Usability of Computerized Lung Auscultation – Sound Software (CLASS) for learning pulmonary auscultation

Ana Machado[a, d]; Ana Oliveira[a, d], MD; Cristina Jácome[a], MD, PhD; Marco Pereira[a, d]; José Moreira[b], MD, PhD; João Rodrigues[b], MD, PhD; José Aparício[c], MD, PhD; Luis M. T. Jesus[a,b], MD, PhD; Alda Marques[a,d], MD, PhD

[a]Lab 3R – Respiratory Research and Rehabilitation Laboratory, School of Health Sciences (ESSUA), University of Aveiro, 3810-193 Aveiro, Portugal

[b]Institute of Electronics and Informatics Engineering of Aveiro (IEETA), University of Aveiro 3810-193 Aveiro, Portugal

[c]Paediatrics Emergency Department, Hospital Lusíadas, 4050-115 Porto, Portugal

[d]Institute for Research in Biomedicine (iBiMED), University of Aveiro, Aveiro, Portugal

Corresponding author: Alda Marques, PT, MSc, PhD, Senior Lecturer, School of Health Sciences, University of Aveiro (ESSUA), Agras do Crasto - Campus Universitário de Santiago, Edifício 30, 3810-193 Aveiro, Portugal. Email: amarques@ua.pt

Number of words of the manuscript: 3419

Number of words of the abstract: 238

Number of figures: 3

Number of tables: 4

Abstract

Rationale: The mastering of pulmonary auscultation requires complex acoustic skills. Computer-assisted learning tools (CALTs) have potential to enhance the learning of these skills; however few have been developed for this purpose and do not integrate all the required features. Thus, this study aimed to assess the usability of a new CALT for learning pulmonary auscultation. **Method:** Computerized Lung Auscultation – Sound Software (CLASS) usability was assessed by 8 physiotherapy students using computer screens recordings, think aloud reports and facial expressions. Time spent in each task, frequency of messages and facial expressions, number of clicks and problems reported were counted. The timelines of the 3 methods used were matched/synchronized and analyzed. **Results:** The tasks *exercises* and *annotation of respiratory sounds* were the ones requiring more clicks (median 132, interquartile range [23-157]; 93 [53-155]; 91 [65-104], respectively) and where most errors (19%; 37%; 15%, respectively) and problems (n=7; 6; 3, respectively) were reported. Each participant reported a median of 6 problems, with a total of 14 different problems found, mainly related with CLASS functionalities (50%). *Smile* was the only facial expression presented in all tasks (n=54). **Conclusion:** CLASS is the only CALT available that meets all the required features for learning pulmonary auscultation. The combination of the 3 usability methods identified advantages/disadvantages of CLASS and offered guidance for future developments, namely in annotations and exercises. This will allow the improvement of CLASS and enhance students' activities for learning pulmonary auscultation skills.

Key Words: Computer-assisted learning; Computerized respiratory sounds; Interface usability; Lung auscultation

1. Introduction

Pulmonary auscultation is an essential part of the physical examination of patients with respiratory conditions [2]. Although auscultation is commonly used among health professionals [1], the mastering of this procedure requires complex acoustic skills to distinguish between different respiratory sounds (RS) with similar frequencies, intensities and timings [36,27]. Currently, health students are taught these skills by repeatedly listening to recordings of typical RS [36,15] and visualizing their waveforms [28]. However, these methods offer limited interaction and provide students with a narrow representativeness of RS and conditions. Thus, to improve health students' skills to detect/discriminate RS, it is crucial to develop innovative teaching methods [20].

Computer-assisted learning tools (CALTs) aim to provide students with complementary activities on a computer, related with the material being taught. Use of CALTs have been shown to allow a more self-directed learning, having the potential to improve teaching and learning skills [34]. Such tools show great potential to be used in the teaching of auscultation, as they would allow students to interact with a diversity of RS recorded in clinical environments, from patients with different conditions and test the knowledge acquired. However, only few have been developed in the area of respiratory medicine [36]. *CompuLung* [20,19] and *R.A.L.E.* [35] are two of the CALTs available, however they are not open source, the first does not allow students to record RS or test their acquired knowledge (e.g., via RS exercises that incorporate solutions created by a panel of RS experts), and the second is only available for *Windows* operative system. Hou et al. [13] have also developed a CALT, aiming to assist nursing education on auscultation, however it does not include a practical component (i.e., it does not allow knowledge testing via exercises/tests resolution). *LungSounds@UA* [25], *RSAS@UA* [8] and *MARS Database* [10] are other available tools but neither of those include simultaneous recording and analysis of RS, nor exercises to test knowledge acquisition. *LungSounds@UA* [25] only allows to record and store RS, whilst with *RSAS@UA* [8] users can analyze but cannot record RS nor have feedback about the analyzes performed. *MARS* [10] is a database of RS that allows users to listen to different RS acquired in real patients, however it does not allow knowledge testing. Thus, new CALTs integrating simultaneously all the required features, i.e., record, storage, playback and analysis of RS, knowledge testing and tutorials about RS; are needed to enhance health students' skills on pulmonary auscultation.

