AGE AND GENDER DIFFERENCES IN EVALUATING THE PEDAGOGICAL USABILITY OF E-LEARNING MATERIALS

Liubomir Djalev, New Bulgarian University

Stanislav Bogdanov, New Bulgarian University

Abstract

The purpose of the study is to examine the pedagogical usability of interactive e-learning materials for foreign language practice. It is based upon two studies of the expected between-group and within-group differences among participants in the educational process. The sample consists of two groups – lecturers and students, a total of 100 participants, each evaluating four materials specifically prepared for this study. Two consecutive repeated measures ANOVA were conducted in which the gender/age, the position of the participants in the educational process, and usability dimensions were the independent variables. Results indicated that all independent variables and their interactions have a significant effects on the evaluations of the pedagogical usability. Women tend to assign higher values than men. Age groups generally differ in their evaluations, although there is a tendency to give similar ratings for the individual dimensions of pedagogical usability. The 31-40 years age group evaluates the materials higher while the lowest evaluations are given by the groups of 21-30 and 50+ year old participants. Students tend to rate the pedagogical usability systemically higher than the lecturers. Usability dimensions also have a significant effect on evaluations. The most prominent feature of the materials, by a great margin, is their Applicability. The findings corroborate previous research which show age and gender differences in web usability do exist. We conclude that these differences exist as much in pedagogical usability as in technical usability. Further investigations are suggested to explore more deeply the differences in the perceived pedagogical value of e-learning materials as this has implications for instructional designers, teachers and learners alike.

Keywords: pedagogical usability, e-learning, foreign language teaching, instructional design, age differences, gender differences

Ljubomir Djalev, PhD in Psycholgy, is Chief Assistant Professor at the Department of Cognitive Science and Psychology, New Bulgarian University. His research interests are in psychometrics, psychological measurement, research methods in psychology and education, statistical data analysis, e-learning.

E-mail: ldjalev@nbu.bg   https://orcid.org/0000-0002-8016-8051

Stanislav Bogdanov holds a PhD in Educational Technology in FLT. He is currently Chief Assistant Professor at the Department of Foreign Languages and Cultures, New Bulgarian University. His main research interests lie in Educational technology, Pedagogical Usability, Language localisation and Translation, Applied Linguistics and Foreign Language Teaching Methodology, Teacher education and development.

E-mail: stanbogdanov@nbu.bg   https://orcid.org/0000-0002-7144-1044

Technical and pedagogical usability

Usability is a concept born in recent years in the field of computer science, which initially focuses on human-computer interaction. In this mode of thinking, the starting point for the development of software applications is not so much their technological aspect, but consumers. Very soon, due to the fierce competition on the market for goods and services, this concept crosses the boundaries of the computer sphere, a testimony of which is the definition of usability in the international ISO standard. According to ISO 9241-11 (1998; 2018) for the Ergonomic requirements for working with visual terminals (later transformed into the more general Ergonomics of human-system interaction), usability is defined as "the extent to which a product can be used by specific users for achieving specific goals with efficiency, productivity and satisfaction in a specific context of use". According to this definition, usability is a quality that manifests in the interaction of humans with products, systems or services created by the human mind and hands. While the concept is still primarily used as a feature of software applications and web sites, it is also applicable to appliances, tools, and anywhere where people interact with objects, products or services. The word "specific" emphasizes that the product used is easy, convenient, and satisfying; it enables users to do the job it is intended for. There is no product that is usable or not usable in itself - this quality depends on the goals and needs of specific users in a particular context.

Despite the widespread adoption of the concept of usability, empirical research in this field, particularly the practical application of the ISO standard, is too scarce (Jokela et al., 2004). The reason for this is perhaps less in the weak theoretical or pragmatic interest than in the novelty of the concept, which is still seeking its quantitative and qualitative parameters. For example, Quesenbery (2001), noting that usability as "ease of use" of the product is a simplification of the problem, sets a series of challenging questions about what people really understand by "usability". Is it the result of any activity (for example, a software product), or is it a customer-centered development process, or is it a set of techniques (e. g. heuristics), or a product development philosophy designed to meet the needs of users? (Quesenbery, 2001)

Strictly speaking, none of these questions has a correct answer including the last one which tempts us to prefer it because of its generalization. Usability is a characteristic of the product, but it cannot be considered as its inherent, human-independent feature. In

accordance with Heider's (1958) early ideas in his Attribution theory, one can say that in the process of interaction with the product individuals attribute the characteristics of "usability" to the product. This property, however, exists only in their minds.

On the other hand, usability can be seen as a quantitative result, which is a desirable consequence of the application of a holistic user-oriented product development philosophy. Here specific techniques are used to increase the level of this feature. The authors draw attention to the multidimensionality of the concept of usability, and Quesenbery (2001) ingeniously identifies five "Es" which collectively describe consumer requirements of the product: it must be Efficient, Effective, Engaging, Error free and Easy to learn. We may notice the versatility of this view of usability.

The concept of pedagogical usability fits perfectly into this conceptual framework. It arises as a result of the natural evolution of Information and Communication Technologies (ICT) and their use for educational purposes. In this setting, researchers begin to pay attention to the pedagogical aspects of the usability of these technologies. Teaching materials used in the learning process are products designed to support the achievement of certain educational goals and they must be subject to at least the same usability requirements. However, the understanding of pedagogical usability is not limited to simply replacing the term "product" with "teaching material" in any of the usability definitions.

Many researchers distinguish between "technical" and "pedagogical" usability of educational technologies. According to Hadjerrouit (2012), technical usability refers to the seamless interaction between the person and the product (or the system, in particular certain educational software), which aims to reduce the cognitive load (Sweller, 1988) in the interaction and to free up cognitive resources. These resources can be used in the learning process, the support of which is the main purpose of pedagogical usability. Specifically, in the context of e-learning, technical usability is expressed in terms of how much an educational site is convenient, practical, and useful for learners as regards shaping the content, information structure, navigation and overall site design. Several systems of criteria have been developed to assess technical usability (see Nielsen, 1990; Nielsen, 2000, Nokelainen, 2005, Hadjerrouit, 2012; Djalev & Bogdanov, 2013). Pedagogical usability refers to the learning aspects of educational technologies and their potential to support the learning process (Melis, Weber &

Andrès, 2003). To evaluate these aspects of e-learning systems, Nokelainen (2004, 2005) has developed a comprehensive system of criteria that can be applied to any e-learning material (see also Hadjerrouit, 2012; Djalev & Bogdanov, 2013).

