Perspectives in the Medical Latin Language E-testing

Miervaldis KARULIS
Language Centre, Rīga Stradiņš University
Riga, LV-1007, Latvia

ABSTRACT

The study is dedicated to e-testing of knowledge of Medical Latin terminology among international medical students at Rīga Stradiņš University (Latvia). It offers examples of e-test models and evaluation algorithms using Moodle environment. The research population were 149 first year medical students who took the end-of-semester test in the study course "Medical Terminology in Latin". E-test results were expressed in percentage (n = 140; 1st quartile – 61% , 2nd – 71%, 3rd – 77%). Structured questionnaire and open-ended survey questions were used to reveal complexity and usefulness of e-testing from students' point of view. No statistically significant difference of opinions was found among e-test takers – those who considered e-tests useful for checking their knowledge and those who did not – on such e-test taking factors as immediate feedback, paper and time economy, stress and additional influencing factors (computer literacy, time restriction, technical failures). Statistically slight significant difference of opinions was found among e-test takers – those who considered e-tests more complicated to perform and those who did not – on such e-test taking factors as immediate feedback and additional influencing factors (computer literacy, time restriction, technical failures). Students who did not considered e-tests more complicated to perform were more critical of these factors. Students acknowledged e-testing as a modern testing mode.

Keywords: e-testing, e-tests, Medical Latin terminology, Moodle, medical students, summative assessment.

INTRODUCTION

Tempora mutantur et nos mutamur in illis. - Times change, and we change along with them.

International students of the study programmes "Medicine" and "Dentistry" have been acquiring the compulsary course "Medical Terminology in Latin" for more than 20 years at Rīga Stradiņš University (RSU). The number of international students from European and some Asian countries has been increasing steadily since 2006 and has reached nearly 600 in 2017. All medical students take the course of Human Anatomy where they have to acquire terms in Latin according to *Terminologia anatomica* (1998)*,* which is the international standard on human anatomic terminology published online in 2011 [1]. Latin is also used in specialised courses later in clinical practice. Latin and Greek are source languages that unify international medical terminologies reducing confusions and facilitating forming new concepts with clear semantics and grammatical norms [2]. In [3], there is a reference to six main groups of medical terms: anatomical terms (e.g., *cranium*), names of symptoms and syndromes (e.g., *apoplexia*), names of diseases (e.g., *typhus*). names of medical materials (e.g., *opium*),

tools and instruments used in medical procedures (e.g., *forceps*). verbs connected with medical activities, processes, and physiological phenomena (e.g., *exitus*). It is desirable that medical professionals have comprehension of all these term groups.

During the Latin course at RSU, the very basics of the Latin language and its grammatical structures used in medicine are studied [4] so that being medical doctors they could appreciate "the history and original meaning of words" that offer "a new dimension to their professional language" [5].

In order to modernize the way of checking students' knowledge in Latin anatomical and medical terminology, electronic tests (e-tests) were introduced in the Autumn Semester 2016. Taking maximum advantage of Moodle (Modular Object-Oriented Dynamic Learning Environment) e-learning environment enriches the studying process to a greater extent by tying together student's individual work, grades and feedback, which aims to develop student autonomy and independence – one of the aspects of the student-centred approach, which is the main direction of the university's policy. Thus, by implementation a new knowledge-testing system in the Latin course, students' interests are taken into account.

E-ASSESSMENT/E-TESTING

The term "e-assessment" was introduced in the information society at the very end of the 20th century and was defined as "the use of electronic systems for the development, operation and delivery of accredited qualification assessment or the collection of performance evidence" [6]. In the USA, the advent of using computers began in 1994 with computer adaptive testing. E-assessment is a broad term that includes any usage of computer systems of a wide spectrum in assessment. Almost a decade later, in 2006, the European Council documents called for objective testing of skills in foreign languages recommending to adopt e-testing to administer language tests [7].

