Development and Validation of a Concept Inventory for Interpreting Kinematics Graphs in the Tanzanian Context

Beni Mbwile, Celestin Ntvuguruzwa, and K. K. Mashood

Abstract: This paper discusses the development and validation of a concept inventory for interpreting kinematics graphs in the Tanzanian context. The study involved 61 participants comprising physics pre-service teachers, secondary school teachers, diploma college tutors, and a university lecturer from Tanzania. We developed 25 multiple-choice questions for interpreting kinematics graphs. The different steps in the development process used are selecting the topic, setting objectives, constructing questions, validating questions, and reliability testing. We carried out descriptive and inferential statistical analysis by using Statistical Package for Social Science (SPSS) version 22 followed by item analysis for pre-and post-piloting. Findings revealed normal distribution scores with a mean and standard deviation of 39.28±10.893 for pre-piloting and 40.16±8.08 for post-piloting. It also revealed no significant difference between pre-and post-piloting results with a p-value of 0.414. In addition, correlation coefficients for test re-test reliability were .783 and .878 for single and average measures respectively. Moreover, item analysis in terms of difficulty index, discrimination index, and distractor efficiency agreed with the published standards. Based on these findings, the study recommends the use of developed and validated kinematics graphs concept inventory by physics educators in both research and classroom instructions in the Tanzanian context.

Keywords: Concept inventory, kinematics graphs, Physics teachers, Tanzania context.


Introduction

Educators, strive to improve the conceptual understanding of their students. To succeed, educators must assess students’ understanding and address their misconceptions. One of the tools commonly used for assessing student understanding and misconceptions is concept inventory (Knight, 2010). Concept inventory refers to a set of Multiple-Choice Questions (MCQs) carefully crafted on a topic or a concept aimed at identifying students’ misconceptions or alternative conceptions and addressing them accordingly (Mashood & Singh, 2013). Extensive validation is required when developing concept inventories such that we can discern what learners think when selecting each response option (Adams & Wieman, 2011). One of the benefits of concept inventories is that they enable us to rapidly diagnose misconceptions that students may have before instruction on the subject matter (Libarkin et al., 2011).

Already Concept inventories have been developed for several concepts and topics in physics which include; force (Hestenes et al., 1992), kinematics (Mashood & Singh, 2013), electricity and magnetism (Ding et al., 2006) among others. Currently, there are no published concept inventories for kinematics in the Tanzanian context. Kinematics describes the motion of points, objects, and groups of objects without considering the causes of motion (Antwi, 2015). It includes the concepts of position, displacement, space, speed, velocity, and acceleration. Kinematics concepts are very essential because they build a foundation that is required by students to learn the next concepts (Amin et al., 2020). A lack of understanding of these concepts may result in students having a poor understanding of further physics concepts (Manurung et al., 2018). Kinematics concepts are displayed in the form of graphs and presented as formulas.

*Corresponding author: Beni Mbwile, University of Rwanda-College of Education (UR-CE), African Centre of Excellence for Innovative Teaching and Learning Mathematics and Science (ACEITLMS), Rwanda / University of Dar es Salaam-Mkwawa University College of Education, Tanzania benimbwile1@gmail.com

© 2023 The Author(s). Open Access - This article is under the CC BY license (https://creativecommons.org/licenses/by/4.0/).
Kinematics graphs have position, velocity, and acceleration as the ordinate and time as the abscissa. However, fundamental errors are found when students plot and interpret these graphs. For example, in illustrating the position-time graph students often make mistakes in plotting graphs of rest objects, they demonstrate the movements observed directly on the graph spatially without noticing the time axis (Amin et al., 2020). Other studies by Antwi et al. (2018), Phage et al. (2017), and Zavala et al. (2017) also have revealed student difficulties to interpret kinematics graphs.

In the Tanzanian context, kinematics graphs are taught to form two secondary school students under the topic called motion in a straight line. Students' performance on this topic is low for several years. For example, in the years 2019 and 2020, students’ average performance on the topic in Form Two National Assessment (FTNA) was 11.7% and 12.5% respectively (National Examination Council of Tanzania [NECTA], 2019a, 2020a). Figure 1 below is a vivid example of students’ performance on the topic in 2019.

![Figure 1. Students’ Performance per Topic for FTNA in 2019 (NECTA, 2019a)](image)

Figure 1 above entails how students performed on various topics in 2019, motion in a straight line included. Only 12.5% passed the question on motion in a straight line from 568,305 students who sat for Physics FTNA in 2019. This signifies difficulties facing students on the topic which requires extra effort to improve their performance. Apart from the low performance of students on motion in a straight line, the topic is also less researched in the context of Tanzania. Therefore, we decided to develop the concept inventory on kinematics graphs (position, velocity, and acceleration versus time graphs) which is the most challenging part of this topic.

**Research Questions**

1. How to develop a valid and reliable kinematics graphs concept inventory assessment tool in the context of Tanzania?
2. Is there any statistically significant difference in performance between pre- and post-piloting of kinematics graphs concept inventory?

**Research Hypothesis**

H₀ There is no statistically significant difference in performance between pre- and post-piloting of kinematics graphs concept inventory

**Literature Review**

Several studies have suggested some of the stages employed when developing a valid and reliable assessment tool, concept inventory included. For example, Yang (2014) based on Item Response Theory (IRT), described five stages involved when developing an assessment tool. The stages include understanding, constructing, testing, analyzing, and modifying. In addition, Adams and Wieman (2011) proposed four phases of developing a valid test instrument which includes; a) delineation of the purpose of the test and the scope to be measured, b) develop the test specifications, c) Development, evaluation, field testing, and selection of the items and scoring procedures and guides, and d) assembly and evaluation of a test for operational use. The study by Mashood and Singh (2013) used theoretical analyses and
iterative empirical investigation steps to develop a concept inventory in rotational kinematics. The process was facilitated by methodologies such as think-aloud protocol, semi-structured interviews, and retrospective probing.