The two concepts are interconnected. Furthermore, the pedagogical aspects of e-learning materials can be considered as added value to their technical aspects. Kukulska-Hulme & Shield (2004) examine the usability of e-learning websites as a multi-layered feature in which technical usability forms the bottommost layer without being self-sufficient.

Research objective

In previous publications (Djalev & Bogdanov, 2013; Bogdanov, 2013a), we presented some results from an empirical study of the pedagogical usability of e-learninig materials (EDM) used in English language teaching settings. We examined the problem from the viewpoints of the two main groups of participants in the learning process - teachers and students. In particular, we presented the results of the analyses of (1) the latent structure underlying the evaluation of the pedagogical usability of the EDM; (2) the reliability and consistency of the pedagogical usability ratings of the two groups of participants; as well as (3) some expected differences depending on the role of the participants in the educational process; whether the teachers were native or non-native English speakers; the university where students are trained, and the type of EDM. In another paper Bogdanov (2013) presented results from a study of the effect of the intensity (frequency of exposure) of e-learning on the evaluation of the pedagogical value of the e-materials.

The aim we set out in this paper is to complement the emerging complex picture of the pedagogical usability of the EDM with two studies of the expected between-group differences by gender and age among participants in the educational process. The interaction of these factors with the role of individuals in the educational process and with the dimensions of pedagogical usability of EDM will also be analyzed.

Method

Design

This study focuses on the search for between-group differences in evaluating the pedagogical usability of the EDM used in English language training at C1 level of Bulgarian

students from two universities in the country. For this purpose two consecutive analyses of variance were conducted in which the gender and age of the participants were taken as the main independent variables. As an additional between-group variable with two levels - teachers and students, the position of the participants in the educational process was used. The study also included a within-group independent variable with 10 levels, which were the dimensions of pedagogical usability. Thus, the design of each of the analyses was mixed, with two between-group and one within-group variables, with the latter two (position and usability) being used to examine only their interaction with gender and age. The dependent variables were the participants' evaluations of the individual EDM pedagogical usability dimensions, expressed by their composite scores on the subscales of the Pedagogically meaningful eLearning questionnaire (PMLQ) (Nokelainen, 2006) presented later. The method applied for statistical analyses of the data was repeated measures ANOVA.

Materials

For the purposes of the study, four sets of e-learninig materials were developed, designated as EDM 1, EDM 2, EDM 3 and EDM 4. The materials were developed in accordance with the advanced level C1 curriculum according to the Common European Framework of Reference for Languages developed by the Council of Europe (2001) in the general English language courses at New Bulgarian University. EDM 1 is designed to introduce a new grammar structure and consists of 20 screens with embedded images; EDM 2 is designed as revision of thematic vocabulary and consists of 5 screens with interactive exercises and built-in audio recordings; EDM 3 consists of a 9-screen sequence of a grammatical structure, divided into small portions in the following mode: presentation - practice - elaboration; EDM 4 is a continuation of EDM 3 and consists of 11 screens for extra practice for listening comprehension, finding key vocabulary, and word order practice. (for more details about the purpose, structure and the development of the EDM, see Djalev & Bogdanov, 2013).

Measurement tool

The evaluation of the pedagogical usability of the EDM was performed using the Pedagogically Meaningful eLearning Questionnaire (PMLQ), developed by P. Nokelinen and a team (2006) from the University of Tampere, Finland. The questionnaire has two forms - for students (with 56 questions) and for teachers (with 60 questions). The two forms are equivalent with the exception of the last four questions from the teacher form

which only concern this group of participants in the educational process. Therefore, these four questions (57, 58, 59 and 60) were removed from the questionnaire, as well as three other questions (16, 17 and 18) because of their inapplicability to the developed EDMs. The responses to the questions are on a 6-point Likert-type scale in which values range from 1 (disagree) to 5 (completely agree), and 6 - N/A (not applicable to this material). Prior to data processing, the last point (6) was transformed to a zero value, giving the scale at least ordinal meaning without altering the verbal meaning of that point.

The questions are randomized but organized into ten subscales of pedagogical usability - *1. Learner control, 2. Learner activity, 3. Cooperative/ Collaborative learning, 4. Goal orientation, 5. Applicability, 6. Added value, 7. Motivation, 8. Valuation of previous knowledge, 9. Flexibility, and 10. Feedback*. Further information on the PMLQ and its scale structure, as well as the English version of the questionnaire, can be found in Nokelainen (2006), and in Djalev & Bogdanov (2013).

Participants

The study included two categories of participants: English language teachers and students attending general English language courses, with a total sample size of 100 subjects. The first group consisted of 20 lecturers recruited for the study because of their participation in various interest groups on the Internet, that are involved in developing interactive learning content. Participants in this group were 10 women and 10 men. With respect to age, all participants were divided into four groups (21-30, 31-40, 41-50 and 50+ years). Among the teachers, the representatives of the first age group (21-30) were 5 (25%), 4 teachers (20%) were between 31 and 40 years old, 5 teachers (25%) were aged 41-50 years and another 6 teachers (30%) were 50 and more years old. In terms of nationality, 8 of the teachers (40%) were Bulgarians, the rest belonged to eight other nationalities as follows: United Kingdom - 5 participants (25%), and Australia, The Netherlands, France, Germany, New Zealand, Poland and USA – 1 teacher (5%) each. Teachers for whom English is mother tongue were 8 (40%). The rest are speakers of their national languages respectively. The teachers' profiles suggest that the sample is well balanced with regard to gender and age, with a predominant share of the participants in the higher age groups, therefore with greater teaching experience.

The group of students included 80 participants, 48 of whom were male (60.00% of all students) and 32 female (40.00%). The largest subgroup fell within the age range

of 21-30 years (n=67, 83.75%). The other participants were relatively evenly distributed in the higher age groups, with 4 students in the 31-40 age group (5%), 3 in the 41-50 age group (3.75%) and 6 of the participants were over 50 years old (7.50%). Most of the participants in this group were students at New Bulgarian University (NBU) (n=63, 78.75%) and the rest were students at the University of National and World Economy (UNWE) (n=17, 21.25%).

