Users' preferences for pedagogical e-Content: A utility/usability survey on the Greek Illustrated Science Dictionary for School

Ioannis Lefkos, Maria Mitsiaki

University of Macedonia, Democritus University of Thrace

Abstract

This chapter draws on empirical data from a web questionnaire developed to evaluate the utility and usability of ELeFyS, an e-Content/online dictionary for the integrated approach of Language and Science within the plurilingual classroom. A 11-item questionnaire was developed and distributed to 84 informants (in-service & future teachers and primary school students) in school and university settings of Northern Greece. The data collected were both quantitative and qualitative: each quantifiable item scored on a 5-point Likert scale and it was followed by an open-ended question where the users justified their selected score. All 11 items were grouped within three (3) factors. Confirmatory factor analysis was used to determine the factor structure (construct validity). The instrument's reliability was examined by internal consistency. For the data analysis comparisons between groups were made by means of non-parametric statistics. The results confirm the model's goodness of fit and they reveal its high reliability. They also advocate the participants' positive view of ELeFyS. At the same time, they reveal the necessity to establish an e-dictionary/-Content culture and to adopt an expanded notion of scientific literacy in the Greek school practice.

Keywords: Science Dictionary, Scientific Literacy, Usability

1. Introduction

Web-based education has been around for quite some time, but gained much attention and massive participation worldwide during the last ten years (Kidd, 2010). Several months ago, the coronavirus lockdown gave even more impetus to e-learning practices and contents for younger learners either in the first (L1) or in the second/foreign language (L2), testing, thus, their efficacy and readiness. At the same time, numerous scholars and educators stress the urgent need to handle an avalanche of new digital skills embedded to a handful of new literacies (a.o. Cope & Kalantzis, 2000; Luzón et al., 2010; Henderson & Romeo, 2015), so that distance education and e-learning practices are not fragmentary but holistic (Tzifopoulos, 2020, Williamson et al., 2020), leading to learning approaches "not about the computer" but "with the computer" (Koutsogiannis, 2011).

Such multiliteracies are characterized by (a) *embeddedness*, since they are embodied into different discourses/texts, and (b) *interconnectedness*, since several combined literacy types are required for the redefined 21st century competences. The embedded learning perspective shifts the emphasis from the "narrow in-the-mind" vision to "a broader person-in-the-world" vision (Chee, 2007: 14) whereas the interconnected perspective attempts to get the most out of the newly emerging web-genres, allowing for more interdisciplinary practices.

Scientific literacy, as viewed in the recent study of the Committee on Culture and Education (Siarova et al., 2019), can be considered to be such a 'literacy-cluster' embodying several competences, skills, and literacies (Figure 1). The expanded notion of scientific literacy serves 3 Visions: *Vision I* places the emphasis *within Science*, *Vision II* views *Science in relation to society*, and *Vision III* involves learners *in Science within society* (Liu, 2013: 29).

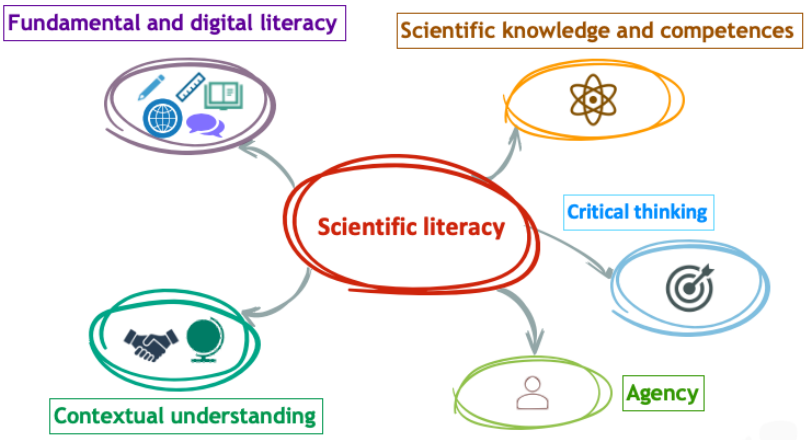

Fig. 1. Key-components of scientific literacy (adapted from Siarova et al., 2019)

Simply put, scientific literacy calls for:

- *basic* reading, writing, listening, speaking, math *skills*,
- *digital skills*, such as getting involved in critical learning offered via e-tools, in a way that learners master ideas and not keystrokes (Gilster, 1997),
- *hands-on skills* (experimentation),
- *critical information skills*, such as inquiry, analysis, synthesis, report, explanation, and argumentation (Osborne, 2002),
- *core scientific skills*, such as understanding the language and content of Science (Wellington & Osborne, 2001),
- *cultural skills*, such as contextualizing scientific concepts and phenomena (Plakitsi, 2010),
- *civic literacy skills*, such as public understanding of Science for active citizenship (Miller & Pardo, 2000), and

- *media literacy skills*, such as being able to assess the meaning of any kind of messages connected to the public understanding of Science and conveyed through media (Potter, 2001, Hobbs, 2011).

Under such a dynamic perspective, e-Contents are expected to keep pace with a more grounded approach; they have to be flexible, multifunctional, and tailored to the complex needs of the 21st century netizens, providing, thus, a challenging environment for learning and inspiring the development of multiple skills.

Among the different e-Contents delivered through the Internet, online dictionaries are tightly associated with the notion of pedagogy (cf. Chi, 2013: 165) in multiple ways. As resources of lexical knowledge, they can accompany any practice or situation that involves learners in autonomous/individual or collaborative productive and receptive tasks (a.o. Scholfield, 1982; Rundell, 1999) both in language and in discipline courses. The most widely acknowledged merit of online dictionaries is the easier and faster access they provide to lexical resources compared to their print predecessors (Nesi, 2013: 70). A further highly important advantage of e-dictionaries is their massive storage capacity compared to the print ones, allowing for a considerable bulk of information, i.e. lexicogrammatical, multilingual, visual, audio, etc. Their extended capacity makes it possible to exploit multiple modes other than text in its traditional sense, e.g. sound, illustrations, animated pictures/images, multimedia, usage boxes, etc. Moreover, online multifunctional dictionaries may cater for the differentiated needs of learners with different access points, i.e. basic/elementary, independent/intermediate, and proficient/advanced learners (Leech and Nesi, 1999: 296-303). Last but not least, online dictionaries are flexible, adaptive, and easy to update e-Contents (Lew & de Schryver, 2014: 345), with potential interactive extensions, especially when they allow users to contribute new lemmas or exchange ideas/comments. This is a really crucial asset that enables compilers to accommodate the users' needs.

