	.5
Contented	.5
Decent	.5
Effective	.5
Excellent	1.0
Fair	.5
Good	21.5
Great	4.4
Lacking	.5
Moderate	.5
Nice	2.0
Okay	5.9
Pleased	.5
Really Satisfied	1.0
Satisfied	58.8
Understanding	.5
Very Good	.5
Total	100.0

When participants were surveyed regarding their satisfaction with the instructional approach employed during the Cellular Respiration laboratory activity, most of them perceived it as satisfactory (58.8%). Most of the remaining responses were favorable terms. However, a recorded response indicated a need for improvement, stating that the method was lacking, with a minimal percentage of 0.5%.

Table 3. Perception of Grade 12 STEM Students from S.Y. 2021-2022 on the competencies from the video presentation method used for the conduct of the General Biology laboratory activity, Cellular Respiration

Response	Percentage
Adequate	.5
Attentive	.5
Clear	1.0
Communication	.5
Competent	.5
Countless	.5
Encouraging	.5
Enough	.5
Excellent	1.5
Good	38.1
Great	26.3
Helpful	1.5
Informative	.5
Memorable	.5
Nice	2.0
Okay	19.5
Outstanding	.5
Practical	1.5
Remarkable	.5
Sufficient	.5
Timely	.5
Useful	2.5
Total	100.0

Upon looking into the respondents' perspective regarding the learning competencies they have gained from the method used for the laboratory activity about Cellular Respiration, 38.1% responded Good, 26.3% mentioned that it was Great, and 19.5% said Okay. No response has reached a majority percentage, and all responses are considered positive feedback.

Validity and Indicators Formulated from the Proposed Tool

Table 4. Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy Test and Bartlett's Test of Sphericity Test

Test		Value	Interpretation
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.912	Excellent sampling adequacy
Bartlett's Test of Sphericity	Approx. Chi-Square	11154.413	
	df	325	
	Sig.	.000	significant

Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy tests the strength of the correlation between the variables, a necessary assumption for Factor Analysis. It can be observed that the result is at KMO = 0.91, interpreted as excellent sampling adequacy. This verbal interpretation for the KMO interpretation was derived from the study of Blbas, Hazhar & Kadir, and Dler (2019) entitled "An Application of Factor Analysis to Identify the Most Effective Reasons that University Students Hate to Read Books."

This result from the KMO implies that it is plausible to conduct the analysis. Upon observing the results of Bartlett's Test of Sphericity, where $p<0.01$, this implies that the data set is not an identity matrix, thus, is suitable for data reduction. Because of satisfying the assumptions, a Principal Factor Analysis was conducted.

Table 5. Communalities Table

	Initial	Extraction
Item1	1.000	.876
Item2	1.000	.822
Item3	1.000	.350
Item4	1.000	.698
Item5	1.000	.870
Item6	1.000	.885
Item7	1.000	.887
Item8	1.000	.864
Item9	1.000	.928
Item10	1.000	.829
Item11	1.000	.819
Item12	1.000	.897
Item13	1.000	.911
Item14	1.000	.589
Item15	1.000	.740
Item16	1.000	.910
Item17	1.000	.920
Item18	1.000	.907
Item19	1.000	.904
Item20	1.000	.894
Item21	1.000	.922
Item22	1.000	.908
Item23	1.000	.922
Item24	1.000	.910
Item25	1.000	.903
Item26	1.000	.895

Extraction Method: Principal Component Analysis.

Looking at the communalities Table 5, it can be observed that one item is at 0.350 (item 3). This outlier indicates that this item should be removed to increase the instrument's validity and reliability. Based on the communalities table, a Principal Component Analysis shall still yield a valid result and be used to determine which items are deemed for removal.

