MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2023, 19(11), em2360
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/13812

Development of learning path map of work and energy for high schoolers by using cognitive diagnostic assessment

Kittitas Wancham [1*], Kamonwan Tangdhanakanond [1]

[1] Faculty of Education, Chulalongkorn University, Bangkok, THAILAND

Received 04 August 2023 ▪ Accepted 11 October 2023

Abstract
The conception of work and energy is fundamental to learning physics and is essential to learning other subjects. However, most students still lack knowledge and understanding of work and energy. This may be due to previous research that aimed to develop students using similar teaching methods without considering the individual knowledge state of each student. We, thus, sought to develop the mastery test on work and energy and the learning path map of work and energy using cognitive diagnostic assessment. Participants were 537 tenth graders in Bangkok, Thailand, which were chosen by the multistage random sampling. The mastery test on work and energy developed is divided into six attributes, i.e., (1) work, (2) power, (3) kinetic energy, (4) gravitational potential energy, (5) elastic potential energy, and (6) law of conservation of energy. The test exhibited good psychometric properties, which were evaluated based on item parameters, content validity, construct validity, concurrent validity, classification consistency index, and classification accuracy index. The significant finding was the development of the learning path map of work and energy. The map illustrates students' learning progression in different attribute profiles regarding work and energy. It proves to be highly beneficial for teachers in designing personalized learning methods for individual students. Additionally, it allows for tracking the learning progress of students until they have a comprehensive understanding of work and energy in all its attributes.

Keywords: learning path map, work and energy, cognitive diagnostic assessment, G-DINA model framework

INTRODUCTION

The conception of work and energy is highly significant in physics and serves as a fundamental basis for learning other topics, such as momentum and collisions, electricity, and heat and gases. Additionally, work and energy are essential concepts in STEM (science, technology, engineering, and mathematics) education, where a deep understanding of work and energy is necessary for effective learning (Pellegrino & Hilton, 2012). Furthermore, work and energy are applied in other branches of learning, such as biology, chemistry, and geography (Saglam-Arslan & Kurnaz, 2011). Work and energy are also crucial topics for explaining everyday phenomena. Students, thus, need to have a good understanding of work and energy to learn physics effectively. However, work and energy are challenging topics for students at all levels due to their abstract nature, causing many students to lack a grasp of work and energy (Kassiavera et al., 2019; Pramesti et al., 2020). Mustofa et al. (2019) found that 28.33%-55.00% of 11th graders have mastery about work and energy. Takaoglu (2018) found that 6.96%-16.18% high schoolers have mastery about work and energy. Afif et al. (2017) found that 29.63% of high schoolers have mastery about work and energy. In addition, Rivaldo et al. (2020) found that only 11.26%-38.03% of undergraduates have mastery of work and energy.

In addition to the abstract nature that makes it difficult for students to understand the concepts of work and energy, the discrepancy between the everyday meaning of work and energy and their scientific meaning is another obstacle that prevents students from grasping the concepts. This is because students have

✉ wkittitas@outlook.com **(*Correspondence)** ✉ tkamonwan@hotmail.com

Contribution to the literature

- The mastery test on work and energy developed comes in a short-answer test format, created based on cognitive diagnostic assessment, to measure the mastery of work and energy in six specific attributes.
- The learning path map of work and energy illustrates the development of students' learning in work and energy based on their individual attribute profiles.
- It provides information to teachers to design learning methods suitable for each student's knowledge state.

misconceptions, meaning that they learn about work and energy based on incorrect understanding derived from their daily life experiences prior to studying this topic. As a result, their misconceptions hinder their accurate comprehension of the subject matter (Elisa et al., 2019; Mustofa et al., 2019; Rivaldo et al., 2020).

Due to the difficulties that come with learning about work and energy, there have been numerous research studies that aim to promote students' knowledge in this area. However, the majority of students still lack a significant understanding of work and energy (Kassiavera et al., 2019; Mustofa et al., 2019). This may be attributed to the fact that these research studies often focus on overall scores and overlook the mastery of work and energy in each attribute. In other words, students with the same overall scores may possess different levels of knowledge in various attributes, requiring different approaches to promote their learning. Therefore, it is crucial for teachers to design learning strategies that are suitable for the knowledge state of each individual student (Bai, 2020; Wu et al., 2023). In order to effectively develop students and cater to their individual needs, it is important to consider students' knowledge in each specific attribute by applying cognitive diagnostic assessment (CDA). CDA is able to assess students' thinking processes and provide detailed diagnostic information about their learning (Huang et al., 2022). Additionally, it enables the creation of a personalized learning path map, which can guide teachers in designing learning experiences that are suitable for each individual student.

A learning path map represents all possible learning paths for a student, indicating the sequential stages of beneficial learning development and providing guidance for tracking the progress of student's learning. It also establishes directions for individualized learning progression, starting from the initial point until achieving the specified goals (Bai, 2020; Chen et al., 2017; Wu et al., 2023). It should be noted that each student's learning path possesses unique characteristics based on their attribute profile or knowledge state. Moreover, the learning path map provides crucial information for teachers to design appropriate teaching strategies tailored to each student, including learning activities, tasks, teaching materials, communication methods, and assessment methods, ensuring that every student can learn and acquire knowledge in all attributes (Wu et al., 2022, 2023).

We, therefore, developed the learning path map of work and energy by using CDA to present information to physics teachers as a guideline for designing focused and appropriate learning experiences on the work and energy topic for each student. The objective was to ensure that students have comprehensive knowledge of work and energy, which will contribute to effective learning outcomes in physics. The study aimed to achieve two objectives, i.e.,

(1) developing and examining the psychometric properties of the mastery test on work and energy using CDA and

(2) developing the learning path map of work and energy using CDA.

LITERATURE REVIEW

The review of relevant literature and research is divided into two parts, i.e., CDA, and conceptions of work, and energy. The details are, as follows:

Cognitive Diagnostic Assessment

CDA is an educational assessment approach designed to diagnose specific attributes of individual students and provide detailed diagnostic information. It aims to identify students' strengths and weaknesses by providing personalized diagnostic information. Additionally, it provides teachers with information for instructional planning to enhance students' learning. The process of CDA involves five main steps, i.e.,

(1) determining assessment goals,

(2) identifying and validating the cognitive model,

(3) constructing and validating the Q-matrix,

(4) selecting cognitive diagnostic models (CDMs) for data analysis, and

(5) reporting the assessment results.

These steps are detailed in Ravand and Baghaei (2020), Rupp et al. (2010), and Wancham et al. (2023).

Determining assessment goals

Teachers must establish clear assessment goals that align with teaching and learning objectives. They should be able to effectively describe the attributes being measured and specify the number of attributes to be assessed, as well as the number of proficiency levels for each attribute, such as mastery or non-mastery.

Identifying & validating cognitive model

Teachers must identify all the attributes being measured in the test and define the attribute specification, also known as the cognitive model. The cognitive model represents the thought process that students use to answer the questions. The construction of the cognitive model considers hierarchical relationships between attributes based on relevant theories and research findings. Hierarchical relationships in the cognitive model functions as guidelines for creating the Q-matrix. The validation of the cognitive model can be done through three methods, i.e.,

(1) think-aloud protocol,

(2) eye-tracking studies, and

(3) expert judgment.

The think-aloud protocol is particularly important as it can be used to examine, verify, or modify the relationships between attributes in the cognitive model. Furthermore, teachers need to specify the details of the attributes being measured, and the required details to be specified are, as follows:

(1) The construct refers to the target of measurement, which is an internal characteristic of human beings.

(2) The definitional grain size refers to the depth of defining the measured attribute, depending on the desired level of depth among students. If the depth is low, the attribute will have a broader scope but less precision in its definition. The depth should align with the assessment goals.

(3) The attribute label refers to the words or phrase that identifies the significant meaning of the measured attribute.

(4) The attribute definition refers to text that describe the various aspects of the attribute in detail.

(5) The code for the attribute refers to details that use to identify the characteristics of the test items that measure the particular attribute.

Constructing & validating Q-matrix

After identifying the details regarding the attributes to be measured, the next step is to temporarily create a Q-matrix, done by considering the relationships between attributes in the cognitive model. Then, experts evaluate the accuracy and appropriateness of the Q-matrix creation.

Once the Q-matrix has been refined based on the recommendations from experts, the construction of items based on the tentative Q-matrix is carried out. Subsequently, the validation and adjustment process is conducted until a complete Q-matrix is obtained. Q-matrix is a table that indicates the attributes to be measured by each test item, with the test items listed in rows and the attributes listed in columns. The table consists of number 1 and 0 in which 1 represents the measurement of a specific attribute, and 0 represents the absence of measurement for that attribute. Each test item can measure one or multiple attributes. The process of creating a Q-matrix involves three steps, as follows:

Constructing of incidence matrix: The incidence matrix represents the number of possible test items used to measure all combinations of attributes when they are independent. It is a $K \times I$ matrix, where K is the number of attributes, and I is the total number of possible test items. The number of possible items is 2^K-1, excluding the item that do not measure any attribute. The incidence matrix consists of the numbers 1 and 0, indicating whether an attribute is measured or not by the corresponding test item, respectively.