Computerized Lung Auscultation – Sound Software (CLASS) [27] was developed to simultaneously record, analyze and interpret RS. CLASS had a preliminary validation in which its utility and potential to be used in academic and clinical environments were highlighted [27]. However, only users' personal perceptions, through questionnaires and focus group, were assessed and other recommended procedures for usability testing were not performed (i.e., computer screen and facial expressions videos, and think aloud reports) [18,17]. Therefore, this study aimed to evaluate the usability of CLASS for learning pulmonary auscultation, according to the international standards for software validation [14,17].

2. Methods

2.1. CLASS description

CLASS has been based on two previously developed applications: *LungSounds@UA* [25] and *RSAS@UA* [8]. This CALT allows recording and analyses of RS in a single application and aims to be used by health students for learning purposes.

CLASS is organized in four tabbed document interfaces: main, recordings, annotations and tutorials. It allows RS recording with a digital stethoscope or microphone (Fig. 1), storage, playback of files and analyses, practice of RS exercises, which have been developed and solved by a panel of RS experts (AM, CJ and AO) to form RS gold standards, and further knowledge consolidation using the available tutorials (Fig. 2) on RS definition, acoustic properties and clinical interpretation.

Detailed description of CLASS can be found elsewhere [27].

Fig. 1 – CLASS window for respiratory sounds recording.

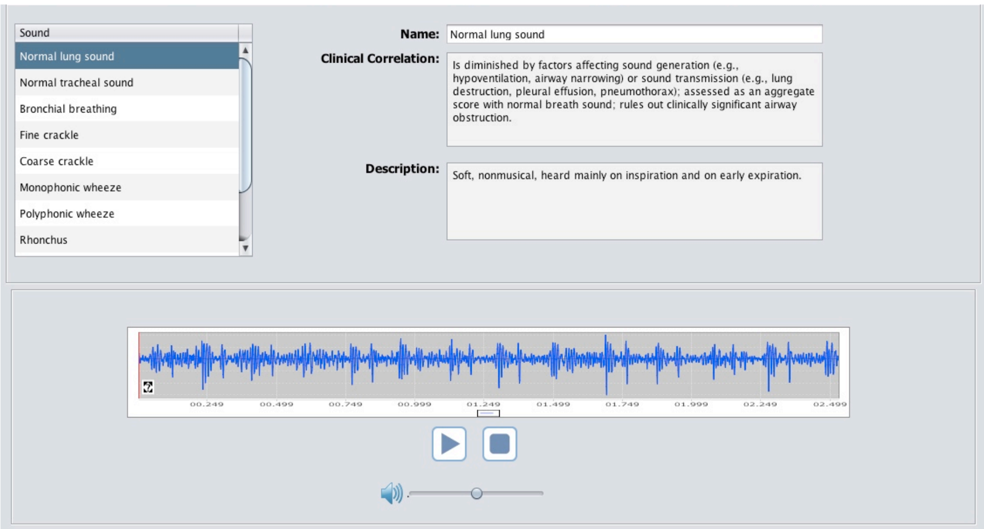


Fig. 2 – CLASS tutorials.

2.2. Study design

A cross sectional study was developed to test CLASS in eight individual evaluation sessions conducted on the same day at the University of Aveiro.

Eight physiotherapy students with previous education in respiratory physiotherapy were informed about the study and asked about their willingness to participate. Sample size was selected based on previous studies reporting up to 80% of sensitivity in detecting interface's problems using 8-10 participants [17] and on the definition of usability according to the ISO 9241-11 [14].

Ethical approval was previously obtained from School board Ethics Committee and written informed consents were collected from all participants.

Participants' interaction with CLASS was recorded simultaneously with three different usability testing methods: computer screen videos (CSVs), think aloud (TA) reports and facial expression videos (FEVs). These methods were selected as they have been described as the most effective to evaluate participants' interaction with a system, while performing the same tasks [17].

CSVs is one of the most recommended methods to test usability [7]. This method consists in recording the user's computer screen while interacting with a system [7], thus allowing to collect objective data of users performance, such as the time spent in each task and the number of errors occurred [18].

TA involve the audio recording of users verbalizing their thoughts when using a system, which informs on the problems found during the interaction [17]. This method assesses users' thought processes or decision making when performing a specific task [5].

FEVs consist in video-recording users' facial expressions when interacting with a system [17]. This method captures participants' focus of attention, their interaction with the environment and specially their emotions, as they are primarily communicated through facial expressions [24].

2.3. Procedures

Two days prior to the validation sessions, participants attended a 60 minute group training session [6]. Participants received a user-manual describing the general structure of CLASS and were encouraged to explore the application on a computer without talking with each other. No further contact with CLASS was provided to participants until the validation session.