For all the above reasons, online dictionary access has recently signaled the depreciation of the print dictionary, to the extent that many publishers discontinue printing and move entirely to the digital medium (Lew & de Schryver, 2014: 352). In the same line, it is no wonder that scientific works in lexicography are steadily replacing 'looking-up words' with 'searching for words' (de Schryver, 2012: 488, 492). However, "going on-line" cannot be considered as a dictionary merit by itself; in several cases on-line dictionaries are just the digital versions of their print counterparts (Lew & de Schryver, 2014: 352). Such a limited view underestimates the power and creativity both of lexicographic products and e-Contents in the digital school era.

The first print pedagogical dictionaries were introduced into the Greek educational system as official schoolbooks 15 years ago, and yet dictionary use is not a widespread practice within the school setting, especially for young learners. All school dictionaries have been uploaded on the Digital School Platform (ebooks.edu.gr), however in a static non-interactive form (only text and images). Not to mention that pedagogical specialized e-dictionaries with terminology are far from being embedded to the school practice.

To establish an online dictionary culture, new e-Contents/ dictionaries need to be compiled and to be consistent with the needs dictated by the digital age. However, real change has to be managed as to the content and form of future e-dictionaries, so that they move away from being viewed as "isolated islands of knowledge" (Robert Amsler, cited in Lew & de Schryver, 2014: 352). Amsler suggests that the future of e-dictionaries lies in the "new ways to display existing dictionary information and in connecting dictionary information to other knowledge". He continues that "It's a matter of either having lexical knowledge that nobody else has or displaying lexical knowledge in ways that are so convenient that other means of access are less attractive". At the same, it is timely to acquaint educators and students with the benefits of dictionary use within the multicultural classroom for all subjects and convince them to embrace it (Ranalli, 2013; Lew, 2013).

Taking into consideration the previous discussion on the need for flexible e-Contents that function as lexical knowledge resources and foster the multidimensional view of literacies, two years ago we proceeded with the compilation of the online *Greek Illustrated Science Dictionary for School* (ELeFyS, www.elefys.gr), in an attempt to provide young primary school learners with a tool that may help them in developing their academic and scientific literacy within an integrated approach of Language and Science Learning. From its initial conceptualization ELeFyS was intended both as a lexicographic product and a multifunctional e-Content, seeking to provide learners with stimuli relevant to all key-constituents of scientific literacy (Figure 1). In order to keep up with the evolving nature of scientific literacy -as of any type of literacy- and the differentiated learner needs in e-dictionary use, from the compilation of ELeFyS's alpha edition we sought to eavesdrop on users' perceptions, which we considered to be a prerequisite for the development of flexible and adaptive e-Contents.

This study is in line with feedback research on e-Content delivery, as it reports on a pilot web questionnaire survey with 84 participants, conducted to evaluate users' perceptions on the design features of ELeFyS. It can be considered to be a user-related study aspiring to shed light on the utility of ELeFyS's macro- and microstructure and its usability, by recording both qualitative and quantitative data. We developed a three-dimensional instrument based on the relevant theoretical models and despite the small size of the sample, we attempted a confirmatory factor analysis, in an effort to test the model's goodness of fit and predict any problems with the dataset. In the same vein, concerning reliability the internal consistency measured by Cronbach's alpha was tested for all the items and for each factor separately. For the quantifiable responses descriptive and non-parametric statistics are offered whereas for the qualitative aspects a critical discussion is made. The results indicate a good model fit, high internal consistency, and a high utility/usability rate for most of the features. Moreover, the qualitative comments reveal (a) a positive attitude towards the multifunctionality of ELeFyS, (b) adequate understanding of its constituents, and (c) interesting suggestions on future improvements. At the same time crucial questions are posed as to the establishment of an online dictionary and scientific literature culture in Greek schools.

The structure of the chapter is as follows: in section 2, we proceed with a brief description of ELeFyS. Section 3 provides the reader with some background on the concepts that are central to the analysis: dictionary and e-Content user-related studies and tools. The research methodology follows in section 4. Finally, the paper concludes with a discussion of the findings followed by some implications for further research.

2. ELeFyS: an e-Content for Science Education and Language Learning[1]

2.1. Scope & user profile

ELeFyS attempts to capture the broad conceptualization of future online dictionaries, since it has been compiled as an e-Content for the integrated development of scientific and linguistic literacy in the school context, grounding lexical knowledge in the school discipline of Science/Physics. To fulfil such an objective, generic entries include scientific terms that fall within the school subject of Physics and are likely to be encountered in the upper grades of primary and the lower grades of secondary school; however, the dictionary's coverage is not restricted to terminology, but is also expanded to the terms' respective general sense(s) and use(s).

In sum, ELeFyS constitutes a novel endeavor of combining pedagogy and specialization in order to meet the complex linguistic and cognitive/scientific needs of young school learners (native Greek or second/foreign language learners). It caters for several types of uses that target the school children's receptive and productive skills.

2.2. Innovation

ELeFyS innovates in several aspects (cf. Mitsiaki & Lefkos, 2018), as it is:

- the first *Greek specialized Science dictionary* for school that fosters content-based language learning, thus promoting reception and production both of scientific terms and their respective everyday use, e.g. *ενέργεια* 'energy' in Science, but also 'energy' in general vocabulary,
- a *pedagogical dictionary* intended to cover the specific cognitive, cultural, linguistic, and encyclopedic needs of primary and secondary school students (11-14 years old),
- a monolingual dictionary with *multilingual lexical information*, since it establishes interlingual equivalence of scientific terms in 5 languages: English, Standard Arabic, Russian, Turkish and Chinese, thus being a useful reference tool for L2 learning,

[1] For a comprehensive analytical description of ELeFys from a lexicographic point of view, see Mitsiaki & Lefkos (2018).

- an *illustrated dictionary*, as it provides visual tools (images, animations, etc.),
- an *online dictionary* freely accessible on the Internet that circumvents the common dictionary conventions in terms of space limitation and makes imaginative use of new technologies in order to ensure flexibility, user-friendliness, and a pedagogy-oriented format, and
- an *e-Content* with multimedia that can function complementarily to the schoolbooks or other educational resources.