The NBU students' profile is very diverse due to the specifics of the organization of foreign language learning at the university, where students from different majors, years of study and of different ages may be enrolled in the same English language course at level C1. The profile of UNWE students is much more homogeneous. Another significant difference between the students of the two universities can be seen in their experience with EDM and in e-learning in general. While for NBU students, much of the curricular content of the courses in the various majors, including those in full-time education, is also available through the institutional Moodle as an e-learning platform, for the students of the UNWE, e-learning is rather "incidental" and "extra", which is provided to them when possible and desired by teachers.

Procedure

The interactive e-learning materials developed as objects of evaluation, along with instructions and the PML questionnaire, were uploaded to a designated web site. The instructions, the same for teachers and students, required the participants to do the interactive online exercises and, after each EDM, to evaluate it using the online questionnaire. Responses to questions were stored in an online database[1]. Time constraints on the task were not specified because it was of no interest to the present study. The data were collected from April to June during the second semester of the academic year.

Results

Two consecutive ANOVAs with repeated measures were performed with different configurations of independent and dependent variables to examine between-group and within-group differences.

---

[1] Data is available as open data on Mendeley; see References.

Between-group gender differences

The first analysis is designed to compare the results of the evaluation activity of the participants in the study, differentiated into two gender groups. As noted, the sample of participants is fairly balanced (a total of 58 men and 42 women). In the study participants were asked to evaluate each EDM, but not all did the task correctly. Therefore, the analysis was performed on data from 393 responses. Two more variables were included in it. A second independent variable reflects the position of the participants in the educational process with two levels (lecturers and students), and the latter represents the various aspects (dimensions) of the pedagogical usability of the EDM, presented through the subscales of the PML questionnaire. Dependent variables were the composite scores on these subscales. This outlines the complex between-within design of the study, which includes two between-group and one within-group factor. Repeated measures ANOVA analysis was performed on the data. The results are presented in the following Table 1.

Table 1
Results of the Repeated measures ANOVA with factors Gender, Position of the participant in the educational process and Dimensions of pedagogical usability

Effect	SS	df	MS	F	p	Partial eta-squared	Observed power (alpha=0.05)
Intercept	918369.76	1	918369.76	18120.17	0.00	0.98	1.00
Gender	3691.10	1	3691.10	72.83	0.00	0.16	1.00
Position	3337.64	1	3337.64	65.85	0.00	0.14	1.00
Gender*Position	571.79	1	571.79	11.28	0.00	0.03	0.92
Error	19715.36	389	50.68				
Usability	268327.5	9	29814.17	2543.30	0.00	0.87	1.00
Usability*Gender	1496.7	9	166.29	14.19	0.00	0.04	1.00
Usability*Position	766.7	9	85.19	7.27	0.00	0.02	1.00
Usability*Gender* Position	680.4	9	75.60	6.45	0.00	0.02	1.00
Error	41041.0	3501	11.72				

The table contains a lot of evidence about the existence of statistically significant dependencies between the variables included in the analysis. First of all, we will look at the results of the general comparison of the gender-based groups, which express the main efect of this factor, as well as its interaction with the factor, representing two groups of participants in the educational process and with the pedagogical usability subscales of the PML questionnaire.

Considering gender as a single between-subjects factor, the difference between the mean scores of the two gender groups was statistically significant at $p = 0.00$. Women tend to assess the pedagogical usability of the EDMs systematically higher than men - the mean scores of the groups were M = 20.19 and M = 17.78 respectively, with equal standard deviations (SD = 0.20 for both groups). The difference between the two means estimated by the effect size is large ($\eta_p^2 = 0.16$) (Cohen, 1988; Miles & Shevlin, 2001) at the maximum power of the test ($1-\beta = 1.00$).

As an independent single factor, the position of the participant in the educational process also affects the evaluation ($p = 0.00$), i. e. lecturers and students differ significantly in their evaluations of the pedagogical usability of the EDMs. In this comparison, students tend to rate the pedagogical usability systemically higher than the lecturers. The respective means scores were M = 20.13 and M = 17.84, with standard deviations shows greater homogeneity in students' evaluations (SD = 0.13) compared to the teachers' (0.25). The difference between the two means estimated by the effect size is large ($\eta_p^2 = 0.14$) at the maximum statistical power ($1-\beta = 1.00$).

The latter factor also has an effect of interaction with the gender of the participants. Although with a small effect size ($\eta_p^2 = 0.03$), it nevertheless has a high statistical power ($1-\beta = 0.92$). The following graph illustrates the systematically higher ratings that women (both lecturers and students) give to the pedagogical usability of the EDMs.
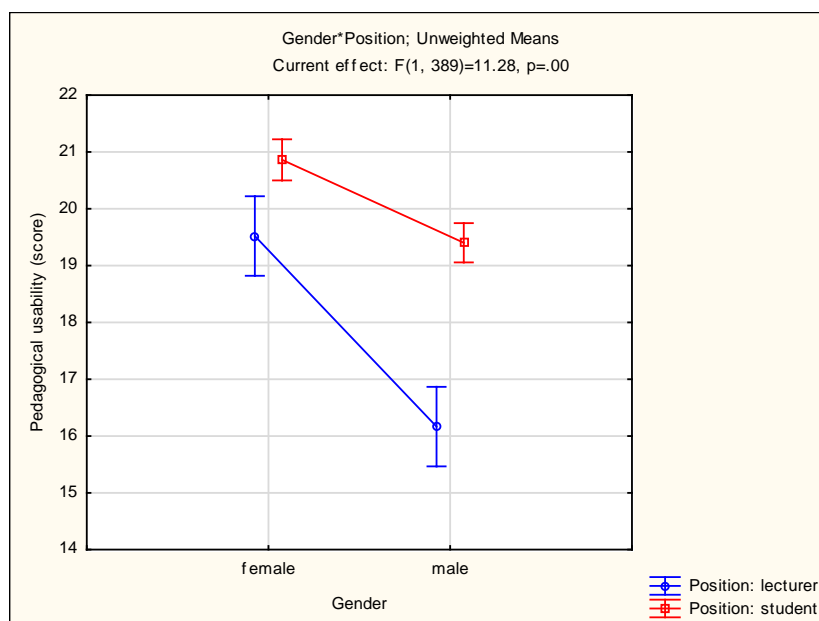

Fugure 1. Unweighted mean scores of pedagogical usability evaluations - interaction of the factors Gender and Position of the participants in the educational process

However, the differences between the mean scores of the four categories of participants are not the same and have no equal significance. The confidence intervals around the means and the Bonferroni post-hoc test showed statistically significant differences between each two categories ($p < 0.05$), with the exception of the female lecturers and male students ($p = 1.00$).