The validation sessions occurred in two rooms, set up according to Kushniruk and Patel [17]. Participants were seated in front of a desk with a laptop with CLASS and the TipCam Screen Recording Software installed [32], an audio-recorder, a digital stethoscope (WelchAllyn Meditron 5079-400), headphones and a video camera pointed to their faces (Fig. 3). Two researchers involved in the development of CLASS conducted the sessions, however they only intervened to clarify participants' doubts. All participants received an user-manual and a case-study developed according to Kushniruk et al. [18]. The case-study aimed to guide participants to perform the same tasks, representative of the real use of the application, i.e., *create a patient*, *record two RS* files (in the researcher or themselves), *annotate the recorded RS* (i.e., identification of respiratory phases, abnormal RS), perform one *beginner exercise* and one *advanced exercise* (i.e., identification of respiratory phases, abnormal RS, addition/removal of annotations, comparison of annotations with the gold standard) and *consult tutorials*.

One researcher read the case-study aloud and participants were given enough time to read it by themselves and clarify any doubts before starting the tasks.

Then, the researchers turned on the recorder software, video camera and audio-recorder.

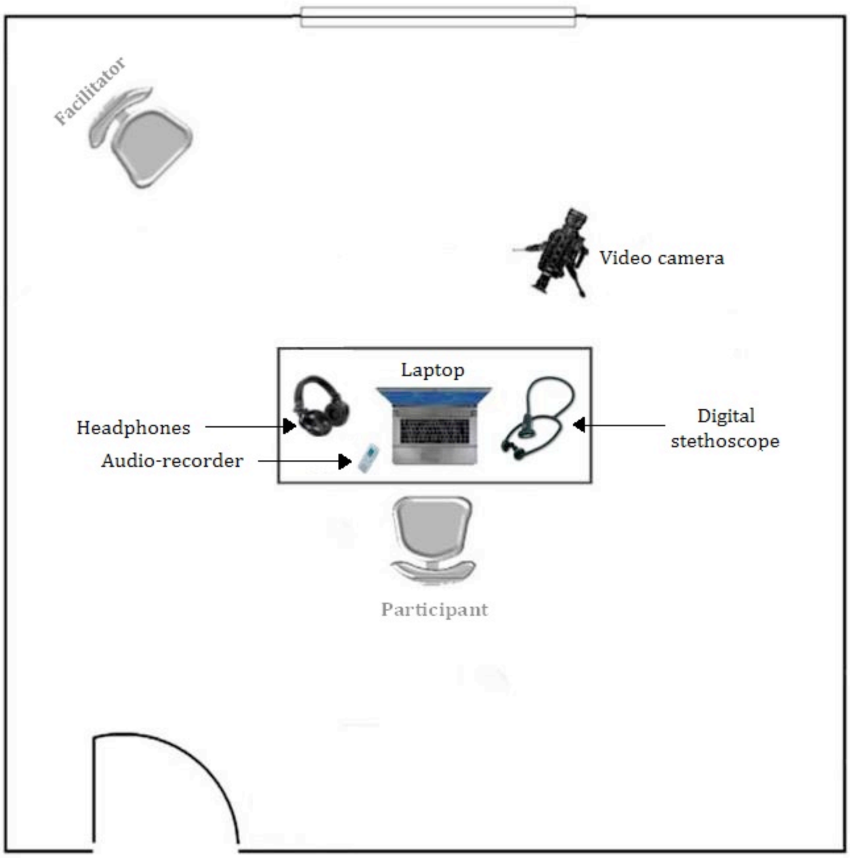


Fig. 3 – Room setup.

3. Data analysis

Descriptive statistics were used to describe the sample. Data is presented as mean (standard deviation), median [interquartile range] or number (percentage).

3.1. Analysis of Computer Screen Videos

Two researchers independently observed and analyzed the CSVs using the Noldus The Observer XT 10.5 software (Noldus International Technology, Wageningen, Netherlands) [21]. This software has been developed to manage and analyze observational data, and its use in human-computer interaction studies has previously been validated [23,38]. The time spent in each task, frequency of warning, error and success messages and number of clicks per task were counted.

3.2. Analysis of Think Aloud reports

Three researchers listened and transcribed eight audio files. Then, two researchers conducted a thematic analysis [33], codifying it under the following themes, previously agreed in a consensus meeting:

- *Report* – commentaries describing which tasks were being performed in the interface.
- *Doubt* – commentaries reporting doubts in understanding the case-study or performing tasks.
- *Problem* – commentaries reporting problems/difficulties when interacting with the interface.

The *problems'* theme was further grouped into 3 categories:

- *Layout* – commentaries about the interface design and presentation*.*
- F*unctionalities* – commentaries reporting difficulties/problems with interface functions*.*
- *Unfamiliarity* – commentaries reporting difficulties using the interface due to participants' lack of familiarity using it*.*

Disagreements in data coding and grouping were solved by consensus and when consensus could not be reached, a third researcher was consulted. The frequency of each theme/category was analyzed.

3.3. Analysis of the Facial Expression Videos

Two researchers analyzed participants' facial expressions using the Noldus software.

Facial expressions were studied by analyzing the frequency and duration of a list of behaviors (ethogram), derived from: the literature [3,23]; preliminary observations of the video recordings [22]; and the facial acting coding system [4].

