DATA NOTE

REVISED Dataset for measuring the conceptual understanding of optics in Rwanda [version 2; peer review: 2 approved]

Kizito Ndihokubwayo[1*], Michael Ralph[2*], Irénée Ndayambaje[1*], Jean Uwamahoro[1*]

[1]African Center of Excellence for Innovative Teaching and Learning Mathematics and Science (ACEITLMS), University of Rwanda College of Education (URCE), Kayonza, P.O Box 55 Rwamagana, Rwanda
[2]Department of Educational Psychology, University of Kansas, Joseph R. Pearson Hall, Rm. 621, 1122 West Campus Rd. Lawrence, Kansas, 66045-3101, USA

* Equal contributors

v2 **First published:** 28 Jul 2021, **10**:679
https://doi.org/10.12688/f1000research.53135.1
Latest published: 10 May 2022, **10**:679
https://doi.org/10.12688/f1000research.53135.2

Abstract

This dataset is an accumulation of data collected to test Rwandan physics students' conceptual understanding of light phenomena and to assess instructional tools for active learning of optics. We collected and analysed data from 251 grade 11 (senior 5) students using our Light Phenomena Conceptual Assessment (LPCA) tool and from 136 grade 10 (senior 4) students using Geometric Optics Conceptual Understanding Test (GOCUT) in 2019. Before collecting data, we designed and validated LPCA and GOCUT, and tested their reliability. Data were collected before and after students learnt about the unit of light. Both day and boarding schools in rural and urban areas were included in our sampling. Data collected were test scores from students after performing a 30-item LPCA test or 25-item GOCUT test in 40 minutes. The data may be reused to extend students' understanding of optics concepts through item analysis, analysis of school characteristics such as location and school type, or by analysing students' characteristics such as subject combinations.

Keywords

light phenomenon, optics, conceptual understanding, Rwandan students

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 10 May 2022	✓ view	✓ view
version 1 28 Jul 2021	? view	? view

1. **Hassen Ghalila**, University of Tunis El Manar, Tunis, Tunisia
University of Tunis El Manar, Tunis, Tunisia

2. **Imelda Kemeza**, Mbarara University of Science and Technology, Mbarara, Uganda

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Kizito Ndihokubwayo (ndihokubwayokizito@gmail.com)

Author roles: Ndihokubwayo K: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ralph M**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Ndayambaje I**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Uwamahoro J**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The African Centre of Excellence for Innovative Teaching and Learning Mathematics and Science (ACEITLMS), based at the University of Rwanda College of Education (URCE), fully funded the first author for studies, data collection, student exchange program, and publications.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Ndihokubwayo K, Ralph M, Ndayambaje I and Uwamahoro J. **Dataset for measuring the conceptual understanding of optics in Rwanda [version 2; peer review: 2 approved]** F1000Research 2022, **10**:679
https://doi.org/10.12688/f1000research.53135.2

First published: 28 Jul 2021, **10**:679 https://doi.org/10.12688/f1000research.53135.1

REVISED **Amendments from Version 1**

Some typos were amended. The major content remained the same.

Any further responses from the reviewers can be found at the end of the article

Introduction

Assessment inventories data provide insights into the classroom atmosphere and show students' progress in grasping certain concepts, and these are essential for teachers, educationists, educational evaluators, and researchers. Such inventories may be used to test students' understanding of a certain concept or may be used to test the effectiveness of a particular teaching approach or instructional tool. This dataset is an accumulation of data collected for the first author's doctoral research project "*Assessment of Instructional Tools for Active Learning of Optics at Advanced Level Secondary Schools in Rwanda*".[1–6] In this project, there was a need to first assess students' conceptual understanding of light phenomena and, second, to assess the effectiveness of instructional tools (such as University of Colorado Boulder's interactive PhET simulations, and YouTube videos) to improve the learning of optics. Thus, this article presents data from two different inventories or tests that were designed. (i) the Light Phenomena Conceptual Assessment (LPCA) and (ii) Geometric Optics Conceptual Understanding Test (GOCUT). Both these datasets are useful to researchers that will use LPCA, or GOCUT data, or to those who want to understand Rwandan physics students' performance. This will enable researchers to reanalyse the data in different contexts, such as item analysis theory, comparing school characteristics such as students' performance in day schools compared to boarding schools, comparing rural schools to urban schools, analysing subject combinations, etc. In this vein, LPCA data are discussed in detail to guide research practitioners on how students' performance and test item performance-related data are analysed.