Table 6. Principal Component Analysis

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
ss1	20.103	77.318	77.318
2	1.860	7.152	84.470
3	.899	3.456	87.926
4	.605	2.327	90.253
5	.405	1.558	91.811
6	.310	1.192	93.003
7	.246	.946	93.949
8	.223	.857	94.806
9	.214	.825	95.631
10	.176	.678	96.309

11	.146	.561	96.869
12	.142	.546	97.415
13	.129	.496	97.912
14	.125	.482	98.394
15	.085	.328	98.721
16	.073	.281	99.003
17	.060	.229	99.232
18	.051	.195	99.427
19	.037	.144	99.571
20	.028	.109	99.680
21	.027	.103	99.783
22	.016	.060	99.843
23	.014	.055	99.898
24	.013	.050	99.947
25	.008	.030	99.978
26	.006	.022	100.000

The Principal Component Analysis was used as the extraction method in determining the number of components being measured by the instrument. The table reflects two identifiable components in which the computed eigenvalues are more significant than 1: λ=20.103 and λ=1.86. The variance of the first component is 77.318%, while the second is at 7.152%. The cumulative percentage for which the identified factors account for variance is 84.47%.

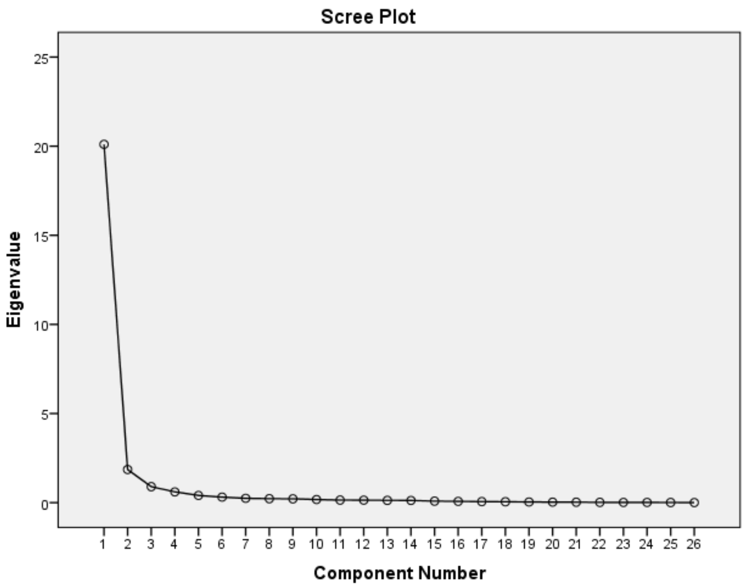

Figure 1. Scree Plot of Eigenvalues and Component Number Items

The Scree plot reflects the eigenvalues when plotted as a line graph. It can be seen that the dip is on the second component, thus supporting the claim that the instrument measures two components, accounting for 84.47% of the total variance obtained.

Table 7. Component Matrix

	Component	
	1	2
Item1	.180	.918
Item2	.877	-.228
Item3	-.564	.180
Item4	.592	.590
Item5	.932	-.011
Item6	.940	-.029
Item7	.940	-.063
Item8	.929	.013
Item9	.963	-.029
Item10	.911	.004
Item11	.905	.022
Item12	.947	-.013
Item13	.955	-.015
Item14	.767	-.015
Item15	.426	.747
Item16	.953	-.047
Item17	.959	-.024
Item18	.952	-.021
Item19	.950	-.040
Item20	.944	-.050
Item21	.959	-.039
Item22	.953	-.015
Item23	.959	-.047
Item24	.953	-.052
Item25	.949	-.047
Item26	.945	-.042

Extraction Method: Principal Component Analysis.
a. 2 components extracted.

The factor loading shows that Component 1 comprises most of the loadings. However, it can be noted that items 3 and 4 have lower values as compared to the other items for both components. The computed factor loading for the second component suggests that this component may be disregarded since the variance accounted for by component 1 is still a significant percentage of 77.318%.

Table 8. Rotated Component Matrix

	Component	
	1	2
Item1		.934
Item2	.907	
Item3	-.591	
Item4	.425	.719
Item5	.906	
Item6	.918	
Item7	.926	
Item8	.897	
Item9	.939	
Item10	.881	
Item11	.871	
Item12	.920	

Item13	.928	
Item14	.746	
Item15		.830
Item16	.935	
Item17	.935	
Item18	.927	
Item19	.930	
Item20	.927	
Item21	.939	
Item22	.926	
Item23	.941	
Item24	.935	
Item25	.931	
Item26	.926	

Table 8 shows the rotated component matrix of the research tool's original twenty-six (26) items. It can be said that most of the loadings are from component one (1) since 23 of the actual items are categorized under it. Only three (3) items are ordered under component two (2). In addition, this varimax rotation was conducted to redistribute the factor loadings to address the cross-loadings. It can be observed that only item 3 has a negative loading on component 1. Only item 4 has cross-loading for components 1 and 2. Again, most of the items measure component 1. Items 1 and 15 have component 2 loadings, implying that these measure the second component.