The following categories composed the ethogram: *eye contact with the screen*; *verbal communication*; *look away*; *read the case*; *smile*; and *other*, such as *frown*, *confusion*, *head shake*, *consult the manual* or *hand gestures* (Table 1).

Table 1 –Ethogram of the facial expressions.

Categories	Description
Eye contact with the screen	The user looks directly to the screen, clearly focused on reading, searching or understanding something in the interface.
Verbal communication	The user communicates using words and/or sentences, to clarify any doubt about the system or report his/her thoughts or problems found.
Look away	The user looks away from the screen, looking around for nothing in particular.
Read the case	The user looks directly to the case study, to read or understand something in it.
Smile	Facial expression where the lips slide back or move away slightly (mouth can be half opened) as indicative of agreement, comprehension and accomplishment.
Other	
Frown	The user corrugates both eyebrows as indicative of frustration/dislike for not understanding the interface or not finding what he/she is looking for.
Confusion	Facial expression where the eyes are wide open and the face shows confusion, as indicative of a mistake or misunderstanding.
Head shake	The user shakes his/her head in a negative way as indicative of disagreement.
Consult the manual	The user looks directly to the user-manual, to understand how to perform one or more tasks.
Hand gestures	The user moves with his/her hands while trying to accomplish a task, to support the thinking process.

After the individual analysis of the CSVs, TA reports and FEVs, two researchers matched their timelines to relate the facial expressions, problems reported at TA and error messages with the performed tasks.

4. Results

Eight physiotherapy students (37.5% males; age 20.5 (0.5) years) completed the training and validation sessions. During the analysis, one screen recording was found to be corrupted due to a technical problem and was excluded. Hence, 23 video and audio files were analyzed: 7 CSVs, 8 TA reports and 8 FEVs. Each participant took on average 32 (12) minutes to complete the tasks.

4.1. Computer Screen Videos

Participants spent more time in the *advanced exercise* (6.3 min [1.8–8.4 min]), followed by the *annotation of the recorded RS* (5.8 min [4.5–7.3 min]) and the *beginner exercise* (4.5 min [3.4–13.1 min]). The tasks with the shortest duration were *create a patient* (2.4 min [0.5–3.1 min]) and *consult tutorials* (2.1 min [0.9–2.8 min]) (Table 2).

Regarding to the number of clicks needed to accomplish a task, the shortest task (i.e., *consult tutorials*) was associated with fewer clicks (12 [1–29]). Similarly, the most time-consuming task (*advanced exercise*) was associated with the highest number of clicks (132 [23–157]).

Considering the messages displayed by the interface, each participant found a median of 15 [12–19] messages, of which 46.6% (n=62) were *success messages*, 33.1% (n=44) *warning messages* and 20.3% (n=27) *error messages*. Most *error messages* occurred at *beginner* (n=10; 37%) and *advanced exercises* (n=5; 18.5%).

Table 2 – Events found in the computer screen videos.

Task	Time spent (min)	Error messages	Warning messages	Success messages	Number of clicks
Create a patient	2.4 [0.5–3.1]	4 (14.8%)	0 (0%)	5 (8.1%)	40 [6–45]
Record RS	4.2 [3.1–6.9]	3 (11.1%)	0 (0%)	20 (32.3%)	32 [25–67]
Annotate RS	5.8 [4.5–7.3]	4 (14.8%)	10 (22.7%)	8 (12.9%)	91 [65–104]
Beginner exercise	4.5 [3.4–13.1]	10 (37%)	21 (47.7%)	20 (32.3%)	93 [53–155]
Advanced exercise	6.3 [1.8–8.4]	5 (18.5%)	10 (22.7%)	9 (14.5%)	132 [23–157]
Consult tutorials	2.1 [0.9–2.8]	0 (0%)	3 (6.8%)	0 (0%)	12 [1–29]
Total	31.3 [23.3–41.7]	27 (100%)	44 (100%)	62 (100%)	394 [242–549]

Data is presented as median [interquartile range] or number (percentage).

4.2. Think Aloud reports

A total of 447 interventions were found at TA transcriptions: 283 (63.3%) *reports*, 74 (16.6%) *problems* and 67 (15.0%) *doubts*. Each participant intervened approximately 6 times [2.3–12.0] to report a problem, resulting in 14 different interface *problems* detected: 7 (50.0%) related to the interface *functionalities*, 5 (35.7%) due to *unfamiliarity* with the interface and 2 (14.3%) related to the interface *layout*.

Regarding interface *functionalities*, 7 participants (87.5%) reported difficulties hearing/annotating the recorded RS due to noise/interference and 4 participants (50%) claimed that the sound presented in the *advanced exercise* was too low in volume to be clearly heard. Other *functionalities'* problems such as impossibility of navigating between tabbed document interfaces without losing the previous inserted information (n=1; 12.5%), interface crashing during the tasks (n=1; 12.5%), difficulties selecting and removing the respiratory events from the annotation panel (n=1; 12.5%), difficulties adding respiratory events (n=1; 12.5%) and difficulties in selecting the required patient (n=1; 12.5%) were also reported.