The study describing the development of the LPCA tool and its implementation was published in Physics Education (PED)[6] and the LPCA study instrument is available on protocols.io and Physport platform. The LPCA is a conceptual understanding test composed of 30 items addressing geometric and physical optics. It was designed based on students' misconceptions related to the everyday understanding of light phenomena. The data connected to this tool are available in *Underlying data*[7] and were listed and analysed in a Microsoft Excel file titled 'Pre-Post-Test LPCA Data - Senior 5 Rwandan physics students'. This file contains three sheets; the first sheet presents the pre-test data, the second sheet presents the post-test data, while the third sheet contains filtered data (students who performed both the pre- and post-test). The data comprises various students' backgrounds; rural and urban schools, boarding and day schools, and different subject combinations (see Table 1 in Methods section).

The study describing the development and implementation of the GOCUT tool was published in the African Journal of research in Mathematics, Science and Technology education (AJRMSTE).[3] The revised protocol where rote learning-related items were removed, is also available in.[2] The GOCUT is a conceptual understanding test composed of 25 items of geometric optics. It was designed based on various existing inventories. The data connected to the GOCUT study are available in *Underlying data*[8] and were listed and analysed in a Microsoft Excel file titled 'Pre-Post-Test GOCUT Data - Senior 4 Rwandan physics students'. This file contains seven sheets; the first sheet introduces the data collected, while other sheets present pre-test and post-test data for three groups of instructional tools of intervention (control group, PhET simulations group, and YouTube videos group). The data comprises various students' backgrounds; rural and urban schools, boarding and day schools, and different subject combinations (see Table 2 in Methods section).

Methods

Data collection

LPCA

A total of eight Rwandan secondary schools were involved in the study. We selected two districts in Kigali city, and two districts in the rural Eastern Province. We listed the schools in those four districts, and chose two schools from each district that accommodated physics in their subject combinations. School characteristics, location, and type of school (School 1 to School 4 are from Kigali, while School 5 to a School 8 were from the eastern province, see Figure 1) were considered during the selection process. These school characteristics, location, and type of school were considered during the selection process so as to include a diverse group of students and to avoid any potential sources of bias.

We employed a pre- and post-test design[9] to collect the data for measuring students conceptual understanding of optics-related concepts. The LPCA was administered twice to the students via paper form, before and after learning about the unit of light in senior-5.[10] A total of 251 students from grade 11 or senior 5 (S5) were considered the final sample after

Table 1. Characteristics of data for the Light Phenomena Conceptual Assessment (LPCA) implemented at S5 or Grade 11. Note: PCM: Physics-Chemistry-Mathematics, PCB: Physics-Chemistry-Biology, MPC: Mathematics-Physics-Computer science, MPG: Mathematics-Physics-Geography

s/n	School	Class	Combination	No of students at Pre-test	No of students at Post-test	Filtered
1	School 1	3	PCM, PCB, MPC	23, 41, 43	25, 43, 39	19, 39, 36
2	School 2	1	PCB	26	30	23
3	School 3	1	MPG	32	31	30
4	School 4	1	MPG	16	19	16
5	School 5	1	PCB	27	24	23
6	School 6	1	PCB	44	40	38
7	School 7	1	MPG	19	16	16
8	School 8	1	MPG	12	11	11
Total				283	278	251

Table 2. Characteristics of data from the Geometric Optics Conceptual Understanding Test (GOCUT) implemented at S4 or Grade 10. Note: PCM: Physics-Chemistry-Mathematics, PCB: Physics-Chemistry-Biology, MPC: Mathematics-Physics-Computer science, MPG: Mathematics-Physics-Geography

School	School location	School type	Subject combination	Teaching intervention	No of students at Pre- and Post-test
School 1	Urban	Boarding	PCM	YouTube videos	30
School 1	Urban	Boarding	PCB	PhET simulations	36
School 2	Urban	Day	PCB	Traditional methods	29
School 3	Urban	Boarding	MPG	Traditional methods	16
School 4	Urban	Day	MPG	PhET simulations	9
School 7	Rural	Boarding	MPG	YouTube videos	16
Total					136