2.3. Macrostructure & Microstructure

To estimate the dictionary's coverage, we were based on school textbook corpora and equivalent pedagogical Science dictionaries. We opted for a systemic presentation of the entries, by arranging the concepts according to their semantic interconnectedness. Such an arrangement led to a grouping of terms into their hypernym concepts of Science, i.e. *Heat, Electricity*, etc., which is also in accordance with the taxonomy portrayed in the Greek Science textbooks. Up to now the beta edition contains 200 multi-lemmas.[2]

Each dictionary page corresponds to a distinct lemma (Figure 2). Navigation through the dictionary is facilitated by hyperlinks to other layers of information and navigation buttons. Moreover, a user-friendly search function is provided. Finally, there is an accessibility widget overlay providing aids for users with physical, visual or hearing disabilities, like text read aloud, text size, color contrast control, etc.

The pedagogical role of ELeFyS is ensured by the use of lexicographic symbols instead of metalanguage. The main lemma consists of sub-lemmas organized in nests. Equal weight is given to all dictionary-relevant features, such as collocational properties, word families, relationships of synonymy and hyponymy, contextual preferences, grammar, register, and etymology, to help learners replace the apparent linguistic randomness with systematicity. To assist L2 learners, recorded pronunciation files are stored for each lemma.

The definitions of scientific terms are promoted to appear at the top left side of the entry, and they are followed by the corresponding definitions of general vocabulary. Besides conventional defining formulae, contextual defining formats are used, such as full-sentence definitions, embedded in a rich microstructure. Scientific definitions are of graduated difficulty, following a ranking from the simplest (suggested for a primary observation/understanding of the phenomenon) to the most complex (leading to academic wording).

A broad spectrum of examples for every lemma is offered at the right side of the page, so that its syntactic and collocational behavior is fully illustrated. Both authentic and lexicographer-made examples are used, in order to reveal the words' patterning. Each sense and use is accompanied by illustrations, selected by specific

2 The compilation of an online dictionary is a dynamic process; thus, more lemmas are about to be added in the future.

criteria, such as the target-group's age and their cultural background as well as the type of licensing (CC-BY).

Lastly, for every single lemma thought-provoking encyclopedic, critical, experimental, and cultural stimuli are provided:

- suggestions for experimentation that enhance critical thinking or/and intercultural sensitivity,
- hyperlinks to (a) Wikipedia, for a deeper understanding of physical phenomena and their history, (b) videos in YouTube, (c) the Digital Educational Resources from *Photodentro (the Greek National Aggregator of Educational Content)*, (d) multimedia available at *Noesis* (Thessaloniki Science Center and Technology Museum), and
- suggestions for dictionary use that involve students in reading, listening, speaking, and writing tasks both in Science and in Language courses.

Fig. 2. ELeFyS page, lemma *βρασμός* 'boiling'

3. Dictionary/e-Content use and user-related studies

Dictionary use is lately experiencing an upsurge of interest, especially when it comes for online dictionaries delivered for specific user groups in specific regions (a.o. Nesi 2013; Wingate 2004; Lew & Galas 2008; Welker, 2010; Gavriilidou 2013). Such research is emanating from the need to gain feedback both for the dictionaries' ease of consultation, usefulness or functional quality and for the identification of users' needs, preferences and dictionary reference skills. To that end, different research methods are being employed, originating from either positivistic or naturalistic approaches (Cohen et al. 2007), such as questionnaire surveys (a.o. Chatzidimou, 2007; Gavriilidou, 2013), interviews (East, 2008), log files (Hult, 2012), eye-tracking (Tono, 2011), etc.; in some cases mixed or triangulated methods are used.[3]

Questionnaire surveys are still the most widely-used method, despite the criticism they receive as to their reliability, accuracy and the fact that users and compilers do not always share the same language (Lew, 2002, Nesi, 2013; Chi, 2013). Several objections are also raised to the sample size and nature, i.e. usually small and convenient samples. However, questionnaire studies for e-dictionary evaluation or use seem to face the same dilemmas as in all fields of empirical studies. What can change the disposition towards such surveys is a more careful implementation, so that reliability and validity is ensured during the design and data collection process.

The present study falls within the scope of research evaluating the utility and usability of online dictionaries (cf. Ball & Bothma, 2018). Swanepoel (2001: 167) relates dictionary quality and dictionary design, in a way that the evaluation of functional quality is not a detached or cut-off process, but "it goes hand in hand with the design process", working like a "thermostat", and thus revealing the modifications that can be made. However, utility/usability assessment of online dictionaries is not a 'one-size-fits-all' process. In his extensive review of literature concerning dictionary evaluation, Swanepoel (2008) makes it clear that the evaluation criteria may vary in scope, as they range from covering all dictionary types to being dictionary genre-specific, aiming at all the design features of a dictionary or focusing on only one specific feature. In a more recent study, Ball & Bothma (2018) identify 7 evaluation criteria for e-dictionaries: *content, information structure, navigation, access (searching and browsing), help, customization, and innovative technologies used to manage information*.

Since ELEFYS combines the features of both an online dictionary and a wide-scope e-Content, we should make reference to usability evaluation as a general concept that embraces different kinds of digital applications. Usability is a concept mainly derived from Information Science (Heid & Zimmermann, 2012) and its evaluation can be formative (Sauro & Lewis, 2012, p.10), when it aims to reveal users' perceptions about the under-development material.

3 For a comprehensive approach, see Nesi (2013), Lew (2013) and Lew & de Schryver (2014).

Several types of standardized usability tests are available and despite their different approaches, they all engage users in a task or scenario with the under-test material and then record their subjective opinions (Sauro & Lewis, 2012, p.186). One of the most widely used instruments seems to be the System Usability Scale – SUS (Brooke, 1996), which comprises 10 items; although it was initially assumed as a unidimensional tool, it was later found (Lewis & Sauro, 2009) that it actually has two factors: (1) *Usable* (8 questions) and (2) *Learnable* (2 questions). Quite similarly, one of the most highly appreciated and reliable (Revythi & Tselios, 2019) tools is the Technology Acceptance Model – TAM questionnaire, proposed by Davis (1989) and based on the idea that a user's intention to use a product is primarily affected by two factors: (1) *Perceived Usefulness* (6 items) and (2) *Perceived Ease of use* (6 items). However, most of the generic usability tests omit important information specific to an interface type (Sauro, 2015), hence falling out of the scope of our investigation.