Constructing of reachability matrix: The reachability matrix represents direct and indirect relationships among the specified attributes in a cognitive model. It is a $K \times K$ square matrix, where K is the number of attributes. The attributes are both in the rows and columns. The matrix consists of the numbers 0 and 1, where the values in the diagonal are 1. This means that if there is a direct or indirect relationship between the attribute in the corresponding row and the attribute in the corresponding column, it is represented by 1. If there is no relationship at all, it is represented by 0.

Constructing of reduced Q-matrix: The reduced Q-matrix represents the number of possible test items based on the relationships between the specified attributes in the cognitive model. The reduced Q-matrix is obtained by removing items or columns that do not adhere to the relationships between the specified attributes from the incidence matrix or by eliminating items that do not align with the columns in the reachability matrix. In this case, the reduced Q-matrix used as a guideline for item generation should be transposed in order to get the new Q-matrix that contains the items in the rows and the attributes in the columns.

Q-matrix validation can be performed using three methods, i.e.,

(1) think-aloud protocol,

(2) expert judgment, and

(3) empirical data analysis based on CDMs, such as examining mesa plots.

Mesa plots are line graphs that show the best q-vector (Q-matrix row vector) for each test item at the edge of the mesa. This method is effective and advantageous as it provides information about alternative q-vectors for revising the tentative Q-matrix. In this case, teachers should use multiple methods to validate the Q-matrix in order to gather strong evidence to support the accuracy of the Q-matrix.

Selecting cognitive diagnostic models for data analysis

This step involves analyzing the individual test item responses of students using CDM that appropriate to the data, the measured attributes, and the complete Q-matrix. The estimation of item parameters and examinee parameters, such as attribute profiles and mastery statuses for each attribute, will vary depending on the model used for data analysis. Before interpreting both item parameters and examinee parameters, teachers need to check the convergence of parameter estimation and assess the fit between the model and empirical data by considering fit indices. Additionally, the analysis of CDM provides classification consistency index and classification accuracy index, which are used to evaluate the reliability and validity of the diagnostic test. The chosen model for research purposes is G-DINA model framework, as it is a general model and allows for parameter constraints to get various popular sub-models, which differ based on the applied condensation rule.

G-DINA model framework was developed to alleviate the assumption of DINA model due to the fact that examinees who have non-mastery at least one attribute are assumed to have the same probability of answering the item correctly. G-DINA model framework, which can effectively differentiate the examinees, encompasses DINA model, DINO model, and A-CDM model. These models are used for analyzing dichotomous responses. The details of G-DINA model can be summarized, as follows: (de la Torre, 2011; de la Torre & Minchen, 2019)

G-DINA model divides examinees who respond to each item into $2^{K_j^*}$ groups, where K_j^* represents the number of attributes required to answer item j correctly. The probability that an examinee with attribute profile α_{ij}^* will answer item j correctly can be determined, as follows:

$$P\left(X_{ij}=1\,|\alpha_{ij}^*\right)=\delta_{j0}+\sum_{k=1}^{K_j^*}\delta_{jk}\alpha_{ik}+\sum_{k'=k+1}^{K_j^*}\sum_{k=1}^{K_j^*\text{-}1}\delta_{jkk'}\alpha_{ik}\alpha_{ik'}+\ldots+\delta_{j=1,2,\ldots,K_j^*}\prod_{k=1}^{K_j^*}\alpha_{ik}, \tag{1}$$

where α_{ij}^* is the attribute profile for the attributes used to answer item j by examinee i. δ_{j0} is the intercept for item j, meaning the probability that an examinee who does not possess any of the attributes required to answer item j correctly will answer the item correctly. δ_{jk} is the main effect due to α_k for item j, meaning the change in probability when an examinee with mastery in attribute k, which is used to answer item j, answers the item correctly. $\delta_{jkk'}$ is the interaction effect due to α_k and α_k for item j, meaning the change in probability when an examinee with mastery in both attribute k and k', which are used to answer item j, answers the item correctly. $\delta_{j=1,2,\ldots,K_j^*}$ is the interaction effect due to α_1, …, $\alpha_{K_j^*}$ for item j, meaning the change in probability when an examinee with mastery in attribute k up to K_j^*, which are used to answer item j, answers the item correctly.

When setting parameters in G-DINA model to 0, except for δ_{j0} and $\delta_{j=1,2,\ldots,K_j^*}$ G-DINA model becomes DINA model. The probability that an examinee with attribute profile α_{ij}^* answers item j correctly can be illustrated, as follows:

$$P(X_{ij}=1\,|\alpha_{ij}^*)=\delta_{j0}+\delta_{j12\ldots K_j^*}\prod_{k=1}^{K_j^*}\alpha_{ik}. \tag{2}$$

The result is $g_j=\delta_{j0}$ and $1-s_j=\delta_{j0}+\delta_{j=1,2,\ldots,K_j^*}$, where g_j is the guessing parameter, which refers to the probability that an examinee without mastery of at least one attribute measured in item j can answer the item correctly. While s_j refers to the slip parameter, which refers to the probability that an examinee with mastery of all the attributes measured in item j answers the item incorrectly. Therefore, DINA model has two item parameters per item (de la Torre, 2009).

When setting the parameters in G-DINA model to equal values, they are assigned, as follows: $\delta_{jk}=-\delta_{jk'k''}=\ldots=(\text{-}1)^{K_j^*+1}\delta_{j=1,2,\ldots,K_j^*}$ for $k=1,2,\ldots,K_{jh}^*$, $\acute{k}=1,2,\ldots,K_{jh}^*-1$ and $k''>k',\ldots,K_{jh}^*$ with only the parameters of δ_{j0} and δ_{jk} left. G-DINA model will transform into DINO model, where the probability that an examinee with attribute profile α_{ij}^* answers item j correctly can be determined, as follows:

$$P(X_{ij}=1\,|\alpha_{ij}^*)=\delta_{j0}+\delta_{jk}\alpha_{ik}. \tag{3}$$

The result is $g_j=\delta_{j0}$ and $1-s_j=\delta_{j0}+\delta_{jk}$ in which g_j is the guessing parameter and s_j is the slip parameter, but the meaning is different from DINA model. The guessing parameter refers to the probability that an examinee without mastery of any of the attributes measured to answer item j answers the item correctly. On the other hand, the slip parameter refers to the probability that an examinee with mastery of at least one of the attributes measured to answer item j answers the item incorrectly. Therefore, DINO model has two item parameters per item (Templin & Henson, 2006).

When all the interaction parameters in G-DINA model are set to zero, G-DINA model becomes the additive CDM or A-CDM model. In A-CDM model, it is possible to determine the probability that an examinee with attribute profile α_{ij}^* answers item j correctly, as follows:

$$P(X_{ij}=1\,|\alpha_{ij}^*)=\delta_{j0}+\sum_{k=1}^{K_j^*}\delta_{jk}\alpha_{ik}. \tag{4}$$

A-CDM model, therefore, has a total of K_j^*+1 parameters.

G-DINA, DINA, DINO, and A-CDM models have different condensation rules. In DINA model, the condensation rule is based on a conjunctive condensation rule, which estimates the probability of answering an item correctly when a student possesses all the required attributes. However, if the student lacks at least one attribute, they will answer the item incorrectly. In DINO model, the condensation rule is based on a disjunctive condensation rule, which estimates the probability of answering an item correctly when a student possesses at least one of the required attributes. However, if the student lacks all attributes, they will answer the item incorrectly. On the other hand, A-CDM model utilizes an additive condensation rule, where the probability of answering an item correctly depends on the cumulative mastery of each attribute measured in the item. Each attribute's contribution is considered independently, meaning that the more mastery a student has in a specific attribute required for the item, the higher the probability of answering the item correctly. In the case of G-DINA model, being a general model, it categorizes students into different groups based on the attributes measured in the item. Students with different levels of mastery in the required attributes will have different probabilities of answering the item correctly. It is worth noting that G-DINA model incorporates all three condensation rules (Ravand & Baghaei, 2020). We, thus, compared the fit of these four models with empirical data to select the model that best fits the data.

Reporting assessment results

The final step of CDA is reporting the attribute profile to each individual student based on the data analysis using CDMs. The score report should be tailored to each student and include information about mastery statuses in each attribute, strengths, weaknesses, suggestions for learning improvement, and interpretation guidelines. Additionally, the score report should be presented in the form of graphs and appropriate text that is easily understandable. Moreover, the information obtained from data analysis can be used to create a learning path map for each student, which serves as a guide for teachers in designing personalized learning experiences for students.

Conceptions About Work & Energy

Based on the review of additional physics textbook at the upper secondary level in Thailand conducted by The Institute for the Promotion of Teaching Science and Technology (2020), as well as international physics textbooks such as Hewitt (2015)'s and Serway and Vuille (2018)'s physics textbooks, the gist of work and energy can be summarized, as follows.

Work

Work is the result of exerting force on an object, causing the object to undergo displacement in the direction of the force. Work can be calculated using the formula $W = F\Delta x cos\theta$.

Work can have positive, negative, or zero values depending on the angle between the force $\vec{F}$ and the displacement $\Delta\vec{x}$, as follows:

(1) When the force $\vec{F}$ and the displacement $\Delta\vec{x}$ have the same direction (θ=0°) or form an acute angle (0°<θ<90°), the work done by the force will be positive.