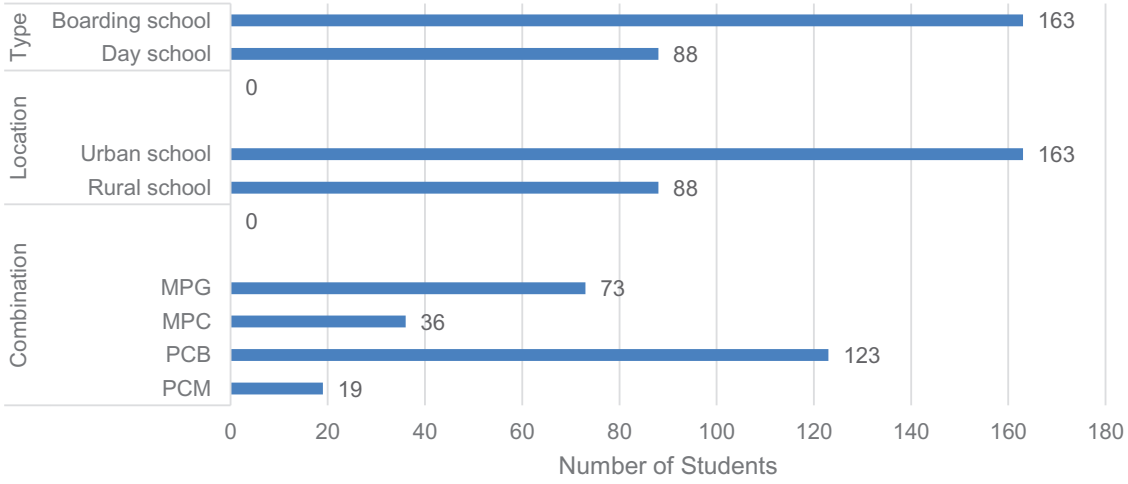

Figure 1. Number of students who carried out the Light Phenomena Conceptual Assessment (LPCA), according to type and location of school, and subject combination characteristics. PCB: Physics-Chemistry-Biology, MPG: Mathematics-Physics-Geography, PCM: Physics-Chemistry-Mathematics, MPC: Mathematics-Physics-Chemistry.

removing those who sat for pre-test and missed post-test, and vice versa (see Table 1). The methods for the data coding are presented in the Analysis section. These students had no other teaching interventions offered apart from usual teaching.

GOCUT

The boarding and day secondary schools chosen to be involved in the GOCUT were the same as for the LPCA (schools from rural areas were sampled from Eastern Province, while those from urban areas were sampled from Kigali city). However, three schools were excluded due to ineffectiveness of implementing the designed intervention. Thus, researchers were not able to implement the intervention at these schools. Students were from grade 10 or senior 4 (S4), with various subject combinations. PCB: Physics-Chemistry-Biology, MPG: Mathematics-Physics-Geography, PCM: Physics-Chemistry-Mathematics. Table 2 displays characteristics of school and students in which the instructional tools were implemented and GOCUT was administered.

Teaching interventions of PhET simulations and/or videos compiled on YouTube were offered (see Table 2) to the students. Details of the YouTube videos, including the names of any companies/institutions responsible for creating the materials are available in[3] p. 257). GOCUT was administered twice to the students via paper form, before and after learning about geometric optics via the teaching interventions in senior-4.[10] A total of 136 students from grade 10 or senior 4 (S4) were involved in the study (see Table 2).

The data were initially (pre-test) collected in January 2019 and finally (post-test) at the end of March 2019. The answer choices for GOCUT are A, B, C, and D. These choices measure the students' conceptual understanding of optics, where one is stem (correct answer) while other three choices are distractors (wrong answers). Where the student did not answer, N is coded, while where the student answered more than one answer, T is coded. For the drawing question (item 13), C was coded for students who correctly drew, while W was coded for those who wrongly drew. For the explanatory question (item 9), the extended explanation was provided in the column after AH, after the drawing question.