Also, Kirya et al. (2021) used the Delphi technique to develop a concept inventory in rotational kinematics based on the Ugandan context. The Delphi technique facilitated the process of finding experts who ranked each concept inventory item based on the relevant criteria. Furthermore, Taylor et al. (2020), suggested steps to be used when developing a concept inventory such as establishing topics, administering open-ended questions, converting to MCQs, validation through interviews, and statistical validation. Moreover, Kalas et al. (2013), used stages such as the identification of students’ misconceptions, designing MCQs, and validation of MCQs to develop a Meiosis concept inventory. This implies the importance of following systematic procedures when developing valid and reliable assessment tools.

**Concept Inventory in Physics**

Several concept inventories have been developed in physics subject since 1985 when Halloun and Hestenes (1985) developed the Mechanics Diagnostic (MD) test of conceptual understanding of Newtonian mechanics. Some notable concept inventories which have been developed in physics so far include; Force Concept Inventory (FCI) (Hestenes et al., 1992). FCI becomes a successor to MD and it opens the eyes of many traditional instructors across the globe on how to develop and use concept inventory for classroom instructions. It consists of 30 MCQs used to probe learners’ conceptual understanding of Newtonian mechanics. It is the most often used assessment tool for measuring the effectiveness of classroom instruction. Other inventories include Force and Motion Conceptual Evaluation (FMCE) (Thornton & Sokoloff, 1998), and Brief Electricity and Magnetism Assessment (BEMA) (Ding et al., 2006). Specifically, for kinematics graphs, some of the concept inventories developed include the Test for Understanding Kinematics Graphs (TUKG) (Beichner, 1994) and an inventory on Modifying Test for Understanding Kinematics Graphs (MTUKG) (Zavala et al., 2017) among others.

**Challenges of Kinematics Graphs to Learners**

The study conducted in Ghana by Antwi (2015) revealed that Ghanaian science students were facing difficulty to describe the shapes of kinematics graphs, convert from one to the other form of kinematics graphs, failing to get the meaning of the slopes, and difficulty in interpreting areas under kinematics graphs. These challenges were a result of students not practicing these approaches in their classrooms. Another study by Amin et al. (2020) in Indonesia revealed students’ difficulties to interpret kinematics graphs such as problems to read graphs, how to use formulas in solving kinematics questions, and how to solve tests in graphical form. Meanwhile, in South Africa, Phage et al. (2017) found that students fail to interpret kinematics graphs due to insufficient contextual knowledge of mathematics and physics which is foundational in graph representations. According to Beichner (1994), students have difficulty in differentiating the meanings of position, velocity, and acceleration versus time graphs. For example, learners tend to think a graph is a literal picture and fail to interpret the area under the graph. Moreover, McDermott et al. (1987) revealed two difficulties students encounter with kinematics graphs which include connecting graphs to physical concepts and to the real world. Failure to master kinematics graphs results in students’ low performance on the topic of kinematics and overall performance in physics.

**Importance of Kinematics Graphs**

The knowledge of kinematics graphs helps to build a foundation for drawing graphs that are beyond kinematics graphs (Phage et al., 2017). This knowledge of graph construction is very important in learning physics concepts because graphs are an integral part of experiments, the heart of physics (Beichner, 1994). Graphs are used by physicists as a language of physics because of their ability to summarize a large amount of information. Through graphs, the use of complex formulas can be minimized and some equations can be obtained from graphs. A student who can interpret kinematics graphs is likely to perform better in both theory and practical (Amin et al., 2020). This will improve not only the performance of students in the topic of kinematics and in experiments but will also raise student overall performance in physics.

**Methodologies**

This study adopted a sequential explanatory design during the process of data collection. Sequential explanatory is a design in mixed research approach whereby the outcomes of quantitative data collection methods inform qualitative data collection methods (Bowen et al., 2017; Creswell & Clark, 2017). Thus, research findings obtained from the pre-piloting of kinematics graphs concept inventory informed data collections through focus group discussions. Also, a design-based approach was used to design kinematics graphs concept inventory. Generally, there are several methods used for developing and validating concept inventories and may vary from one person to another (Lindell et al., 2007). This study employed five stages including selecting the topic, setting objectives, constructing MCQs, validating MCQs, and testing the reliability of MCQs.
Study Participants

The study had a sample size of 61 comprising 50 diploma pre-service physics teachers, 4 secondary school physics teachers, 4 physics tutors from diploma teachers’ colleges, 2 university pre-service physics teachers, and 1 university physics lecturer from Tanzania. Diploma college physics tutors, secondary school physics teachers, university pre-service physics teachers, and a university lecturer formed a panel of experts for developing kinematics graphs concept inventory. Also, 50 diploma college pre-service physics teachers participated in a pilot study and focus group discussions. Experts were selected based on the following criteria: i) secondary school teachers should have experience of not less than 5 years in teaching physics; ii) diploma college tutors should be teaching physics content knowledge and pedagogical knowledge; iii) university lecturers should be the one who started teaching physics in ordinary level secondary schools before joining the university; iv) University pre-service teacher should be a second-year student with an overall performance of not less than a Grade Point Average (GPA) of 3.5 for first-year results; and v) all participants should be willing to participate.