E-testing relates to a specific purpose of assessment (diagnostic, placement or summative), mainly as an administrative procedure using computer software. Pereira and Scheuermann in [7] describe e-testing from perspectives of advantages and challenges for three sides: the test taker, test administrator and test developer, who are directly involved in this process. General benefits of e-testing for all the sides can be briefly summarised as cheapness, immediate response and feedback, flexibility in terms of location and timing, high-security, objectivity, large data storage on a server, using various forms of multimedia and interactivity, using a variety of electronic devices [8]. Yet, the main issues in research in e-testing are the effect of anxiety of the test takers, inconsistency of scoring

algorithms, the Internet insecurity, a time consumimg and complicated process of designing tests, which may influence the test performance [7]. Much research has been done to determine the mode (computerized and paper-based testing) effect on language testing performance. In one of the recent articles on English language testing [9], no significant difference was found between scores on two modes and there was no positive interaction between testing mode preference and test performance. Other findings in [9] reveal that paper-based testing "was more comfortable and less fatiguing than computerized while computerized was easier (..) to record and change (..) answers". However, Stănculea [8] points out that " the main expense is not technical, but providing high quality valuation products, although this cost is identical to that used for the assessment paper."

E-TEST DESIGN

The design of e-tests comprises two components: an evaluation system and a questions bank. Moodle provides a wide range of assigning points to questions and the category aggregation strategies, which can be used flexibly [10]. It also allows to create quizzes that can be useful for language testing: multiple choice, true/false, matching, embedded answers (cloze), short answer, essay, numerical, description. For marking questions, in language e-testing these quizzes work better "with questions producing convergent answers (..) – a limited set of correct answers" [6]. However, in e-testing of Medical Latin, they can be used successfully because no sentence and paragraph responses are tested. Besides, many combinatorial varieties of terms and synonymy are not possible. Therefore, multiple choice and cloze question types were chosen for e-testing.

In order to increase the maximum of scoring, a partial answer marking was introduced. To do this, some possible variations of forming a Medical Latin term and a partially correct attempt were taken into account. For instance, for translation of 'the deep vein of the femur' the following syntax for Moodle short answer was created:
{3:SA:%0%vena * * *~=vena profunda femoris~%100%vena femoris profunda~%67%vena profunda~%67%vena femoris~%67%vena profunda *~%67%vena femoris *~%67%vena * profunda~%67%vena * femoris~%33%vena~%33%vena *~%0%* * profunda *~%0%* * * profunda *~%0%* * * profunda~%0%* profunda * *~%0%profunda *~%0%* femoris *~%0%* * * femoris *~%0%* * * femoris~%0%* femoris * *~%0%femoris *~%0%* vena~%0%* vena *~%67%* profunda femoris~%67%* femoris profunda~%33%* profunda~%33%* femoris~%0%* * profunda femoris * ~%0%* * profunda femoris~%0%* profunda femoris*~%0%* femoris profunda *~%0%* * femoris profunda~%0%*femoris profunda*~%33%* profunda *~%33%* femoris *}.
Table 1 demonstrates marking of the translation of the term 'the deep vein of the femur'.

Table 1. Marking algorithm for a sample of short answer question.

No.	Attempt	Mark
1	*vena femoris profunda*	3
2	*vena profunda femoris*	3
3	*vena femoris*	2
4	*vena profunda*	2

Table 1(continued).

No.	Attempt	Mark
5	*vena femoris x*	2
6	*vena profunda x*	2
7	*vena x femoris*	2
8	*vena x profunda*	2
9	*vena*	1
10	*vena x*	1
11	*femoris x*	0
12	*profunda x*	0
13	*x vena*	0
14	*x vena x*	0
15	*x femoris profunda*	2
16	*x profunda femoris*	2
17	*x femoris*	1
18	*x profunda*	1
19	*x femoris profunda x*	0
20	*x profunda femoris x*	0
21	*x femoris x*	1
22	*x profunda x*	1

In Table 1, x means any combination of letters or numbers or a misspelled word/words that is not accepted as the right answer. If an attempt contains more than 3 any words or symbols, the mark is 0. Figure 1 shows response templates of a translation attempt of the term 'the deep vein of the femur'.
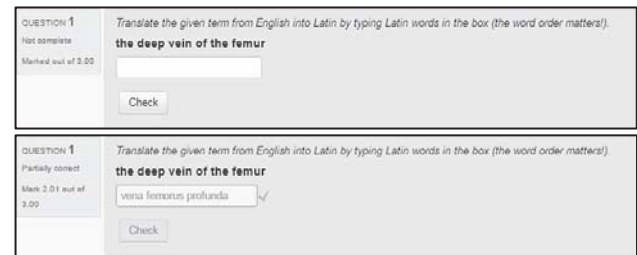

Figure 1. Moodle response templates.