(2) When the force $\vec{F}$ and the displacement $\Delta\vec{x}$ are perpendicular to each other (θ=90°), the work done by the force will be zero.

(3) When the force $\vec{F}$ and the displacement $\Delta\vec{x}$ have the opposite directions (θ=180°) or form an obtuse angle (90°<θ<180°), the work done by the force will be negative.

Power

Power is the rate at which work is done or the amount of work that occurs per unit of time. It is used to indicate the ability to perform work efficiently within a specific time interval. Generally, power refers to average power, which can be calculated using the formula $P_{av} = \frac{W}{\Delta t}$.

Mechanical energy

Energy refers to the quantity that indicates the ability to perform work, resulting in changes such as altering the state of motion, transforming into other forms of energy, or changing states. Mechanical energy, on the other hand, refers to the energy generated by the motion or movement of an object, which is the sum of kinetic energy and potential energy. Mechanical energy can be divided into two types, i.e., kinetic energy and potential energy, as detailed below.

Kinetic energy: Kinetic energy is the energy of an object in motion or with velocity and can be calculated using the formula $E_k = \frac{1}{2}mv^2$. The relationship between work and kinetic energy is, as follows: The work done by the net force acting on an object is equal to the change in its kinetic energy. This change in kinetic energy can either increase or decrease, depending on the direction of the net force acting on the object. Specifically, if the net force acts in the same direction as the object's motion, the kinetic energy of the object increases. On the other hand, if the net force acts in the opposite direction to the object's motion, the kinetic energy of the object decreases. Additionally, if the net force is zero, the kinetic energy of the object remains constant.

Potential energy: Potential energy is the energy associated with the position or shape of an object. In mechanics, potential energy is related to two types of potential energy, i.e., gravitational potential energy and elastic potential energy. Details are, as follows:

Gravitational potential energy: Gravitational potential energy is the energy associated with the position of an object in a gravitational field. It can be calculated using the formula $E_p = mgh$. The work done and gravitational potential energy are related, as follows: The work done in lifting an object vertically at a constant speed is equal to the change in gravitational potential energy, which can either increase or decrease depending on the direction of the object's motion. Specifically, when lifting an object to a higher level from its original level, the gravitational potential energy increases. Conversely, when lowering an object from its original level, the gravitational potential energy decreases. For an object at a fixed position, the gravitational potential energy remains constant.

Elastic potential energy: Elastic potential energy is the energy associated with the position under the influence of an elastic force, such as a spring. In other words, elastic potential energy is the energy stored in a spring when it is stretched or compressed from its equilibrium position. It can be calculated using the formula $E_{p_s} = \frac{1}{2}kx^2$. The work done and elastic potential energy are related, as follows: the work done by a spring force is equal to the change in elastic potential energy, which can either increase or decrease depending on the difference between the initial and final positions. Specifically, if the distance between the initial and final positions increases, the elastic potential energy increases. Conversely, if the distance decreases, the elastic potential energy decreases. For an object at the same position, the elastic potential energy remains constant.

Law of conservation of energy

The law of conservation of mechanical energy stipulates that the mechanical energy of an object remains constant in all situations when only conservative forces act on the object. Conservative forces are forces that, when applied to an object, result in work that does not depend on the path of motion. Examples of conservative forces are gravitational force and spring force. On the other hand, nonconservative forces are forces that, when applied to an object, result in work that depends on the path of motion. If a nonconservative force acts on an object, the total kinetic energy and potential energy of the system will not remain constant. For example, when an object moves under the influence of friction, which is a nonconservative force, the work done by friction will cause a loss of mechanical energy in the system. The amount of energy lost is equal to the work done by friction. However, when the lost energy is combined with the remaining mechanical energy, the total energy will remain constant, following the law of conservation of energy, which states that the total energy of a system is conserved and can be transformed from one form to another.

METHOD

Informants

The informants are divided into three groups, as follows:

(1) Experts for validating the cognitive model of work and energy, as well as related details. This group consists of five experts in physics education.

(2) Experts for validating the Q-matrix, including five experts in physics education and two experts in educational measurement and assessment. The physics education experts should have completed their studies in the field of education with a major in physics and have at least five years of teaching experience. The educational measurement and assessment experts should have completed their doctoral studies in the field of educational measurement and assessment, have at least five years of teaching or working experience, and have experience in research related to CDA.

(3) 10 10^{th} graders, consisting of five male and five female students. They will be involved in data collection using the think-aloud protocol while taking the mastery test on work and energy. The collected data will be used to revise the Q-matrix and the cognitive model of work and energy.

Participants

The participants consist of 537 tenth graders in Bangkok, Thailand. The participants are divided into 264 male students (49.16%), and 273 female students (50.84%). This sample size is sufficient for analyzing students' responses using G-DINA model framework. It is recommended to have a sample of at least 500 individuals to accurately estimate the parameters (Hu et al., 2016).

The participants were obtained through multistage random sampling. In the first stage, schools were randomly selected using simple random sampling, resulting in five selected schools. In the second stage, tenth graders within each selected school were randomly sampled with 120 students selected from each school, equally divided between males and females (60 students per gender). This resulted in a total of 600 students. However, only 537 students completed all test items and agreed to participate in the study.

Materials

The materials consist of a validation form and a think-aloud record form. The details are, as follows.

Validation form

The validation form is used to assess the alignment of the test items with the Q-matrix. It is completed by five experts in physic education who identify the attributes of work and energy measured by each test item. The form consists of 18 items that require experts to specify the attributes.

Think-aloud record form

The think-aloud record form is used to record students' cognitive processes while they are thinking aloud when taking the mastery test on work and energy. It is applied in a pilot study involving a small sample of ten students who exhibit similarities to the research participants. The purpose of this data collection is to gather data for refining the Q-matrix and the cognitive model of work and energy. Additionally, the form is used to record the results of diagnosing the students' mastery of work and energy in the research participants, which includes 20 students. The diagnostic results are employed to analyze the concurrent validity of the test. The attributes to be diagnosed include:

(1) work,

(2) power,

(3) kinetic energy,

(4) gravitational potential energy,

(5) elastic potential energy, and

(6) the law of conservation of energy.

Procedure

We conducted the following steps to develop the mastery test on work and energy and the learning path map of work and energy:

(1) Specify the attributes of work and energy to be measured.

(2) Identify the details of the attributes to be measured, which are divided into four components, i.e.,

(a) construct,

(b) attribute label,

(c) attribute definition, and

(d) code for the attribute.

(3) Develop a cognitive model of work and energy by considering the hierarchical relationships among the attributes to be measured.

(4) Have five experts in physics education review the accuracy and appropriateness of the specified attributes, the details of the attributes, and the cognitive model of work and energy through interviews. Then, revise the specification of the attributes, the details of the attributes, and the cognitive model based on the experts' suggestions.

(5) Create the Q-matrix using data from the revised cognitive model of work and energy.

(6) Have three experts in physics education and another two in educational measurement and assessment review and verify the accuracy and suitability of the Q-matrix through interviews. Then, revise the Q-matrix according to the experts' suggestions.

(7) Develop the mastery test on work and energy that aligns with the revised Q-matrix.

(8) Have five experts in physics education identify the attributes of work and energy to be measured in each test item using the validation form, in order to use the data to further improve the Q-matrix.

(9) Pilot the mastery test on work and energy with a small sample group that resembles the research participants. The group consists of ten students, divided equally into five males and five females. The think-aloud protocol is employed, where students verbalize their thoughts while taking the test. Collecting data through think-aloud protocols typically requires a minimum sample size of five-10 participants (Trenor et al., 2011). We examine the cognitive processes of students in answering the test items and record them in the think-aloud record form. The data is then analyzed by using the content analysis and so that the Q-matrix and the cognitive model of work and energy can be improved. Additionally, the understanding of language usage in the items is also checked.

(10) Administer the mastery test on work and energy with the research participants to utilize the item responses in assessing the psychometric properties of the test and validating the Q-matrix. The Q-matrix validation is done by examining the mesa plot. Then, consult with five experts in physics education to make necessary adjustments to ensure its accuracy.

(11) Diagnose the mastery of work and energy of the research participants using the think-aloud protocol. Randomly select 20 students as the samples, consisting of 10 males and 10 females. Analyzing the congruence of the diagnostic results using Cohen's kappa requires a sample size of at least 10-30 individuals (Bujang & Baharum, 2017). We record the diagnostic results in the think-aloud record form to identify students' mastery statuses for each attribute.

(12) Construct the learning path map of work and energy by considering the attribute profiles of the research participants and the hierarchical relationships of the measured attributes in the cognitive model of work and energy.

Data Analysis

(1) Perform a content analysis to analyze the data collected by interviewing experts in physics education about the accuracy and suitability of the identification of the measured attributes, the specification of the details of the measured attributes, and the cognitive model of work and energy.

(2) Perform a content analysis to analyze the data collected by interviewing experts in physics education and experts in educational measurement and assessment about the accuracy and appropriateness of the Q-matrix.

(3) Analyze the agreement among experts' judgment results about identifying attributes to be measured in each item using the proportion of agreement according to Fleiss' (1971) formula.