The data in Table 1 show significant differences in the ratings of the participants on the individual subscales of the PML questionnaire, which means differences with respect to the dimensions of pedagogical usability. The level of statistical significance ($p = 0.00$), the large effect size ($\eta_p^2 = 0.87$), as well as a hgh power of the test (1-$\beta = 1.00$) show unequivocally that the individual aspects of pedagogical usability of the EDMs, as a within-subject factor, have a significant effect on evaluations.

Although statistically significant, differences were observed among all subscale pairs ($p < 0.05$) except for subscales *6. Added value* and 9. *Flexibility* ($p = 1.00$), most of the mean subscale scores were in the relatively narrow range of 10.12 (subscale 8. *Valuation of previous knowledge*) and 24.81 (subscale 2. *Learner activity*). In contrast to these relatively low values, the mean score on the subscale 5. *Applicability* reaches 48.07, which demonstrates the high quality of EDMs developed for the purposes of the study with respect to this important dimension of pedagogical usability.

An interaction of within-group factor and each of the two between-group factors (Usability*Gender, and Usability*Position) was observed, with significance levels at $p = 0.00$, with medium to low effect sizes but with a high power of the test (1-$\beta = 1.00$). In the interaction between usability dimensions and gender, Bonferroni post-hoc test showed statistically significant differences in almost all combinations of the levels of the two factors ($p < 0.05$) with few exceptions. Overall, the level of the scores depends on the gender of the subjects and on the concrete dimension of usability, with women giving higher ratings than men.

In the second combination of usability dimensions and position of the participants factors, Bonferroni post-hoc test showed almost the same type of relations between the levels of the two independent variables ($p < 0.05$). Similarly, the mean scores depend on the participant's position and on the particular dimension of usability, with a clear tendency for students to rate usability higher than lecturers.

Especially interesting is the observed interaction of the three independent variables (Usability*Gender*Position). The following graph illustrates the strength and direction of their interaction.
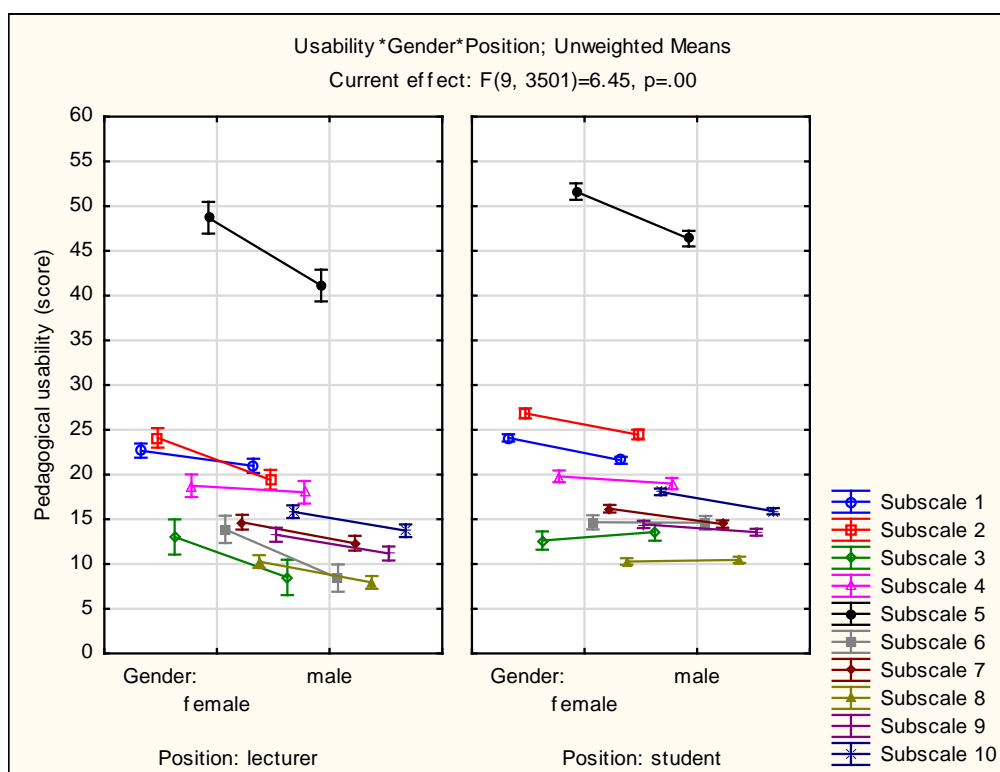

Figure 2. Unweighted mean scores of pedagogical usability evaluations - interaction of the factors Dimensions of pedagogical usability, Gender and Position of the participant in the educational process

The graph shows the almost ubiquitous dominance of female ratings (both teachers' and students' ones) of the individual pedagogical usability dimensions over those of male subjects. Some exceptions are also observed, for example in the ratings of subscales 6. *Added value*, 8. *Valuation of previous knowledge* and 3. *Cooperative / Collaborative learning*, and only for students.

There is also a particular stratification of the subscales, more clearly expressed in the students' ratings. Considearble higher ratings of all subjects on the subscale *5. Applicability* are observed. After the top layer, which includes only this subscale, follows a second layer formed by two subscales - 1. *Learner control* and 2. *Learner activity* which were rated higher both by teachers and students, women and men. For students, this layer is separated, and for lecturers, it is diffusely connected to the lowest layer in which all other subscales are located.

Between-group differences by age

The next research question is related to the search for between-group differences in the age factor, as well as its interaction with other factors affecting the evaluation of EDMs. The design of this study is similar to that of the previous one. As the first independent variable, the age of the participants was selected at four levels: 21-30, 31-40, 41-50 and over 50 years old. As a second independent variable, the position of the participants in the educational process with 2 levels (lecturers and students) was included. Another variable was also included, related to the various aspects of the pedagogical usability of the EDMs, which is presented through the subscales of the PML questionnaire. Dependent variables were the composite scores on the individual subscales of the same questionnaire.

As with the results of the previous analysis, this design also features a number of significant effects, some of which have large effect sizes and high statistical power. Such effects on pedagogical usability evaluations are exerted by all three independent variables, as well as by all interactions between them. The data from the analysis are presented in the following Table 2. Since the main effects of the position and usability factors were presented in the previous study, we will consider here only the main effect of the age factor and its interactions with other factors.