Several Medical Latin term formation models and their marking algorithms were designed by using Moodle cloze syntax, for instance,
ABC
A_1/A_2BC
AB_1/B_2C
ABC_1/C_2
$AB \longleftrightarrow C$
$AB \longleftrightarrow C_1/C_2$
ABCD
$ABCD_1/D_2$
$ABCD_1/D_2/D_3$
ABCD =ACDB
$AB \longleftrightarrow C \longleftrightarrow D$
ABCDE
$AB \longleftrightarrow CDE$,
where A, B, C, D, E are Latin words in medical terms and A_1, A_2, B_1, B_2, $C1$, C_2, D_1, D_2, D_3 are their synonyms; the two direction arrows mean that the words can be interchangeable.

For questions where translation from Latin into English is required the syntax for Moodle multiple choice was used, for instance, for translation of '*crista tuberculi maioris humeri*' it looks like this:
{1:MC:=crest~of~greater~tubercle~humerus~tubercule~crests~for~big~majority~majorly~tuber~tuberous~tubercles

~humeri~crested}
{1:MC:crest~=of~greater~tubercle~humerus~tubercule~cr
ests~for~big~majority~majorly~tuber~tuberous~tubercles
~humeri~crested}
{1:MC:crest~of~=greater~tubercle~humerus~tubercule~cr
ests~for~big~majority~majorly~tuber~tuberous~tubercles
~humeri~crested}
{1:MC:crest~of~greater~=tubercle~humerus~tubercule~cr
ests~for~big~majority~majorly~tuber~tuberous~tubercles
~humeri~crested}
{1:MC:crest~=of~greater~tubercle~humerus~tubercule~cr
ests~for~big~majority~majorly~tuber~tuberous~tubercles
~humeri~crested}
{1:MC:crest~of~greater~tubercle~=humerus~tubercule~cr
ests~for~big~majority~majorly~tuber~tuberous~tubercles
~humeri~crested}.

Figure 2 shows response templates of a translation attempt of the term '*crista tuberculi maioris humeri'*.

(The list of words to choose from: petrified, major, nerve, stony, majority, canalized, nervous, canals, of, stone, for, nerves, majorly, channel, canal, petty)
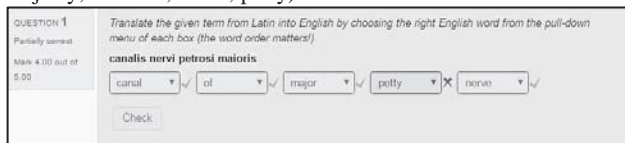

Figure 2. Moodle response templates.

From the pull-down menu list where words appear randomly each time when it opens, the testee chooses the right word in the right sequence. Multiple choice excludes using English synonyms and making spelling mistakes, which can be misinterpreted or evaluated subjectively in paper-based testing.

AIMS AND METHODS

The aims of the present study were to analyze the achievements of students' test results and to find out their opinions on certain aspects in relation to e-testing implementation in the course "Medical Terminology in Latin" for international students.

To achieve these aims, quantitative and qualitative methods were applied. These methods included: students were directly observed when doing e-tests and they were surveyed by structured questionnaire and open-ended questions at the very beginning of the second semester in the students' groups in order to get a feedback from those who had already passed the e-test. The survey was conducted in February 2017, involving 149 first year international students (130 women, 19 men). From 149 questionnaires received 1 of them did not provide the answer to 1 question, 4 - to 2 questions, 1 - to 3 questions. Responses to the closed questions of the questionnaire were analyzed using IBM SPSS Statistics 20; for hypotheses testing the Chi-Square test was used with the 95% confidence level in the analysis of 2x2 cross tabulations.

Descriptive statistics were carried out with IBM SPSS Statistics 20 to analyse the e-test results.

