

PAPER • OPEN ACCESS

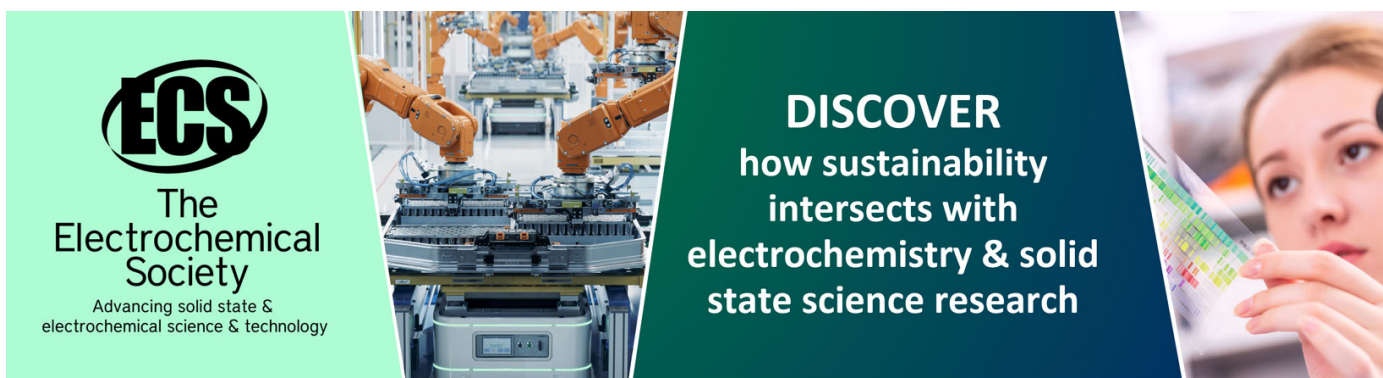
The validity of Four-Tier's misconception diagnostic test for Work and Energy concepts

To cite this article: S Anggrayni and F U Ermawati 2019 *J. Phys.: Conf. Ser.* **1171** 012037

View the [article online](#) for updates and enhancements.

You may also like

- [An investigation of students' conceptual understanding levels on fluid dynamics using four-tier test](#)
F H Dewi, A Samsudin and M G Nugraha
- [Four-Tier Diagnostic Test With Certainty of Response Index on The Concepts of Fluid](#)
R Diani, J Alfin, Y M Anggraeni et al.
- [The Development of a Four Tier-Based Diagnostic Test Diagnostic Assessment on Science Concept Course](#)
F. Fakhriyah and S. Masfuah



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

The validity of Four-Tier's misconception diagnostic test for Work and Energy concepts

S Anggrayni¹, F U Ermawati^{1*}

¹Physics Department, Faculty of Mathematics and Natural Science, Universitas Negeri Surabaya, Surabaya 60231, Indonesia

*E-mail: frida.ermawati@unesa.ac.id

Abstract. Misconceptions has been commonly found during teaching-learning process in senior high school, including in Physics subjects. These misconceptions may affect students' understanding of subsequent concepts. Therefore misconception should be detected earlier. This paper reports the work to develop a misconception diagnostic test instrument in the four-tier format for Work and Energy concepts and to analyze the validity of the instrument. The instrument was being tested to 50 students in one of the Science Classes, 11th grade in Senior High School 4 Sidoarjo. Based on the trials, it can be known that the developed instrument is valid, both internally and externally. The reliability index is also bigger than the reference value. This also implies that the intended instrument is reliable. Therefore it can be concluded that the developed four-tier format diagnostics test is proper to use.

1. Introduction

An effective learning process is a learning process that can condition students to gain their optimum achievement which correspond to their own capabilities. A good teacher should always try to create an effective learning in his/her classroom so that his/her students can easily understand the concepts delivered by their teacher. In fact, only some students can meet the above expectation. Students often faced the problems or learning difficulties so that they need help and support from the surrounding environment to solve such problems or difficulties. The most common educational problem is misconceptions. Suparno [1] defined that "Misconceptions is one's understanding of a concept that isn't consistent with the true concept". Misconceptions may also be meant as conceptual misunderstandings, as a result of the preconceptions mismatched between the student's preconception and the scientist's concept [2-3].