Table 2
Results of the Repeated measures ANOVA with factors Age, Position of the participant in the educational process and Dimensions of pedagogical usability

Effect	SS	df	MS	F	p	Partial eta-squared	Observed power (alpha=0.05)
Intercept	666092.31	1	666092.31	15459.26	0.00	0.98	1.00
Age	2237.70	3	745.90	17.31	0.00	0.12	1.00
Position	6899.07	1	6899.07	160.12	0.00	0.29	1.00
Age*Position	3024.29	3	1008.10	23.40	0.00	0.15	1.00
Error	16588.48	385	43.09				
Usability	200076.30	9	22230.70	2020.01	0.00	0.84	1.00
Usability*Age	2789.82	27	103.33	9.39	0.00	0.07	1.00
Usability*Position	2022.11	9	224.68	20.42	0.00	0.05	1.00
Usability*Age* Position	1672.30	27	61.94	5.63	0.00	0.04	1.00
Error	38133.25	3465	11.01				

The significance of the effect of factor Age as a single demographic variable is at $p = 0.00$, with the power of the test $1-\beta = 1.00$. Figure 3 shows interesting dynamics in

evaluations of the pedagogical usability of EDMs with increasing age of the subjects. The subjects of the lowest and the highest age groups are more critical of the qualities of the EDMs, and those aged 31-40 tend to rate them the highest. One common downward trend can be observed in the level of mean EDMs usability evaluations with the increasing age of the participants. This overall tendency is also pronounced, as the effect size is rather strong ($\eta_p^2 = 0.12$).
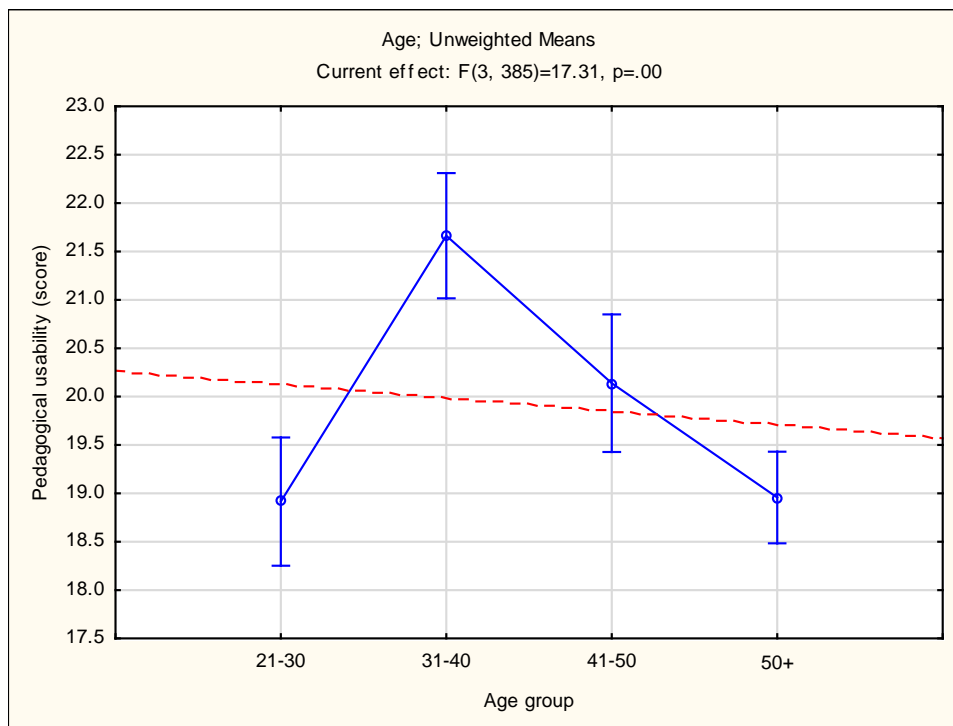

Age; Unweighted Means
Current effect: F(3, 385)=17.31, p=.00

Figure 3. Unweighted means of pedagogical usability evaluations by age groups

The application of Bonferroni post-hoc test indicates that statistically significant differences exist only between some of the age groups. More generally, the lower levels of scores of the two end groups (ages 21-30 and 50+) contrast with the higher levels of the two middle age groups (31-40 and 41-50 years). In particular, the statistical significance of the differences between the mean scores of the 21-30 age group with the next two is respectively $p = 0.00$ and $p = 0.03$. The 41-50 age group holds a special place among the others. In terms of its scores, its representatives are close to both the 31-40 years old group ($p = 0.37$) and the 50+ age group ($p = 0.45$).

The interaction of factors age and position also has a significant effect on participants' ratings, with a level of significance $p = 0.00$, effect size $\eta_p^2 = 0.15$ and a high power of the test $1-\beta = 1.00$. The analysis of the mean scores of the different categories of participants presented in Figure 4 shows that, overall, students rate the pedagogical usability of EDMs

more highly than lecturers. However, a significant effect of the interaction between the two factors was observed only in the first two age groups (21-30 and 31-40 years old).
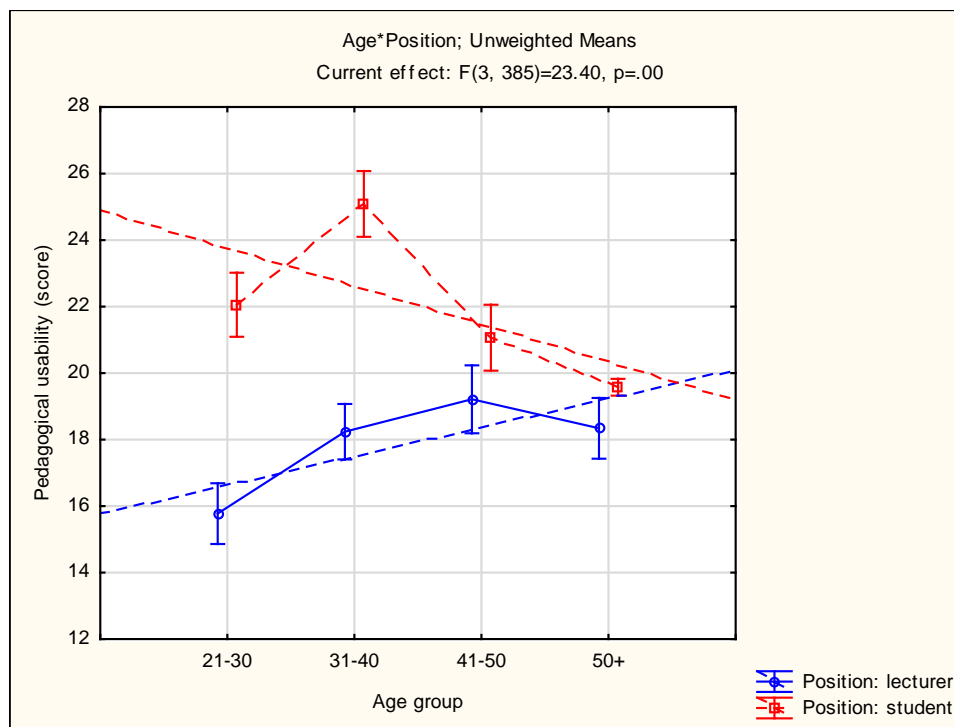