Analysis

This section presents the step-by-step analysis of the LPCA data. We took the case of the first inventory (light phenomena conceptual assessment, LPCA) to extend the description of analysis to help research practitioners in educational research get insight into performance and conceptual understanding test analysis. Please note that unlike the LPCA data file, the file for the GOCUT does not provide accumulated or detailed analysis. Nevertheless, LPCA and GOCUT are similar in manner; their data were recorded and arranged in the same way, so the explanation of how we analysed LPCA data may be used to analyse the GOCUT data.

We used Microsoft Excel 2016 to analyse the data. Since the LPCA test was a multiple-choice test (except for item 11 which requests a supporting explanation), each item has four choices—from A to D. We recorded this data in an Microsoft Excel sheet by putting an assigned letter to each item (A, B, C, or D). Where a student assigns more than one answer, we recorded "T" while where the student selects nothing or skips the question; we recorded "N."

The first analysis was to use "COUNTIF" function to count the number of students who answered each letter; the sum should be the total number of students (see, for example, in pre- or post-test sheet, column F, row 4-10). The second analysis was to mark students by giving a score of "1" to everyone who answered each item correctly (who chose the right answer) and by giving a score of "0" to those who selected the wrong answer, did not answer, or selected more than one answer. We use "IF" and "EXACT" functions (see, for example, column AM, row 15). After computing these functions for each student, we summed the total scores for each student (see column BR) and the corresponding percentage scores (see column BS). These percentage scores show the students' performance (scores received by every student over the whole LPCA test). A histogram was computed to check the normal distribution of the test scores (number of students in each assigned interval of scores, please see column BU-CG). The significance of performance before and after learning optics was computed in the filtered sheet (see column W-AA).

The third analysis was item analysis. See the bottom of "IF" and "EXACT" analysis on row 299 in pre-test sheet, for example. The sum of scores for each item of LPCA was computed to reveal the difficulty of the test. A graph was generated showing all 30 items; among them, some are difficult (performed by few students), and others are easy (performed by most of the students). In other words, it was more difficult to perform well in some of the items, and that these items were answered by fewer students. For this analysis, further analysis may generate a graph showing the answer choice for each item (please refer to the *Underlying data*.[7] It shows the number of students who selected every letter for each item. It shows how the correct answer varies from alternative choices and analyses the students' misconceptions. This figure is generated using the records of the first analysis (counted numbers of answers using "COUNTIF" function).

In the filtered sheet, we have filtered the students who sat for both pre- and post-test. This helps for side-by-side analysis of the results and helps to keep each student's scores parallel so that the difference between both test scores is clear. It tracks the performance along with both tests, i.e., whether the students performed better in the post-test or the inverse. If it is inverse, analysis of misconceptions and a revisit of the instructions may be further studied (to understand why the student failed after learning, performing even more worse than he/she performed before learning). We have shown how Cohen's D effect size and Normalised learning gains <g> are computed to measure the impact of instruction (see column W-AA, row 259-269). Effect size is computed by taking the difference of means of post-test and pre-test dividing by the average of standard deviations (see cell Y263). Cohen,[11] Sawilowsky,[12] and Mangnusson[13] interpret "d" of 0.20 as small, 0.50 as a medium, and 0.80 as large. Normalised learning gain <g> is calculated by taking the difference of means of post-test and pre-test, dividing by the maximum mean. The maximum mean is the difference of 100% or highest score and the mean of pre-test scores (check out cell Y264). Hake[14] interprets a <g> of <.3 as small, <g> of.3 to.6 as medium, and large and <g> of >.7 as large.

Validation

The data from both tools are valid and reliable as the tools underwent a rigorous validation and a test-rested reliability was checked before the official use. We first searched the literature for possible misconceptions that students had on the topic of optics and available tests to remedy them. We then drafted questions, using our experiences from the classroom, Rwandan textbooks (in case of LPCA), and existing tests, research articles and textbooks (in the case of GOCUT). We shared the survey questions with four university professors in physics education for content validation (i.e. to check that the questions were testing the real constructs/concepts we intend to evaluate) and to 38 students–selected from two schools from elsewhere, i.e. schools not included in this study—for face validation (i.e. to check the difficulty of questions so as to identify any confusion that may rise). The initial number of questions for each test was above 50 items, after improving them using suggestions from both validators, we reached 30 LPCA items and 25 GOCUT items.

