applied sciences

MDPI

Article

User Experience Questionnaire in sign language for deaf and hard of hearing

Klemen Pečnik [1]*, Žana Juvan [2], Gregor Dolinar [3], and Matevž Pogačnik [4,*]

1 Laboratory for Multimedia, Faculty of Electrical Engineering; Klemen.Pecnik@fe.uni-lj.si
2 Laboratory for Multimedia, Faculty of Electrical Engineering; Zana.Juvan@fe.uni-lj.si
3 Laboratory of Applied Mathematics and Statistics, Faculty of Electrical Engineering; Gregor.Dolinar@fe.uni-lj.si
4 Laboratory for Multimedia, Faculty of Electrical Engineering; Matevz.Pogacnik@fe.uni-lj.si

* Correspondence: klemen.pecnik@fe.uni-lj.si; matevz.pogacnik@fe.uni-lj.si; Tel.: +386 31 758 885 K.P.; +386 1 4768 101 M.P.

Abstract: Accessibility of interactive services and applications is an important part of an inclusive society, even more so in the light of the ongoing digitaliation process. While legal and regulatory frameworks are being implemented and the services are adopting accessibility standards, we still lack user experience evaluation approaches, which would be adapted to deaf participants. Therefore, the standard User Experience Questionnaire (UEQ) was translated to Slovenian Sign Language (SSL) and its appropriateness and reliability evaluated in a research study. For this purpose, two interactive TV applications for deaf viewers were developed and were evaluated using 1) a standard textual UEQ and 2) a UEQ translated to Slovenian Sign Language. The evaluation was done by 17 deaf participants and 4 sign language interpreters. A thorough statistical analysis of the results was performed on the level of individual UEQ statements as well as on the level of UEQ scales. The results are very promising and show that there were no statistically significant differences in evaluation results when gathered with the UEQ in sign language compared to results gathered with the textual standard UEQ. The participants also expressed a strong preference for UEQ in sign language as it allows them to participate more independently in user experience evaluation studies.

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Sensors* **2022**, *22*, x. https://doi.org/10.3390/xxxxx

Academic Editor: Firstname Lastname

Received: date
Accepted: date
Published: date

Keywords: UEQ, Deaf and Hard of Hearing, User Experience, User Testing, Accessibility, Slovenian Sign Language, Sign Language, interactive multimedia, HbbTV,

1. Introduction

Technology development has significantly contributed to quality-of-life improvements for Deaf and Hard of hearing, especially in the fields of communication and education. There are many organisations contributing to such improvements, with one of most important ones being W3C Web Accessibility Initiative (WAI), standardizing the Web Content Accessibility Guidelines (WCAG) [1]. Many countries are striving towards improvement of the accessibility situation, mostly through legal and regulatory frameworks, while there is still a lot of room for improvement in terms of funding the development of accessibility solutions. The Republic of Slovenia, being an EU member state, is following EU directives, enacting accessibility laws, and gradually progressing towards accessible public services. The World Health Organisation (WHO) has projected that by 2050 nearly 2.5 billion people will have some degree of hearing loss [2], while at the moment there are 75 million deaf people in the world. In Slovenia there are around 1500 deaf people with 1000 of them using the Slovenian sign language as their primary language [3]. While a laudable improvement on accessibility regulation, technological solutions, and service development has been made in the past 10 years, a lack of user-friendly approaches adapted to deaf participants for user testing and user experience evaluations, is still noticeable. A growing awareness, that the evaluations and testing of new technologies and services are of paramount importance, has resulted in many new evaluation methods and approaches, but these are adapted only to the general population without disabilities. In A Case Study about Usability, User Experience and Accessibility Problems of Deaf Users with Assistive Technologies [4] researchers were evaluating the usability and user experience of assistive technologies and one of the research findings was the conclusion that users are not adopting assistive technologies due to a bad user experience. We believe that bad user experience with assistive technologies as well as any accessible solutions could be significantly improved if the evaluation procedures themselves are made accessible and adapted for the deaf population.