Figure 4. Unweighted mean scores of pedagogical usability evaluations - interaction of the factors Position of the participat in the educational process and Age

There is a clear and opposite tendencies of participants' evaluations of the pedagogical usability of EDM depending on their position in the educational process and age. While with increasing age teachers tend to rate more EDM highly, older students become more critical of the qualities of EDM than their younger counterparts. Thus, older participants in the educational process (teachers and students aged 41-50 and 50+ years) converge in their opinions, which makes the differences between their ratings non-significant.

The interaction between the factors Dimensions of usability and Age also has a significant effect on participants' ratings at p = 0.00, with a medium effect size η_p^2 = 0.07 and a statistical power 1-β = 1.00.

As can be seen in Figure 5, participants from different age groups have generally similar ratings across the individual subscales. For most of the dimensions of pedagogical usability of the EDMs, for example subscales 1. *Learner control*, 4. *Goal orientation*, 7. *Motivation*, 8. *Valuation of previous knowledge* and 9. *Flexibility,* the mean age group evaluations are close or, if there is a significant differences between some of

them, they have a small effect size. For the other dimensions, such as 2. *Learner activity* or 6. *Added value*, and in particular 3. *Collaborative learning* and 5. *Applicability*, there are more considerable differences, with greater effect size; and the highest scores of the EDMs being assigned by the 31-40 years age group and the lowest – by 21-30 years age group. Perhaps the most interesting result of this analysis is the contrast scores of participants from the four age groups on the quality of EDMs on subscale 5. *Applicability*, which oppose their ratings on all other subscales. Obviously applicability is the most valuable quality of the EDMs developed for the purpose of this study.
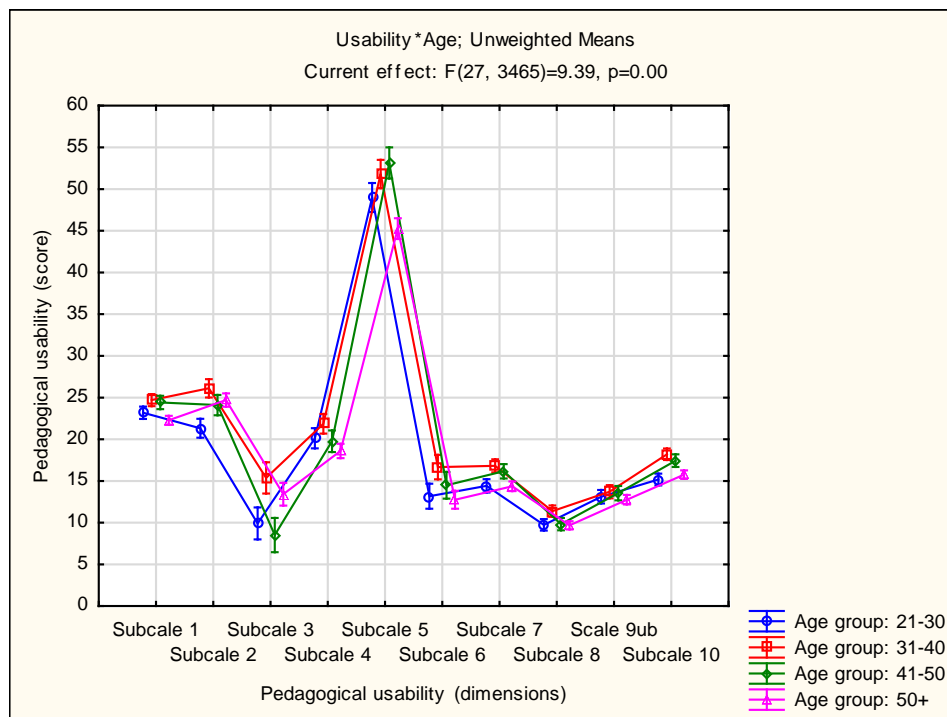

Figure 5. Unweighted mean scores of pedagogical usability evaluations - interaction of the factors Dimensions of pedagogical usability and Age.

The interaction between the three independent variables (Usability*Age*Position) is also characterized by statistical significance at $p = 0.00$, with a rather medium effect size $\eta_p^2 = 0.04$ and the power of the test $1-\beta = 1.00$. The data presented in Figure 6 show some similarities in the evaluations of the different categories of participants, determined by their position in the educational process and their age.

Here again, similar to the analysis of the interaction between the factors dimensions of usability, gender and position of the participants, a three-layer vertical structure of the evaluations of the pedagogical usability subscales are noticed. The topmost layer, which includes only 5. *Applicability* subscale, is clearly distinguishable