Half of the diploma pre-service physics teachers were in 1st year and the other half were in 2nd year majoring in physics, education, and one more subject. The average ages of diploma pre-service physics teachers ranged from 21 to 23 years. Diploma pre-service physics teachers were targeted because apart from majoring in physics, they are future teachers in ordinary-level secondary schools. Also, each year they participate in teaching practice in ordinary-level secondary schools where kinematics graphs are taught. Participants were randomly picked by using YES/NO cards to get a representative sample for the pilot test. YES/NO cards were distributed to be picked by first-year and second-year pre-service teachers. Pre-service teachers who picked NO cards were excluded from the pilot test. Similar sampling technique was used to obtain participants for focus group discussions. However, only pre-service teachers who sat for the pilot test were considered for focus group discussions.

Steps in Developing Kinematics Graphs Concept Inventory

Figure 2 below shows a summary of methodological steps used in designing and validating kinematics graphs concept inventory.

1. Selection of the Topic
2. Setting Specific Objectives
3. Constructing MCQs
   a). Meeting of Experts
   b). Pre-Piloting
4. Validating MCQs
   c). Focus Group Discussions
   d). Post-Piloting
5. Testing Reliability of MCQs

Figure 2. Steps in Developing Kinematics Graphs Concept Inventory.

1. Selection of topic for developing concept inventory

The topic of motion in a straight line was selected from among physics topics with low students’ performance in ordinary-level secondary school national examinations. Student’s item response analysis reports on the FTNA and candidates’ response item analysis reports on the Certificate of Secondary Education Examination (CSEE) both prepared NECTA were used in identifying topics in which students are not performing well for several years. The topics identified were motion in a straight line, equilibrium, waves, electromagnetism simple machines, and newton’s law of motion (NECTA, 2019a, 2019b, 2020a, 2020b). For the past three years (2018, 2019, and 2020), students’ average performance for these topics in FTNA and CSEE was below 20%. Motion in a straight line was chosen to develop concept inventory due to the low performance of students in 2019 and 2020 in FTNA which was 11.7% and 12.5% respectively (NECTA, 2019a, 2020a). Also, the topic is less researched in the context of Tanzania. Moreover, it consists of kinematics concepts...
which build the foundation that is required by students to learn the next concepts (Amin et al., 2020). Therefore, we decided to develop kinematics graphs (position, velocity, and acceleration versus time graphs) concept inventory which is the most challenging part of this topic.

2. Setting specific objectives

Efforts were made to make sure that all necessary specific objectives required to be known by the students in the Tanzanian context were included in designing the kinematics graphs concept inventory. Efforts which were made include consulting physics syllabus for ordinary-level secondary school and physics textbook for ordinary-level secondary school form two. These two documents are produced by the ministry and institute of education in Tanzania as curriculum materials authorized to be used in ordinary-level secondary schools (Ministry of Education and Vocational Training, 2007; Tanzania Institute of Education, 2021). Also, previous scholarly works on kinematics graphs were used in formulating specific objectives (Amin et al., 2020; Antwi, 2015; Beichner, 1994). Specific objectives were formulated by considering the position, velocity, and acceleration versus time graphs as summarized in Table 1 below.

Table 1. Outline for Designing Kinematics Graphs Concept Inventory

<table>
<thead>
<tr>
<th>S/N</th>
<th>Kinematics-graphs concepts</th>
<th>Specific Objectives intended for the Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Position-time graphs</td>
<td>Identify different states of an object (i.e., forward, backward, and stationary)</td>
</tr>
<tr>
<td>02</td>
<td>Position and velocity-time graphs</td>
<td>Find the distance covered by the object</td>
</tr>
<tr>
<td>03</td>
<td>Position, velocity, and acceleration-time graphs</td>
<td>Determine the velocity of the object</td>
</tr>
<tr>
<td>04</td>
<td>Velocity-time graphs</td>
<td>Find the acceleration of an object</td>
</tr>
<tr>
<td>05</td>
<td>Position-time graphs</td>
<td>Interpret the slope of kinematics graphs</td>
</tr>
<tr>
<td>06</td>
<td>Velocity and acceleration-time graphs</td>
<td>Interpret the areas under kinematics graphs</td>
</tr>
<tr>
<td>07</td>
<td>Position, velocity, and acceleration-time graphs</td>
<td>Convert between kinematics graphs</td>
</tr>
</tbody>
</table>

3. Construction of MCQs

We constructed the first draft of our inventory which had 35 MCQs. Among the materials which helped in designing the inventory were physics textbooks for ordinary-level secondary school form two (TIE, 2021), form two and four national examination past papers, and teachers’ made classroom examinations. Also, previously published works on kinematics graphs such as Amin et al. (2020), Antwi (2015), and Beichner (1994) were used for restructuring some of the test items to avoid unnecessary repetition. For example, the item in Figure 3 below was adapted from the Test of Understanding Graphs in Kinematics (TUG-K) (Beichner, 1994).

Qn. The displacement versus time graph for five objects is given below. Which object is moving fast in the forward direction with constant velocity?

(A) II, (B) III, (C) I, (D) V, (E) IV

Figure 3. An Example of a Position-Time Graph Item for Testing Constant Velocity

Therefore, from the same perceptions, we constructed 15 test items for position-time graphs, 12 test items for velocity-time graphs, and 8 items for acceleration-time graphs. More weight was given to position-time graphs because velocity and acceleration time graphs are derived from position-time graphs.