(4) Compare the fit of CDMs used for analyzing response data, including G-DINA, DINA, DINO, and A-CDM models, to select the model that best fits with the empirical data. This is done by considering absolute fit indices, i.e., $RMSEA_2$ and SRMSR, which should have values not exceeding 0.05 (Liu et al., 2018). Then, choose the model that meets the fit criteria. If multiple models meet the fit criteria, consider relative fit indices, i.e., -2LL (-2 log likelihood), AIC (Akaike information criterion), and BIC (Bayesian information criterion) for model selection.

(5) Validate the Q-matrix by considering the mesa plot using GDINA R package.

(6) Estimate item parameters of the mastery test on work and energy using GDINA R package.

(7) Analyze the psychometric properties of the mastery test on work and energy, including reliability and validity, as follows:

   (a) Analyze the classification consistency index and classification accuracy index for each attribute using GDINA R package.

   (b) Analyze the construct validity of the test using confirmatory factor analysis with Mplus software and assess the fit between the measurement model and the empirical data using absolute fit indices. The criteria for evaluation are based on Kline (2016) and Weston and Gore Jr (2006).

      i. Chi-square value is not statistically significant.

      ii. CFI has a value of at least 0.95.

      iii. TLI as a value of at least 0.95.

      iv. RMSEA has a value not exceeding 0.06.

      v. SRMR has a value not exceeding 0.08.

   (c) Analyze the concurrent validity by examining the agreement between diagnostic results using G-DINA model framework and the think aloud protocol. Use Cohen's kappa to analyze the agreement and utilize the R package.

(8) Estimate examinee parameters, i.e., attribute profile, using GDINA R package.

RESULTS

Results of Developing & Examining Psychometric Properties of Mastery Test on Work & Energy

In the process of developing a mastery test on work and energy, we have defined the targeted attributes, constructed a cognitive model for work and energy, specified details regarding the measured attributes, and created and validated the Q-matrix as a guideline for test development. Subsequently, the psychometric properties of the mastery test on work and energy were examined. The presentation of the research findings is divided into four topics, i.e.,

(1) defining the targeted attributes and constructing the cognitive model for work and energy,

(2) specifying details regarding the measured attributes of work and energy,

(3) creating and validating the Q-matrix, and

(4) examining the psychometric properties of the mastery test on work and energy.

Defining targeted attributes & constructing cognitive model for work & energy

The developed mastery test on work and energy is a short answer test, where a score of one is given to a correct answer and zero to an incorrect answer. The mastery test contains six attributes, which are

(1) work,

(2) power,

(3) kinetic energy,

(4) gravitational potential energy,

(5) elastic potential energy, and

(6) law of conservation of energy.

Based on hierarchical relationships among these attributes, a cognitive model for work and energy can be constructed. The model represents the knowledge that understanding one attribute requires prior knowledge of other attributes, as detailed below.

The attribute of work serves as a foundation for learning the other four attributes, which are power, kinetic energy, gravitational potential energy, and elastic potential energy. Additionally, the attributes of kinetic energy, gravitational potential energy, and elastic potential energy are foundational for learning the concept of the law of conservation of energy. Therefore, hierarchical relationships among all six attributes, or cognitive model of work and energy, can be depicted as shown in **Figure 1**. The cognitive model of work and energy has been validated through expert review in physics education and the think-aloud protocol.

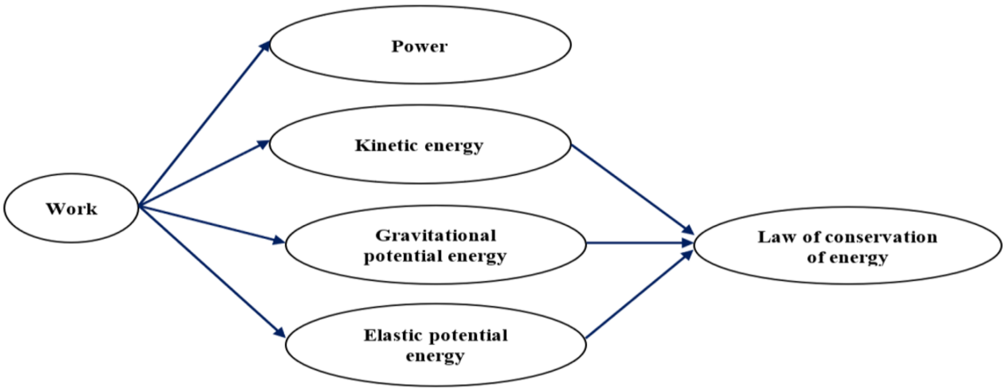

Figure 1. Cognitive model of work & energy (Source: Authors' own elaboration)

Table 1. Details in measuring mastery about work

Variable	Definition
Construct	Work & energy
Attribute label	Work
Attribute definition	Work is result of exerting force on an object, causing object to undergo displacement in direction of force. It can be calculated, as follows: $W=F\Delta x\cos\theta$, where W is work done by a constant force $\vec{F}$ measured in joules (J), F is magnitude of force measured in newtons (N), Δx is magnitude of displacement measured in meters (m), & θ is an angle between force $\vec{F}$ & displacement $\Delta\vec{x}$.
Code for attribute	Item that measures this attribute assigns a code of 1, while the item that does not measure this attribute assigns a code of 0. Item that measures this attribute has one of following formats: (1) Calculating work done by a single force acting on an object moving on a horizontal surface, an inclined plane, or in vertical direction under gravity, at any angle to displacement. Force may include frictional force or air resistance, or it may not. (2) Calculating work done by a resultant force acting on an object moving on a horizontal surface, an inclined plane, or in vertical direction under gravity, at any angle to displacement. Resultant force may include frictional force or air resistance, or it may not. (3) Calculating quantities related to work using formula for an object moving on a horizontal surface, an inclined plane, or in vertical direction under gravity. Calculations may involve frictional force or air resistance, or not.

Table 2. Details in measuring mastery about power

Variable	Definition
Construct	Work & energy
Attribute label	Power
Attribute definition	Power is rate at which work is done or amount of work that occurs per unit of time. Power refers to average power. It is calculated, as follows: $P_{av} = \frac{W}{\Delta t} = Fv_{av}$, where P_{av} is average power measured in watts (W), W is work done measured in joules (J), Δt is time measured in seconds (s), F is magnitude of a constant force measured in newtons (N), & v_{av} is average speed measured in meters per second (m/s).
Code for attribute	Item that measures this attribute assigns a code of 1, while item that does not measure this attribute assigns a code of 0. Item that measures this attribute has one of following formats: (1) Calculating power of an object moving on a horizontal surface, an inclined plane, or in vertical direction under gravity. It may involve frictional force or air resistance, or it may not. (2) Calculating quantities related to power using formulas for an object moving on a horizontal surface, an inclined plane, or in vertical direction under gravity. Calculations may involve frictional force or air resistance, or they may not.

Specifying details regarding measured attributes of work & energy

We have defined the details regarding the attributes being measured in work and energy, comprising all six attributes, with four key points, i.e.,

(1) construct,

(2) attribute label,

(3) attribute definition, and

(4) attribute code.

These details function as guidelines for constructing the test items that align with the definitions of each attribute. We considered additional content from various sources, including the physics textbooks developed by Hewitt (2015), National Science Teaching Association (2017), Serway and Vuille (2018), and The Institute for the Promotion of Teaching Science and Technology (2020). These sources were used to establish the details about the attributes to be measured. The accuracy and suitability of these details have been reviewed and confirmed by physics education experts.

Table 1 shows details in measuring mastery about work. **Table 2** shows details for measuring mastery about power.

Table 3 shows details for measuring mastery about kinetic energy.

Table 3. Details in measuring mastery about kinetic energy

Variable	Definition
Construct	Work & energy
Attribute label	Kinetic energy
Attribute definition	Kinetic energy is energy of an object in motion or with velocity. It is calculated, as follows: $E_k = \frac{1}{2}mv^2$, where E_k is kinetic energy measured in joules (J), m is mass measured in kilograms (kg), & v is speed measured in meters per second (m/s).
Code for attribute	Item that measures this attribute assigns a code of 1, while item that does not measure this attribute assigns a code of 0. Item that measures this attribute has one of the following formats: (1) Calculating kinetic energy of an object moving on a horizontal surface, an inclined plane, or in vertical direction under gravity without any frictional force or air resistance. (2) Calculating quantities related to kinetic energy using formula for an object moving on a horizontal surface, an inclined plane, or in vertical direction under gravity without any frictional force or air resistance.

Table 4. Details in measuring mastery about gravitational potential energy

Variable	Definition
Construct	Work & energy
Attribute label	Gravitational potential energy
Attribute definition	Gravitational potential energy is energy accumulated in an object within a gravitational field. It is calculated, as follows: E_p=mgh, where E_p is gravitational potential energy measured in joules (J), m is mass measured in kilograms (kg), g is gravitational acceleration measured in meters per second squared (m/s^2), & h is height of an object above a reference level measured in meters (m).
Code for attribute	Item that measures this attribute assigns a code of 1, while item that does not measure this attribute assigns a code of 0. Item that measures this attribute has one of following formats: (1) Calculate gravitational potential energy of an object moving on an inclined plane or in vertical direction under gravity without any frictional force or air resistance. (2) Calculate quantities related to gravitational potential energy using formula for an object moving on an inclined plane or in vertical direction under gravity without any frictional force or air resistance.