(with means in the range 42.60-57.88). The second layer is comprised of two (or three) subscales: 1. *Learner control* (with mean values in the range 20.70–28.06) and 2. *Learner activity* (with mean values in the range 17.80–30.18). Near these subscales, especially in the profile of the 21-30 age group, the subscale 4. *Goal orientation* is located (with mean values in the range 16.75–23.76). The other subscales form the lowest, denser layer where the subscales of the pedagogical usability are clustered and are more difficult to differentiate.
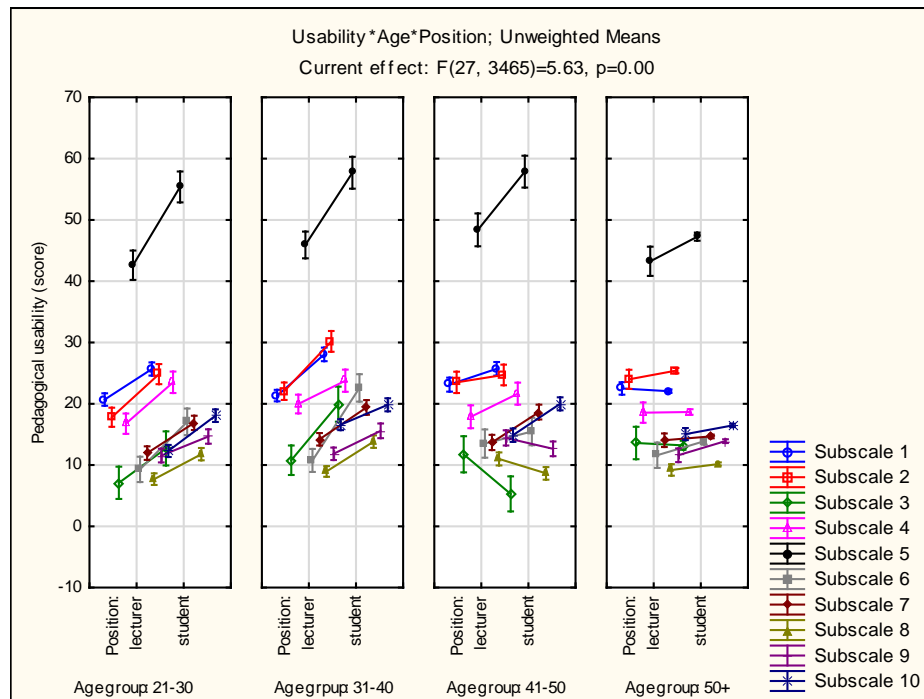

Figure 6. Unweighted mean scores of pedagogical usability evaluations - interaction of the factors Dimensions of pedagogical usability, Age and Position of the participant in the educational process

This stratification of the profiles of the individual subscales illustrates the effect of the factor Dimensions of usability. The overall vertical position of the profiles in the individual plots reflects in turn the effect of the age factor. It is best seen in the positions of the profiles of the subscale 5. *Applicability*. The overall mean subscale scores of the four age groups are respectively 18.92, 21.66, 20.14 and 18.96. The effect of the position factor is expressed in the slope of the profiles of the individual subscales. In most profiles, students' ratings are higher than those of the lecturers', especially in the first two age groups. Also, in most of them the differences between the mean scores of the two groups of participants are statistically significant. The following two higher age groups show mixed trends. In the 41-50 age group, students also score higher on

subscales with overall higher ratings, and lower scores on scales in the lower layers. In the 50+ years group, the scores of the two groups of participants are more aligned, with more non-significant differences between the respective means.

Discussion

The major aim of this study was to explore the gender and age differences in the evaluation of the pedagogical usability of interactive e-learning material designed for English language teaching. The position of the participants in the educational process and the dimensions of pedagogical usability were observed as additional independent factors. Numerous previous studies have shown that gender preferences do exist in web usability, for instance, web design in e-busines (Cyr & Bonanni, 2015). However, even studies in e-learning contexts, such as the study by Cuadrado-García et al. (2010), who investigated gender differences in e-learning use and assessment, have studied technical usability features. Such studies have examined, for instance, perceived ease of use and perceived usefulness (Okazaki & Renda dos Santos, 2012; Ong & Lai, 2006), while we found no studies with a focus on pedagogical usability.

The results of this study showed that all independent variables, as well as all possible interactions between them, have a signifficant effect on the pedagogical usability evaluations of the EDM. Generally, women tend to assess the pedagogical usability systematically more highly than men. In the interaction of gender with the position of the participants, both female teachers and female students rate usability more highly than their male colleagues. The level of ratings depends on the interaction between the participants' gender and the usability dimensions, with women more likely to give higher ratings than men on the individual dimensions. The interaction between the three independent factors also has a significant effect, within which several tendencies emerge - women tend to rate usability more highly than men; and students more highly than teachers; and that different dimensions have different levels of usability. The study shows that Subscale 5. *Applicability* is the most prominent dimension of pedagogical usability regardless of age, gender and the position of the participant. An explanation of these findings can be based on the claim that women are at an advantage in language education (Astleitner & Steinberg, 2005) and all materials in the present study are designed for foreign language practice. Further investigations are suggested to understand more deeply why a particular feature, in our case

Applicability, is given preference and under what conditions.

As a separate factor, age also has a significant effect on evaluation of pedagogical usability. There is an interesting curve in the usability evaluations which shows that the two age groups of participants, youngest and oldest, tend to rate usability lower than the participants in the two middle groups. The general tendency is for decreasing evaluations with increasing age. However, observing the interaction of the factor Age with the position of the participants in the educational process gives a more accurate picture of this dynamics. In general, students give higher usability ratings than faculty members. On the other hand, with increasing age, students tend to give lower usability scores and teachers tend to give higher scores. The general difference in the levels of evaluation between the two groups could be explained by experience with e-learning materials. Teachers have experience in both the development and application of EDMs for educational purposes, which keeps them at a certain distance from these concrete EDMs and leads to more criticisms of their qualities. Students, probably not all, have experience only as EDM users and have not yet developed the sensitivity to the qualities of EDMs that will make them more critical. The different directions of trends for students and teachers are probably due to generational differences related to their overall experience of working in an online environment and with web applications. Younger generations of students are more active Internet users who accept easily and appreciate highly the various web-based technologies. In contrast, older students would probably prefer some of the classic methods of presenting training materials. Younger teachers, probably for the same reasons, are more critical than their older colleagues who can be satisfied even with smaller steps towards new Internet technologies. All participants value almost equally some usability dimensions such as *Learner control, Goal orientation, Motivation, Valuation of previous knowledge* and *Flexibility*, and *Feedback,* with the age groups giving close ratings across all EDMs. For the other dimensions, such as *Learner activity* or *Added value*, and in particular *Collaborative learning* and *Applicability*, there are more considerable differences and the highest scores of the EDMs being assigned by the 31-40 years age group and the lowest – by 21-30 years age group.