4. Validation of MCQs

a) Meeting with physics experts

We conducted content validity of kinematics graphs concept inventory with eleven physics experts. Validation of kinematics graphs through the panel of experts was presented as PowerPoint slides show. Presenting test items as PowerPoint decreases the use of paper, preserves the integrity of the instrument, and ensures that all people see the
test items in the same order (Kalas et al., 2013). Each of the 35 test items was displayed to 11 physics experts only once and the instructions about how to answer were included to avoid possible confusion. This was followed by intensive discussions about the relevance of each test item presented. The time provided to discuss each test item was 5 minutes and physics experts had the option to request more time if needed.

Through expert meetings, 10 test items were discarded, 8 were modified, and 17 were retained. 10 test items were discarded because 1 test item was wrongly constructed, 5 items were likely to bring guessing because of their similarity to other test items and 4 items were directly copied from previously published findings. For example, one item that was copied directly from Planinic et al. (2013) and discarded is shown in Figure 4 below.

![Figure 4. An Example of Velocity-Time Graph Item for Testing Acceleration](image)

The panel of experts insisted on the need to construct our questions and even adapt test items instead of copying directly from someone. In the same manner, other directly copied test items were discarded. Thus, at this stage through the experts' meeting, the second draft with a total of 25-test items was formed.

b) Pre-piloting

We conducted a pre-pilot to 50 pre-service physics teachers to determine the effectiveness of 25-kinematics graphs test items at the beginning of the semester. The teachers' college which was used for the pilot study didn't take part in the actual process of data collection. According to Gorman et al. (2005), a pilot study means taking up and initiating the draft study strategy at a neutral location that will not be used in real fieldwork at the location(s) from which data is to be collected. After pre-piloting, descriptive statistical analysis such as mean, standard deviation, and frequencies was computed. It followed with item analysis in terms of difficulty index, discrimination index, and distractor efficiency. Findings from the item analysis were the bases for pre-service teachers' focus group discussions.

c) Focus group discussions.

We conducted focus group discussions with pre-service physics teachers. Only pre-service teachers who sat for the pilot test were considered for focus group discussion. The purpose of the focus group discussion was not to assess whether the performance was high or low but to capture their thoughts on the quality of test items by looking at the difficulty index, discrimination index, and distractor efficiency revealed from pre-piloting. Any divert from the published standards was addressed. A total of 20 pre-service physics teachers (10 from the first year and 10 from the second year) participated in focus group discussions. Pre-service teachers who participated in focus group discussions were not aware if they will be required to sit for the post-pilot. Through their comments, some modifications were made and included in the third draft of the inventory.

d) Post-piloting

We administered a post-pilot at the end of the semester after 3 months and 3 weeks to the same group of pre-service physics teachers. This period was long enough to prevent influence from pre-pilot and focus group discussions. Post-piloting aimed at assessing whether there was an improvement in the effectiveness of questions in terms of difficulty index, discrimination index, and distractor efficiency in comparison with pre-pilot. Also, aimed at assessing performance improvement. Test items in the post-pilot had minor modifications for distractors and stems.

5. Reliability testing of MCQs

A test re-test reliability was carried out to ensure the test is not only valid but also reliable. The test re-test reliability was aimed at assessing whether the kinematics graphs concept inventory items can be repeated in the same order if they can be given to the population with similar or equivalent cognitive abilities or characteristics. After reliability analysis, the final version of our concept inventory was produced as indicated in the appendix section.
Data Analysis

Descriptive and inferential analyses were carried out by using Statistical Package for Social Science (SPSS) version 22. We started with descriptive analysis such as mean, standard deviations and frequencies of choice selections for the first pilot test. Thereafter, item analysis was conducted in terms of difficulty index, discrimination index, and distractor efficiency (effective, ineffective and non-functioning distractors). After the results of the second pilot, inferential statistics such as paired T-test was carried out.

Ethical Considerations

To fulfil the ethical considerations of the study, we obtained permits from the University of Rwanda and the Ministry of Education, Science and Technology in Tanzania. Also, participants signed consent forms for their willingness to contribute to the study. Moreover, real names were avoided to maintain privacy and confidentiality instead arbitrary coding was used. For example, pre-service teachers who participated in the focus group discussions were represented using P1 to P20.

Results

Descriptive and Inferential Analysis

Statistical descriptive analysis revealed that pre-service physics teachers had a mean performance of 39.28 and a standard deviation of 10.893 (39.28 ±10.893) for pre-pilot. The performance was slightly higher for post-piloting with mean scores of 40.16 and a standard deviation of 8.08 (40.16±8.08). Also, before we opted on whether we should conduct parametric analysis or non-parametric analysis, a normality test was carried out. The normality test was conducted by using the Shapiro-Wilk normality test which is considered to be the most appropriate test for normality (Ghasemi & Zahediasl, 2012). Table 2 below presents normality statistical analysis for pre-and post-pilot findings.

<table>
<thead>
<tr>
<th>Pilot</th>
<th>N</th>
<th>Significance</th>
<th>Data distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-pilot</td>
<td>50</td>
<td>0.599</td>
<td>Normal</td>
</tr>
<tr>
<td>Post-pilot</td>
<td>0.315</td>
<td>Normal</td>
<td></td>
</tr>
</tbody>
</table>

The significance values of 0.599 and 0.315 for pre-and post-pilots presented in Table 2 above, signifies that the two pilots were normally distributed. Thus, providing the need to conduct parametric analysis. Parametric statistical analysis performed for paired sample t-test showed a t-value of -0.824 and a p-value of 0.414. These descriptions are summarized in Table 3 below.