Reliability of the Proposed Tool

Table 9. Cronbach's Alpha Analysis on Pilot Test Results

Cronbach's Alpha	Verbal Interpretation
0.92	Very Reliable

An initial analysis was conducted using Cronbach's alpha to check the instrument's reliability, treating the data from the pilot test. It can be seen from the table that the value obtained is α=0.922, which is verbally interpreted as Very Reliable. This implies that all instrument items have an acceptable reliability level. The verbal interpretation of the result for Cronbach's alpha is derived from Ahdika and Atina (2017).

Table 10. Cronbach's Alpha Analysis on Pilot Test Results

Cronbach's Alpha	Verbal Interpretation
0.973	Very Reliable

Cronbach's alpha was again conducted to test the reliability of the post-test results using the instrument. The table shows that the computed α=0.973 can be interpreted as Very Reliable. This indicates that the existing items on the instrument contribute to its general reliability. Similarly, from the pre-test results, the verbal interpretation of the result of Cronbach's alpha of the post-test is derived from Ahdika and Atina (2017).

Reliability and Validity of the Proposed Tool as a Tool in Assessing the Video Demonstration Presentation Laboratory Activity in General Biology

Table 11. Speculative Figures

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Item1	107.4537	220.435	.220	.975
Item2	107.2049	210.968	.832	.972
Item3	109.8439	242.544	-.535	.985
Item4	107.3805	203.796	.590	.975
Item5	107.3220	205.827	.918	.971
Item6	107.2634	206.568	.917	.971
Item7	107.2878	206.628	.918	.971
Item8	107.3122	205.245	.918	.971
Item9	107.2780	206.212	.948	.971
Item10	107.3415	205.157	.903	.971
Item11	107.3415	205.353	.894	.971
Item12	107.3171	205.365	.940	.971
Item13	107.3073	205.655	.943	.971
Item14	107.4439	206.081	.753	.972
Item15	107.4098	215.145	.463	.974
Item16	107.2732	206.229	.936	.971
Item17	107.2927	205.953	.944	.971
Item18	107.3024	205.516	.939	.971
Item19	107.2829	206.037	.938	.971
Item20	107.2780	206.202	.928	.971
Item21	107.2683	206.501	.942	.971
Item22	107.2732	205.886	.934	.971
Item23	107.2829	206.380	.940	.971
Item24	107.2537	206.386	.936	.971
Item25	107.2780	206.143	.932	.971
Item26	107.2683	206.354	.929	.971

Although all items in the test are deemed acceptable, a further analysis was conducted to see if there were items that may be considered for revision or, at best, for removal. Based on the generated values above, it can be said that if Items 1, 3, 4, and 15 are revised or removed, there is a slight increase in the value of Cronbach's alpha. If Item 1 is removed, $\alpha=0.975$ will be the value of Cronbach's alpha. For Item 3, $\alpha=0.985$. Whereas for Item 4, the value will be $\alpha=0.975$. Revising or removing Item 15 will yield a new value at $\alpha=0.974$. These values are supported by the observed Corrected Item-Total correlation, which is lower than the average values of the other items on the research instrument. Item 1=0.22; Item 3=-0.535; Item 4=0.590; and Item 15=0.463. The action to revise or remove these items can be anchored from Sauro, Jeff (2017) in the article How to Know Which Items to Remove in a Questionnaire. The conduct of Factor Analysis to determine which items to remove or revise is essential for this study. In addition, Scree plots are necessary to show the visual representation of the items included in the proposed research tool.

Table 12. Cronbach's Alpha Analysis, excluding Items 1,3,4 and 15

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.993	.993	22

A Cronbach's Alpha reliability test was further conducted after removing items 1, 3, 4, and 15. It can be observed from the table that the computed value of α=0.993, which is Very Reliable, is higher than the pilot test and the initial post-test result values of α=0.922 and α=0.973, respectively. This implies that removing these items increased the tool's reliability.