Overall, the data from this study did not provide sufficient information to explain the difference in the perceived value of any of the pedagogical dimensions of e-learning

materials. On the other hand, it corroborates previous research that shows differences in age and gender when assessing usability. Our conclusion is that age and gender differences in pedagogical usability exist as much as they do in technical usability. Obviously, the gender and age gap has not yet closed, though the Internet and computers are more widely available at home and in institutions. Meanwhile, computer-based education is spreading fast. Therefore, it could be that male and female, as much as young and mature users as learners, have agreed on the saliency of certain features of educational software and systems.

One of the limitations of the study is that we did not look at the differences between young and mature female learners and young and mature male learners, which may have offered insight into the differences in the evaluations of these subgroups. Further research is necessary to try to find what motivates such differences since pedagogical usability has implications for instructional designers, teachers in online settings and ultimately (language) learners in online and hybrid learning environments.

References

Astleitner, H., & Steinberg, R. (2005). Are there gender differences in web-based learning? An integrated model and related effect sizes. *AACE Journal, 13*(1), 47-63. Retrieved from http://www.learntechlib.org/p/18902

Bogdanov, S. (2013). Latent Structures of Teachers and Students in Evaluating the Pedagogical Usability of E-Learning Materials for Language Teaching. In *Innovacionnye Processy v Issledovatel'skoj i Obrazovatel'noj deâtelnosti* [Innovative Processes in Research and Educational Activity] (pp. 3-6). Perm: Perm National Research Polytechnic University, https://doi.org/10.6084/m9.figshare.4204866.v1

Bogdanov, S. (2013a). Group differences in the evaluation of pedagogical usability of e-learning materials. In: *Novye tehnologii v obrazovatel'nom prostranstve rodnogo i inostrannogo âzyka* [*New technologies in the educational space of native and foreign languages*] (pp. 8-16). Perm: Perm National Research Polytechnic University.

Bogdanov, S. (2019). *Age and Gender Differences in Evaluating the Pedagogical Usability of E-Learning Materials* (v1) [Data set]. Mendeley Data. https://doi.org/10.17632/tzfdvz366g.1

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N. J.: Erlbaum.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)*. Council of Europe. Retrieved from https://rm.coe.int/1680459f97

Cuadrado-García, M., Ruiz-Molina, M.-E., Montoro-Pons, J.D. (2010). Are there gender differences in e-learning use and assessment? Evidence from an interuniversity online project in Europe. *Procedia - Social and Behavioral Sciences, 2*(2), 367-371. https://doi.org/10.1016/j.sbspro.2010.03.027

Cyr, D., & Bonanni, C. (2005). Gender and website design in e-business. *International Journal of Electronic Business, 3*(6), 565-582. https://doi.org/10.1504/ijeb.2005.008536

Djalev, L., & Bogdanov, S. (2013). *Dve gledni tochki kam pedagogicheskata polzvaemost na elektronni didakticheski materiali* [*Two Viewpoints on Pedagogical Usability in eLearning for Languages*]. Retrieved from http://www.lulu.com/shop/stan-bogdanov-and-ljubomir-djalev/two-viewpoints-on-pedagogical-usability-in-elearning-for-languages/paperback/product-21066553.html

Hadjerrouit, S. (2012). Investigating technical and pedagogical usability issues of collaborative learning with Wikis. *Informatics in education, 11*(1), 45–64.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley & Sons. https://doi.org/10.1037/10628-000

International Organization for Standardization. (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability* (ISO 9241-11:1998). Retrieved from https://www.sis.se/api/document/preview/611299

International Organization for Standardization. (2018). *Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts* (ISO 9241-11:2018). Retrieved from https://www.iso.org/standard/63500.html

Jokela, T., Iivari, N. & Tornberg, V. (2004). *Using the ISO 9241-11 definition of usability in requirements determination: case studies*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=73139C6B269D886867B122351A021ABB?doi=10.1.1.628.5652&rep=rep1&type=pdf

Kukulska-Hulme, A.M. & Shield, L.E. (2004). The keys to Usability in e-Learning Websites. *Proceedings of Networked learning conference 2004*. Lancaster University, UK, 5-7.04.2004. Retrieved from http://www.networkedlearningconference.org.uk/past/nlc2004/Proceedings/individual_papers/kukulska_shield.htm

Melis, E., Weber, M. & Andrès, E. (2003). Lessons for (Pedagogic) Usability of eLearning Systems. In A. Rossett (Ed.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2003* (pp. 281-284). Chesapeake, VA: AACE. Retrieved from http://www.editlib.org/p/14936

Miles, J., Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. Sage:London.

Nielsen, J. (1990). Evaluating Hypertext Usability. In Jonassen, D. H. & Mandl, H. (Eds.), *Designing Hypermedia for Learning*, Berlin: Springer-Verlag, 147-168. https://doi.org/10.1007/978-3-642-75945-1_9

Nielsen, J. (2000). *Designing Web Usability: The Practice of Simplicity*. Indianapolis: New Riders.

Nokelainen, P. (2004). Conceptual Definition of the Technical and Pedagogical Usability Criteria for Digital Learning Material. In L. Cantoni & C. McLoughlin (Eds.), *Proceedings of ED-MEDIA 2004--World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 4249-4254). Lugano, Switzerland: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/primary/p/11688/

Nokelainen, P. (2005). The technical and pedagogical usability criteria for digital learning material. In P. Kommers & G. Richards (Eds.), *Proceedings of ED-MEDIA 2005--World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 1011-1016). Montreal, Canada: Association for the Advancement of Computing in Education (AACE). Retrieved from https://www.learntechlib.org/primary/p/20212

Nokelainen, P. (2006). An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Educational Technology & Society, 9*(2), 178-197.

Okazaki, S., & Renda dos Santos, L. M. (2012). Understanding E-learning Adoption in Brazil: Major Determinants and Gender Effects. *International Review of Research in Open and Distributed Learning*, *13*(4), 91-106. https://doi.org/10.19173/irrodl.v13i4.1266

Ong, C. S. & Lai, J. Y. (2006). Gender Differences in Perception and Relationships among Dominants of E-Learning Acceptance. *Computers in Human Behavior, 22*(5), 816-826. https://doi.org/10.1016/j.chb.2004.03.006

Quesenbery, W. (2001). What does usability mean: Looking beyond 'ease of use'. *Proceedings of the 48th Annual Conference, Society for Technical Communication*. Retrieved from http://www.wqusability.com/articles/more-than-ease-of-use.html

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4