In the *unfamiliarity* with the interface category, 2 participants (25%) showed difficulties understanding the aim of the *advanced exercise* (i.e., correct the annotations already performed in a sound file) and the concept of gold standard. Other *unfamiliarity* problems such as difficulties in identifying the right patient to record the RS in the patient list (n=1; 12.5%), difficulties in identifying the different colors corresponding to each respiratory phase, crackles and wheezes annotation in the annotation panel (n=1; 12.5%) and difficulties in understanding the annotation process (n=1; 12.5%) were also reported.

Concerning to the interface *layout*, 2 participants (25%) referred that the application should have a timeline that follows RS reproduction, and 1 (12.5%) reported he/she missed a toolbar which allowed scrolling throughout the table presenting the respiratory events annotated.

After matching the problems reported in TA with the tasks participants performed, it was observed that most problems occurred at *advanced* (n=7; 46.7%) and *beginner exercises* (n=5; 33.3%). *Create a patient* and *consult tutorials* tasks did not present any problems reported. Additionally, *annotate RS* was the task were most participants found problems (n=5; 62.5%) (Table 3).

Table 3 – Number of participants reporting a problem per task.

Problems	**Tasks**						Total[a]
	Create a patient	Record RS	Annotate RS	Beginner exercise	Advanced exercise	Consult tutorials	
Recorded RS	-	3 (37.5%)	5 (62.5%)	-	-	-	7 (87.5%)
RS at *advanced exercise*	-	-	-	-	4 (50%)	-	4 (50%)
Navigation between TDI	-	-	-	-	1 (12.5%)	-	1 (12.5%)

Interface crashing	-	-	-	1 (12.5%)	1 (12.5%)	-	1 (12.5%)
Select/remove annotation	-	-	-	-	-	-	1 (12.5%)
Add annotation	-	-	-	1 (12.5%)	-	-	1 (12.5%)
Keeps backing to other patient	-	-	1 (12.5%)	-	-	-	1 (12.5%)
Understand *advanced exercise*	-	-	-	-	2 (25%)	-	2 (25%)
Choose wrong patient	-	1 (12.5%)	-	-	-	-	1 (12.5%)
Understand gold standard	-	-	-	1 (12.5%)	1 (12.5%)	-	2 (25%)
See respiratory phases' lines	-	-	-	1 (12.5%)	-	-	1 (12.5%)
Difficulties annotating	-	-	-	1 (12.5%)	-	-	1 (12.5%)
Missing timeline	-	-	1 (12.5%)	-	1 (12.5%)	-	2 (25%)
Missing scroll bar	-	-	-	-	1 (12.5%)	-	1 (12.5%)
Total[b]	0 (0%)	4 (50%)	5 (62.5%)	4 (50%)	4 (50%)	0 (0%)	8 (100%)

Data is presented as number (percentage).
[a] Participants reporting the same problem at more than one task were counted only once.
[b] Participants reporting more than one problem at the same task were counted only once.

4.3 Facial expressions

Eye contact with the screen was the behavior category with the highest duration (mean duration 28 (10) min) whilst *verbal communication* was the category most frequently observed (48.5 [9.5–81.8]).

Other categories frequently observed were *look away* from the screen (34 [13–45.8]) and *read the case* (18.5 [15.5–22.5]). *Smile* was the less observed category (5 [1.25–10.5]).

In the *others* category, 5 participants showed *confusion* (1 [0–4.75]), 5 *shook* their *heads* (1 [0–2]), 4 *frown* their brows (0.5 [0–2]), 3 *consulted the manual* (0 [0–2.5]) and 3 presented *hand gestures* (0 [0–1]).

After matching the facial expressions with the tasks performed, it was observed that *smile* appeared mainly when *recording a RS* (n=10; 18.5%) and *creating a patient* (n=9; 16.7%), *confusion* was observed mostly at *advanced exercise* (n=6; 35.3%) and *head shake* occurred generally when *annotating a RS* (n=4; 40%) (Table 4).

Table 4 – Facial expressions observed when the participants performed specific tasks.

Facial expressions	**Tasks**							Total
	Create a patient	Record RS	Annotate RS	Beginner exercise	Advanced exercise	Consult tutorials	Other	
Smile	9 (16.7%)	10 (18.5%)	3 (5.6%)	1 (1.9%)	5 (9.3%)	1 (1.9%)	25 (46.3%)	54 (100%)
Frown	1 (11.1%)	1 (11.1%)	0 (0%)	0 (0%)	2 (22.2%)	0 (0%)	5 (55.6%)	9 (100%)
Confusion	0 (0%)	1 (5.9%)	2 (11.8%)	3 (17.6%)	6 (35.3%)	0 (0%)	5 (29.4%)	17 (100%)
Head shake	1 (10%)	2 (20%)	4 (40%)	0 (0%)	2 (20%)	0 (0%)	1 (10%)	10 (100%)
Hand gestures	0 (0%)	2 (40%)	2 (40%)	0 (0%)	0 (0%)	0 (0%)	1 (20%)	5 (100%)

Data is presented as number (percentage).