<table>
<thead>
<tr>
<th>Pilots</th>
<th>N</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>T</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-pilot</td>
<td>50</td>
<td>39.28</td>
<td>10.893</td>
<td>-0.824</td>
<td>49</td>
<td>0.414</td>
</tr>
<tr>
<td>Post-pilot</td>
<td>40.16</td>
<td>8.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Findings from Table 3 above indicate the majority of pre-service physics teachers were scoring below 50%. Although test scores were normally distributed which means pre-and post-pilots were relevant to pre-service physics teachers but this is an indication of the presence of some difficulties facing them to interpret kinematics graphs. Meanwhile, the p-value of 0.414 is above the value which indicates a statistical significance difference between pre-and post-pilots. When p-value ≤ 0.05 indicates a significant difference in performance between tests and when p-value > 0.05 indicates no significant difference between the two tests. Therefore, the narrow hypothesis was accepted because the findings, revealed there was no statistically significant difference in performance between pre-and post-piloting.

Item Analysis for Pre-Pilot

Item analysis in terms of difficulty index, distractor efficiency (effective distractor, ineffective distractors, and non-functioning distractor), and discrimination index for pre-piloting were computed from the frequency of choice selection. Findings revealed 11 ineffective distractors, 1 non-functioning distractor, and 88 effective distractors. It signifies that 88% of distractors were effective because a distractor is considered effective if selected by not less than 5% of participants, ineffective if selected by less than 5% and non-functioning if not-selected by anyone (Hingorjo et al., 2012). Also, the difficulty index and discrimination index ranged from 0.14 to 0.78 and 0.21 to 0.86 for MCQs respectively. The optimal range of the difficulty index is 0.2 to 0.8 while the minimum required discrimination index is 0.3 (Taib et al., 2014). Only 1 of 25 items was below the required range of difficulty index and 6 of 25 questions had discrimination indexes below 0.3. Therefore, items with a difficulty index below 0.2, discrimination index below 0.3, distractors selected by less than 5%, and distractors not selected by any participants were identified as items that need to be improved. These items were taken to the pre-service physics teachers’ focus group discussion, the next stage of validating the kinematics graphs concept inventory.
Analysis of Pre-Service Teachers’ Explanations

It was important to analyze pre-service physics teachers’ oral explanations during focus group discussions on item analysis which divert from published standards. Pre-service teachers provided adequate explanations that gave us significant insight into the way they reason on different items. Since it is extensive, the full report on Pre-service teachers’ explanations will be given in a subsequent paper. Findings from focus group discussions with pre-service teachers were based on items with a discrimination index below 0.3, difficulty index below 0.2, ineffective distractors, and a non-functioning distractor. Findings indicated some changes to be done to improve our inventory. For example, when pre-service teachers were asked why distractor A in item number 17 as shown in Figure 5 below attracted only 2% of pre-service teachers.

17. You are given a graph as shown in the figure below. What does the area represent?
   (A) Momentum (B) Speed (C) Retardation (D) Total distance travelled (E) Velocity

![Figure 5. An Example of Test-Item for Testing Area Under Velocity-Time Graph](image)

One pre-service teacher P7 responded that:

I think the problem is coming from the word momentum which is used in distractor A, it doesn’t relate to other words. This is because, in the topic of motion in a straight line, we learn about distance, displacement, speed, velocity, and acceleration. Therefore, it is better to replace the word momentum with one of the words I have mentioned.

The above quotation signifies the need of using distractors that are more related to the correct response to attract students who are in doubt of the correct choice or do not know the right answer. The comment from P7 helped to replace the word momentum with the word acceleration. More comments received from questions asked during pre-service teachers’ focus group discussions led to several modifications of test items. Below are some examples of modifications made due to the non-functioning distractor, ineffective distractors, discrimination indexes and difficulty index.

The non-functioning distractor “C” in question 21 was replaced as shown in Figure 6 below

![Figure 6. An Example of a Changed Non-Functioning Distractor](image)
Ineffective distractor “D” in question 25 was replaced as shown in Figure 7 below.

![Figure 7. An Example of a Changed Ineffective Distractor](image)

Due to the low discrimination index and difficulty index, the stem of the question 10 was modified by replacing the words “at any point between time (t)= 2 seconds to time (t)= 6 seconds” with the words “at a time (t)= 4 seconds”.

Due to the low discrimination index, the stem of the question 18 was modified by replacing the words “do not indicate acceleration” with the words “indicates motion with zero acceleration”.

Modifications through pre-service teachers’ focus group discussions resulted in producing the 3rd draft of the concept inventory for interpreting kinematics graphs.