Moreover, distinct criteria have also been identified for the digital learning material according to which usefulness is a two-value concept embracing (1) *pedagogical usability*: the extent to which the "functions of a system correspond with the needs of the users", and (2) *technical usability*: "how well the users are able to use the functions offered by the system" (Nokelainen, 2006: 180). More recently, Papadakis et al. (2020) proposed an evaluation tool for educational applications with 13 items in 4 factors, i.e. Usability, Efficiency, Parental Control, and Security.

Drawing insight from all fields of user-related studies on online dictionaries, educational e-Content and general digital applications, we follow Swanepoel's (2008) argumentation that software usability could be correlated to the functional approach methodology for the design and evaluation of dictionaries, since they are both user-oriented, focusing their evaluation on users (while performing certain actions in the context of using a product and rating it on a functionality/usability scale).

The questionnaire survey carried out in this study is a means of formative assessment of ELeFyS. Since ELeFyS is still being compiled with more lemmas being added and decisions on content, structure, and layout being made, we conducted this pilot study in order to find out to what extent it meets the needs of its target users.

4. Methods

4.1. Research questions

The objectives of the current research are three-fold: (1) to provide some preliminary insight into the users' perceived usefulness of ELeFyS's features and the usability of its functions, (2) to develop an instrument for the assessment of ELeFyS's usefulness and usability based on previous relevant dictionary and e-

Content usability tools, and (3) to seek evidence for the instrument's dimensional structure and internal consistency.

We opted for a small-scale pilot study with a restricted sample size, so as to draw valuable feedback at an early stage and to improve the quality and efficiency of a future long-term and triangulated study on ELeFyS' effectiveness to enhance the end-user's dictionary/ scientific literacy/ e-Content skills.

The web questionnaire (https://forms.gle/dXRqkVxuQp9Zc7Sd7) was administered to 84 informants in 3 subgroups: in-service teachers, future teachers, and primary school learners from schools and universities in Northern Greece (September 2018, March 2020), a sample size considered to be adequate for a pilot study.

The following research questions were shaped:

RQ1: Is the designed tool valid and reliable? (dimensional structure, internal consistency)

RQ2: To what extent do ELeFyS's users find its constituents useful, its layout attractive and its functions easy to use?

RQ3: Does the perceived usefulness and usability of ELeFyS vary significantly between the 3 subgroups in respect to their different characteristics or roles?

4.2. Instrument

Since no other questionnaire is available, to our knowledge, for a multifunctional Language and Science lexicographic e-Content, we had to construct a new instrument, adopting, though, the generic dimensions and wording of relevant validated e-Content and lexicographic tools (see Section 3 of this Chapter). Thus, both the dimensional structure and the item specification emerge from the aforementioned theoretical constructs on user-related studies on e-dictionaries / (educational) e-Contents and multiliteracies.

First, the construct's dimensional structure draws from the usability criteria for digital learning material, i.e. *pedagogical usability* and *technical usability*. Second, the underlying factors postulated for the instrument are heavily dependent on the conceptual model of scientific literacy (as portrayed in Fig. 1); such an expanded notion of scientific literacy embraces (a) *academic (scientific) and fundamental (communicative) competences*, (b) *stimuli for linguistic and scientific engagement, contextual understanding, critical thinking, and learner agency*, and (c) *digital competences*. Third, the pre-conceptualized factor structure is inspired by the current need to develop and view digital dictionaries as novel pools of interdisciplinary knowledge and skills taking into consideration the following criteria: *content, information structure, access, and navigation*.

In an attempt to bridge the gap between theory and observation, we designed an 11-item web questionnaire. Drawing from the previous discussion, we grouped the 11 items within the following 3 factors (F1, F2, F3):

(F1) *academic and communicative lexical information* (5 items)

(F2) *stimuli for further linguistic and scientific engagement* (4 items),

(F3) *technical usability: Ease of navigation and attractiveness* (2 items).

Both the first and the second factor fall within the *pedagogical usability* criterion and they reflect the dictionary's *content* and *information structure*. However, we considered them to be 2 distinct factors, as the second one reflects the novelties not found in other specialized pedagogical dictionaries, offering users food for thought and engagement in interdisciplinary tasks. The third factor falls within the *technical usability criterion* and measures ease of *access/navigation* and attractiveness. Of course, the conceptualized 3-factor model deviates from the 2-factor models found in generic application studies; however, the inherent multiple functionalities of ELeFyS and the specificity of its potential users favor our decision.

Table 1. The instrument's quantifiable items

Questions	Targeted ELEFYS's features
F1 *(Academic and communicative lexical information)*:	
To what extent do you find useful…	
Q1: the multiple scientific definitions?	graded scientific definitions
Q2: the additional everyday word definitions?	everyday word definitions
Q3: the scientific terms' examples?	examples for scientific terms
Q4: the everyday words' examples?	the terms' respective use in everyday language
Q5: the pictorial illustrations?[4]	pictorial & animated illustrations
F2 *(Stimuli for further linguistic and scientific engagement)*:	
To what extent do you find useful…	
Q6: the etymological information notes?	etymology boxes
Q7: the grammatical information notes?	grammar boxes
Q8: the scientific terms' equivalents in other languages?	interlingual equivalence tables
Q9: the encyclopedic & critical thinking notes?	encyclopedic, experimentation, critical notes
F3 (*Technical usability*)**:** *How would you rate…*	
Q10: the overall presentation and graphical interface?	attractiveness of layout
Q11: the search & navigation (was it easy to find what you were looking for)?	ease of navigation

Each item (Q1-11) comprises 2 interrelated questions, a quantifiable one (1a-11a, Table 1), and a complementary one eliciting open-ended responses that justify

[4] Pictorial illustrations are considered to be a crucial aid to lexical meaning and exemplification.

the users' score (1b-11b). The items are followed by a final section with recommendations for improvement.

Each quantifiable item (Q1-11) scored on a 5-point Likert scale anchored with 1=Not at all useful, and 5=Extremely useful. The qualitative judgments are not systematically analyzed in this report, but they are mentioned sporadically as an aid in the interpretation of the quantitative data.