RESULTS

E-testing included six tasks that had to be accomplished within 90 minutes. E-tests were designed in the Moodle environment to allow to test any Medical Latin terminology, to randomize questions, to specify a time frame for availability, to set a time limit on task performance, to determine how many times a task could be taken and how it was to be scored. Table 2 summarizes the results of the e-test. To assess the reliability or internal consistency of test items, a Cronbach's alpha coefficient was calculated and for the six items it was 0.715, which is considered "acceptable". Reliability of the e-test can be increased by excluding multiple choice questions or adding more short answer ones.

Table 2. The e-test task content, weights and results.

Tasks	Number of questions	Weight; %	Cronbach's alpha if item deleted	Median (n=140)
Task 1: Translation of the given one-word anatomical terms into Latin and typing their dictionary forms. (Short answer)	20	20	0.646	83.5%
Task 2: Choosing the correct form of an adjective that can follow the given noun. (Multiple choice)	5	10	0.709	100.0%
Task 3: Translation of the given terms from English into three- or four-word Latin terms.(Short answer)	15	25	0.610	67.6%
Task 4: Translation of the given three- or four-word terms from Latin into English.(Multiple choice)	15	20	0.635	75.0%
Task 5: Choosing the correct noun form that can precede the given adjective. (Multiple choice)	10	10	0.706	100.0%
Task 6: Analysis of each word in the given three-word Latin term by determining its declension, gender, case and number. (Multiple choice)	3	15	0.717	97.2%

Both the Kolmogorov-Smirnov and Shapiro-Wilk tests indicate that total scores are not normally distributed. The histogram (Figure 3) shows that the total scores are skewed left. The quartiles are as follows: 1st – 61%, 2nd – 71%, 3rd – 77%. This suggests that the extreme scores (100%) should be eliminated, i.e., a number and weight of each question type should be commensurable, which requires further research.

Figure 4 illustrates the performance of task 3 (translation of the given terms from English into three- or four-word Latin terms –

see Table 2). Presumably due to the lack of time, scores were lower for the majority of students. On the contrary, for instance, the multiple choice task (choosing the correct form of an adjective that can follow the given noun – see Table 2) were completed ahead of time (see Figure 5). This proposes to revise the time limits for each type of task and to change them.
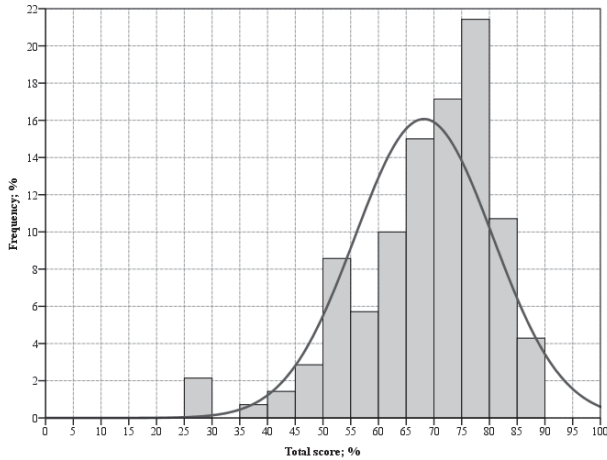

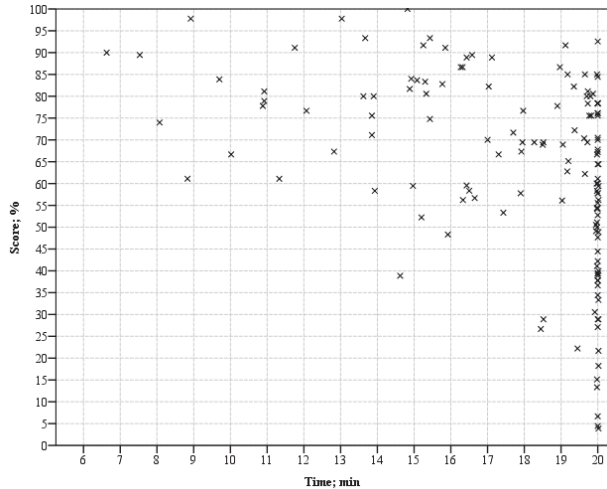Figure 3. Histogram of e-test total scores.