Table 13. Standard Deviation Table

	Mean	Std. Deviation	Verbal Interpretation
Item1	4.2488	.65795	Strongly Agree
Item2	4.4976	.57415	Strongly Agree
Item3	1.8585	1.11781	Disagree
Item4	4.3220	1.18560	Strongly Agree
Item5	4.3805	.71511	Strongly Agree
Item6	4.4390	.68774	Strongly Agree
Item7	4.4146	.68512	Strongly Agree
Item8	4.3902	.73691	Strongly Agree
Item9	4.4244	.67909	Strongly Agree
Item10	4.3610	.75176	Strongly Agree
Item11	4.3610	.75176	Strongly Agree
Item12	4.3854	.71591	Strongly Agree
Item13	4.3951	.70362	Strongly Agree
Item14	4.2585	.84973	Strongly Agree
Item15	4.2927	.70168	Strongly Agree
Item16	4.4293	.68680	Strongly Agree
Item17	4.4098	.69162	Strongly Agree
Item18	4.4000	.71125	Strongly Agree
Item19	4.4195	.69283	Strongly Agree
Item20	4.4244	.69338	Strongly Agree
Item21	4.4341	.67287	Strongly Agree
Item22	4.4293	.70093	Strongly Agree
Item23	4.4195	.67853	Strongly Agree
Item24	4.4488	.68138	Strongly Agree
Item25	4.4244	.69338	Strongly Agree
Item26	4.4341	.68728	Strongly Agree

This table reflects the mean and standard deviation of each item in the tool. It can be observed that there is an outlier value with a mean of $\bar{x}$=1.85 interpreted as Disagree. This is the obtained value for item 3, which is congruent with the other findings that state it needs to be removed. The remaining items are all interpreted as Strongly Agree.

Discussion

The perceptions of Grade 12 STEM students toward the video demonstration method for the General Biology laboratory activity on Cellular Respiration were examined across three dimensions: facility of use, student satisfaction, and perceived laboratory competencies. In this study, facility of use refers to the extent to which students found the video demonstration clear, accessible, and easy to follow when conducting the laboratory task independently. Results showed that a majority of responses were classified as good

(62.5%), suggesting that most students were able to understand procedural steps and replicate the activity using household materials. This finding aligns with previous studies indicating that video demonstrations support procedural clarity and reduce cognitive load in online science laboratories (Babincakova & Bernard, 2020).

Student satisfaction pertains to learners' overall approval of the instructional approach, including engagement, clarity of instruction, and perceived usefulness. More than half of the respondents (58.8%) rated their experience as excellent. The small proportion of lacking responses (0.5%) suggests that while the method was generally effective, some students may have experienced limitations such as insufficient interactivity or limited opportunities for feedback, consistent with findings reported by Gross (2020).

Perceived laboratory competencies refer to students' self-assessment of science process skills such as observing, following procedures, interpreting results, and connecting theory with practice. Although responses were distributed across categories, the highest proportions were classified as good (38.1%) and great (26.3%). The absence of a dominant category indicates varied learner experiences, which is expected in self-reported competency measures and reinforces the importance of using validated instruments when evaluating laboratory instruction (Bretz, 2019).

Construct Validity and Factor Structure of the Tool

Construct validity testing demonstrated strong psychometric properties. The Kaiser–Meyer–Olkin (KMO) value of 0.91 indicated excellent sampling adequacy, and Bartlett's Test of Sphericity was significant ($p < .01$), confirming the suitability of the data for factor analysis. Exploratory factor analysis revealed two components accounting for 84.47% of the total variance, with the first component explaining 77.32%, indicating a dominant underlying construct related to the evaluation of video-based laboratory activities.

Items 1, 3, 4, and 15 demonstrated lower communalities and weaker factor loadings, indicating limited contribution to construct clarity. The identification of these items supports evidence-based refinement of the instrument rather than arbitrary exclusion.

Reliability of the Evaluation Tool

The reliability analysis further supported the robustness of the instrument. Cronbach's alpha coefficients indicated high internal consistency for both the pilot test ($\alpha = 0.92$) and the post-test ($\alpha = 0.97$). These values exceed commonly accepted thresholds for educational research, indicating that the evaluation tool consistently measures the targeted dimensions. Removal of items with weak psychometric performance further improved internal consistency, reinforcing the strength of the finalized instrument.