4.3. Participants & data collection

The sample comprises 59 educators (37 school teachers and 22 final-year undergraduate students/future teachers engaged in teaching practices) and 25 primary school students (convenience sample, Table 2). We made an effort to vary our sample and reach beyond the common in research university-student sample, by expanding our research to teachers and primary school learners. This was a conscious decision on our part for several reasons: learners are the end-users of the e-Content, in-service teachers are the ones to introduce novel material within the school classroom, and undergraduate final year students of Teacher Education University Departments are the ones who are mostly acquainted with innovative interdisciplinary approaches in teaching and learning.

Fourteen participants were excluded from the analyses, as their responses were only partially filled or their comments seemed to be out of context. Written consent to participate was obtained from all participants (adult teachers and the young learners' parents).

Table 2. Survey participants by role (*n*=84)

Participant role	*n*	%
In-service Teachers	37	44
Future Teachers	22	26
Primary School Students	25	30
Total	84	100

All the participants filled in the 11-item questionnaire administered after their involvement in (a) training seminars on scientific literacy (in-service & future teachers) or (b) exposure to Content (Physics) and Language Instruction (students). Before completing the questionnaire, all subjects were familiarized with ELeFyS's features during a 1-hour guided (in-person & distance) session of browsing. In this way, we attempted to acquaint users with ELeFyS and eliminate the possibility that responses are influenced by unclear wording or other inconsistencies.

4.4. Data analysis & Results

4.4.1 Content validity

Content validity was established after consulting external expert reviewers, both from the field of Lexicography and from the field of Science Education.

4.4.2 Construct validity

A confirmatory factor analysis using AMOS 26.0 was run, in order to test the fit of the 3-factor model. Six indices were used to evaluate the model's goodness of fit: chi-square, chi-square/*df* ratio, *p*-value, Goodness-of-Fit Index (GFI), Comparative Fit Index (CFI), and Root-Mean-Square Error of Approximation (RMSEA). As displayed in Table 3, a non-significant *p*-value is obtained (.152), indicating an acceptable fit (Byrne, 2001). GFI and CFI are above 0.9 or close to 1 and the RMSEA value is less than 0.06, which is further evidence for a good model fit (Schumacker & Lomax, 2012, Hu & Bentler, 1999). The previous results insure the validity of our construct.

Table 3. Goodness-of-Fit indices of the 3-factor model

χ^2	df	χ^2/df	p	GFI	CFI	RMSEA
50.291	41	1.22	0.152	0.902	0.97	0.05

4.4.3. Internal consistency and reliability

To check the instrument's reliability, we calculated: (a) the Cronbach's alpha coefficients, and (b) the correlation between the individual items and the total score for all items (Table 4). High internal consistency is ensured for all factors except for the third one (technical usability) which is at the cut-off point (0.689), a finding to be further discussed.

Table 4. Indices of Internal Consistency of the administered questionnaire

	Total	F1	F2	F3
Correlation with total score	1	0.908	0.905	0.663
Alpha coefficient	0.896	0.842	0.842	0.689

4.4.4. Descriptive statistics and between-group differences

Table 5 displays descriptive statistics (means and standard deviations) for each factor and for each group of informants.

Table 5. Means (*M*) and standard deviations (SD) per group (*n*=84) and factor (F1-F3)

	M (SD)		
Factors	In-service teachers (n=37)	Future teachers (n=22)	School students (n=25)
F1: Academic and communicative lexical information	4.44 (0.48)	4.55 (0.50)	4.18 (1.03)
F2: Stimuli for further linguistic and scientific engagement	3.99 (0.62)	4.00 (0.87)	4.33 (0.98)
F3: Technical usability	3.84 (0.74)	4.22 (0.65)	4.22 (1.02)

Notably, the overall picture does not exhibit intense variation. In almost all cases, the mean values are greater than 4.00. This can be interpreted as a very positive user perspective, but at the same time it can be indicative of a ceiling effect, since accumulated percentage of scores 1, 2 & 3 is between 10-30%, thus leaving a spacious 70-90% for scores 4 & 5. Despite the well-acknowledged impact of a ceiling effect on reliability, the bunching of scores at the upper level could be acceptable for an instrument that assesses a novel e-Content which integrates dictionary, encyclopedia, school textbook, multimedia and other functions, and might differ from the conventional unifunctional print school material the users are used to.

As displayed in Fig. 3, the differences in scores between the three groups are significant at $p<.05$ for two out of three factors.

More specifically, all three groups seem to share similar perceptions on the usefulness of F1. It is worth mentioning that the primary school learners exhibit a lower mean (4.18) compared to both the teacher groups (around 4.5) and more divergent opinions as denoted by the high *SD* value (1.03). For F2 the in-service and future teachers' preferences seem to converge again (M=3.99, M=4.00); however, the primary school students seem to be more enthusiastic (M=4.33), despite their high divergence of opinion (SD=0.98). A quite reverse finding was obtained for F3; this time the future teachers and the primary school students seem to agree on the usability of ELeFyS (M=4.22), whereas in-service teachers are more reluctant (M=3.84).

As already mentioned, an interesting pattern observed in Table 4 concerns the standard deviation values. A closer look at the data, shows a tendency for increasingly diverging opinions as we move from the In-service Teachers to Future Teachers and finally to Primary Students. This finding might be correlated with the fact that In-service Teachers are greatly influenced by their teaching experience, while Future Teachers are just beginning to get involved in teaching practices. Another influential factor might also be the obvious age difference between these three sample groups.

We proceeded with statistical analyses using SPSS software (ver. 25.0) to examine whether the recorded preferences of the three surveyed samples display significant differences. The results of the one-way Kolmogorov-Smirnov Test for each participant group suggested that the item scores differ from the normal distribution (p<0.05).

Since our data were not normally distributed, we adopted the non-parametric independent samples Kruskal Wallis test to check for between-group significant differences in Likert scale scores.

As displayed in Fig. 3, the differences in scores between the three groups are significant at p<.05 for two out of three factors.

More specifically, the differences in scores for F1 are not significant. The between-group differences appear to be significant for both F2 (p=.036) and F3 (p=.043).