Based on the preliminary work held in Senior High School 4 Sidoarjo, it has been found that some misconceptions took place on the Work and Energy concepts. For example, in our daily activity, there are three conditions where someone are called or assumed to do a work (see Figure 1-3 for those three conditions). However, based on Physics concept, the intended work is zero. At the same time, in class, students have assumed that the work carried out by each person in Figure 1-3 are not zero. Therefore the different justification between the common people assumption and the Physics concept is oftenly confusing the students and therefore leading to misconception among them.



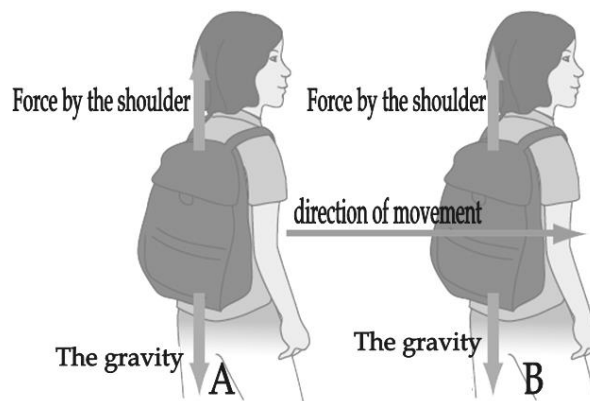


Figure 1. Condition 1 Someone walked from position A to position B while carrying a backpack [4].

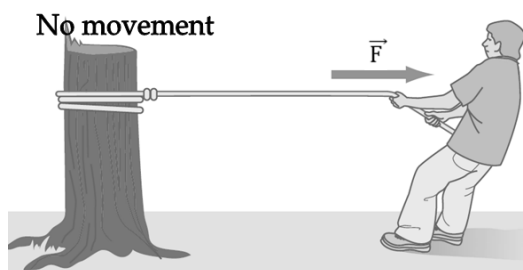


Figure 2. Condition 2 Someone pulls a tree with a force $\vec{F}$. However, the tree doesn't move [4].

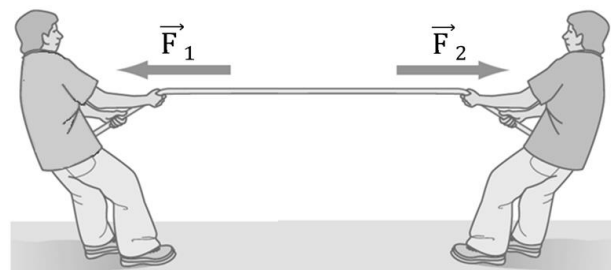


Figure 3. Condition 3 Two peoples are doing a tug of war, where $\vec{F}_1 = \vec{F}_2$, adapted from [4].

According to the Physics concepts, a Work is formulated by Equation (1) below,

$$W = \vec{F} \cdot \vec{s} = |\vec{F}| |\vec{s}| \cos \theta \quad (1)$$

Where, W is Work (scalar); $\vec{F}$ is a force (vector); $\vec{s}$ is a displacement (vector); and θ is the angle between the force and the displacement of the object.

In Figure 1, the lifting force on the shoulder directed perpendicular to the direction of the movement. Thus, in Eq. 1 $\cos \theta$ is zero because the value of $\cos 90^\circ$ is zero, so that the work carried out by the person in Figure 1 is also zero. Meanwhile the work resulting from the displacement from A to B positions is not zero. In Figure 2, the force to pull the tree gives rise to a work, however, the tree doesn't move. Therefore, based on the Eq. (1), the total work carried out by someone in Figure 2 is zero. Figure 3 shows that $\vec{F}_1 = \vec{F}_2$ so that $\sum \vec{F} = 0$. Thus, based on Equation 1 the work will be equals zero. Thus, the three mentioned conditions which causes to zero work are (1) when the force direction is perpendicular to the movement direction, (2) when there isn't movement on the system, and (3) when the total forces equal to zero.

As illustrated above, these misconceptions can be found from daily student's experiences [5-6]. These experiences will form a preconception in a student's mind. Preconception often formed automatically according to student's level of thinking, so the truth isn't guaranteed [2-3, 6-7]. When the preconception is wrong, this will affect their understanding to the subsequent concepts [3] because misconception is believed to be inhibiting to acceptance and integration of new knowledge or skills [8]. Therefore, it is necessary for a teacher to detect any misconception on his/her students's mind earlier to prevent any further effect on the subsequent Physics concept [3].