Table 5. Details in measuring mastery about elastic potential energy

Variable	Definition
Construct	Work & energy
Attribute label	Elastic potential energy
Attribute definition	Elastic potential energy is energy stored in an object subjected to elastic force, such as a spring. It is calculated, as follows: $E_{p_s} = \frac{1}{2}kx^2$, where E_{p_s} is elastic potential energy measured in joules (J), k is a spring constant measured in newtons per meter (N/m), & x is distance that spring is stretched or compressed away from equilibrium position measured in meters (m).
Code for attribute	Item that measures this attribute assigns a code of 1, while item that does not measure this attribute assigns a code of 0. Item that measures this attribute has one of following formats: (1) Calculate elastic potential energy of an object attached to a spring moving on a horizontal plane, an inclined plane, or in vertical direction without any frictional force or air resistance, while applying Hooke's law if applicable. (2) Calculate quantities related to elastic potential energy using formula for an object attached to a spring moving on a horizontal plane, an inclined plane, or in vertical direction without any frictional force or air resistance, while applying Hooke's law if applicable.

Table 4 shows details for measuring mastery about gravitational potential energy.

Table 5 and **Table 6** shows details for measuring mastery about elastic potential energy and law of conservation of energy, respectively.

Creating & validating Q-matrix

We conducted the following steps to create a Q-matrix as a guideline for constructing the mastery test on work and energy. There are four steps involved, i.e.

(1) creating an incidence matrix,

(2) creating a reachability matrix,

(3) creating a reduced Q-matrix, and

(4) generating a Q-matrix to serve as a guideline for constructing the test items by converting the reduced Q-matrix in order to get the new Q-matrix that contains the items in the rows and the attributes in the columns, as shown in **Table 7**.

Accuracy of Q-matrix was validated by experts in educational measurement and assessment. Furthermore, the feasibility of constructing test items based on the Q-matrix was evaluated by experts in physics education.

We validated the Q-matrix using three methods, i.e.,

(1) expert judgment,

(2) think-aloud protocol, and

(3) mesa plot analysis.

The results of each method are, as follows.

Table 6. Details in measuring mastery about law of conservation of energy

Variable	Definition
Construct	Work & energy
Attribute label	Law of conservation of energy
Attribute definition	Law of conservation of energy stipulates that total energy of a system remains constant, but it can be transformed from one form of energy to another.
Code for attribute	Item that measures this attribute assigns a code of 1, while item that does not measure this attribute assigns a code of 0. Item that measures this attribute has one of following formats: (1) Apply law of conservation of energy to calculate total energy, kinetic energy, gravitational potential energy, elastic potential energy, or related quantities of an object moving on a horizontal plane, an inclined plane, in vertical direction under gravity, or an object attached to a spring, considering at least two positions. Only conservative forces act on object, & there is mechanical energy involved of at least two types. (2) Apply law of conservation of energy to calculate total energy, kinetic energy, gravitational potential energy, elastic potential energy, work done by external forces, or related quantities of an object moving on a horizontal plane, an inclined plane, in vertical direction under gravity, or an object attached to a spring, considering at least two positions. Both conservative & nonconservative forces act on object, & there is mechanical energy involved of at least two types. Additionally, there is work done by at least one external force.

Table 7. Q-matrix for constructing mastery test items about work & energy

Item	Work	Power	KE	GPE	EPE	LCE
1	1	0	0	0	0	0
2	1	1	0	0	0	0
3	1	0	1	0	0	0
4	1	1	1	0	0	0
5	1	0	0	1	0	0
6	1	1	0	1	0	0
7	1	0	1	1	0	0
8	1	1	1	1	0	0
9	1	0	0	0	1	0
10	1	1	0	0	1	0
11	1	0	1	0	1	0
12	1	1	1	0	1	0
13	1	0	0	1	1	0
14	1	1	0	1	1	0
15	1	0	1	1	1	0
16	1	1	1	1	1	0
17	1	0	1	1	1	1
18	1	1	1	1	1	1

Note. KE: Kinetic energy; GPE: Gravitational potential energy; EPE: Elastic potential energy; & LCE: Law of conservation of energy

For expert judgment, five physics education experts were involved in assessing test items, consisting of 18 items, constructed based on the Q-matrix. They identified the targeted attributes for each item. The evaluation indicated that out of the 16 items, all five experts agreed that the test items measured the targeted attributes that aligned with the Q-matrix, with a proportion of agreement equal to 1.00 (P_i=1.00). However, for items 14 and 15, one expert had a different opinion compared to the other four experts. These items still aligned with the Q-matrix, but the proportion of agreement was 0.60 (P_i=0.60).

The think-aloud protocol was conducted by having students verbalize their thoughts while taking the mastery test on work and energy. This was done to examine the thought processes students used in answering the questions and to consider the attributes being assessed by each test item. The evaluation revealed that all 18 test items measured the targeted attributes according to the Q-matrix.

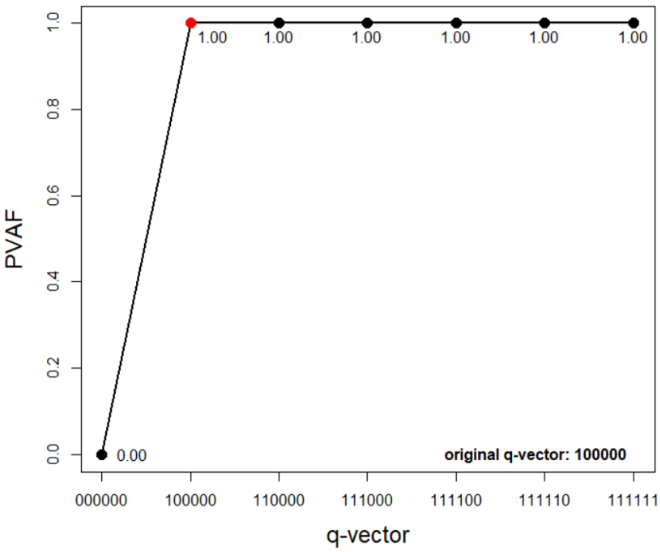

Figure 2. Mesa plot for item 1 (Source: Authors' own elaboration)

Mesa plot analysis was conducted by examining the best q-vectors for each test item at the mesa edges. An example is shown in **Figure 2**, which is the mesa plot for test item 1. The interpretation of the results indicates that the appropriate q-vector for item 1 is 100000 (measuring the attribute of work), which aligns with the Q-matrix that was created. The evaluation found that out of the 17 test items measured the attributes according to the Q-matrix.

However, for test item 10, the mesa plot analysis revealed that the suitable q-vector is 100010, indicating that it measures the attributes of work and elastic potential energy. When the five physics education experts examined the mesa plot analysis results, they all agreed that item 10 should assess the attribute of power, in line with the original Q-matrix or having a q-vector of 110010. Therefore, we retained the q-vector for item 10 as originally assigned.

Table 8. Fit indices for DINA, DINO, A-CDM, & G-DINA

Model	$RMSEA_2$	SRMSR	-2LL (*df*)	AIC	BIC
DINA	0.05	0.07	7,341.02 (99)	7539.02	7963.34
DINO	0.05	0.12	8,124.55 (99)	8322.55	8746.87
A-CDM	0.08	0.05	7,183.74 (140)	7463.74	8063.78
G-DINA	-	0.03	6,902.87 (321)	7544.87	8920.68

Note. For G-DINA, it is not possible to estimate the value of $RMSEA_2$ due to insufficient *df* for estimation

Table 9. Guessing parameters & slipping parameters of mastery test

Item	Guessing	Slip	Item	Guessing	Slip
1	0.09	0.03	10	0.00	0.06
2	0.07	0.11	11	0.02	0.10
3	0.02	0.08	12	0.06	0.04
4	0.07	0.09	13	0.03	0.06
5	0.06	0.10	14	0.00	0.06
6	0.00	0.06	15	0.00	0.07
7	0.04	0.08	16	0.00	0.07
8	0.03	0.07	17	0.00	0.17
9	0.00	0.07	18	0.00	0.17

Table 10. Classification consistency indices & classification accuracy indices of mastery test

Index	Attribute					
	Work	Power	Kinetic energy	Gravitational potential energy	Elastic potential energy	Law of conservation of energy
Classification consistency	0.99	0.88	0.88	0.89	0.88	0.63
Classification accuracy	0.99	0.90	0.89	0.90	0.90	0.64

Based on the validation of the Q-matrix using all three methods, it can be concluded that the constructed Q-matrix is accurate and can be effectively used to estimate the parameters of the test items and examinees with precision. Furthermore, the test items developed in accordance with the Q-matrix align with it, providing evidence to support the content validity of the mastery test on work and energy.