**Item Analysis of Post-Piloting**

Item analysis for post-pilot revealed improvement of difficulty index, discrimination index, and distractor efficiency (effective distractor, ineffective distractor, and non-functioning distractor). These analyses were possible through the frequency of choice selection as shown in Table 4 below.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Frequency of a choice selection</th>
<th>Difficulty index</th>
<th>Discrimination index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>14 24 05 03 04</td>
<td>0.48</td>
<td>0.43</td>
</tr>
<tr>
<td>Question 2</td>
<td>07 12 04 06 21</td>
<td>0.42</td>
<td>0.50</td>
</tr>
<tr>
<td>Question 3</td>
<td>29 03 09 03 06</td>
<td>0.58</td>
<td>0.36</td>
</tr>
<tr>
<td>Question 4</td>
<td>05 10 05 04 26</td>
<td>0.52</td>
<td>0.43</td>
</tr>
<tr>
<td>Question 5</td>
<td>05 09 04 27 05</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>Question 6</td>
<td>12 08 17 07 06</td>
<td>0.24</td>
<td>0.36</td>
</tr>
<tr>
<td>Question 7</td>
<td>13 07 18 05 07</td>
<td>0.36</td>
<td>0.43</td>
</tr>
<tr>
<td>Question 8</td>
<td>09 14 05 19 03</td>
<td>0.38</td>
<td>0.64</td>
</tr>
<tr>
<td>Question 9</td>
<td>08 28 04 04 06</td>
<td>0.56</td>
<td>0.43</td>
</tr>
<tr>
<td>Question 10</td>
<td>09 06 20 03 12</td>
<td>0.24</td>
<td>0.36</td>
</tr>
<tr>
<td>Question 11</td>
<td>05 09 27 05 04</td>
<td>0.54</td>
<td>0.64</td>
</tr>
<tr>
<td>Question 12</td>
<td>07 06 18 07 12</td>
<td>0.36</td>
<td>0.36</td>
</tr>
<tr>
<td>Question 13</td>
<td>25 05 07 06 07</td>
<td>0.50</td>
<td>0.86</td>
</tr>
<tr>
<td>Question 14</td>
<td>06 26 04 08 06</td>
<td>0.52</td>
<td>0.72</td>
</tr>
<tr>
<td>Question 15</td>
<td>03 19 11 13 04</td>
<td>0.26</td>
<td>0.43</td>
</tr>
<tr>
<td>Question 16</td>
<td>25 03 11 07 04</td>
<td>0.22</td>
<td>0.43</td>
</tr>
<tr>
<td>Question 17</td>
<td>03 04 05 35 03</td>
<td>0.70</td>
<td>0.36</td>
</tr>
<tr>
<td>Question 18</td>
<td>10 13 08 07 12</td>
<td>0.26</td>
<td>0.57</td>
</tr>
<tr>
<td>Question 19</td>
<td>08 15 15 05 07</td>
<td>0.30</td>
<td>0.36</td>
</tr>
<tr>
<td>Question 20</td>
<td>04 08 10 05 23</td>
<td>0.46</td>
<td>0.50</td>
</tr>
<tr>
<td>Question 21</td>
<td>11 09 05 04 21</td>
<td>0.42</td>
<td>0.50</td>
</tr>
<tr>
<td>Question 22</td>
<td>22 04 10 06 08</td>
<td>0.44</td>
<td>0.57</td>
</tr>
<tr>
<td>Question 23</td>
<td>16 07 07 12 08</td>
<td>0.24</td>
<td>0.43</td>
</tr>
<tr>
<td>Question 24</td>
<td>19 14 04 07 06</td>
<td>0.28</td>
<td>0.36</td>
</tr>
<tr>
<td>Question 25</td>
<td>11 12 08 05 14</td>
<td>0.22</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Note: The correct choice is bolded.
From the table above, frequencies of choice selection for each item indicate that all distractors for post-pilot were effective because were selected by not less than 5% of participants. It signifies an increase in distractor efficiency from 88% for pre-pilot to 100% for post-pilot. Also, item analysis in terms of the difficulty index and discrimination index for MCQs ranged from 0.22 to 0.70 and 0.36 to 0.86 respectively. These findings fall within the required range of published standards of 0.2 to 0.8 for the difficulty index and ≥0.3 for the discrimination index (Taib et al., 2014). Therefore, the inventory can discriminate effectively among pre-service physics teachers, secondary school students and their teachers.

**Reliability Analysis**

Analysis of test re-test reliability revealed there was a high interclass correlation between pre-and post-piloting results. The correlation coefficients were .783 and .878 for single and average measures respectively. As the correlation coefficient approaches 1 it signifies the presence of strong relationships between the variables (Bonett & Wright, 2015). Findings reveal that our inventory can produce consistent results if admitted to different groups of people with similar characteristics. At this stage, the concept inventory for interpreting kinematics graphs in the Tanzanian context was considered reliable and valid. Therefore, the third draft of the kinematics graphs concept inventory assessment tool was converted to the final version as indicated in the appendix section.

**Discussions**

This paper aimed only at designing a valid and reliable assessment tool for interpreting kinematics graphs. It does not reveal the misconceptions/alternative conceptions pre-service physics teachers had with kinematics graphs, rather it provides a tool that can be used by educators in the Tanzanian context. Pre-service physics teachers’ performance for pre-and post-piloting indicates their scores were normally distributed with the mean performance of 39.28 and 41.16 respectively. The performance was from participants who studied the topic of motion in a slight line during their time in ordinary-level secondary schools. They passed physics in both FTNA and CSEE and were physics majors when pursuing diplomas in teachers’ education programs. The mean performance obtained of 50% was below the expectation from people who were familiar with the topic.

Pre-service teachers’ low performance aligns with findings from Basic Education Statistics of Tanzania reports prepared by the President’s Office - Regional Administration and Local Government (PO-RALG) which indicate students’ physics performance in CSEE to be below 50% for consecutive three years (PO-RALG, 2019, 2020). It also concurs with findings from NECTA (2019a, 2020a) which reveal students’ mean performance was below 13% for the topic of motion in a straight line. Moreover, the findings agree with Beichner (1994) who found an average performance of 40% when testing students’ ability to interpret kinematics graphs. Moreover, Amin et al. (2020), Antwi (2015), and Phage et al. (2017) revealed low performance of students on kinematics graphs beyond expectations despite learning the topic in previous years. This signifies that the challenge of kinematics is universal and requires the use of proper assessment strategies.

