

GERGANA PETKOVA

Medical University of Plovdiv

gergana.petkova@mu-plovdiv.bg

ORCID: 0000-0001-7209-9765

VANYA IVANOVA

Plovdiv University

vantod@uni-plovdiv.bg

ORCID: 0000-0002-0043-8606

USING ONLINE PRACTICE TESTS BASED ON AUTHENTIC SHORT STORIES IN ENGLISH TO ENHANCE STUDENT LANGUAGE SKILLS

Keywords: e-learning, practice tests, digital tests, short stories, test criteria

Abstract

The purpose of this article is to explain the principle behind practice tests based on authentic short stories in English, describe their construction, and outline their application which aims to increase student knowledge and skills in a foreign language. The methodology employs digital practice tests created following a set of criteria for test development based on Bloom's taxonomy. The results of the study indicate certain benefits of the practice tests: short stories introduce new vocabulary and the tests provide reinforcement, which both facilitates and motivates the students.

Introduction

The aim of this article is to describe certain benefits resulting from practice tests based on authentic short stories in English and to provide examples in an attempt to illustrate how they can be applied to increase student knowledge and skills in English when studied as a foreign language. For this reason, it is necessary to briefly outline the process behind the practice tests and their benefits, the stages involved in the test construction, the specific criteria relating to the test content based on

Bloom's taxonomy, as well as examples of test questions used in practice tests based on literary sources.

Practice tests, also known as self-study tests, are considered to be among the most efficient methods of studying (Dunlosky et al. 2013). The tests are practical, specifically developed by the teacher on the basis of the study material covered in the classroom, while at the same time adhering to precise rules regarding their construction and administration. Some of the advantages of practice tests are that they support learners of different ages and abilities; they help in memorizing content over a longer period of time and increase student achievement (Karpicke and Blunt 2011; McDaniel et al. 2012). Testing not only reveals the level of student knowledge but also enhances it, thus improving the long-term retention of information, which is known as the testing effect. In other words, as demonstrated in Roediger and Karpicke (2006), Roediger et al. (2011a), and Roediger et al. (2011b), a practice test not only assesses what is known, but also enhances the subsequent retention of this knowledge.

To create a practice test, the test developer must undertake a number of steps in order to ensure its quality is of the required standard (Fulcher and Davidson 2007; Douglas 2010; Carr 2011; Eustace et al. 2014; ALTE 2018).

Test development stages

First of all, an **overall plan** is created. During this stage, test authors need to consider exactly what they wish to measure and clearly define the purpose of the test, because this is a prerequisite in increasing the test validity. They also need to make a decision about the most appropriate test format. The most widely used closed-ended questions are multiple-choice questions (MCQ). These have a number of advantages both for the teacher and the student. Although MCQs take longer to create than open-ended questions, much less the time is needed to conduct and mark the tests. They are objective and, given a key, easy to grade. As far as the student is concerned, closed questions such as MCQs are quicker and easier than open essay-type questions, and additionally within the same amount of time a larger number of structured-answer questions can be completed. As a rule, MCQs are more reliable than open questions as the reliability of a test is related directly to its objectivity and the number of items it contains. Regarding the number of distractors (answer options) in a test, the quality of the distractors is more important than their number; nevertheless, the standard for teacher-generated tests is considered to be four- or five-option items (Stoyanova 1996).

The next step in the test construction procedure is the **content definition** – at this stage test authors determine what content is to be tested.

Following this, a **test specification** is written, that is a plan of the test. This is a detailed, practical document indicating what data the test will contain, namely a document which is intended to assist in the test construction. Test specifications include information about the purpose of the test, and the learners who will be taking the test (age, gender, level of proficiency, first language, country of origin,

level of education, reasons for taking the test, etc.); the number of test sections, their length, and manner of differentiation between them; the text type (written/spoken; sources, topics, degree of authenticity, etc.); a description of the language skills to be tested; the types of test tasks; the number of test items in each section and their relative weighting; the test methods (MCQs, gap filling, matching, transformations, picture descriptions, essay writing, etc.); the rubrics; the assessment criteria, etc. (Davidson and Lynch 2002).