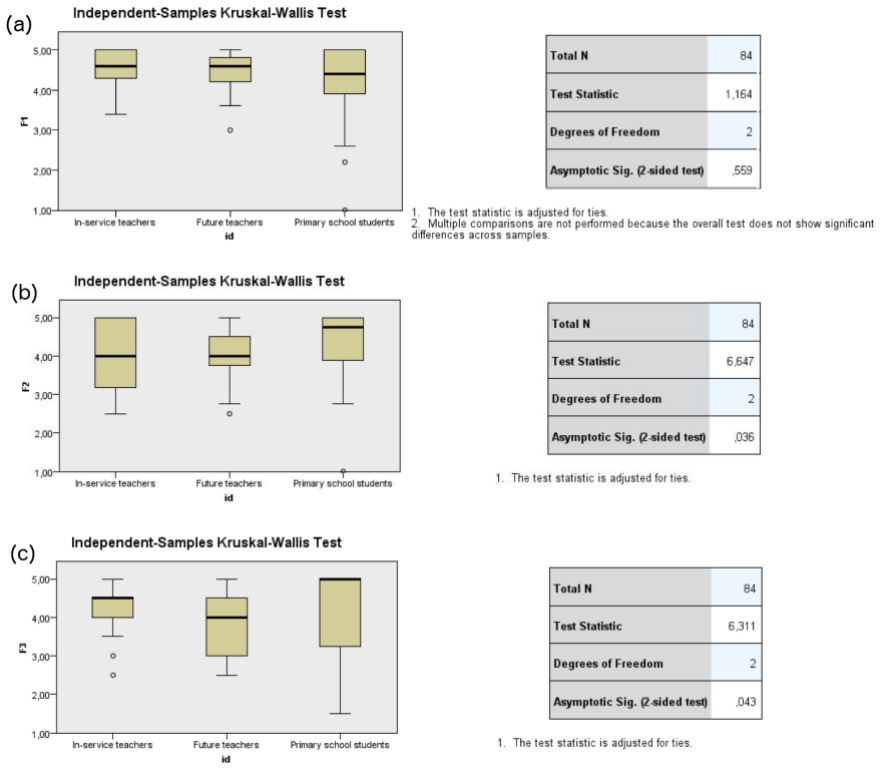

Fig. 3. Independent-samples Kruskal-Wallis test displaying between-group differences by factor (a-F1, b-F2, c-F3)

4.4.5. Qualitative responses

The data collected from the survey's open-ended questions were subjected to qualitative content analysis, in order to identify response clusters and trends. Although the qualitative data analysis is not part of this chapter, mentioning some of the participants' judgements will provide support for the quantitative data presented earlier.

For example, the design feature investigated in Q3a (examples for the scientific definitions, F1) was highly appreciated from all participants. In Q3b, the informants had to comment on the perceived purpose of this feature. Most of the comments exhibit a very positive to almost enthusiastic implied acceptance, e.g. "the scientific definitions are crucial for constructing the meaning", "they are fostering the ability to incorporate scientific terms in written and oral expressions", "they can help in retaining related meanings for much longer", "they can help us (students) to understand the scientific terms".

Another feature worth mentioning is the perceived purpose of the pictorial illustrations, investigated in Q9a (F1). Some typical comments from Q9b were "illustrations create a friendly environment and trigger the interest", "illustrations can be much helpful for the visual type learners", "illustrations can be an aid for the understanding of the concepts".

All the aforementioned comments are more or less being repeated in all three sample groups. Even students express similar views in their own wording.

Finally, on the additional free comment section participants express their overall positive attitude (consistent with the overall picture of the qualitative data). Comments like "keep it up this way!", or "thank you for your effort" were the most common ones. Some informants would make suggestions like "I would like more lemmas to be included", or "it would be nice to add the pronunciation function in all presented languages".

5. Discussion

Despite its small-scale pilot nature, this study (a) reveals several interesting considerations on the users' perceived utility and usability of ELeFyS, and (b) provides valuable feedback for the development of the ELeFyS questionnaire but also for instruments that assess the usefulness and usability of multifunctional lexicographic e-Content.

As far as the development of the ELeFyS questionnaire is concerned, the procedure was not free from restrictions. In the first place we ran an exploratory factor analysis using principal axis factoring with oblique rotation to investigate the factor structure and see how the variables relate and group based on inter-variable correlations. The analysis revealed that the instrument has a two-factor structure grouping together Q1 to Q9 and Q10 to Q11. Unfortunately, in this way F1 explained about 50% of the total variance (61%). This pitfall is also reflected in the lower Cronbach's a coefficient for the factor of technical usability (0.689). However, as the sample size of this initial pilot study is small, we decided to follow our theoretical rationale, thus to postulate three correlated factors. Moreover, we are

aware that our sample is marginally sufficient for a confirmatory factor analysis and an investigation of the tool's internal consistency. Of course, since the completion of the questionnaire follows engagement in tasks, seminars, and practice with ELeFyS, a much bigger sample could be a difficult endeavor. In any case, although the analyses revealed a good fit of the model and adequate to high reliability, these results point at a slight revision of the instrument (more specific items for F3), a subsequent revalidation and implementation (bigger sample).

Moving to the analysis of the users' scores and preferences, the overall picture is striking in their positive stance towards ELeFyS, as proved by the high means for all factors and sample groups. To go in more depth, we should mention that all groups seem to reward both the theoretical and the practical / more experiential constituents of ELeFyS (definitions, examples and pictorial illustrations, F1). On the other hand, both In-Service and Future Teacher groups appear to appreciate the linguistic aids of ELeFyS (usage boxes for etymology, grammar, F2) less than young learners, as if language may not interfere with Science Education. Such a finding might be indicative of a more dissociated cut-off approach of language and discipline courses in the Greek school, despite the voices of scholars who argue for an integrated approach of language and content. At the same time, this finding can be interpreted in terms of the dictionary referential skills of both (In-service / Future) teachers and learners. Therefore, it seems that either the teachers view ELeFyS as an e-Content for Physics, where language takes less space, or they are less aware of the benefits that a dictionary's constituents can offer to the learners. This takes us undoubtedly to the necessity to establish an e-dictionary use culture in Greek schools.

We should also comment on the more conservative view put forward by teachers of an e-Content that functions also as a lexical knowledge resource. A striking finding is that young learners reward the existence of interlingual equivalents and critical stimuli (F2), whereas In-service / Future Teachers seem to be rather reluctant in acknowledging their utility. A disappointing admission to be made is that e-Content with multilingual references is not a common practice in the Greek mainstream classroom yet. Maybe it is the case that teachers see no point in offering multilingual scaffolding for emergent bilinguals, when they are far from understanding and producing more complex academic/scientific language. At the same time, they appear to be less informed on or less convinced of the utility of stimuli for the students' critical, experimental, encyclopedic, and cultural, engagement. It is possible, though, that teachers view such an extension as distractive from the content-oriented curriculum that is heading to the acquisition of scientific concepts and phenomena. On the contrary, the students that participated in the research appreciate the most these features of ELeFyS, possibly revealing their need for a plurilingually and critically oriented school reality.