Involving pre-service teachers in validating kinematics graphs concept inventory through focus group discussions helped to shape questions to the desired level of quality. Their involvement in developing and validating concept inventory concurs with Mashood and Singh (2013) who involved students during the development of the rotational kinematics concept inventory and found that involving them, helped significantly in constructing good test items. Similarly, Kalas et al. (2013) involved students through interviews during the development of the Meiosis concept inventory whereby contributions provided by students were used to modify the inventory. Moreover, Shoji et al. (2021) reported significant contributions from students during the validation of FCI test items. Therefore, learners are key stakeholders because they provide useful information when developing test items.

To obtain a valid and reliable kinematics graphs concept inventory, test items were pre-and post-piloted. Piloting inventory instruments aligns with Mashood and Singh (2013) who conducted a total of four pilots to get a desired rotational kinematics concept inventory. Also, the findings align with (Ndihokubwayo et al., 2020) who piloted twice the inventory for light phenomena. Although piloting an instrument more than once provides the opportunity of testing both the validity and reliability, some developers may pilot the instrument at once or do otherwise without piloting.

Although there was a slightly different between the pre-pilot and post-piloting, the results from statistical analysis indicated there was no statistical difference in performance between pre-and post-piloting. This is because the calculated p-value was 0.414. Findings concur with the explanation provided by Andrade (2019) who contended that when the p-value is ≥0.05 means there is no statistically significant difference in performance between the variables.

Item analysis in terms of distractor efficiency has indicated that 88 out of 100 distractors were effective for pre-piloting while 100 out of 100 were effective for post-test because were selected with not less than 5% of respondents. These findings go parallel with the description provided by Hingorjo et al. (2012) who pointed out that a distractor is considered effective if selected by at least 5% of the students, ineffective if selected by less than 5%, and considered to be a non-functioning distractor if not selected by any student. In addition,
Also, difficulty indexes ranged from 0.16 to 0.78 with one item having a difficult index below 0.2 for pre-piloting while none of the items had a difficult index below 0.2 for post-piloting. This signifies that test items were of optimal difficulties for the targeted group. The findings concur with Taib et al. (2014) argued that the optimal range of difficulty index range from 0.2 to 0.8, any range index below 0.2 means the item is very difficult and above 0.8 means that the item is too easy for students and needs to be improved through revision or rejected. Moreover, item analysis in terms of discrimination index ranged from 0.17 to 0.58 with 6 items having discrimination indexes below 0.3 for pre-piloting and all items above 0.3 for post-piloting. The findings align with Hingorjo et al. (2012) argued that a test item with a discrimination index ranging from 0.40 and above is considered a very good test item, 0.30 to 0.39 is considered a good item, 0.20 to 0.29 is considered a marginal test item which requires improvement, and 0.19 or less is poor test items which should be rejected or improved by a revision. Therefore, in our inventory, items were of the required range discrimination index.

Moreover, test re-test reliability coefficients of .783 and .878 for single and average measures respectively align with the required standards. The findings are in line with Bonett and Wright (2015) who argued that the correlation of the test re-test is regarded as strong if the correlation coefficient approaches 1. Thus, the concept inventory had the strong correlation.

**Conclusion**

We developed a concept inventory for interpreting kinematics graphs (position, velocity, and acceleration versus time graphs) for physics teachers in the context of Tanzania. The different steps in the development process included selecting the topic, setting objectives for the intended concepts, constructing MCQs, validating MCQs, and reliability testing. MCQs were validated through the panel of experts, pre- and post-piloting, and focus group discussions. Test re-test reliability conducted indicated a strong interclass correlation. Therefore, the developed concept inventory is valid and reliable to be used for interpreting kinematics graphs in the context of Tanzania.

**Recommendations**

Although this paper discusses the development and validation of kinematics graphs (position, velocity, and acceleration vs time graphs) in the context of Tanzania, kinematics concepts are also presented as formulas. We recommend the development of kinematics concept inventories assessment tools which will include both graphs and formulas. Also, the validated 25-test items on kinematics graphs can be used by many physics educators in both research and classroom instructions to identify students’ misconceptions or alternative conceptions of kinematics graphs in the context of Tanzania. Moreover, there is a need to conduct similar studies on other topics of physics such as equilibrium, waves, electromagnetism, simple machines, and newton's law of motion among others whereby according to NECTA (2019a, 2019b, 2020a, and 2020b) students’ performance in these topics for several years is low. Thus, coming up with more findings on challenges facing students and providing alternative solutions for improving their performance.

**Limitations**

As a part of the validation process of MCQs, we aimed at conducting an initial pilot and final pilot for pre-service physics teachers at two different diploma teachers' training colleges. Unfortunately, the administration from one of the targeted teachers' colleges for the final pilot was not willing to participate in the study despite the research permit they received. According to them, time was not enough to allow pre-service teachers to participate in the study at the same time prepare themselves for teaching practices. As a result, we decided to conduct a delayed post-pilot to the same group of pre-service physics teachers at the teachers' college where the initial pilot was conducted.

**Acknowledgements**

We salute the African Centre of Excellence for Innovative Teaching and Learning Mathematics and Science (ACEITLMS) which is based at the University of Rwanda for the inspiration, encouragement, and financial support it provided towards the successful completion of this paper.

**Authorship Contribution Statement**

Mbwile: Study conception and design, data collection, data analysis. Ntivugurizwa: Study conception and design, data analysis, critical revision of the manuscript, supervision and final approval. Mashood: Study conception and design, data analysis, critical revision of the manuscript, supervision and final approval.

**References**


Mashood, K. K., & Singh, V. A. (2013). Development of a concept inventory in rotational kinematics: initial phases and some methodological concerns. In G. Nagarjuna, A. Jamakhandi & E. M. Sam (Eds.), *Proceedings of EpiSTEME 5*
International Conference to Review Research on Science, Technology and Mathematics Education (pp. 139-144). Homi Bhabha Centre for Science Education-Tata Institute of Fundamental Research.