5. Discussion

This is the first study reporting on the combination of CSVs, FEVs and TA reports to validate a CALT. Comprehensive and more objective results on the validation of CLASS have been found, namely observing that the most time-consuming tasks, were the ones associated with the display of more error messages, higher number of clicks, problems reported and negative facial expressions (e.g., *confusion* and *head shake*). Despite these drawbacks, it was also found that *smile* was the only facial expression present in all tasks performed, which indicates that, overall, participants were satisfied with CLASS functionalities and performance, and therefore it has potential to be integrated in students' learning activities.

The CSVs data allowed observing that error messages were the less frequent type of message found, and that *beginner* and *advanced exercises* were the tasks with the highest prevalence of error messages. This is a positive result towards the implementation of CLASS, since it overcomes the drawbacks of similar previously validated CALTs, reported as less intuitive [25], and follows the literature recommendations on error rates and error prevention (i.e., low error rates and error prevention, are desirable in human computer interfaces) [11,12,30]. Also, it emphasizes previous results which showed that CLASS is easy to navigate and perform tasks [27].

Nevertheless, it should be noted that *beginner* and *advanced exercises* need to be further improved, as it is known that a high prevalence of errors affects task's outcomes and lead to more usability problems [30]. These improvements will be accomplished by: i) adding a timeline that follows RS reproduction, ii) using RS of better quality, iii) simplifying the process of selecting, add and removing respiratory events, iv) adding an extra simpler and easier level of exercises.

Although more error messages have been reported at *beginner exercise,* participants needed more time to complete the *advanced* than the *beginner exercise*. It has to be noted that, although exercises had different complexity levels, they were similar in terms of the interface commands. Therefore, after performing the *beginner exercise*, participants may have developed a better understanding of which steps to perform in the interface to complete the *advanced exercise* (e.g., how to add/remove a respiratory event and start/pause the RS file). This phenomenon has been previously described by Davis et al. [6], who claimed that prior knowledge may aid in learning a succeeding task, and can also be supported by the few number of unfamiliarity problems found at *advanced exercise* relatively to *beginner exercise*.

Literature has shown that the number of clicks during a given task are indicative of users' behavior [11], being the tasks which require higher number of clicks associated with higher levels of effort needed to accomplish it [31]. In the study of Krall and Sittig [16], participants suggested a reduction in the number of mouse clicks to increase system's efficiency. This information can be directly applied to the improvement of CLASS by showing that it is essential to reduce the number of mouse clicks needed to fulfill *beginner* and *advanced exercises,* and *annotation of RS* [31]. Similar to what has been found with the CSVs analysis, analysis of FEVs further emphasized that *beginner/advanced exercises* and *annotation of RS* where the tasks most associated with negative (e.g., *confusion*) [26] and disagreement expressions (e.g., *head shake*) [9].

TA reports have shown potential to collect very detailed and exceptionally revealing data in real-time use [5]. Although it was impossible to distinguish TA reports from communication with the researcher at transcription, the association of these reports with FEVs allowed this distinction for the majority of data. Combination of these data was essential to distinguish between problems found by users that could be solved by themselves, from problems requiring the researcher intervention.

Problems found at TA were mainly related with interface *functionalities*. The most reported problem was difficulties hearing/annotating the recorded RS due to noise/interference. Although some of the problems found were similar to the one's previously reported [27], this study uncovered new important difficulties, such as the low volume of the RS file presented at *advanced exercise*, the need for a toolbar that allows scrolling throughout the table presenting the events annotated and the misunderstanding of the gold standard. This information is essential to rethink the presentation of exercises and especially of the gold standard, as this has been considered a crucial feature of educational respiratory CALTs [8] and is one of the major improvements of CLASS when compared with previously validated systems [8].

The tasks *create a patient* and *consult tutorials* were the ones where *smile* was mainly observed. It is known that this facial expression is usually linked to happiness, agreement and accomplishment [37,23] and thus, may reflect the importance that participants attribute to these tasks and the pleasure felt when accomplishing them with success. Nevertheless, although being more frequent in these tasks, *smile* was the only facial expression observed in all tasks which shows that, although improvements are needed, the interface was overall friendly to use [23,37]. These results are similar to those found by Semedo et al. [27] in the questionnaires filed by participants after the validation sessions.

Finally, it should be noted that each participant only reported approximately 6 problems in their interaction with the interface, and almost half of the problems were due to *unfamiliarity* with the interface or interface *layout*. Therefore, it seems that with users' experience and incorporation of layout suggestions, CLASS has great potential to be easily incorporated in students' academic activities.

5.1. Limitations

Some limitations need to be acknowledged. Firstly, CLASS was tested only with physiotherapy students, leaving aside other health students which could potentially benefit from its use. It should be noted that this was a preliminary validation and according to the current guidelines these students were representative of the target user population [18]. After implementing all the required improvements, it is planned to test CLASS with a broader sample including other students, health professionals and researchers. Secondly, the presence of external observers in the testing rooms might have led to psychological, physiological and emotional changes [29]. Nevertheless, the interaction with the researcher has been reduced to the essential minimum and the organization of the testing room followed standardized rules [17] to prevent participants' distraction and distress. Thus, it is believed that researchers' influence was not significant to the results found.