There is a plethora of quantitative and qualitative research methods that help the developers with evaluations of different solutions' user experience and its aspects, such as usability, task difficulty, effort, mental demand, learnability, attractiveness, novelty, trust, VR sickness, etc. In this paper we will focus on the quantitative methods as we are trying to get an objective insight into the usability of our solutions. One of the oldest and widely accepted methods is the System Usability Scale (SUS), which is quick and easy to use, and is often used as a benchmark method [5]. It consists of a 10-item questionnaire with 5-point scale answers, resulting in a single score ranging from 0 to 100 indicating the usability and learnability of the evaluated solution. A similar method to SUS, but somewhat more targeted and with more sub scores is the Post Study System Usability Questionnaire (PSSUQ) [6]. It has a 19-item questionnaire with 7-point scale and provides scores for system usefulness, information quality, interface quality and an overall score. Similarly, to SUS, it can be done using either a paper or a digital version. Another well-known method is NASA Task Load Index (NasaTLX), that rates perceived workload to assess a task, system, or team's effectiveness [7]. It consists of 21 item questionnaire, 15 pairing questions and 6 scales, giving three different results. It is more complex than SUS and quite demanding to set up. Measurements of software quality from the users' perspective is often performed using the Software Usability Metric (SUMI) methodology [8]. It has been used for over 25 years and consists of 50-item questionnaire, with 3-point scale, providing results for efficiency, affect, helpfulness, control, learnability, and a global SUMI score. It produces reliable results with small sample sizes and allows for their comparison with the global SUMI repository of over 2000 evaluated software solutions. However, a 50-item questionnaire can be quite tedious for the participants and requires a lot of participants' time and effort. For target evaluation method the User Experience Questionnaire (UEQ) [9,10] was selected since it is a quick and reliable method and includes both classical usability measurements as well as user experience aspects. It consists of a 26-item questionnaire with 7-point Likert scale and returns scores on 6 scales: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty, thus covering the pragmatic and hedonistic aspects of evaluated solutions. It also contains many useful tools to help calculate the results' statistical reliability, identify inconsistencies, calculate scores using benchmark values and finally produce reports.

It is apparent that measuring user experience (UX) is a general practice with the goal of creating a satisfying experience for users of products and designs. There are many different UX methods, however they are all based on the existing and standardized models, making them hard to apply in certain use cases. This is especially true in the case of trying to evaluate UX with the deaf and hard of hearing, due to the general assumption that all deaf can read and write. About 0.1% of the population becomes deaf before learning any language, so sign language is the main form of communication for these people. As a result, those people do not know the phonetics of spoken languages, they do not understand them, and they cannot read written words as they do not understand the structure of sentences and other linguistic peculiarities of spoken languages [11,12]. Since UX evaluations have to include the actual target users, when developing solutions for deaf and hard of hearing, alternative approaches of measuring UX have to be used, and whenever one of the standard methods is applied, there is a risk of getting wrong data due to misunderstanding. In order to provide equal rights and opportunities, our main goal was to adapt the UEQ, a standardized method for measuring user experience, for the deaf and hard of hearing.

The goal of our research was to evaluate the appropriateness and usefulness of a UEQ questionnaire translated into the Slovenian Sign Language (SSL) [13]. If proven suitable, it would enable more independent evaluations of interactive services by the deaf users, with minimized need for a sign language interpreters' involvement. This has not only positive practical consequences, but also a high symbolic significance. Namely, it would allow for more independent involvement of deaf users, which is one of the important principles of the "United Nations Convention on the rights of the persons with disabilities" [14].