To diagnose student's misconceptions, a teacher usually use diagnostic tests. A diagnostic test is an assessment instrument which related to continuous or recurrent learning difficulties that are not solved

and as a cause of student's learning difficulties [7]. A suitable diagnostic test used for a large number of students is a multiple-choice diagnostic test. There are four formats of the multiple-choice diagnostic tests, i.e. (1) one-tier (2) two-tier, (3) three-tier, and (4) four-tier [9-10]. The four-tier diagnostic test consists of four components, namely answers of the questions, the level of answer confidence, reasons for choosing a particular answer and the level of reason confidence. This particular diagnostic test is the refinement of the similar diagnostic test with three-tier format which only consists of the first three components above. Similarly, the three-tier diagnostic test is an improvement of the same but two-tier diagnostic test. The intended improvement is the component of the reason for choosing an answer [10]. For that reason, a four-tier multiple choice diagnostic test can be assume to be the most accurate diagnostic test to identify misconceptions [9-10, 19]. Table 1 recapitulates the student's conception level that varies from understand a concept (U), partial understanding (PU), misconception (M), not understand a concept (NU), and un-code (UC) and the combination of possible answers (1st tier) and its level of confidence (2nd tier), possible reasons (3rd tier) and its level of confidence (4th tier).

Table 1. The combination of four-tier answers modified from [9].

Conception Level	1 st tier	2 nd tier	3 rd tier	4 th tier
Understand a concept (U)	Correct	Sure	Correct	Sure
Partial Understanding (PU)	Correct	Sure	Correct	Not sure
	Correct	Not sure	Correct	Sure
	Correct	Not sure	Correct	Not sure
	Correct	Sure	Incorrect	Sure
	Correct	Sure	Incorrect	Not sure
	Correct	Not sure	Incorrect	Sure
	Correct	Not sure	Incorrect	Not sure
	Incorrect	Sure	Correct	Sure
	Incorrect	Sure	Correct	Not sure
	Incorrect	Not sure	Correct	Sure
	Incorrect	Not sure	Correct	Not sure
Misconception (M)	Incorrect	Sure	Incorrect	Sure
Not Understand a concept (NU)	Incorrect	Sure	Incorrect	Not sure
	Incorrect	Not sure	Incorrect	Sure
	Incorrect	Not sure	Incorrect	Not sure
Un-code (UC)	There is "tier" which aren't answered or the answers is more than one available option			

A student is called to understand a concept if the answer of the question is correct and he/she sure for the answer, then the reason for choosing the reason is correct and he/she also sure for the reason. There are eleven of combinations that a student is called to partial understanding, as given in Table 1. A student is called to misconception if the answer of the question is incorrect but he/she sure for the answer, then the reason for choosing the reason is also incorrect but he/she sure for the reason. And there are three combinations that a student is called not understand a concept, as given in Table 1.

This paper is intended to develop a misconception diagnostic test instrument in the four-tier format for Work and Energy concepts and analyze the validity of the instrument. The function of this developed test is to assist the teacher in identifying the student's misconceptions on these concepts, so that the teacher can determine the appropriate learning strategy [8].

2. Research Method

A draft instrument of misconception diagnostics test is written (some of them is available in Appendix A) and tried to 30 new students (year 2017) in Physics Department The State University of Surabaya to collected kind of possible answers and the reasons. The developed instrument was then tested to the real student tasty (50 students in one of the Science Classes, 11th grade in Senior High School 4 Sidoarjo) to obtain the external validity and reliability of the test. The total number of the test is 25.

In addition to that, the internal validity of the developed instrument was also examined by asking three independent physics lectures. They feel out the questioner and asked to give feedback. Based on the feedback, the second draft of the instrument was developed (some of the developed misconception diagnostic test is provided in Appendix B). This validation includes three aspects, i.e. (a) content, (b) compliance with basic competencies (construction), and (c) languages. Equation (2) is employed the internal validity of the developed instrument.

$$P = \frac{S_R}{N \cdot P_A \cdot R} \cdot 100\% \quad (2)$$

Where, P is the percentage of internal validity; S_R is the total of respondent's scores; N is maximum score in the questioner; P_A is the total items of the questioner; and R is the total of respondent.

The obtained percentage of internal validity then was interpreted on Table 2, so that we can know the criteria of validity.

Table 2. The Percentage Interpretation of internal validity results [12].