Examining psychometric properties of mastery test on work & energy

Before examining the psychometric properties of the mastery test on work and energy, we compared the fit of CDMs with the item responses by considering the absolute fit index, specifically SRMSR. We found that only G-DINA model demonstrated fit with the item responses, as its SRMSR was less than 0.05, as shown in **Table 8**. Therefore, we used G-DINA model to examine the psychometric properties of the mastery test on work and energy and analyzed the data in other aspects.

Based on the estimation of the item parameters of the 18 items using G-DINA model, it was found that there were a total of 258 item parameters, including the intercept parameters, the main effect parameters, and the interaction effect parameters. These item parameters do not directly indicate the quality of the item. Moreover, the large number of parameters makes it challenging to consider each parameter individually. Therefore, we assessed the quality of the items based on the guessing parameter and slipping parameter, each consisting of only two parameters. The guessing parameter refers to the probability that a student with non-mastery of all the specific attributes measured in that item will answer the item correctly.

The slipping parameter, on the other hand, refers to the probability that a student with mastery of all the specific attributes measured in that item will answer the item incorrectly. A high-quality item should have the value of the guessing and slipping parameters not exceeding 0.20 (Akbay & de la Torre, 2020). The estimation of the guessing and slipping parameters showed that the value of the guessing parameter ranged from 0.00 to 0.09, while the value of the slipping parameter ranged from 0.03 to 0.17, as shown in **Table 9**. This indicates that the 18 items are of sufficient quality and can be used to diagnose students' mastery about work and energy.

Based on the analysis of the classification consistency index and the classification accuracy index of the mastery test on work and energy, it was found that the test had a classification consistency index ranging from 0.63 to 0.99 when categorized by the measured attributes. Additionally, it had a classification accuracy index ranging from 0.64 to 0.99 when categorized by the measured attributes, as shown in **Table 10**.

Based on the examination of the construct validity of the mastery test on work and energy using confirmatory factor analysis, it was found that the measurement model of work and energy demonstrated a good fit with the empirical data (χ^2[4, N=537]=3.07, p=.55, CFI=1.00, TLI=1.00, RMSEA=0.00, SRMR=0.01). To improve model-data fit, the measurement errors were allowed to correlate. For standardized factor loadings, the factor loadings were statistically significant at a .05 level for all indicators and had positive values at a high level.

The factor loadings of the five indicators, i.e., work, power, kinetic energy, gravitational potential energy, and elastic potential energy, were similar in magnitude and higher than the factor loading of the law of conservation of energy, as shown in **Table 11** and **Figure 3**. To conclude based on above results, the mastery test on work and energy demonstrates construct validity.

Based on the results of analysis of concurrent validity considering the congruence between the diagnoses of mastery about work and energy using the think-aloud

Table 11. Confirmatory factor analysis results of measurement model of work & energy

Indicator	Unstandardized factor loading		Standardized factor loading		R^2
	b (SE)	t	β (SE)	t	
1. Work	1.00	-	0.99 (0.01)	566.95	0.98
2. Power	1.01 (0.01)	118.16	0.97 (0.01)	330.39	0.95
3. Kinetic energy	1.06 (0.01)	82.01	0.97 (0.01)	364.17	0.95
4. Gravitational potential energy	1.06 (0.01)	78.56	0.96 (0.01)	262.49	0.93
5. Elastic potential energy	1.06 (0.01)	72.85	0.96 (0.01)	281.52	0.93
6. Law of conservation of energy	0.75 (0.04)	21.31	0.72 (0.02)	33.94	0.52

Note. χ^2(4, N=537)=3.07, p=.55; CFI=1.00; TLI=1.00; RMSEA=0.00; SRMR=0.01; & all factor loadings were significant at .05

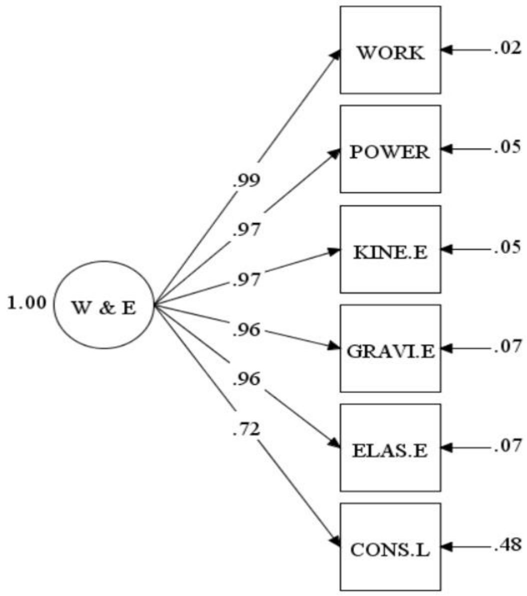

Figure 3. Measurement model of work & energy (Source: Authors' own elaboration)

Table 12. Attribute profiles of participants

Profile	Frequency	Percentage (%)
100000	122	22.72
111111	97	18.06
000000	87	16.20
111110	68	12.66
111000	39	7.26
110010	36	6.70
111100	19	3.54
110110	16	2.98
100010	14	2.61
101010	14	2.61
111010	10	1.86
110100	9	1.68
101110	5	0.93
100110	1	0.19
110000	0	0.00
101000	0	0.00
100100	0	0.00
101100	0	0.00
101111	0	0.00

Note. There were six binary digits (i.e., 1 indicating mastery & 0 indicating non-mastery) in each attribute profile that were sorted in order of all six attributes, which were (1) work, (2) power, (3) kinetic energy, (4) gravitational potential energy, (5) elastic potential energy, & (6) law of conservation of energy, respectively

protocol and the analysis from G-DINA model, it was found that the attribute profiles of students obtained from both methods were highly consistent (Rodrigues et al., 2019). Cohen's kappa coefficient was 0.82.

Based on the examination of the psychometric properties of the mastery test on work and energy from various sources of evidence, it can be concluded that the test has good psychometric properties. It can be effectively used to accurately diagnose mastery about work and energy.

Result of Developing Learning Path Map of Work & Energy

The total number of possible attribute profiles, when measuring six attributes, is 64 profiles. However, there are 19 attribute profiles that are possible according to the relationships among the attributes in the cognitive model of work and energy. These 19 profiles align with the Q-matrix created, as shown in **Table 7**, along with attribute profiles, where students lack mastery in all attributes, represented as "000000". When estimating the attribute profiles of the participants using G-DINA model, it was found that they had a total of 14 attribute profiles that corresponded to the relationships in the cognitive model of work and energy. Among these, the most prevalent profile was "100000" (mastery in work), with 122 individuals (22.72%), followed by the profile "111111" (mastery in all six attributes) with 97 individuals (18.06%), and the profile "000000" (lack of mastery in all six attributes) with 87 individuals (16.20%). However, five attribute profiles were not found in the participants, i.e.,

(1) 110000,

(2) 101000,

(3) 100100,

(4) 101100, and

(5) 101111, as shown in **Table 12**.

These findings support the accuracy of the relationships among the attributes in the cognitive model of work and energy.

When considering all 19 possible attribute profiles based on the relationships among the attributes in the cognitive model of work and energy, including hierarchical relationships, it is possible to create the learning path map of work and energy, as shown in

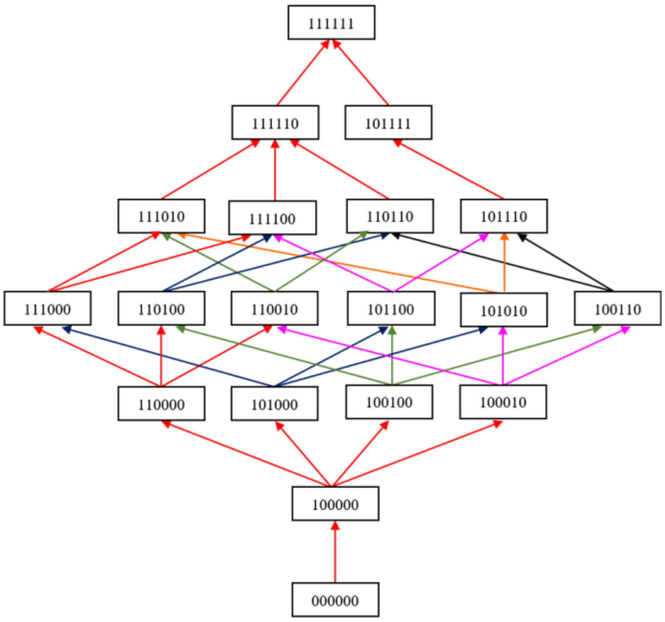

Figure 4. Learning path map of work & energy (Source: Authors' own elaboration)

Figure 4. This map helps determine the direction of students' learning development from the starting point until they acquire knowledge in all six attributes of work and energy. Some learning paths offer multiple options for learning development. Students can choose any path because ultimately, all paths lead to mastery in all attributes. For example, students with the profile "101100" (mastery in work, kinetic energy, and gravitational potential energy) have two possible learning paths. The first path is, as follows: 101100→ 111100→111110→111111. This means that students with the profile "101100" need to develop their knowledge in power, resulting in the profile becoming "111100". They then further develop their knowledge in elastic potential energy, resulting in the profile becoming "111110".