The test specification stage is followed by **item development**, in which test authors produce an initial set of test items. This set should comprise a larger number of questions than will actually be present in the final test because some items will be removed after the pretest, usually due to their poor quality. The recommended ratio of the total number of items and the number of items included in the final test is 3:2. The number of test items depends on various factors such as time constraints, the students' age, the type of items, etc.

After this is the stage of **test design and assembly**. Test authors have to ensure that the content tested corresponds to the content in the specification.

The next stages include **test production** and **test administration** (identical time limits and proctoring for all students is essential).

Besides grading the tests, the stage of **scoring test responses** also involves a final item analysis, including the raw score mean, the standard deviation, the mean item difficulty and discrimination, the range of the raw score, a distractor analysis, etc., and any inconsistencies, identified by the final item analysis, must be investigated and resolved prior to reporting the test scores.

Then, **passing scores** are established – in criterion-referenced tests, usually the passing standard is 50% correct answers.

The next stage is **reporting test results** in a manner that avoids any misunderstanding (the format of the score presentation needs to be considered – as a percentage of the total maximum score and/or a grade, etc.). **Item banking** is the process of storing test items for potential future use.

Last but not least is the **technical report**, which describes the important aspects of the test development, administration, scoring, and reporting, as well as the test analysis and evaluation (Ivanova 2011).

Since its creation, Bloom's taxonomy (Bloom et al. 1956) has been used as a basis for test construction. The taxonomy describes the main levels in the cognitive domain, which are hierarchically arranged on the principle of "from simple to complex." The levels can be considered as varying degrees of difficulty – learning at the higher levels depends on having attained the prerequisite knowledge and skills at the lower levels. This approach enables the measurability of student knowledge, skills, and attitudes. The test criteria range from a lower-order to a higher-order level within the cognitive domain.

Digital practice tests have been used in English language education in the Faculty of Mathematics and Informatics (FMI) at the University of Plovdiv for eight academic years. These tests are constructed in line with the five basic criteria for evaluating learning outcomes in English language training, criteria adapted from

Bloom's taxonomy. The test construction criteria are: retention of information, comprehension of vocabulary, application of lexical and grammatical elements in practice, analysis of the use of lexical and grammatical units in context, and evaluation and creation of a text (Ivanova and Terzieva 2016).

Criteria for language test construction

Criterion I. Reproduction of information

Test tasks are intended to check the students' knowledge of vocabulary and grammatical constructions; consequently, those taking the test are required to recognize, recall, and reproduce lexical and grammatical items. For instance, the students may be required to choose the correct word in order to complete an idiomatic expression in a sentence (the correct answers in the examples are given in *italic*).

Example. His parents asked him to get a job until they were in the face but he ignored them.

- a) red
- b) brown
- c) *blue*
- d) black

Criterion II. Understanding the meaning of a word, expression or phraseological unit, and identifying a corresponding element

Test assignments check the learners' understanding of the meaning of specific vocabulary through the recognition of synonyms and antonyms of words and expressions, matching words to their definitions, paraphrasing, etc.

Example. Definitions: Match the idiom to its meaning.

in the pink

- a) *to be in good health*
- b) not to have enough money
- c) to be very quiet and shy
- d) to have great admiration for oneself

Criterion III. Detection and correction of errors in various contexts

The learners are required to detect and/or correct different types of mistakes in a given context. True or false and multiple choice test questions are used for finding and/or correcting spelling and grammatical errors.

Example. Choose the part of the sentence which contains a spelling or grammar mistake.

He thought he should report the incident but then he decided to let sleeping dogs lying.

- a) thought
- b) should report
- c) decided to let
- d) *sleeping dogs lying* (lie)

Criterion IV. Analysis of the use of words or expressions and selection of the relevant grammatical form of verbs in context

The students put their theoretical knowledge into practice by comparing and then distinguishing between various options, before selecting the most appropriate in particular instances. Multiple choice, cloze, or short-answer open test questions are used to:

- Select the sentence which most accurately paraphrases the one given, in order to illustrate the meaning of a word, expression, or phraseological unit in context;
- Choose the appropriate grammatical form of verbs in a particular setting;
- Add the most appropriate lexical or grammatical unit in a specific context.