Ethics statement

The study procedure was approved by the ethical committee in the University of Rwanda College of Education's research unit and innovation (permit number: 01/P-CE/483/EN/gi/2018). Ethical clearance was provided after reviewing our research proposal. Our data collection involved secondary school students aged between 16 and 23 years old. Parental consent was not obtained for students under 18 (adult age in Rwanda); however, the study was considered low risk. We explained the purpose of our study to teachers and asked teachers, as well as the students, to sign an informed consent form before partaking in our tests and study. We assured them that the voluntary participation and publication of data would not reveal individual participants' identities. Data were treated confidentially, and we have deleted the students' names from our data to maintain their anonymity. Since the first protocol (LPCA) was fully designed by authors and the second protocol (GOCUT) was designed based on existing tests, there was no special approval obtained from developers, however, we fully credited their sources and works.

Data availability

Underlying data

Mendeley Data: Pre-Post-Test LPCA Data: Senior 5 Rwandan physics students. https://data.mendeley.com/datasets/dbvh59jg7j/1.[7]

This project contains the following underlying data:

- LPCA.pdf (copy of the light phenomena conceptual assessment (LPCA), an inventory test of 30 items)

- Pre-Post-Test LPCA Data - Senior 5 Rwandan physics students.xlsx (MS Excel file that contains the data)

Mendeley Data: Pre-Post-Test GOCUT Data: Senior 4 Rwandan physics students. https://data.mendeley.com/datasets/mmtpw5nvg3/1.[8]

This project contains the following underlying data:

- GOCUT.pdf (copy of the geometric optics conceptual understanding test (GOCUT), an assessment test of 25 items)

- Pre-Post-Test GOCUT Data - Senior 4 Rwandan physics students.xlsx (MS Excel file that contains the data)

Acknowledgments

The LPCA and GOCUT were content validated by Prof Scott Franklin, Rochester Institute of Technology, NY, US, and Prof Eleanor Syre, Kansas State University, KS, US and face validated by students from GS Mukarange, Kayonza, Rwanda, and GS Saint Aloys, Rwamagana, Rwanda. This data article was commented on by Ms. Josiane Mukagihana and Ms. Celine Byukusenge.

References

1. Ndihokubwayo K, Uwamahoro J, Ndayambaje I: **Classroom observation data collected to document the implementation of physics competence-based curriculum in Rwanda.** *Data Br.* 2021; vol. **36**, no. June, p. 107055.
PubMed Abstract | Publisher Full Text | Free Full Text

2. Uwamahoro J, Ndihokubwayo K, Ralph M, *et al.*: **Physics Students' Conceptual Understanding of Geometric Optics: Revisited Analysis.** *J. Sci. Educ. Technol.* 2021; **30**(0123456789): 1–13.
Publisher Full Text

3. Ndihokubwayo K, Uwamahoro J, Ndayambaje I: **Effectiveness of PhET Simulations and YouTube Videos to Improve the Learning of Optics in Rwandan Secondary Schools.** *African J. Res. Math. Sci. Technol. Educ.* 2020; **24**(2): 253–265.
Publisher Full Text

4. Ndihokubwayo K, Uwamahoro J, Ndayambaje I: **Usability of Electronic Instructional Tools in the Physics Classroom.** *EURASIA J. Math. Sci. Technol. Educ.* 2020; **16**(11): 1–10.
Publisher Full Text

5. Ndihokubwayo K, Uwamahoro J, Ndayambaje I: **Implementation of the Competence-Based Learning in Rwandan Physics Classrooms: First Assessment Based on the Reformed Teaching Observation Protocol.** *EURASIA J. Math. Sci. Technol. Educ.* 2020; **16**(9): 1–8.
Publisher Full Text

6. Ndihokubwayo K, Uwamahoro J, Ndayambaje I, *et al.*: **Light phenomena conceptual assessment: an inventory tool for teachers.** *Phys. Educ.* 2020; **55**(3): 035009.
Publisher Full Text