Finally, they develop their knowledge in the law of conservation of energy, resulting in the profile becoming "111111". As for the second path, it is, as follows: 101100→101110→101111→111111. In other words, students with the profile "101100" need to develop themselves to acquire knowledge in elastic potential energy, resulting in the profile becoming "101110". They then further develop their knowledge in the law of conservation of energy, resulting in the profile becoming "101111". Finally, they develop their knowledge in power, resulting in the profile becoming "111111". From both learning paths, it can be observed that students need to develop knowledge in elastic potential energy before learning about the law of conservation of energy. As for power, it can be developed either before or after learning about elastic potential energy and the law of conservation of energy.

Upon examining the learning paths of the majority of participants, which reflect the main learning development paths in work and energy, it was found that main learning paths are, as follows: 000000→ 100000→100010→110010→110110→111110→111111.

DISCUSSION

Discussion on Results of Developing & Examining Psychometric Properties of Mastery Test on Work & Energy

We developed the mastery test on work and energy according to CDA approach, with a strong emphasis on setting assessment goals, identifying and validating the cognitive model, constructing and validating the Q-matrix, and selecting CDMs for data analysis (Ravand & Baghaei, 2020; Rupp et al., 2010; Wancham et al., 2023). As a result, the mastery test on work and energy demonstrates good psychometric properties. Specifically, the 18 items exhibit high quality, and the test exhibits content validity, construct validity, and concurrent validity. Furthermore, the test exhibits high classification consistency index and classification accuracy index categorized by the measured attributes, except for the attribute of law of conservation of energy, which shows relatively low values for both indices. This may be due to the fact that the attribute of law of conservation of energy is measured by only two test items. Therefore, it is recommended to have at least three test items measuring each attribute to ensure higher classification consistency index and classification accuracy index (Javidanmehr & Sarab, 2017). Consequently, the test can be effectively used to accurately diagnose students' mastery in the area of work and energy.

The developed cognitive model of work and energy is based on hierarchical relationships between the attributes being measured. This hierarchical structure is known as the attribute hierarchy and follows a convergent attribute hierarchy pattern. The attribute hierarchy is specifications of the relationships among attributes that suggest that mastering an attribute requires mastering others first. It determines the attribute profile of the examinees, which is categorized into five types.

First, linear attribute hierarchy presents a hierarchy in which all attributes are arranged in a linear sequence, meaning that mastery in the final attribute requires mastery in all preceding attributes.

Second, convergent attribute hierarchy involves a hierarchical structure, where one parent attribute leads to multiple child attributes. In other words, examinees must possess mastery in at least one preceding attribute to demonstrate mastery in the lower-level attributes of the same branch. Third, divergent attribute hierarchy consists of multiple branches stemming from the same parent attribute. This means that mastery in any attribute within a branch requires mastery in the preceding attributes of that branch. Fourth, unstructured attribute hierarchy presents a hierarchy in which one attribute is a prerequisite for other distinct attributes. In other words, examinees can have mastery in lower-level attributes without having mastery in other lower-level attributes. Fifth, mixed structure comprises multiple sets of attribute hierarchies, where each set has no interrelationship. Additionally, the cognitive model may also include an independent structure, which allows mastery in one attribute without requiring prior knowledge in other attributes (Tu et al., 2019; Wancham et al., 2022). We also ensured the accuracy of the cognitive model of work and energy by conducting expert judgment and think-aloud protocols to verify the relationships between attributes. This process helped in creating an accurate Q-matrix, as the correctness of the Q-matrix depends on the accuracy of the cognitive model, which ultimately leads to the estimation of reasonable attribute profiles (Cai et al., 2018).

We created and validated a Q-matrix with rigorous scrutiny. We examined the accuracy through three methods, i.e., expert judgment, think-aloud protocol, and mesa plot analysis. These methods ensured confidence in the correct specification of the Q-matrix. Incorrect Q-matrix specification, known as Q-matrix misspecification, affects the accuracy of estimating item parameters and classifying examinees based on attribute profiles (Rupp & Templin, 2008). Q-matrix misspecification can occur at the attribute level or the item level. At the attribute level, it involves defining an attribute that is inappropriate. At the item level, it involves assigning incorrect values of one and zero in the Q-matrix. There are three types of misspecification, i.e.,

(1) under specification, where the value is set as zero when it should be one, indicating that the item should require that attribute for a correct response,

(2) overspecification, where the value is set as one when it should be zero, indicating that the item does not require that attribute for a correct response, and

(3) mixed specification, which combines both under specification and overspecification errors (Chen, 2017).

We, thus, provided details, including the definition of the targeted attributes, and sought expert assistance to ensure the accuracy and appropriateness.

Additionally, we validated the Q-matrix through all three methods to gather strong evidence supporting its accuracy, minimizing potential errors. This process ensures the confidence that the constructed Q-matrix is accurate, leading to accurate estimation of item parameters and attribute profiles.

Comparing the fit of CDMs with item responses, it was found that G-DINA model exhibited the best model-data fit. This indicates that attributes of work and energy possess compensatory attributes, meaning that lack of mastery in one attribute can be compensated by mastery in other attributes. Consequently, students with different attribute profiles have varying probabilities of answering test items correctly, depending on their knowledge of each specific attribute (de la Torre, 2011).

Discussion on Result of Developing Learning Path Map of Work & Energy

Generally, teachers aim to develop students' learning in the area of work and energy based on their overall scores. However, students with the same overall score may have different attribute profiles or knowledge states (Bai, 2020). This mismatch in learning development results in students not reaching their full potential, as evidenced by their lack of understanding in the majority of work and energy concepts (Afif et al., 2017; Mustofa et al., 2019; Rivaldo et al., 2020; Takaoglu, 2018). Since students with different attribute profiles require different learning methods, creating the learning path map of work and energy helps teachers plan their instruction. This includes designing learning activities, tasks, teaching materials, communication methods, and assessment methods to cater to individual students' learning development based on their attribute profiles.

Each student will have a unique learning path, starting from their initial attribute profile and progressing until they attain mastery in all attributes (Wu et al., 2022, 2023). It should be noted that the learning path map of work and energy exhibits a nonlinear pattern, which follows hierarchical relationships among work and energy attributes. The learning process begins with basic attributes in the attribute hierarchy and progresses towards the development of higher-order attributes.

If there are multiple paths for learning development, teachers should choose a path based on the students' background, learning environment, and available learning resources in order to design the most suitable learning experience for the students (Wu et al., 2023). Each learning path should adjust learning methods or activities to be appropriate for each individual student. Methods that can be used for developing mastery on work and energy include providing feedback (Wancham & Tangdhanakanond, 2022), simulation-based learning (Putranta & Wilujeng, 2019), and flipped classroom (Astra & Khumaeroh, 2019).

This learning path does not solely depend on hierarchical relationships of attributes but also on other factors, such as curriculum management in each country and curriculum enrichment activities (Wu et al., 2020). Thus, it is advisable to study the learning path map of work and energy in other countries in order to adjust the learning path map for students in each country. This will enhance effective learning of work and energy for students, considering that each country has its own educational model and varying social and cultural environments. Hence, there may be different variations of the learning path map for work and energy.

To ensure effective and well-defined development of students' learning in the area of work and energy, it is necessary to study the guidelines for developing students' learning in each attribute profile. This will create effective learning paths, as different learning paths may require different teaching methods (Bai, 2020). Furthermore, the learning path map of work and energy should be validated through longitudinal studies to analyze in-depth data on students' knowledge state regarding work and energy. This comprehensive understanding of students' learning paths will facilitate accurate adjustments to the learning path based on actual circumstances (Chen et al., 2017).

CONCLUSIONS

This study aimed to develop mastery test on work and energy and learning path map of work and energy using CDA. Developed test is divided into six attributes:

(1) work,

(2) power,

(3) kinetic energy,

(4) gravitational potential energy,

(5) elastic potential energy, and

(6) law of conservation of energy.

It possesses good psychometric properties, making it suitable to accurately diagnose students' mastery of work and energy. Also, it can be used to create a new version of mastery test on work and energy by utilizing information from details of the attributes and a Q-matrix as a guideline for test development. The learning path map of work and energy displays the learning progression of work and energy for students, consisting of 19 attribute profiles. This map is highly beneficial for teachers in designing personalized learning experiences that align with each student's knowledge state and individual needs. Additionally, it allows for monitoring students' learning progress along different learning paths until they achieve complete mastery of all attributes. This approach differs from the traditional one-size-fits-all teaching method that does not consider individual differences among students. Thus, learning path map of work and energy enables students to learn about work and energy more effectively.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: This study was supported by grants for development of new faculty staff, Ratchadaphiseksomphot Fund, Chulalongkorn University, Bangkok, Thailand.