Example. Select the option which describes the meaning of the sentence as closely as possible.

My sister has green fingers.

- a) She is always dropping things and breaking them.
- b) She is very young and inexperienced.
- c) She is a very jealous girl.
- d) *She is good at gardening.*

Criterion V. Text creation and evaluation

The students organize their ideas into a coherent text to summarize their viewpoints and evaluate others' opinions and texts. The learners may be required to write an essay, analyze a text, or comment on a given topic.

Example. Explain the meaning of the idiomatic expression and describe a situation from your own experience in about 100–150 words to illustrate its use.

If you admit that you have told a white lie, you mean that

Using literary texts in foreign language education

Using literature in English language teaching offers a considerable number of advantages for learners such as developing their reading skills, contributing to a better understanding of the foreign language, and providing an authentic model of language use. Literary texts promote greater language proficiency and an awareness of other cultures by stimulating the imagination of the students as well as developing their critical thinking and creative skills. Learners can benefit enormously from exposure to authentic use of specific language, such as phrasal verbs and idiomatic expressions, which often prove challenging. Short stories are especially useful since they are more concise than a novel and can be read without breaks; they usually focus on a single plot, a major character and a central theme, which facilitates understanding. As demonstrated by Ceylan in 2016, reading authentic short stories assists students by broadening their perspectives, encouraging their thinking about reasons rather than the result, illustrating different life styles, providing a comparison between their own and other cultures, increasing their appreciation of literature, and cultivating their love of reading. For all these reasons, we have taken advantage of the numerous benefits in using short stories in language education by combining reading with practice tests based on the plot, the vocabulary, and the grammar in each story (Ceylan 2016).

Materials and methods

The methodology used in this article involves integrating original literary texts in English within the foreign language education of university students in the FMI. The majority of the students in the FMI only learn English during their first year at the university, that is for two trimesters of 10 academic weeks each, except for those majoring in Business Information Technology who also study Business English for a further trimester in their fourth year. Depending on their curriculum, the students have four or five seminars in English per week which they take consecutively once a week. Following the above-mentioned stages of test development, we constructed digital practice tests for first-year full-time and part-time university students majoring in Informatics, Software Engineering, and Software Technology and Design at the B2 level of English according to CEFR (Council of Europe 2001). Our aim was to try to expand the foreign language vocabulary and grammatical proficiency of the university students by using authentic materials. The practice tests were assigned as homework so the students had a week in which to complete them using either a computer or a mobile phone. The self-study tests included a short story to be read first, before undertaking a number of tasks based on the texts. The test items were created according to the five criteria based on Bloom's taxonomy described above; examples of each are provided, with the chosen texts being six unabridged short stories by Saki – Hector Hugh Munro (1870–1916) (Munro 1993).

These stories were selected because of the author's subtle humour and the unexpected twists in the plot, which, we hoped, would make them enjoyable, arouse the students' interest and motivate them to read the stories in their entirety, as well as reflect on them, which, in turn, would lead to an improvement in both their reading and objective analysis skills.

Examples of criteria-based questions in practice tests based on literary texts

Some examples of the test questions based on the five criteria described above, which were actually included in the practice tests taken by students in the FMI during the 2019/2020 and 2020/2021 academic years, are provided below (again, the correct answers are given in *italic*). The test items aim to verify the comprehension and language acquisition after reading the classic short story, "The Lumber Room", a story about a clever but mischievous boy, Nicholas, who seeks to outwit his aunt so that he can gain access to the lumber room with its hidden treasures and curiosities.

- I. Mark the statements as True, False, or Not Given (if the fact is not mentioned or cannot be inferred in the reading) according to the text.
 1. There was a frog in Nicholas's bowl of bread-and-milk.
True
False
Not given

2. Nicholas often asked his soi-disant aunt questions about the lumber room.

True
False
Not given

II. Matching

3. Match the word or phrase to its definition: frivolous

Example: She thinks window shopping is a frivolous activity.

- a) *not important; not deserving serious attention*
- b) involving the possibility of something bad or unpleasant happening
- c) feeling, causing, or showing great happiness
- d) having serious meaning or worth

4. Match the word or phrase to its definition: quiver

Example: I need a quiver for my new hobby.