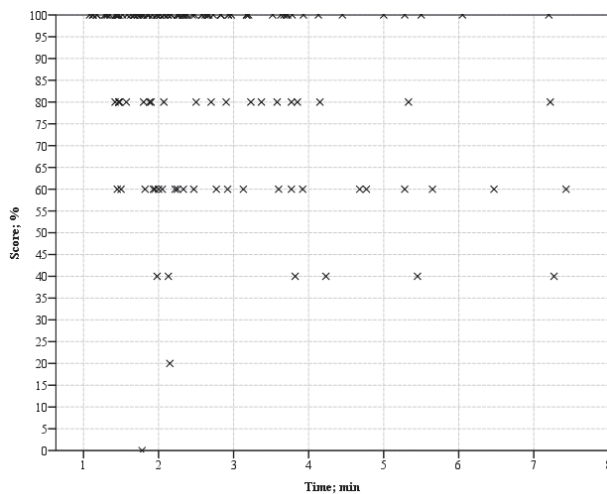Figure 4. Influence of time students spent on task 3 on scores

Figure 5. Influence of time students spent on task 2 on scores.

To find out to what extent opinions on some e-testing aspects in terms of its advantages and disadvantages differ, according to the students' responses to the survey, the two sets of opposite views were identified : 1) the opinions that a) e-tests were more complex to perform and b) e-tests were not more complex to perform compared to paper-based testing; 2) the opinions that a) e-tests were not much more useful and b) e-tests were useful for testing knowledge of Medical Latin terminology compared to paper-based testing.

The following hypotheses were subsequently raised in this study:
Hypothesis 1 (in relation to e-test complexity):
H0: there is no difference between the students' opinions on certain aspects of those who consider e-testing of knowledge in Medical Latin as more complex to perform compared to paper-based testing and those who do not.
H1: there are differences between the students' opinions on certain aspects of those who consider e-testing of knowledge in Medical Latin as more complex to perform compared to paper-based testing and those who do not.
Hypothesis 2 (in relation to e-test usefulness):
H0: there is no difference between the students' views on certain aspects of those who believe that e-tests are useful for testing of knowledge in Medical Latin compared to paper-based testing and those who believe that they are not much more useful.
H1: there are differences between the students' views on certain aspects of those who believe that e-tests are useful for testing of knowledge in Medical Latin compared to paper-based testing and those who believe that they are not much more useful.

The survey questions included such aspects of the e-test in relation to its advantages and disadvantages of 1) receiving immediate feedback, 2) time and paper savings, 3) stress factor impact, and 4) complementary effects (computer literacy, the time frame, technical failures, etc.).
Between the two opposite students' opinions (hypothesis 2), no difference of opinions on all these above mentioned four aspects, respectively: 1) $\chi2$ (df = 1, n = 147) = 0.627, p = 0.429; 2) $\chi2$ (df = 1, n = 147) = 0.549, p = 0.459; 3) $\chi2$ (df = 1, n = 146) = 0.156, p = 0.693; 4) $\chi2$ (df = 1, n = 146) = 1.918, p = 0.166 was found. Between the two opposite students' opinions (hypothesis 1), no difference of opinions on two aspects mentioned above – time and paper saving - $\chi2$ (df = 1, n = 146) = 0.763, p = 0.382 and stress factor - $\chi2$ (df = 1, n = 144) = 0.056 , p = 0.813 was found. However, minor differences between the latter two opinions were found on immediate feedback ($\chi2$ (df = 1, n = 146) = 3.875, p = 0.049) and the influence of complementary effects ($\chi2$ (df = 1, n = 144) = 4.126, p = 0,042), which were considered more essential of those students who considered that e-tests were not complex.

Finally, students' answers to open-ended questions that asked for other opinions can be summarized into three groups:
1. Unsatisfactory aspects:
a) the allocated time of task completion was too short;
b) for multiple choice questions only one possible response was possible, no synonyms given;
c) if a spelling mistake, marking decreased;
d) translated phrases had to have a precise word order (Latin variants were not permitted);
e) the evaluation system was too rigid, no exceptions were possible.
2. Positive aspects:
a) e-tests as self-checking tests are great;

b) e-tests make it possible to practice individually, as well as provide opportunities to prepare for final tests;

c) e-testing is a modern and progressive way of testing knowledge;

d) a positive attitude of the teacher during e-testing reduces anxiety;

e) immediate test results.