Percentage	Criteria
0%-20%	Very low
21%-40%	Low
41%-60%	Valid enough
61%-80%	Valid
81%-100%	Very valid

Meanwhile, the external validity includes content and construction aspects. To determine the content validity of the developed instrument was being calculate the percentage of false positives (FP) and false negatives (FN) [13]. According to Kirbulut & Geban [14] explain that if the student's answers (in 1st tier) is right, the reason (in 3rd tier) is wrong, and the confidence choice (in 2nd and 4th tier) are sure, he/she is called false positives. Whereas if the student's answers (in 1st tier) is wrong, the reason (in 3rd tier) is right, and the confidence choice (in 2nd and 4th tier) are sure, he/she is called false negatives.

The Pearson Product Moment Correlation Coefficient (Equation (3)) was used to analyze the construction validity of the developed instrument [9, 15, 18].

$$r = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}} \quad (3)$$

Where, r is the correlation between x and y variables; x is the difference between the total of the true answer score (1st and 3rd tier) in every items and the average of the true answer score in all of items; and y is the difference between the total of the sure answer score (2nd and 4th tier) in every items and the average of the sure answer score in all of items.

The developed instrument reliability was determined by calculating the reliability index (r) using Alfa Cronbach Coefficient equation (Equation (4)) [9, 16-18].

$$r = \frac{k}{k-1} \left(1 - \frac{\sum s_i^2}{s_t^2} \right) \quad (4)$$

$$s^2 = \frac{\sum (x_i - x)^2}{n-1} \quad (5)$$

Where, r is the reliability index; k is the total of items; $\sum s_i^2$ is the total of the items scores variant; s_t^2 is the variant of the total scores; n is the total of students; x_i is the student's score in every items; x is the average of scores.

The developed instrument is avowed valid and reliable if the r value (r_{count}) respectively in Equation (3) and (4) is bigger than r value of product moment table (r_{table}) [15].

3. Results and Discussion

3.1. The Instrument Validity

3.1.1. Internal Validity

The internal validity results that was examined by asking three independent physics lectures are given on Table 3.

Table 3. Internal validation results of the developed misconception diagnostic test.

Aspects	Indicators	Score		
		Validator 1	Validator 2	Validator 3
Material (content)	Each test items correspond to the Work and Energy concept	4	4	4
	The test items correspond to the indicators	4	3	4
	The items description was sorted by the according material	4	4	3
	Limitation of questions, answers, and expected reasons were explained clearly.	4	3	3
Construction	The diagnostic test directions were explained clearly.	4	4	4
	The test items criteria correspond to Bloom's taxonomy and basics competencies	4	3	4
	Each test diagnostics items can identify student conceptions	3	3	3
	Optional reasons which were presented can reveal the cause of student's misconceptions	4	3	3
	Deception of optional reasons were rational and homogeny with the answers in the first tier	4	4	3
	Table, graphic, figure, and, the other kinds correspond to the presented problems	4	4	4
Language	The test sentences used good and right Indonesian language	3	4	4
	Various sentences or statement didn't cause multiple interpretations	3	3	3
	Each test items statement were explained clearly and communicatively	3	4	4

Based on the result of internal validity in Table 3 above, we can calculate the percentage of instrument validity used Equation (2). The validity of content, construction, and language aspects are obtained respectively 91.67 %, 90.28 %, and 86.11 %. According to Riduwan & Akdon [12], these percentages criteria of the developed instrument validity is very valid, i.e. in the percentage range of 81 % - 100 %. It's mean that the developed instrument is proper to used.

3.1.2. External Validity

The content validity results of the developed instrument are given on Table 4.

Table 4. Percentage of Positives and Negatives Falses.

Question Number	False Positives (FP)	False Negatives (FN)
1	1	4
2	2	0
3	1	2
4	2	1
5	8	1
6	1	2
7	0	0
8	0	0
9	9	3
10	1	2
11	4	5
12	0	0
13	8	4
14	7	1
15	3	0
16	1	1
17	0	4
18	2	0
19	2	2
20	1	0
21	2	0
22	2	3
23	0	3
24	0	3
25	13	2
Total	70	43
Percentage	5.6 %	3.44 %

Based the data on Table 4 above, it is known that the percentage of false positives (FP) and false negative (FN) obtained in this research is less than 10 %. These results have qualified the criteria of an instrument content validity established by Hestenes & Halloun [13]. According to Hestenes & Halloun [13] explain that the instrument content validity is fulfilled if the value of false positives (FP) and false negative (FN) is less than 10 %. Therefore, the contents validity of the developed instrument is valid.