- a) *a case used for carrying arrows*
- b) a leather case that you wear on your body and that holds a small gun
- c) a small folding case that holds paper money, credit cards, etc.
- d) a protective case for a sword that covers the blade

5. Find a synonym (a word with a similar meaning) of the word: wholesome

Example: I would like to have a wholesome meal.

- a) *healthy*
- b) spicy
- c) cooked
- d) delicious

6. Find a synonym (a word with a similar meaning) of the word: alleged

Example: The alleged thief was arrested.

- a) arrogant
- b) confirmed
- c) foreign
- d) *supposed*

7. Find an antonym (a word with the opposite meaning) of the word: hastily

Example: She hastily changed the subject.

- a) suddenly
- b) *slowly*
- c) unexpectedly
- d) strangely

8. Find an antonym (a word with the opposite meaning) of the word: saunter

Example: He sauntered into the store.

- a) crawl
- b) *run*
- c) steal
- d) giggle

III. Choose the part of the sentence, which contains a spelling or grammar mistake.

9. This morning the boy had refused eating his breakfast on the ground that he wasn't hungry.
 a) had refused
 b) *eating* (to eat)
 c) on the ground
 d) wasn't hungry
10. He had pictured to himself what the room might be like and it came up with his expectations.
 a) had pictured
 b) to himself
 c) might be like
 d) *came up with* (to)

IV. Mark the sentence as True if it doesn't have any lexical or grammatical errors, or False if there is a mistake.

11. The tightness of his shoes had had a disastrous effect on his mood all day.
 a) *True*
 b) False
12. She had grown suspicious at his long disappearance and had leapt into the conclusion that he had climbed over the wall.
 a) True
 b) *False* (leapt to the conclusion)

V. Choose the sentence that explains best the meaning of the given one.

13. Nicholas did not admit the flawlessness of the reasoning.
 a) *Nicholas didn't accept that the philosophy was perfect.*
 b) Nicholas agreed that the thinking was correct but didn't wish to speak aloud about it.
 c) Nicholas believed that the conclusions were only partly true.
 d) Nicholas didn't agree that the way of thinking was imperfect.
14. Nicholas made one or two sorties into the front garden, but never able for a moment to evade the aunt's watchful eye.
 a) *Nicholas tried to enter the front garden once or twice but he couldn't avoid being seen by his aunt.*
 b) Nicholas made a few attempts to go inside the front garden; however, he couldn't give his aunt a good reason for being there.
 c) Nicholas persisted with his attempts to access the front garden but he couldn't keep it secret from his observant aunt.
 d) Nicholas attempted to enter the front garden a couple of times; nonetheless, he could not sort his way through there.

VI. Choose the correct word(s) to fill in the sentences.

15. Last week the student on his knowledge of Bulgarian poets but he was late and missed the exam.
- was to be tested*
 - is to be tested
 - was tested
 - had tested
16. The little boy fell from when he was caught stealing sweets.
- kindness
 - grace*
 - mercy
 - goodwill

VII. Answer the questions based on the story “The Lumber Room”.

17. Why didn't the children enjoy the trip to Jagborough?
18. Which three adjectives can you use to describe Nicholas? Give evidence from the text to support your ideas.

At the end of the test, the students were asked for feedback. First of all, they were asked to answer a few questions about the text (for example “What did and didn't you like about the story?” or “If you could change something in the story, what would it be? Why?”). Then, they were asked to evaluate the test with reference to the story, the test as a whole, the number of questions, the variety of questions, and the level of difficulty of the test items (on a scale from 1 to 6). Finally, the learners were encouraged to make recommendations and ask questions, as appropriate.

Google forms were selected as the testing platform for a number of reasons. It is a useful tool for constructing surveys and tests as it permits the creation of different types of questions such as multiple choice, dropdown, linear scale, etc. It is also possible to add content such as text, images, and video. After administering a test, grading is performed by means of the free tool Flubaroo, which enables the quick and easy grading of closed-ended questions and has an option for the hand grading of open-ended questions. Moreover, it offers additional services, such as:

- Computing average assignment scores
- Computing average scores per question, and flagging low-scoring questions
- Showing a grade distribution graph
- Giving the teacher the option to email each student their grade, and an answer key
- Allowing the educator to send individualized feedback to each student.