Appendix

Kinematics Graphs Concept Inventory

1. The displacement versus time graph for five objects is given below. Which object is moving fast in the forward direction with constant velocity?
   (A) II, (B) III, (C) I, (D) V, (E) IV

   ![Displacement versus time graphs](image1)

2. An object starts from rest and then moves as shown in the area of the figure below. The total distance travelled by the object is
   (A) 550m, (B) 400m, (C) 24m, (D) 275m, (E) 300m

   ![Velocity-time graph](image2)

3. Here is the position-time graph of the motion of the object. Which of the following indicates that the object is not moving at all?
   (A) BC and DE, (B) AB and EF, (C) CD and FG, (D) BC and CD, (E) EF and FG

   ![Position-time graph](image3)

4. The displacement versus time graph for five objects is given below. Which object is moving slowly in the backwards direction with constant velocity?
   (A) I, (B) II, (C) IV, (D) III, (E) V
5. Given five velocity-time graphs below. Which graph represents an object’s motion at constant velocity?
(A) IV, (B) V, (C) II, (D) III, (E) I

6. An acceleration-time graph is shown in the figure below. What does the area under the graph represent?
(A) Change in velocity (B) Total distance travelled (C) Retardation
(D) Total velocity (E) Total time spent

7. Two states of an object are shown in a position versus time graph below. How can you describe the states of an object?
(A) An object is moving backwards and then forward,
(B) An object is moving forward and then stopped
(C) An object is moving backwards and then stopped,
(D) An object is stopped and then moves forward
(E) An object is topped and then moves backwards
8. You are provided with a velocity-time graph below. The acceleration of an object between time \((t)=6 \text{ seconds}\) to time \((t)=9 \text{ seconds}\) is
(A) \(20\text{ m/s}^2\) (B) \(10\text{ m/s}^2\) (C) \(50\text{ m/s}^2\) (D) \(-20\text{ m/s}^2\) (E) \(-10\text{ m/s}^2\)

9. Different states of an object are shown in a displacement versus time graph below. How can you describe an object's motion from point A to point D?
(A) Forward, backward, and stationary,
(B) Backward, stationary, and then forward
(C) Stationary, backward and then forward
(D) Stationary, forward and then backwards
(E) Forward, stationary, and then backwards

10. An object was moving as shown in the figure below. What is the velocity of an object at a time \((t) = 4 \text{ seconds}\)?
(A) \(2\text{ m/s}\), (B) \(4\text{ m/s}\), (C) \(10\text{ m/s}\), (D) \(6\text{ m/s}\), (E) \(5\text{ m/s}\)
11. Given graphs below. Identify two graphs representing objects’ motion at constant velocity
(A) II and III, (B) I and II, (C) II and V, (D) III and IV, (E) IV and V

12. The figure below represents the acceleration-time graph. The change in velocity is?
(A) 10m/s, (B) 40m, (C) 40m/s, (D) 80m/s, (E) 20m/s

13. A car was travelling as shown in the figure below. How long does it travel from time(t) = 5 hours to time(t) = 12 hours?
(A) 400km (B)200km (C)600km (D)1000km (E) 800km

14. Given the velocity-time graph below. Which sentence is the best interpretation of the object’s motion?
(A) Deceleration then acceleration.
(B) Acceleration then deceleration
(C) Retardation then acceleration
(D) Deceleration then retardation
(E) The object does not move.
15. An object path is indicated in the displacement time graph below. How long does it journey from time (t) = 10 seconds to time (t) = 50 seconds?
(A) 20m, (B) 40m, (C) 60m, (D) 80m, (E) 100m

16. Displacement-time graph for an object is shown below. The velocity at the time (t) = 2 seconds is about?
(A) 5m/s, (B) 15m/s, (C) 10 m/s, (D) 2m/s, (E) 3m/s

17. You are given a graph as shown in the figure below. What does the area represent?
(A) Acceleration (B) Speed (C) Retardation (D) Total distance travelled (E) Velocity

18. Consider the following graphs, noting the different axes
Identify graphs which indicate motion with zero acceleration

(A) I, II and IV  (B) III and IV     (C) II and V    (D) IV only     (E) V only

19. The following graph is a position-time graph. The distance of the object from time (t) =0 seconds to time (t) = 20 seconds is

(A) 900Km, (B) 600Km, (C) 1200Km, (D) 1000Km, (E) 1500Km

20. An object starts from rest and then moves as shown in the velocity versus time graph below. At which state an object is said to be decelerating?

(A) AB and DE, (B) AB and EF, (C) DE and EF, (D) BC and DE, (E) CD and EF

21. An object starts from rest and moves forward with constant velocity for five seconds. It then stops for five seconds and continues forward with constant velocity for 10 seconds. Which of the following graph correctly describes this situation?

(A) (B) (C) (D) (E)

22. By referring to the slope of the distance-time graph below. Which one of the sentences best describes the motion of object AC?

(A) AC is moving slower than AB
(B) AC and AB have the same velocity
23. The displacement-time graph below represents an object moving motion during a 60s-time interval.

Which one of the following graphs of velocity versus time would best represent the object’s motion during the same time interval?

24. The velocity-time graph below represents an object’s motion during a 50s-time interval.

Which of the following graphs of acceleration versus time would best represent the object's motion during the same time interval?
25. An acceleration graph for an object during a 50s-time interval is represented below.

Which of the following velocity versus time represents the object’s motion during the same time interval?