Meanwhile, the construct validity was obtained by calculating the correlation between x and y variables. The data of x and y variables are given on Table 5.

Table 5. Data of x and y Variables Results.

Question Number	The total of true answer (x_i)	The total of sure confidence answer (y_i)	$x = x_i - \bar{x}$	$y = y_i - \bar{y}$
1	6	17	-2.8	-5.4
2	8	23	-0.8	0.6
3	12	31	3.2	8.6
4	7	34	-1.8	11.6
5	15	33	6.2	10.6
6	7	19	-1.8	-3.4
7	11	19	2.2	-3.4
8	8	14	-0.8	-8.4
9	13	28	4.2	5.6
10	11	20	2.2	-2.4
11	12	20	3.2	-2.4
12	10	18	1.2	-4.4
13	3	22	-5.8	-0.4
14	11	24	2.2	1.6
15	13	27	4.2	4.6
16	5	21	-3.8	-1.4
17	6	24	-2.8	1.6
18	2	16	-6.8	-6.4
19	12	20	3.2	-2.4
20	3	26	-5.8	3.6
21	6	18	-2.8	-4.4
22	5	18	-3.8	-4.4
23	8	19	-0.8	-3.4
24	8	17	-0.8	-5.4
25	18	32	9.2	9.6
Average	$\bar{x} = 8.8$	$\bar{y} = 22.4$		

Based on data in Table 5 above, we can calculate the value of Pearson Product Moment Correlation Coefficient (r) used Equation (3). The obtained value of r (r_{count}) is equal to 0.487, where the value of r_{table} for $n = 50$ with significant level of 5 % is 0.279 [12]. The result can be concluded that the constructs validity of the developed instrument is valid because r_{count} is bigger than r_{table} [15], so that the construct validity of the developed instrument is valid.

3.2. The Instrument Reliability

Table 6 give the data of items variant to calculate the reliability index of the developed instrument (used Equation (4)).

Table 6. Data of items variant.

Question Number	The items scores variant (s_i^2)	The variant of the total scores (s_t^2)	The total of items (k)
1	0.49		
2	0.58		
3	0.74		
4	0.53		
5	0.72		
6	0.54		
7	0.70		
8	0.57		
9	0.68		
10	0.65		
11	0.63		
12	0.66		
13	0.38		
14	0.65	43.33	25
15	0.78		
16	0.46		
17	0.50		
18	0.23		
19	0.72		
20	0.29		
21	0.50		
22	0.46		
23	0.58		
24	0.58		
25	0.67		
Total	14.295		

Based on data in Table 6 above, we can calculate the reliability index used Equation (4). This research obtained the reliability index (r_{count}) is equal to 0.698, where the value of r_{table} for $n = 50$ with significant level of 5 % is 0.279 [12]. From these results, it can be known that the developed instrument is reliable because r_{count} is greater than r_{table} [15].

Based on the analysis above, it is known that the developed four-tier format diagnostic test instrument qualified the criteria of validity and reliability, both internal and external validity. These results correspond with the criteria of a good instrument by Sugiyono [11]. Therefore it can be concluded that the developed instrument be avowed valid and reliable to use.

4. Conclusion

Based on the results which are obtained of this research, it can be concluded that the developed instrument is valid, both internally and externally. The reliability index is also bigger than the reference value. This also implies that the intended instrument is reliable. Therefore it can be concluded that the developed four-tier format diagnostics test is proper to use.