Results and discussion

The six tests at the B2 language level based on the six short stories by Saki were administered to the first year full-time and part-time Bachelor degree students in the Faculty of Mathematics and Informatics at the University of Plovdiv during the 2019/2020 and 2020/2021 academic years.

The Flubaroo test statistics showed that the average scores in all the tests were around 75% (Fig. 1, Fig. 2) and low-scoring test questions were around 20% (Fig.3). The latter had to be revised before publishing a language manual with the stories and six tests (Ivanova 2020).

The learners were asked to take each test in a convenient place and on a convenient day without specified time restrictions. The short story was added to the test for the students to read in advance with footnotes in English explaining the more challenging vocabulary. After answering the test questions, the students were asked to give feedback and grade both the test and the test items from 1 (DON'T LIKE at all) to 6 (LIKE very much). The results are shown in Fig. 4 and Fig. 5.

Although the most common complaint reported by the students was connected with the large number of test questions, the students acknowledged the benefits of the story introducing new vocabulary and the practice tests reinforcing this. They started using the new words not only in the group discussions about the short stories but also in other contexts. Some of the students' opinions about the tests are presented in Fig. 6 and Fig. 7.

Conclusions

Short stories and practice tests based upon the texts were administered to non-linguistics university students majoring in Informatics, Software Engineering, and Software Technology and Design. Not many of these technologically-minded students were used to reading books for pleasure, especially in a language different from their mother tongue – they were used to skimming texts to understand the essential information, so assigning a task that required them to study a comparatively long story intensively and comment upon it was challenging. Some of the learners admitted that this was the first time that they had read a literary text in English, hence the common complaint about the length of the stories and the numerous test questions. Notwithstanding this, the students gradually started using the new vocabulary in other contexts. A large number of students showed genuine interest in our post-test discussions, asking questions about the author, his characters, and other stories by the same writer. As a result of informal conversations with the students and an online survey on their opinion regarding the practice tests based on authentic short stories in English, the overwhelming majority of students wished to continue using online practice tests during their foreign language education and a consensus was reached to reduce the number of test questions as well as include stories by more contemporary writers.

In summary, our experience with using literature in the form of short stories for language practice has suggested that it is beneficial in several ways – it offers varied authentic language material, provides communicative situations, and brings about cultural awareness. Online practice tests created on the basis of these stories help learners with different language abilities to increase their language knowledge and awareness, critical thinking skills, as well as creative writing. For this reason, we are planning to continue to incorporate authentic short stories within our English language teaching and to develop tests based upon them in order to enhance learning.

Acknowledgements

This paper is supported by grant MU21-FMI-015 from the NPD at the University of Plovdiv “Paisii Hilendarski”, financed by the Bulgarian Ministry of Education and Science.

References

- ALTE. 2018. *Guidelines for the development of language for specific purposes tests. A supplement to the manual for language test development and examining*. [available at: <https://www.alte.org/resources/Documents/6093%20LSP%20Supplement%20-%20WEB.pdf>, accessed: 23 February 2022].
- Bloom B.S., Engelhart M.D., Furst E.J., Hill W.H., Krathwohl D.R. 1956. *Taxonomy of educational objectives: The classification of educational goals*. [Handbook 1: Cognitive domain]. New York: David McKay Company.
- Carr N. 2011. *Designing and analyzing language tests*. Oxford: Oxford University Press.
- Ceylan H.O. 2016. Using short stories in reading skills class. – *Procedia – Social and Behavioral Sciences* 232: 311–315.
- Council of Europe. 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Press Syndicate of the University of Cambridge.
- Davidson F., Lynch B. 2002. *Testcraft: A teacher's guide to writing and using language test specifications*. London, New Haven: Yale University Press.
- Douglas D. 2010. *Understanding language testing*. London: Hodder-Arnold.
- Dunlosky J., Rawson K., Marsh E., Nathan M., Willingham D. 2013. Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. – *Psychological Science in the Public Interest* 14.1: 4–58.
- Eustace J., Pathak P., Muntean C.H. 2014. Practice testing: Enhancing student learning with e-assessment. – *13th Annual information technology & telecommunications conference*. [available at: <http://trap.ncirl.ie/1302/>, accessed: 23 February 2022].
- Fulcher G., Davidson F. 2007. *Language testing and assessment. An advanced resource book*. London, New York: Routledge.
- Ivanova V. 2011. Construction and evaluation of achievement tests in English. – Totkov G., Koychev I. (eds.). *Sbornik ot Chetvarta natsionalna konferentsiya "Obrazovaniето v informatsionното onstestvo", 26–27 May 2011*. [Proceedings of the “Education in the information society” 4th national conference]. Plovdiv: Association for the Development of the Information Society and the Institute of Mathematics and Informatics at the