3. Recommendations:

a) e-tests should be useful for summative evaluation in the end of the course;

b) e-tests would be useful for preparations for examinations timely (for formative assessment);

c) it would be helpful to review some of the tasks;

d) it is not necessary after each question to verify whether the answer is correct;

e) it is desirable not to limit the time for each specific task, but to give a certain time for the entire test as a whole.

Responses from students show a diversity of opinion, but they are not strongly negative, so it is believed that a start of the implementation of e-testing was successful.

CONCLUSIONS

1. Usefulness of the introduction of the e-tests is justified; new methods should be used both for knowledge testing and training in regular practical classes.

2. Moodle e-tests can be linked to other computer-based or online tools (for instance, mobile phones, tablets), thus facilitating student independence.

3. Challenge: the above-mentioned introduction requires hard work from the teacher; the work is rather time-consuming, requires great precision and considerable effort in setting up the template answers and training the system on a range of trial responses.

4. Challenges: excluding cases of cheating/ prompting was not always possible because of a fixed arrangement of computer equipment.

5. Challenges: the teaching staff should come to a consensus regarding a comparability of scoring in e-testing and paper-based testing.

REFERENCES

[1] **Terminologia anatomica**,1998, accessed June1, 2017, available at http://www.unifr.ch/ifaa/Public/EntryPage/HomePublic.html

[2] A. M. Marin, "Medical Vocabulary and the ESP Teacher", **Studii şi cercetări filologice – Seria Limbi Străine Aplicate**, No. 8, 2009, pp. 83-89, Universitatea din Piteşti, Facultatea de Litere, Piteşti, accessed June1, 2017, available at http://scf-lsa.info/wp-content/uploads/2016/07/SCF-nr.8.2009-ft.pdf

[3] I. K. Nagy, "The History, Peculiar Terminology and Translation Problems of the Language of Medicine", **Studies on Literature, Discourse and Multicultural Dialogue,** ed.: Iulian Boldea, Published by Editura, Arhipelag XXI, Târgu-Mureş, România, 2013, pp. 179-189, accessed June1, 2017, available at http://www.upm.ro/ldmd/LDMD-01/LDMD%2001%20-%20Language%20and%20Discourse.pdf

[4] V. Viksne, I. Abelite, "Teaching Latin in International Student Groups: Comparative Study", International conference, **Language Centres in Higher Education, Sharing Innovations, Research, Methodology and Best Practices**, Masaryk University, Brno, Czech Republic, 15−17 January 2015, pp. 126-137, accessed June1, 2017, available at http://casajc.ff.cuni.cz/Documents/CASAJC_online/1502/1502.pdf

[5] H. R. Wulff, "The Language of Medicine", **Journal of the Royal Society of Medicine**, Vol. 97, 2004, pp. 187–188, accessed June1, 2017, available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1079361/pdf/0970187.pdf

[6] J. Winkley, "E-assessment and Innovation", Becta, 2010, 22 p., accessed June1, 2017, available at file:///C:/Users/User/Downloads/e-assessment_and_innovation.pdf

[7] Â. G. Pereira, F. Scheuermann, "On E-testing: an Overview of Main Issues", **Joint Research Centre, European Commission**, 2007, 16 p., accessed June1, 2017, available at https://core.ac.uk/download/pdf/38613296.pdf

[8] L. Stănculea, "E-Assessment, Learning and Appreciation", **Journal of Business Management and Applied Economics**, Vol. 1, Issue 6, 2012, 8 p., accessed June1, 2017, available at https://jbmae.scientificpapers.org/category/vol-1-2012/volume-i-issue-no-6/

[9] H. Khoshsimaa, S. M. H. Toroujenib, "Comparability of Computer-Based Testing and Paper-Based Testing: Testing Mode Effect, Testing Mode Order, Computer Attitudes and Testing Mode Preference", **International Journal of Computer**, Vol. 24, No. 1, 2017, pp. 80-99, accessed June1, 2017, available at http://ijcjournal.org/index.php/InternationalJournalOfComputer/issue/view/35

[10] Moodle Manual, Compiled by HowToMoodle, Moodle Version 2.7, accessed June1, 2017, available at http://www.moodleworld.com/how-to-moodle-course-creator-essential-2-7-manual-by-howtomoodle-moodleworld/