7. Ndihokubwayo K, Uwamahoro J, Ndayambaje I, *et al.*: **Pre-Post-Test LPCA Data: Senior 5 Rwandan physics students.** *Mendeley Data.* 2021. vol. **V1**.
Publisher Full Text

8. Ndihokubwayo K, Uwamahoro J, Ndayambaje I: **Pre-Post-Test GOCUT Data: Senior 4 Rwandan physics students.** *Mendeley Data.* 2021; vol. **1**.
Publisher Full Text

9. Fraenkel JR, Wallen NE, Hyun HH: *How to Design and Evaluate Research in Education.* 8th ed. New York: McGraw Hill; 2012.

10. REB: *Advanced level Physics syllabus.*Kigali: Ministry of Education; 2015.

11. Cohen J: *Statistical Power Analysis for the Behavioral Sciences.* 2nd ed. New York, NY: Lawrence Erlbaum Associates, Publishers; 1988.

12. Sawilowsky SS: **Very large and huge effect sizes.** *J. Mod. Appl. Stat. Methods.* 2009; **8**(2): 597–599.
Publisher Full Text

13. Magnusson K: **Interpreting Cohen's d Effect Size: An Interactive Visualization.** 2021.
Reference Source

14. Hake RR: **Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses.** *Am. J. Phys.* 1998; **66**(1): 64–74.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 19 May 2022

https://doi.org/10.5256/f1000research.133699.r137209

✔ **Imelda Kemeza**
Department of Educational Foundations and Psychology, Mbarara University of Science and Technology, Mbarara, Uganda

Going by the ethics statement:

The authors committed to the best practices for ethics when working with human subjects in this case parents, however, the involvement of students under 18 (Rwandan adult age) on the pretext of the study being considered of low risk seems to weigh in less as a waiver when students under 18 have not reached the legal age of consent to participate in a research activity. This, too, raises questions on the help to develop the students' autonomy and what supports communication between the researcher and the students under 18.

I approve the dataset for measuring the conceptual understanding of optics in Rwanda where it meets with international standards on ethical guidelines.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Educational Psychology with a bias to psychometrics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 12 May 2022

https://doi.org/10.5256/f1000research.133699.r137210

Hassen Ghalila
[1] Department of Physics, Faculty of Sciences, University of Tunis El Manar, Tunis, 1060, Tunisia
[2] LSAMA Laboratory, University of Tunis El Manar, Tunis, Tunisia

I would like to thank the authors for their commitment to this very important issue. Even if they do not fully answer the first question (analyses of the LPCA dataset have already been done in detail in the previous article (Ref.5), it would have been preferable to adapt the manuscript to the results obtained with the supplementary tools used in the GOCUT field), which was also pointed out by the other reviewer, I give my approval for the publication of their article.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Physics of plasma and spectroscopy

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 18 March 2022

https://doi.org/10.5256/f1000research.56488.r124417

Imelda Kemeza
Department of Educational Foundations and Psychology, Mbarara University of Science and Technology, Mbarara, Uganda

The authors' manuscript created a dataset for LPCA and GOCUT for optimized conceptual understanding of light phenomena and also assessed instructional tools for active learning of optics in Physics.

Major concerns:

- In my view, the manuscript authors punctuated the LPCA analysis with appropriate and relevant explanations. A similar treatment to GOCUT analysis would have informed comparisons and discussion of the dual analyses for a better-positioned dataset.

- In the ethics statement: "Our data collection involved secondary school students aged between 16 and 23 years old. Parental consent was not obtained for students under 18 (adult age in Rwanda); however, the study was considered low risk." Whether the study was considered low risk, research ethics are clear on what should be done when participants are under-aged. In this case, where some students were aged 16 and 17 years old, an assent

and/or consent is deemed fit from the right population.

Minor corrections:

- In the introduction section, third paragraph, second sentence: “The revised protocol where rote learning-related items where removed, is also available in.” Replace the second 'where' with 'were'.

- In the methods section, GOCUT in the second paragraph, the last sentence, change S5 to S4 for consistency in grade choice.