- Bulgarian Academy of Science: 276–285. [available at: <http://sci-gems.math.bas.bg/jspui/bitstream/10525/1554/1/adis-may-2011-276p-285p.pdf>, accessed: 23 February 2022].
- Ivanova V. 2020. *Saki – six short stories and six tests*. [= *Language Trainer*]. Plovdiv: Koala Press.
- Ivanova V., Terzieva T. 2016. Criteria for the construction of tests for language assessment and evaluation. – *Proceedings of the MIDOC doctoral conference in mathematics and informatics*. Sofia: “St. Kliment Ohridski” University Press: 58–65.
- Karpicke J., Blunt J. 2011. Retrieval practice produces more learning than elaborative studying with concept mapping. – *Science* 331.6018: 772–775.
- McDaniel M., Wildman K.M., Anderson J.L. 2012. Using quizzes to enhance summative-assessment performance in a web-based class: An experimental study. – *Journal of Applied Research in Memory and Cognition* 1.1: 18–26.
- Munro H.H. 1993. *The collected short stories of Saki*. Suffolk: Wordsworth Editions.
- Roediger H.L., Karpicke J. 2006. Test-enhanced learning – taking memory tests improves long-term retention. – *Association for Psychological Science* 17.3: 249–256.
- Roediger H.L., Agarwal P., McDaniel M., McDermott K. 2011a. Test-enhanced learning in the classroom: Long-term improvements from quizzing. – *Journal of Experimental Psychology: Applied* 17.4: 382–395.
- Roediger H.L., McDermott K., McDaniel M. 2011b. Using testing to improve learning and memory. – Gernsbacher M.A., Pew R.W., Hough L.M., Pomerantz J.R. (eds.). *Psychology and the real world: Essays illustrating fundamental contributions to society*. New York: Worth Publishers: 65–74.
- Stoyanova F. 1996. *Testologiya za uchiteli*. [*Testology for teachers*]. Sofia: Atika.



Student Number	Total Points	Percent	Times Submitted	The aunt sent the children on a trip to ...	There was a frog in Nicholas's bowl of b ...	The aunt didn't have anything else to do ...	Nicholas preferred to go into the lumber ...	The gooseberry garden was much better in ...	This was the first time Nicholas had fit ...	Nicholas often asked his soi-disant aunt ...
1901322040	76	88.37%	1	1	1	1	1	1	1	1
1901322001	77	89.53%	1	1	1	1	1	1	1	1
1901322021	72	83.72%	1	1	1	1	1	1	1	1
1901322004	79	91.86%	1	1	1	1	1	1	1	1
1901322020	70	81.40%	1	1	1	1	1	0	1	1
876587977	55	63.95%	1	1	1	1	1	1	1	1
1901682024	59	68.60%	1	1	1	1	1	1	1	1
1901682002	68	79.07%	2	1	1	1	1	1	1	1
1901051003	68	79.07%	1	1	1	1	1	1	1	1
1901261038	60	69.77%	1	1	1	1	1	0	0	0
1901051005	73	84.88%	1	1	1	1	1	1	1	1
1901261049	67	77.91%	1	1	1	1	1	1	1	1
1901261058	36	41.86%	1	1	1	1	1	1	0	0
1901261033	36	41.86%	1	1	1	1	1	0	0	1