Finally, In-service teachers appear to be less enthusiastic on the technical usability of ELeFyS, a finding possibly attributable to many reasons. Luckily enough, the teachers' open-ended responses are explicit in expressing their needs for a more systematic training in the use of e-Contents.

The qualitative data gathered for this web questionnaire survey shed more light on the quantitative analysis. The participants seem to acknowledge in their open responses the benefits of scientific definition gradedness, despite the fact that they do not always assess it as an extremely useful feature (learners). They also seem to understand the advantage of the parallel provision of everyday meanings and use. Moreover, they are expressed very positively on ELeFyS's layout/presentation and navigation features, commenting on its user-friendliness and attractiveness (especially future teachers and learners). In many cases they contribute comments such as "Keep it as it is" or "There is nothing to modify". Even in the case of less popular features (i.e. tables of interlingual equivalents), they do not suggest their wiping out of the dictionary, but they acknowledge they could function effectively in specific situations (teachers).

In sum, the feedback we gained through the current pilot survey is a positive one. At the same time, we are obliged to reflect on the features of ELeFyS that appear to be less perceived as useful. This is a quite complex process for several reasons. The most crucial of them is put forward by Lew (2011: 9-10): the evaluation of a given feature does not assess obligatorily "its inherent fitness of purpose", but it also reveals the extent to which "the users are habituated" to exploiting such a feature. However, if they are not well-acquainted or habituated to it, it is possible that they will not be positively affected by the novelty. Such a statement can be verified, if we take into consideration that both the In-service / Future Teachers and the students of the current study have been partially engaged in training seminars and content and language-oriented courses, which means that they are not still habituated to these novelties.

Hence, the contribution of this research to the field of e-Content evaluation is that it suggests a generalizable design format for specialized pedagogical dictionaries and their usability tools, one that favors the 3 aforementioned factors.

Moreover, these findings that are undoubtedly treated with caution as they arise from a small-scale research provide insights in the future research for ELeFyS. This has to be a longitudinal research that combines the investigation of the informants' scientific literacy and dictionary reference skills and their training/habituating practices. As soon as the reliability of F3 (technical usability) is fixed by adding more specific questions, the questionnaire is planned to be revalidated and correlated with measures of task effectiveness and efficiency. It goes without saying that in such a research several methodological tools are to be exploited apart from questionnaire surveys. Thus, future research has to be informed by the limitations of the current study, i.e. the small convenient sample, the less elaborate factor on technical usability, and the fact that users' expertise in dictionary use and scientific literacy was not surveyed.

6. References

Ball, L. H., & Bothma, T. J. (2018). Establishing evaluation criteria for E-dictionariEs. *Library Hi Tech*, *36*(1), 152-166. https://doi.org/10.1108/lht-02-2017-0031

Brook, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189-194). Taylor & Francis.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Psychology Press.

Chatzidimou, K. 2007. Dictionary use in Greek education: an attempt to track the field through three empirical surveys, *Horizontes de Lingüística Aplicada* (Neste número: O USO DE DICIONÁRIOS), 6(2), 91-103.

Chee, Y. S. (2007). Embodiment, embeddedness, and experience: Game-based learning and the construction of identity. *Research and Practice in Technology Enhanced Learning*, *02*(01), 3-30. https://doi.org/10.1142/s1793206807000282

Chi, M. L. (2013). Researching pedagogical lexicography. In H. Jackson (Ed.), *The Bloomsbury companion to lexicography* (pp. 165-187). Bloomsbury.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge.

Cope, B., & Kalantzis, M. (2000). *Multiliteracies: Literacy learning and the design of social futures* (1st ed.). Routledge.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340. https://doi.org/10.2307/249008

de Schryver, G. (2012). Trends in twenty-five years of academic lexicography. *International Journal of Lexicography*, *25*(4), 464-506. https://doi.org/10.1093/ijl/ecs030

East, M. (2008). *Dictionary use in foreign language writing exams: Impact and implications*. John Benjamins Publishing.

Gavriilidou, Z. (2013). Development and validation of the strategy inventory for dictionary use (S.I.D.U.). *International Journal of Lexicography*, *26*(2), 135-153. https://doi.org/10.1093/ijl/ect007

Gilster, P. (1998). *Digital literacy*. Wiley.

Gouws, R. H., & Tarp, S. (2017). Information overload and data overload in lexicography. *International Journal of Lexicography*, *30*(4), 389-415.

Heid, U., & Zimmermann, J. T. (2012). *Usability testing as a tool for e-dictionary design: collocations as a case in point*. In R. Vatvedt Fjeld (Ed.), *Proceedings of the 15th EURALEX international Congress 2012* (pp. 661-671). University of Oslo.

Henderson, M., & Romeo, G. (2015). *Teaching and digital technologies: Big issues and critical questions*. Cambridge University Press.

Hobbs, R. (2011). *Digital and media literacy: Connecting culture and classroom*. Corwin Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hult, A. (2012). *Old and New User Study Methods Combined – Linking Web Questionnaires with Log Files from the Swedish Lexin Dictionary*. In R. Vatvedt Fjeld (Ed.), *Proceedings of the 15th EURALEX international Congress 2012* (pp. 922-928). University of Oslo.

Kidd, T. T. (2010). A brief history of eLearning. *Web-Based Education*, 1-8. https://doi.org/10.4018/978-1-61520-963-7.ch001

Koutsogiannis, D. (2011). *Εφηβικές πρακτικές ψηφιακού γραμματισμού και ταυτότητες [Adolescents' Digital literacy practices and identities]*. Greek Language Center.

Leech, G., & Nesi, H. (1999). Moving towards perfection: the learners' (electronic) dictionary of the future. In T. Herbst & K. Popp (Eds.), *The perfect learners' dictionary (?) (Lexicographica Series Maior 95.)* (pp. 295-306). De Gruyter.

Lew, R. (2002). *Questionnaires in dictionary use research: A re-examination*. In A. Braasch & C. Povlsen (Eds.), *Proceedings of the 10th EURALEX international congress (Vol.1)* (p. 267–271). Center for Sprogteknologi, Copenhagen University.