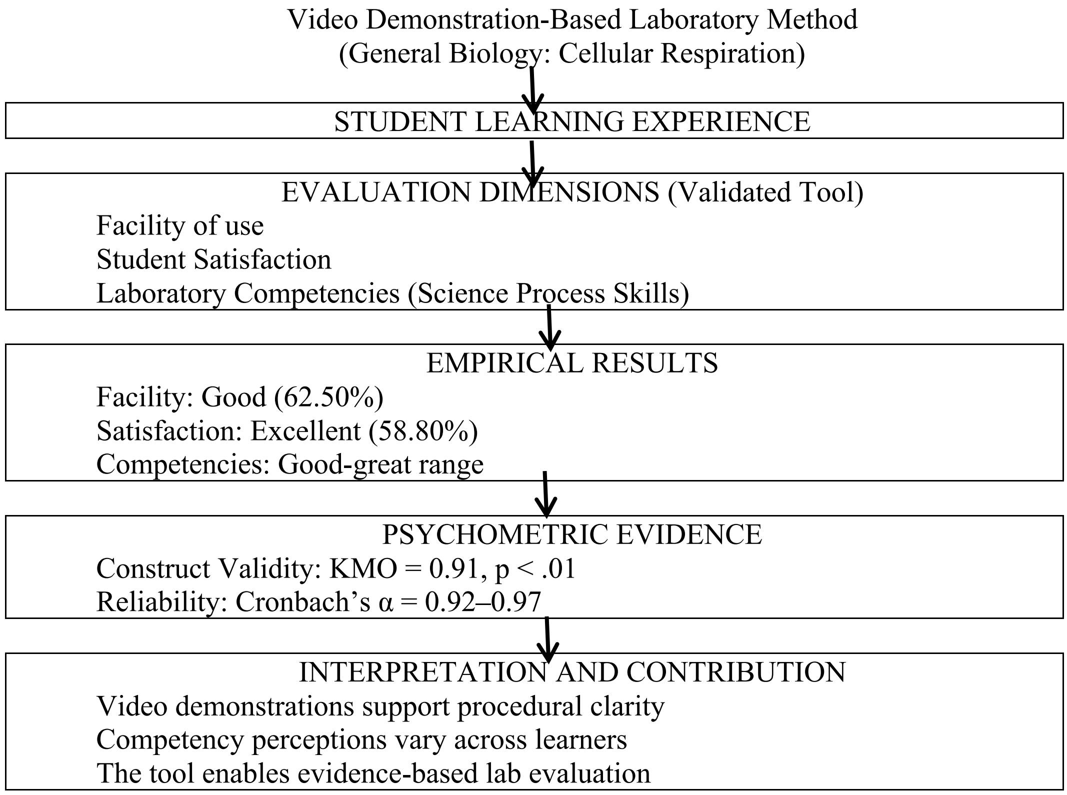

Figure 2. Results–Discussion integration table

CONCLUSION

This study developed and validated an evaluation tool for assessing video demonstration–based laboratory activities in General Biology, specifically for the topic of Cellular Respiration among Grade 12 STEM students at Holy Angel University. Findings show that students generally perceived the method as easy to use, satisfactory, and supportive of laboratory competency development, though variation in competency perceptions highlights the need for systematic evaluation tools rather than reliance on student impressions alone.

Psychometric analyses provided strong evidence of the tool's validity and reliability, as demonstrated by excellent sampling adequacy (KMO = 0.91), a coherent factor structure, and high internal consistency ($\alpha = 0.92–0.97$). These results confirm that the instrument accurately measures the intended dimensions of facility of use, student satisfaction, and science process–based laboratory competencies.

While this study was limited to one laboratory topic and institution, the validated tool may serve as a reference framework for evaluating similar video-based laboratory activities in biology and other STEM disciplines. Future studies may extend this work by applying the tool across multiple topics, institutions, and instructional modalities, as well as by integrating performance-based measures to complement self-reported competencies.