Is the rationale for creating the dataset(s) clearly described?
Partly

Are the protocols appropriate and is the work technically sound?
Partly

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Educational Psychology with a bias to psychometrics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 02 May 2022

Kizito Kizito, University of Rwanda College of Education (URCE), Kayonza, Rwanda

Dear Reviewer,

Thank you for the prodigious comments provided!

- We could not describe GOCUT analysis methods, but since GOCUT data and LPCA data are presented in similar ways and the same analysis discussed for LPCA would be appropriate for GOCUT. This was also highlighted in the manuscript. Actually, when you click on a cell, you can identify the formula used to compute a certain function.

- Thank you so much for the informative instruction on the ethics statement.

- Regarding the typo in the introduction section, “where” was replaced by “were.”

- Regarding the typo in the method section, S5 was replaced by S4.

Competing Interests: No competing interests were disclosed.

Reviewer Report 09 December 2021

https://doi.org/10.5256/f1000research.56488.r101516

?

Hassen Ghalila

[1] Department of Physics, Faculty of Sciences, University of Tunis El Manar, Tunis, 1060, Tunisia
[2] LSAMA Laboratory, University of Tunis El Manar, Tunis, Tunisia

Your manuscript reports on the development and elaboration of a dataset for LPCA and GOCUT to help teachers and the educational system more generally to improve conceptual understanding of optics and associated light phenomena.

Major corrections

1. In my sense, a too great part of this manuscript is dedicated to the LPCA dataset. This is not necessary since you have already done this in your previous article (Ref.5). It would have been preferable to develop more analysis on the results obtained with the supplementary tools used in the GOCUT field. This discussion is totally absent in this manuscript, and I think will help to better address the dataset. Figures inside the excel files (only one for GOCUT?) could be used to help the comparison between the different tools employed with their associated results. It seems (after a very brief look at your data) that lessons using YouTube performed better than the other two? Also, it would be helpful to recall the 'usual teaching' (content and way of teaching) for the S4 students.

2. Your samples are composed of four groups, one from the rural zones with boarding and day schools and another two from Kigali again with boarding and day schools. I didn't see any discussion or comparison on the results obtained for these four groups. Does that mean that there is no difference between them?

3. In the dataset of LPCA, question 7 answers A and C are similar? Is it an error or something missing?

Minor corrections

1. In the Introduction section: One of the two references 7 and 8 is too much and not necessary.

2. In the Methods section, GOCUT: The second to last paragraph, 'A total of 136 students from grade 10 or senior 4 (S5) were involved in the study (see Table 2)'. Replace S5 with S4.

3. In the Ethics statement section: 'Our data collection involved secondary school students aged between 16 and 23 years old.'. Is it 16 or 18 years old?

Is the rationale for creating the dataset(s) clearly described?
Partly

Are the protocols appropriate and is the work technically sound?
Partly

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Physics of plasma and spectroscopy

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 02 May 2022

Kizito Kizito, University of Rwanda College of Education (URCE), Kayonza, Rwanda

Dear Reviewer,

Thank you so much for your insightful comments and suggestions!

- Yes, LPCA was greatly discussed. Both LPCA and GOCUT datasets were described, and the methods used to collect related data were presented. We have clarified in the manuscript that both data sets are similar, although the participants, levels, and tests were different. The data entry was the same. That is why describing the analysis of all of them would be unnecessary. Thus, to avoid duplication of text, we opted to take LPCA as a reference and discuss its analysis procedures. So researchers who want to look at whether GOCUT can use the same analysis methods we presented for LPCA. The figure presented in the GOCUT dataset can be referred to as a plot of others in other sheets since MS Excel formulae are presented in corresponding cells.

- Regarding variables in our sample, the difference between them may be there or not, but this data note is not dedicated to presenting results, discussing, or drawing any conclusion; it just provides room for other researchers to reuse the data. Thus, the collection procedure and proposed data analysis are only presented.

- Regarding question 7 in the LPCA dataset, we could not find a mistake. If you mean answer choices in pre- and post-test, we found them different.

- Regarding the introduction section, since the link for LPCA protocol was provided in the text, reference 7 was deleted.
- Regrading the typo in the methods section, S5 was replaced by S4.
- Regarding the ethical statement section, it is 16 years old.

Competing Interests: None