References

- [1] Suparno P 2013 *Miskonsepsi dan Perubahan Konsep dalam Pendidikan Fisika* (Jakarta: PT. Grasindo Anggota Ikapi)
- [2] Alwan A A 2011 *Procedia Soc. Behav. Sci.* **12** 600.
- [3] Turgut U, Gürbüz F and Turgut G 2011 *Procedia Soc. Behav. Sci.* **15** 1965.
- [4] Bosomworth 2001 *Physics: Concepts and Connections* (Canada: Irwin Publishing)
- [5] Kaltakci D, Gurel, Eryilmaz and McDermott L C 2016 *European J. Phys.* **37** 45705.
- [6] Suprpto N, Syahrul D A, Agustihana S, Pertiwi C A and Ku C H 2016 *Chem.* **25** 718.
- [7] Gurel D K, Eryilmaz A and McDermott L C 2015 *Eurasia J. Mat. Sci & Tech. Educ.* **11** 989.
- [8] Hasan S, Bagayoko D and Kelley E L 1999 *Phys. Educ.* **34** 294.
- [9] KALTAKÇI D 2012 *DEVELOPMENT AND APPLICATION OF A FOUR-TIER TEST TO ASSESS PRE-SERVICE PHYSICS TEACHERS' MISCONCEPTIONS ABOUT GEOMETRICAL OPTICS* (Thesis)
- [10] Hermita N, Suhandi A, Syaodih E, Samsudin A, Isjoni, Johan H, Rosa F, Setyaningsih R, Supriadil and Safitri D 2017 *IOP Conf. Series: J. Phys.: Conf. Series* **895** 12167.
- [11] Sugiyono 2009 *Metode Penelitian Administrasi Dilengkapi dengan Metode R&D* (Bandung: ALFABETA)
- [12] Riduwan 2013 *Rumus dan Data dalam Analisis Statistika* (Bandung: ALFABETA)
- [13] Hestenes D and Halloun I 1995 *The Phyc. Teacher* **33** 502.
- [14] Kirbulut Z D and Geban O 2014 *Eurasia J. Mat. Sci & Tech. Educ.* **10** 509.
- [15] Arikunto 2005 *Dasar-Dasar Evaluasi Pendidikan* (Jakarta: PT Bumi Aksara)
- [16] Nurgiantoro 2010 *Penilaian Pembelajaran Bahasa Berbasis Kompetensi* (Yogyakarta: BPEF-Yogyakarta)
- [17] Mulyasa 2009 *Analisis, Validitas, Reliabilitas, dan Interpretasi Hasil Tes* (Bandung: PT Remaja Rosdakarya)
- [18] Sudjana 2005 *Metoda Statistika* (Bandung: Tarsito Bandung)
- [19] Afif N F, Nugraha M G and Samsudin A 2017 *AIP Conf. Proc.* **1848** 050010.

Appendix A

Some of the draft instrument of misconception diagnostics test.

- 3 Five children are pushing a big wall in the same direction. Their given force is shown in the following table,

Name	Force (N)
Doni	100
Zifa	50
Arya	75
Tri	100
Ari	60

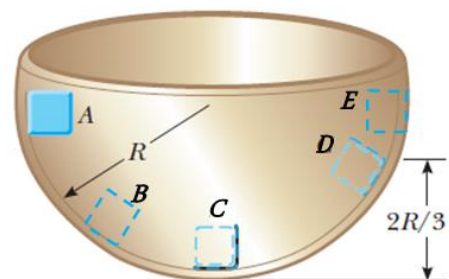
The child who does the greatest Work in pushing the wall is....

Reasons:

.....

- 21 See the picture on the side!

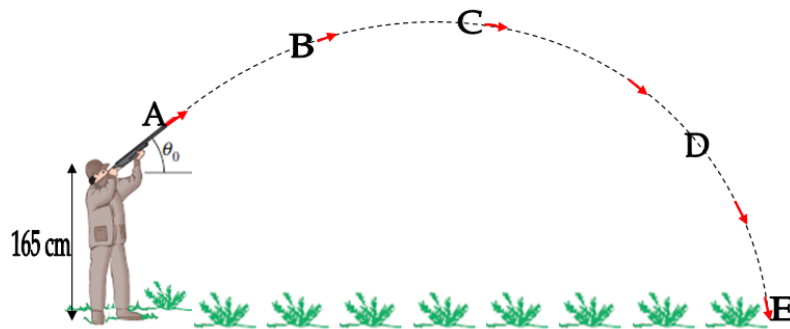
A plastic cube toy with a mass of 100 g is released in a bowl. If the surface of the bowl is slippery, then the plastic toy will have the maximum speed at the position



Reasons:

.....

- 22 Andi fired a 10-gram bullet forming a parabolic trajectory, as shown in the picture below.



The bullet has potential energy and kinetic energy is in the position ...

Reasons:

.....

Appendix B

Some of the developed instrument of misconception four-tier format diagnostics test.

- 3 Five children are pushing a big wall in the same direction. Their given force is shown in the following table,

Name	Force (N)
Doni	100
Zifa	50
Arya	75
Tri	100
Ari	60

The child who does the greatest Work in pushing the wall is....