REFERENCES

Abbey, B. (2020). *Lab experiments moved online or mailed*. La Trobe University. https://www.latrobe.edu.au/news/articles2020/opinion/lab-experiments-moved-online-or-mailed

Ahdika, A. (2017). Improvement of quality, interest, critical, and analytical thinking ability of students through the application of research-based learning in introduction to stochastic processes. *International Electronic Journal of Mathematics Education, 12*, 167–191. https://doi.org/10.29333/iejme/608

Al Natour, S., & Woo, C. (2021). Determinants of learner satisfaction with online video presentation methods. *Internet Research, 31*(1), 234–261. https://doi.org/10.1108/INTR-04-2019-0155

American Chemical Society. (2019). *The importance of hands-on laboratory science*. https://www.acs.org/content/acs/en/policy/publicpolicies/education/computersimulations.html

Aronne, L., Nagle, C., Styers, J., Combs, A., & George, J. A. (2019). The effects of video-based pre-lab instruction on college students' attitudes and achievement. *Electronic Journal of Science Education*. https://files.eric.ed.gov/fulltext/EJ1234449.pdf

Babincakova, M., & Bernard, P. (2020). Online experimentation during COVID-19 secondary school closures: Teaching methods and student perceptions. *Journal of Chemical Education, 97*(9), 3295–3300. https://doi.org/10.1021/acs.jchemed.0c00748

Bretz, S. L. (2019). Evidence for the importance of laboratory courses. *Journal of Chemical Education, 96*(2), 193–195. https://doi.org/10.1021/acs.jchemed.8b00874

Chaytor, J. L., Al-Mughalaq, M., & Butler, H. (2017). Development and use of online prelaboratory activities in organic chemistry. *Journal of Chemical Education, 94*(7), 859–866.

Cobb, S., Corcoran, O., Heaney, R., & Henderson-Begg, S. (2009). Learning gains and student perceptions of a Second Life virtual lab. *Bioscience Education, 13*. https://doi.org/10.3108/beej.13.5

Cresswell, S. L., Loughlin, W. A., Coster, M. J., & Green, D. M. (2019). Development and production of interactive videos for teaching chemical techniques. *Journal of Chemical Education, 96*(5), 1033–1036. https://doi.org/10.1021/acs.jchemed.8b00647

Gamage, K. A. A., Wijesuriya, D. I., Ekanayake, S. Y., Rennie, A. E. W., Lambert, C. G., & Gunawardhana, N. (2020). Online delivery of teaching and laboratory practices during COVID-19. *Education Sciences, 10*(10), 291. https://doi.org/10.3390/educsci10100291

Gross, H. (2020). *With lab courses shifting online, students worry about losing hands-on experience*. The Daily Pennsylvanian. https://www.thedp.com/article/2020/03/online-labs-penn-student-reactions-chemistry-biology-physics

Hofstein, A., & Naaman, R. (2007). The laboratory in science education: The state of the art. *Chemistry Education Research and Practice, 8*(2), 105–107.

Iuliano, E., et al. (2021). Student satisfaction with online workshops replacing practice-oriented activities. *Education Sciences, 11*(10), 600. https://doi.org/10.3390/educsci11100600

Kapilan, N., Vidhya, P., & Gao, X.-Z. (2020). Virtual laboratory as a boon to engineering education during COVID-19. *Higher Education for the Future, 8*(1), 31–46. https://doi.org/10.1177/2347631120970757

Kay, R., Goulding, H., & Li, J. (2018). Assessing the impact of a virtual lab in an allied health program. *Journal of Allied Health, 47*(1), 45–50.

Kovacs, M., Holcombe, A., Aust, F., & Aczel, B. (2021). The importance of tool development for research efficiency. *Information Services & Use, 41*(1–2), 123–130. https://doi.org/10.3233/ISU-210109

Padilla, M. J. (2022). *Science process skills*. NARST. https://narst.org/research-matters/science-process-skills

Petillion, R. J., & McNeil, W. S. (2021). Student satisfaction with synchronous online organic chemistry laboratories. *Journal of Chemical Education, 98*(9), 2861–2869. https://doi.org/10.1021/acs.jchemed.1c00549

Safaah, E. S., Muslim, M., & Liliawati, W. (2017). Teaching science process skills using the learning cycle. *Journal of Physics: Conference Series, 895*, 012106.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55.

Yildirim, F. S. (2021). The effect of virtual laboratory applications on students' achievement. *Journal of Education in Science, Environment and Health, 7*(2), 171–181.