6. Conclusions and Future Work

According to the authors' best knowledge, CLASS is the only available CALT that simultaneously allows RS recording, analysis and evaluation of users' acquired knowledge. The combination of the 3 usability methods allowed a more comprehensive and objective identification of advantages/disadvantages than the conventional single method commonly used and provided guidance for future developments. CLASS seemed friendly to use and therefore, may be integrated in students' activities for learning pulmonary auscultation skills. To enhance CLASS features, improvements should focus on *exercises* and *annotation of RS.* A new version of CLASS that also serves the needs of health professionals and researchers is being prepared.

Acknowledgements

The authors would like to thank to all participants of the evaluation session for their contributions, which will improve the usability of CLASS, and give special thanks to Professor Hans Pasterkamp for his helpful comments in the preparation of this article.

Funding

This work was funded by Fundação Calouste Gulbenkian, Portugal (Project: PGIS ID53 P-136395) and by Fundação para a Ciência e Tecnologia (FCT) (Project: 18/ID/2014). The work was also partially funded by Programa Operacional de Competitividade e Internacionalização - COMPETE, through Fundo Europeu de Desenvolvimento Regional - FEDER (POCI-01-0145-FEDER-016701), Fundação para a Ciência e Tecnologia (PTDC/DTPPIC/2284/2014) and under the project UID/BIM/04501/2013.

Conflict of interest statement

The authors certify that there is no conflict of interest with any financial organization regarding the material discuss in the manuscript.

7. References

1. Andrès E, Reichert S, Gass R, Brandt C (2009) A French national research project to the creation of an auscultation's school: The ASAP project. European journal of internal medicine 20:323-327

2. Bohadana A, Izbicki G, Kraman SS (2014) Fundamentals of lung auscultation. N Engl J Med 370:744-751

3. Branco P, Encarnação LMEd, Marcos AF (2006) It's all in the face: studies on monitoring users' experience.

4. Cohn JF, Ambadar Z, Ekman P (2007) Observer-based measurement of facial expression with the Facial Action Coding System. The handbook of emotion elicitation and assessment:203-221

5. Cotton D, Gresty K (2006) Reflecting on the think‐aloud method for evaluating e‐learning. British Journal of Educational Technology 37:45-54

6. Davis S, Wiedenbeck S (1998) The effect of interaction style and training method on end user learning of software packages. Interacting with Computers 11:147-172

7. Diah NM, Ismail M, Ahmad S, Dahari MKM Usability testing for educational computer game using observation method. In: Information Retrieval & Knowledge Management,(CAMP), 2010 International Conference on, 2010. IEEE, pp 157-161

8. Dinis J, Campos G, Rodrigues J, Marques A Respiratory Sound Annotation Software. In: HEALTHINF, 2012. pp 183-188

9. El Kaliouby R, Robinson P (2005) Real-time inference of complex mental states from facial expressions and head gestures. In: Real-time vision for human-computer interaction. Springer, pp 181-200

10. Gross V, Hadjileontiadis L, Penzel T, Koehler U, Ogelmeier C Multimedia database" Marburg Respiratory Sounds (MARS)". In: Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE, 2003. IEEE, pp 456-457

11. Hilbert DM, Redmiles DF (2000) Extracting usability information from user interface events. ACM Computing Surveys (CSUR) 32:384-421

12. Holzinger A (2005) Usability engineering methods for software developers. Communications of the ACM 48:71-74

13. Hou J, Chen Y-T, Hu L-C, Chuang C-C, Chiu Y-H, Tsai M-S Computer-aided auscultation learning system for nursing technique instruction. In: Engineering in

Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, 2008. IEEE, pp 1575-1578

14. ISO9241-11 (1998) Ergonomic requirements for office work with visual display terminals (VDTs). Part 11: Guidance on usability.

15. Karnath BM, Das Carlo M, Holden MD (2004) A comparison of faculty-led small group learning in combination with computer-based instruction versus computer-based instruction alone on identifying simulated pulmonary sounds. Teach Learn Med 16:23-27

16. Krall MA, Sittig DF Clinician's assessments of outpatient electronic medical record alert and reminder usability and usefulness requirements. In: Proceedings of the AMIA Symposium, 2002. American Medical Informatics Association, p 400

17. Kushniruk AW, Patel VL (2004) Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform 37:56-76

18. Kushniruk AW, Patel VL, Cimino JJ Usability testing in medical informatics: Cognitive approaches to evaluation of information systems and user interfaces. In: Proceedings of the AMIA annual fall symposium, 1997. American Medical Informatics Association, p 218

19. Mangione S, Dennis S CompuLung: a multimedia CBL on pulmonary auscultation. In: Proceedings of the Annual Symposium on Computer Application in Medical Care, 1992. American Medical Informatics Association, p 820