Lew, R. (2011). *User studies: Opportunities and Limitations*. In K. Akasu & S. Uchida (Eds.), *Proceedings of ASIALEX2011* (pp. 7-16). Asian Association for Lexicography.

Lew, R. (2013). *Online dictionary skills*. In I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets, & M. Tuulik (Eds.), *Proceedings of ELex 2013 conference: Electronic lexicography in the 21st century: Thinking outside the paper* (pp. 16-31). Trojina, Institute for Applied Slovene Studies& Eesti Keele Instituut.

Lew, R., & De Schryver, G. (2014). Dictionary users in the digital revolution. *International Journal of Lexicography*, *27*(4), 341-359. https://doi.org/10.1093/ijl/ecu011

Lew, R., & Galas, K. (2008). *Can dictionary skills be taught? The effectiveness of lexicographic training for primary-school-level Polish learners of English*. In E. Bernal & J. DeCesaris (Eds.), *Proceedings of the XIII Euralex international Congress 2008* (pp. 1273-1285). Universitat Pompeu Fabra.

Lewis, J. R., & Sauro, J. (2009). The Factor Structure of the System Usability Scale. In M. Kurosu (Ed.), *Human Centered Design. HCD 2009. Lecture Notes in Computer Science, vol 5619* (pp. 94-103). Springer. https://doi.org/10.1007/978-3-642-02806-9_12.

Liu, X. (2013). Expanding Notions of Scientific Literacy: A Reconceptualization of Aims of Science Education in the Knowledge Society. In N. Mansour & R. Wegerif (Eds.), *Science education for diversity: Theory and practice* (pp. 23-39). Springer. https://doi.org/10.1007/978-94-007-4563-6_2.

Luzón, M. J., Ruiz-Madrid, M. N., & Villanueva, M. L. (Eds.). (2010). *Digital genres, new literacies and autonomy in language learning*. Cambridge Scholars Publishing.

Miller, J. D., & Pardo, R. (2000). Civic Scientific Literacy and Attitude to Science and Technology: a Comparative Analysis of the European Union, the United States, Japan, and Canada. In M. Dierkes & C. V. Grote (Eds.), *Between understanding and trust: The public, science and technology* (p. 81–129). Routledge.

Mitsiaki, M., & Lefkos, I. (2018). *ELeFyS: A Greek Illustrated Science Dictionary for School*. In J. Čibej, V. Gorjanc, I. Kosem, & S. Krek (Eds.), *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts* (pp. 373-385). Ljubljana University Press, Faculty of Arts.

Nesi, H. (2013). Researching users and uses of dictionaries. In H. Jackson (Ed.), *The Bloomsbury companion to lexicography* (pp. 62-74). Bloomsbury.

Nokelainen, P. (2006). An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Journal of Educational Technology & Society*, *9*(2), 178-197. http://www.jstor.org/stable/jeductechsoci.9.2.178

Osborne, J. F. (2002). Science Without Literacy: a ship without a sail. *Cambridge Journal of Education*, *32*(2), 203–218. https://doi.org/10.1080/03057640220147559

Papadakis, S., Vaiopoulou, J., Kalogiannakis, M., & Stamovlasis, D. (2020). Developing and exploring an evaluation tool for educational apps (E.T.E.A.) targeting kindergarten children. *Sustainability, 12(10), 4201*, *12*(10), 4201. https://doi.org/10.3390/su12104201

Plakitsi, K. (2010). Collective curriculum design as a tool for rethinking scientific literacy. *Cultural Studies of Science Education*, *5*(3), 577-590. https://doi.org/10.1007/s11422-010-9288-0

Potter, J. W. (2001). *Media literacy*. SAGE.

Ranalli, J. (2013). Online strategy instruction of integrated dictionary skills and language awareness. *Language Learning & Technology*, *17*(2), 75-99. http://dx.doi.org/10125/44325

Revythi, A., & Tselios, N. (2019). Extension of technology acceptance model by using system usability scale to assess behavioral intention to use E-lEarning. *Education and Information Technologies*, *24*(4), 2341-2355. https://doi.org/10.1007/s10639-019-09869-4

Rundell, M. (1999). Dictionary use in production. *International Journal of Lexicography*, *12*(1), 35-53. https://doi.org/10.1093/ijl/12.1.35

Sauro, J. (2015). SUPR-Q: A Comprehensive Measure of the Quality of the Website User Experience. *Journal of Usability Studies*, *10*(2), 68-86. https://dl.acm.org/doi/abs/10.5555/2817315.2817317

Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann.

Scholfield, P. (1982). Using the Dictionary for Comprehension. *TESOL Quarterly*, *16*(2), 185-194. https://doi.org/10.2307/3586791

Schumacker, R. E., & Lomax, R. G. (2012). *A beginner's guide to structural equation modeling* (3rd ed.). Routledge.

Siarova, H., Sternadel, D., & Szőnyi, E. (2019, September 12). *Research for CULT Committee – Science and scientific literacy as an educational challenge*. Policy Department for Structural and Cohesion Policies, European Parliament. http://bit.ly/2TCc6Uy

Swanepoel, P. (2001). Dictionary quality and dictionary design: A methodology for improving the functional quality of dictionaries. *Lexikos*, *11*, 160-190. https://doi.org/10.5788/11-0-846

Swanepoel, P. (2008). Towards a framework for the description and evaluation of dictionary evaluation criteria. *Lexikos*, *18*, 207-231. https://doi.org/10.5788/18-0-485

Tono, Y. (2011). Application of eye-tracking in Efl learners' dictionary look-up process research. *International Journal of Lexicography*, *24*(1), 124-153. https://doi.org/10.1093/ijl/ecq043

Tzifopoulos, M. (2020). In the shadow of Coronavirus: Distance education and digital literacy skills in Greece. *International Journal of Social Science and Technology*, *5*(2), 1-14. http://www.ijsstr.com/data/frontImages/1._April_2020.pdf

Welker, H. A. (2010). *Dictionary use: A general survey of empirical studies*. Brasilia.

Wellington, J., & Osborne, J. (2001). *Language and literacy in science education*. McGraw-Hill Education (UK).

Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogies and practices: Digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology*, *45*(2), 107-114. https://doi.org/10.1080/17439884.2020.1761641

Wingate, U. (2004). Dictionary use — the need to teach strategies. *The Language Learning Journal*, *29*(1), 5-11. https://doi.org/10.1080/09571730485200031