Fig. 1. Flubaroo statistics on the submitted tests

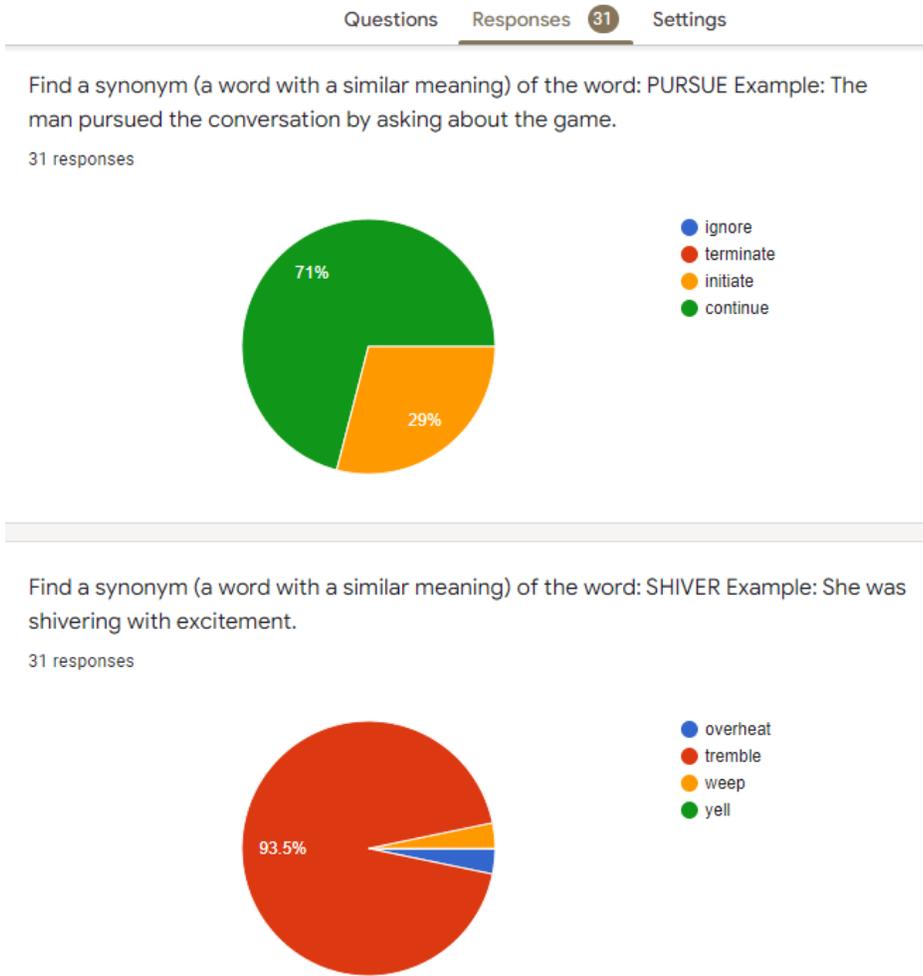


Fig. 2. An example of Flubaroo statistics for average assignment scores

Student Number	The aunt sent the children on a trip to ...	The aunt didn't have anything else to do ...	The gooseberry garden was much better in ...	Nicholas often asked his soi-disant aunt ...	The aunt went into the rain-water tank t ...	Match the word or phrase to its definiti ...	Match the word or phrase to its definiti ...	Match the word or phrase to its definiti ...	Find a synonym (a word with a similar me ...
2101261097	1	1	1	1	1	0	1	1	1
2101681068	1	1	1	0	0	0	1	1	1
2101681015	1	1	1	1	1	0	1	1	1
2101681020	1	0	1	0	1	0	0	1	0
2101681053	0	1	1	0	1	0	1	1	0
2101681048	0	1	0	1	1	1	1	1	1
2101681008	1	1	1	1	1	1	1	1	1
2101681059	1	1	1	0	0	1	1	1	1
2101681064	1	1	1	0	1	1	1	1	1
2101561038	1	1	1	0	1	1	1	1	1
2101681069	1	1	1	0	1	1	1	1	1
2101561065	1	1	1	0	0	1	1	1	1
2101561040	1	1	1	0	1	1	1	1	1
2101561070	0	0	1	1	1	0	1	1	1
2101681034	1	1	1	0	0	0	1	1	1
2101561043	1	1	1	1	0	0	1	1	1
2101561071	1	1	1	1	1	0	1	1	1
2101561059	0	1	1	0	0	1	1	1	1
2101561067	0	1	1	0	0	1	1	1	1
2101561055	0	1	1	0	1	0	1	1	1
	79.39%	91.60%	80.15%	31.30%	58.78%	63.36%	85.50%	96.18%	94.66%