- A. Doni dan Tri
B. Zifa
C. Arya
D. Ari
E. Their Work is same

The level confidence of your answer is



Supportive Reasons:

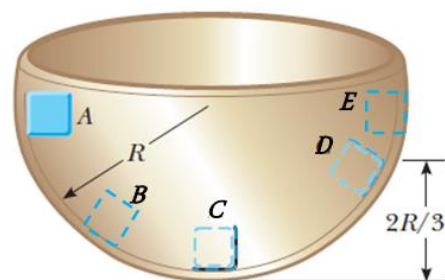
- A. The work depends on the force given.
B. If the force given is greater, the energy required is also greater so that the work is also bigger.
C. Work is proportional to the force. The greater force that is given, the work also being greater.
D. The greater the force that the child does, their energy that being remaining decreases so that the work is made smaller.
E. The greater of the child given force, their hand's feel is heavier due to the reaction force of the wall. Thus, the effort is also made larger.
F. The wall does not move due to the force given to it.
G.

The level confidence of your reason is



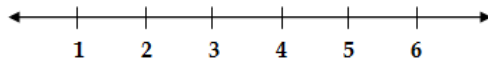
- 21 See the picture on the side!

A plastic cube toy with a mass of 100 g is released in a bowl. If the surface of the bowl is slippery, then the plastic toy will have the maximum speed at the position



- A. A
- B. B
- C. C
- D. D
- E. E

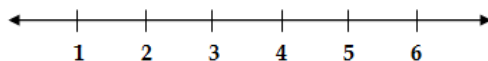
The level confidence of your answer is



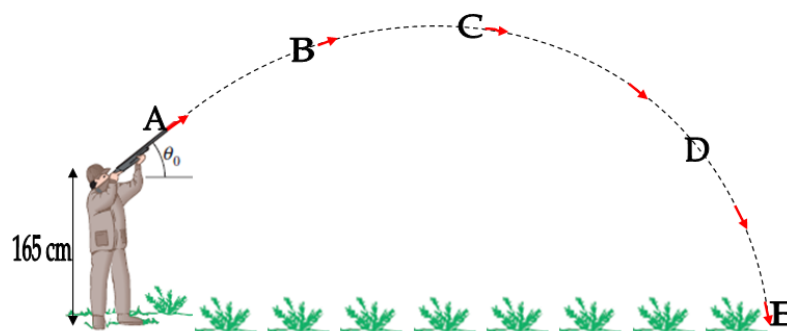
Supportive Reasons:

- A. When the plastic toy moves up, the plastic toy needs a greater energy when compared to when the plastic toy moves down because when the plastic toy moves up, the plastic toy must have a greater speed to reach the point.
- B. When the plastic toy lies at the base position, the plastic toy will stop because the energy is equal to zero, so the plastic toy can not move up.
- C. The mechanical energy at every point is always constant.
- D. The further the plastic toy from its original position, the energy is smaller. Therefore, at the beginning of sliding, plastic toys have maximum speed because it has maximum energy.
- E. When the plastic toy moves down the speed is greater because it is affected by the force of gravity, while when the plastic toy moves up the speed becomes small because it is inhibited by the force of gravity.
- F. In that position, the plastic toy has maximum kinetic energy because it lies at the lowest position so that the plastic toy will move at maximum speed.
- G.

The level confidence of your reason is



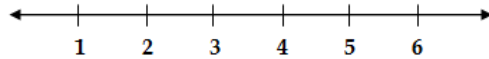
- 22 Andi fired a 10-gram bullet forming a parabolic trajectory, as shown in the picture below.



The bullet has potential energy and kinetic energy is in the position ...

- | | |
|----------------|----------------------|
| A. A, B, dan D | D. A, B, C, dan D |
| B. B dan D | E. A, B, C, D, dan E |
| C. B, C, dan D | |

The level confidence of your answer is



Supportive Reasons:

- A. In the initial position, the maximum potential bullet energy and kinetic energy equal to zero. While in the final position, the maximum kinetic energy of the bullet and the potential energy is equal to zero. Thus, the bullet has potential energy and kinetic energy when the bullet lies in other than the starting position and end position.
- B. In addition to highs and lows, the bullet has potential energy and kinetic energy because the bullet has a speed and a certain height.
- C. In every position, the potential energy is equal to kinetic energy.
- D. At each position, the bullet has a potential energy and kinetic energy because the energies are used to move the bullet.
- E. At the highest position, the bullet has a speed that is v_x so that the kinetic energy of the bullet in that position is not equal to zero.
- F. At the highest position, the bullet does not have a kinetic because the bullet has no speed (stop) then moves down.
- G.

The level confidence of your reason is