Ethical statement: The authors stated that the study was approved by the Research Ethics Review Committee for Research Involving Human Subjects: The Second Allied Academic Group in Social Sciences, Humanities & Fine & Applied Arts at Chulalongkorn University, on Declaration of Helsinki, Belmont Report, CIOMS guidelines, & principle of international conference on harmonization–Good clinical practice (ICH-GCP) on May 16, 2023, with approval number 185/66. Written informed consents were obtained from the participants.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Afif, N. F., Nugraha, M. G., & Samsudin, A. (2017). Developing energy and momentum conceptual survey (EMCS) with four-tier diagnostic test items. *AIP Conference Proceedings, 1848*, 050010. https://doi.org/10.1063/1.4983966

Akbay, L., & de la Torre, J. (2020). Estimation approaches in cognitive diagnosis modeling when attributes are hierarchically structured. *Psicothema* [*Psychothema*], *32*(1), 122-129. https://doi.org/10.7334/psicothema2019.182

Astra, I. M., & Khumaeroh, S. I. (2019). The effect of flipped classroom model on student's physics learning outcome in work and energy concept. *Journal of Physics: Conference Series, 1318*, 012070. https://doi.org/10.1088/1742-6596/1318/1/012070

Bai, S. (2020). Developing a learning progression for probability based on the GDINA model in China. *Frontiers in Psychology, 11*, 569852. https://doi.org/10.3389/fpsyg.2020.569852

Bujang, M. A., & Baharum, N. (2017). Guidelines of the minimum sample size requirements for kappa agreement test. *Epidemiology, Biostatistics, and Public Health, 14*(2), e12267. https://doi.org/10.2427/12267

Cai, Y., Tu, D., & Ding, S. (2018). Theorems and methods of a complete Q matrix with attribute hierarchies under restricted Q-matrix design. *Frontiers in Psychology, 9*, 1413. https://doi.org/10.3389/fpsyg.2018.01413

Chen, F., Yan, Y., & Xin, T. (2017). Developing a learning progression for number sense based on the rule space model in China. *Educational Psychology, 37*(2), 128-144. https://doi.org/10.1080/01443410.2016.1239817

Chen, J. (2017). A residual-based approach to validate Q-matrix specifications. *Applied Psychological Measurement, 41*(4), 277-293. https://doi.org/10.1177/0146621616686021

de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics, 34*(1), 115-130. https://doi.org/10.3102/1076998607309474

de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika, 76*(2), 179-199. https://doi.org/10.1007/S11336-011-9207-7 de la Torre, J., & Minchen, N. D. (2019). The G-DINA model framework. In M. von Davier & Y. Lee (Eds.), *Handbook of diagnostic classification models: Models and model extensions, applications, software packages* (pp. 155-169). Springer. https://doi.org/10.1007/978-3-030-05584-4_7

Elisa, N., Kusairi, S., Sulur, S., & Suryadi, A. (2019). The effect of assessment for learning integration in scientific approach towards students' conceptual understanding on work and energy. *Momentum: Physics Education Journal, 3*(2), 103-110. https://doi.org/10.21067/mpej.v3i2.3761

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378-382. https://doi.org/10.1037/h0031619

Hewitt, P. G. (2015). *Conceptual physics*. Pearson Education Limited.

Hu, J., Miller, M. D., Huggins-Manley, A. C., & Chen, Y. H. (2016). Evaluation of model fit in cognitive diagnosis models. *International Journal of Testing, 16*(2), 119-141. https://doi.org/10.1080/15305058.2015.1133627

Huang, R., Liu, Z., Zi, D., Huang, Q., & Pan, S. (2022). A multi-level remedial teaching design based on cognitive diagnostic assessment: Taking the electromagnetic induction as an example. *Frontiers in Psychology, 13*, 851378. https://doi.org/10.3389/fpsyg.2022.851378

Javidanmehr, Z., & Sarab, M. R. A. (2017). Cognitive diagnostic assessment: Issues and considerations. *International Journal of Language Testing, 7*(2), 73-98.

Kassiavera, S., Suparmi, A., Cari, C., & Sukarmin, S. (2019). Student's understanding profile about work-energy concept based on multi-representation skills. *AIP Conference Proceedings, 2202*, 020060. https://doi.org/10.1063/1.5141673

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford.

Liu, R., Huggins-Manley, A. C., & Bulut, O. (2018). Retrofitting diagnostic classification models to responses from IRT-based assessment forms. *Educational and Psychological Measurement, 78*(3), 357-383. https://doi.org/10.1177/0013164416685599

Mustofa, Z., Sutopo, S., Mufti, N., & Asmichatin, A. (2019). The impact of modeling instruction based on system toward work-energy concept understanding. *Jurnal Penelitian dan Pengembangan Pendidikan Fisika* [*Journal of Physics Education Research and Development*], *5*(2), 145-154. https://doi.org/10.21009/1.05209

National Science Teaching Association. (2017). Topic arrangements of the next generation science standards. *Next Generation Science Standards*. https://static.nsta.org/ngss/AllTopic.pdf

Pellegrino, J. W., & Hilton, M. L. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. National Academies Press.

Pramesti, Y. S., Mahmudi, H., & Setyowidodo, I. (2020). Analyzing students' understanding of work-energy concept. *Journal of Physics: Conference Series, 1521*, 022016. https://doi.org/10.1088/1742-6596/1521/2/022016

Putranta, H., & Wilujeng, I. (2019). Physics learning by PhET simulation-assisted using problem based learning (PBL) model to improve students' critical thinking skills in work and energy chapters in MAN 3 Sleman. *Asia-Pacific Forum on Science Learning and Teaching, 20*(1), 3.

Ravand, H., & Baghaei, P. (2020). Diagnostic classification models: Recent developments, practical issues, and prospects. *International Journal of Testing, 20*(1), 24-56. https://doi.org/10.1080/15305058.2019.1588278

Rivaldo, L., Taqwa, M. R. A., Zainuddin, A., & Faizah, R. (2020). Analysis of students' difficulties about work and energy. *Journal of Physics: Conference Series, 1567*, 032088. https://doi.org/10.1088/1742-6596/1567/3/032088

Rodrigues, I. B., Adachi, J. D., Beattie, K. A., Lau, A., & MacDermid, J. C. (2019). Determining known-group validity and test-retest reliability in the PEQ (personalized exercise questionnaire). *BMC Musculoskeletal Disorders, 20*, 373. https://doi.org/10.1186/s12891-019-2761-3

Rupp, A. A., & Templin, J. (2008). The effects of Q-matrix misspecification on parameter estimates and classification accuracy in the DINA model. *Educational and Psychological Measurement, 68*(1), 78-96. https://doi.org/10.1177/0013164407301545

Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic assessment: Theory, methods, and applications*. Guilford.

Saglam-Arslan, A., & Kurnaz, M. A. (2011). Students' conceptual understanding of energy: Do the learning difficulties in energy concept discovered in the 1990s persist still? *Energy Education Science and Technology Part B: Social and Educational Studies, 3*(1), 109-118.

Serway, R. A., & Vuille, C. (2018). *College physics*. Cengage Learning.

Takaoglu, Z. B. (2018). Energy concept understanding of high school students: A cross-grade study. *Universal Journal of Educational Research, 6*(4), 653-660. https://doi.org/10.13189/ujer.2018.060409

Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods, 11*(3), 287-305. https://doi.org/10.1037/1082-989X.11.3.287

The Institute for the Promotion of Teaching Science and Technology. (2020). *Additional science and technology textbook, physics*. BOWT Printing House.

Trenor, J. M., Miller, M. K., & Gipson, K. G. (2011). *Utilization of a think-aloud protocol to cognitively validate a survey instrument identifying social capital resources of engineering undergraduates* [Paper presentation]. ASEE Annual Conference and Exposition. https://doi.org/10.18260/1-2--18492

Tu, D., Wang, S., Cai, Y., Douglas, J., & Chang, H. H. (2019). Cognitive diagnostic models with attribute hierarchies: Model estimation with a restricted Q matrix design. *Applied Psychological Measurement, 43*(4), 255-271. https://doi.org/10.1177/0146621618765721

Wancham, K., & Tangdhanakanond, K. (2022). Effects of feedback types and opportunities to change answers on achievement and ability to solve physics problems. *Research in Science Education, 52*(2), 427-444. https://doi.org/10.1007/s11165-020-09956-4

Wancham, K., Tangdhanakanond, K., & Kanjanawasee, S. (2022). The construction and validation of the cognitive model of force and motion for a diagnosis of misconceptions. *Journal of Education Naresuan University, 24*(3), 60-70.

Wancham, K., Tangdhanakanond, K., & Kanjanawasee, S. (2023). Sex and grade issues in influencing misconceptions about force and laws of motion: An application of cognitively diagnostic assessment. *International Journal of Instruction, 16*(2), 437-456. https://doi.org/10.29333/iji.2023.16224a

Weston, R., & Gore Jr, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist, 34*(5), 719-751. https://doi.org/10.1177/0011000006286345

Wu, X., Wu, R., Chang, H. H., Kong, Q., & Zhang, Y. (2020). International comparative study on PISA mathematics achievement test based on cognitive diagnostic models. *Frontiers in Psychology, 11*, 2230. https://doi.org/10.3389/fpsyg.2020.02230

Wu, X., Xu, T., & Zhang, Y. (2023). Research on the data analysis knowledge assessment of pre-service teachers from China based on cognitive diagnostic assessment. *Current Psychology, 42*, 4885-4899. https://doi.org/10.1007/s12144-021-01836-y

Wu, X., Zhang, Y., Wu, R., Tang, X., & Xu, T. (2022). Cognitive model construction and assessment of data analysis ability based on CDA. *Frontiers in Psychology, 13*, 1009142. https://doi.org/10.3389/fpsyg.2022.1009142