Fig. 3. Low-scoring questions are highlighted in Google forms

What do you think of the test questions? Grade them from 1 (DON'T like at all) TO 6 (like very much)

 Copy

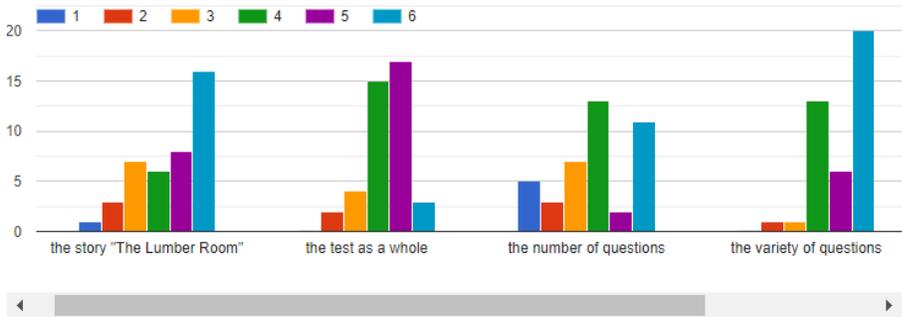


Fig. 4. Feedback by 39 test takers about the test on “The Lumber Room”

What do you think of the test questions? Grade them from 1 (DON'T like at all) TO 6 (like very much)

 Copy

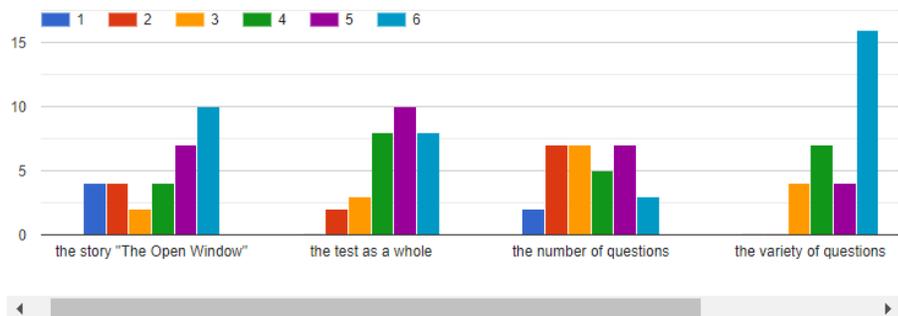


Fig. 5. Feedback by 31 test takers about the test on “The Open Window”

What did you like about the test?

31 responses

- I liked the variety of questions.
- I really liked the variety of questions.
- I really enjoyed the tale and the actual story behind it. I also liked the questions for Synonyms and Antonyms because this is a great way to develop your language skills.
- The variety of questions and that they actually make you think twice before answering.
- I loved the bit with the questions about antonyms. I love antonyms. :) Also all of the complex words used in the story and the exercises.
- Its a chill test.
- What I liked about the test is that there were so many new words that I learnt.
- Matching.

Fig. 6. Students' feedback – what test takers liked about the test based on “The Open Window”

What didn't you like about the test?

31 responses

- It was an interesting test. There is nothing to dislike.
- I didn't like the amount of "writing questions" the test had, but as a whole, it was a delightful challenge.
- The story
- That it was kind of long
- New words.
- I didn't like that the test was too long.
- I didn't like the fact that there were so many questions.
- Maybe that the questions are too much, but generally i think the test is interesting.
- The questions were too much.

Fig. 7. Students' feedback – what test takers didn't like about the test on “The Open Window”