20. Mangione S, Nieman LZ, Gracely EJ (1992) Comparison of computer-based learning and seminar teaching of pulmonary auscultation to first-year medical students. Acad Med 67:S63-65

21. Noldus - The Observer XT. http://www.noldus.com/human-behavior-research/products/the-observer-xt. Accessed 18 May 2017

22. O'Brien HL, Toms EG (2010) The development and evaluation of a survey to measure user engagement. Journal of the American Society for Information Science and Technology 61:50-69

23. Oliveira A, Pinho C, Monteiro S, Marcos A, Marques A (2013) Usability testing of a respiratory interface using computer screen and facial expressions videos. Comput Biol Med 43:2205-2213

24. Pantic M, Rothkrantz LJ (2004) Facial action recognition for facial expression analysis from static face images. IEEE Trans Syst Man Cybern 34:1449-1461

25. Pinho C, Oliveira D, Oliveira A, Dinis J, Marques A (2012) LungSounds@ UA interface and multimedia database. Procedia Technology 5:803-811

26. Rozin P, Cohen AB (2003) High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of Americans. Emotion 3:68

27. Semedo J, Oliveira A, Machado A, Moreira J, Rodrgues J, Aparício J, Pasterkamp H, Jesus LMT, Marques A (2015) Computerised Lung Auscultation – Sound Software (CLASS). Paper presented at the Conference on Health and Social Care Information Systems and Technologies, Vilamoura,

28. Sestini P, Renzoni E, Rossi M, Beltrami V, Vagliasindi M (1995) Multimedia presentation of lung sounds as a learning aid for medical students. Eur Respir J 8:783-788

29. Sonderegger A, Sauer J (2009) The influence of laboratory set-up in usability tests: effects on user performance, subjective ratings and physiological measures. Ergonomics 52:1350-1361

30. Squires D, Preece J (1999) Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. Interacting with computers 11:467-483

31. Tamir D, Komogortsev OV, Mueller CJ An effort and time based measure of usability. In: Proceedings of the 6th international workshop on Software quality, 2008. ACM, pp 47-52

32. TipCam. http://tipcam.en.softonic.com/download. Accessed 12 June 2015

33. Vaismoradi M, Turunen H, Bondas T (2013) Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. Nurs Health Sci 15:398-405

34. Veneri D (2011) The role and effectiveness of computer-assisted learning in physical therapy education: a systematic review. Physiother Theory Pract 27:287-298

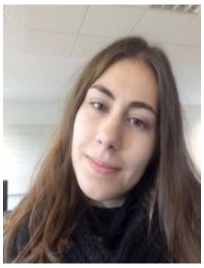
35. Ward JJ (2005) Rale lung sounds 3.1 professional edition. Respir Care 50:1385-1388

36. Ward JJ, Wattier BA (2011) Technology for enhancing chest auscultation in clinical simulation. Respir Care 56:834-845

37. Yeasin M, Bullot B, Sharma R (2006) Recognition of facial expressions and measurement of levels of interest from video. Multimedia, IEEE Transactions on 8:500-508

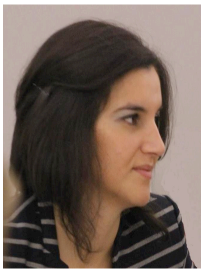
38. Zimmerman PH, Bolhuis JE, Willemsen A, Meyer ES, Noldus LP (2009) The Observer XT: A tool for the integration and synchronization of multimodal signals. Behavior research methods 41:731-735

Authors' biography



Ana Machado holds a degree in Physiotherapy (2016) and is currently taking a master degree in Cardiorespiratory Physiotherapy. She is a research fellow at University of Aveiro.

Ana Oliveira holds a degree in Physiotherapy (2011) and master in Physiotherapy (2014) from the University of Aveiro and currently is a PhD fellow at the University of Aveiro and University of Porto.



Cristina Jácome, PT, MSc has a PhD in Physiotherapy (2016). She is a Lecturer at School of Health Sciences of the University of Aveiro and a Researcher at the General Practice Research Unit in Tromsø.

Marco Pereira has a degree in New Communication Technologies and a master in Multimedia Communication from the University of Aveiro. He is a developer of software at University of Aveiro.

José Moreira holds a PhD in Computer Science and Networks from the ParisTech and the Faculty of Engineering, University of Porto. He is an Assistant Professor and Researcher at University of Aveiro.

João Rodrigues received is PhD degree in 2004 in Electrical Engineering from the University of Aveiro. He is currently an Assistant Professor and Researcher at University of Aveiro.

José Aparício holds a PhD in Pediatrics from the Faculty of Medicine of University of Porto (2002). He is currently working as pediatrician at Hospital Lusíadas, Porto.

Luis Jesus received his Ph.D. degree in Acoustic Phonetics in 2001 from the University of Southampton, UK. Since 2001, he has been a member of the academic staff at University de Aveiro in Portugal.

Alda Marques has a PhD in Physiotherapy (joint project between University of Aveiro and University of Southampton). She is currently a senior lecturer and a researcher at the University of Aveiro.