The evaluations of the SSL translated UEQ questionnaire were performed using two interactive solutions: 1) A Hybrid Broadcast-Broadband Television (HbbTV) application with an adjustable sign language interpreter layout and 2) A virtual 3D sign language interpreter for a limited vocabulary, which has been recognized as a solution for enhancing accessibility for broadcast services [15], also implemented in the form of a HBBTV application. HbbTV as a platform has also been recognized as the technology that may be used to enhance accessibility services [16]. The experiment details are described in the Materials and methods section.

The results are very encouraging and present a basis for future work in this field.

2. Materials and Methods

In this section the approach to the SSL translation of the standardized UEQ questionnaire is explained, followed by a description of the apparatus, designed for the evaluation experiment, and the description of the experiment procedure. Finally, the demographic structure of the experiment's participants is described.

2.1. *Augmentation of the standardized UEQ with sign language*

The questionnaire translations, experiments, measurements, and final analysis are based on the Standardized Slovenian UEQ. The questionnaire consists of 26 questions, which are not actual questions, but rather sets of two opposing statements. The users decide which of the opposing statements suits their experience best, rated on a 7-point Likert scale. All gathered responses are then used in the official UEQ Data Analysis Tool, where they are transformed and the main results are presented on 6 scales: Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty.

In general, sign languages have significantly smaller vocabularies in comparison to spoken languages, therefore the translation to signed UEQ statements had to be adapted accordingly. An extensive study, involving several Slovenian Sign Language (SSL) interpreters, has been conducted, with the goal to find the best matches and most appropriate sign language translations for all 26 statement pairs. All written UEQ statement pairs were first evaluated with the interpreters, using multiple standardized UEQ translations (Slovenian, English and German) to find the statement's best equivalent in SSL. These translations were later tested in a preliminary study. For 18 statement pairs it was possible to find a direct translation from written to sign language. However, for the remaining 8 pairs, a direct translation was not possible, because these expressions do not exist as signs in the SSL vocabulary. An example of such statement is the statement "dull", which was translated to the signed expression "without ideas". Another challenge when translating the statements was the fact that a singular UEQ statement can only appear once in the whole questionnaire, while similar written expressions often translate in the same sign (gesture). Such examples are expressions "conventional" and "dull". During the preliminary study all the originally selected sign expressions were evaluated and based on the experts' feedback four sign expressions were corrected for better distinctiveness, while two expressions were replaced with similar ones, making them more straightforward.

The main idea was to provide the accessible UEQ as a standalone solution in the form of videos in a purpose-built web page, allowing for very simple interaction. All statements were first signed by an SSL interpreter in front of a green screen and recorded with professional video equipment, allowing for further video editing, background adjustments and other adaptations. The videos were then edited in post-production, the green background was removed and the videos, containing corresponding statement pair signs, were placed in their respective positions as presented in Figure 1. Two versions of such statement pair videos were exported, the first, where both statements were signed simultaneously, and the second, where statements in the pair were signed sequentially.

2.2. *Apparatus*

The whole testing procedure was originally designed as a Hybrid Broadcast Broadband Television (HbbTV) application with the backend system working over DVB-T network. However, due to the lack of availability of HbbTV 2.0 compliant TV sets [17], a web-based application running in Chrome with the RedOrbit HbbTV emulator extension [18] was implemented, allowing for a simple and intuitive user interaction using a regular TV remote control. Considering all comments and suggestions from the preliminary evaluations, user comments from the interviews, and recommendation for sign language interpreter positioning and sizing [19,20] the whole research framework (experiments and evaluations) was designed as an application with predefined structure and workflow and with the assumption that most of the potential users would be deaf or hard of hearing. All available navigation options were always shown on top of the application in the form of a standardized HbbTV legend, making use of the standard TV remote buttons.

The UEQ evaluation part of the experiment was designed as a simple animated questionnaire with intuitive navigation and the same design approach for all evaluations. The pre-recorded videos of SSL interpreters signing the statement pairs would be shown in full screen for each respective UEQ question, with a simple 7-point scale positioned in the middle of the screen, and a progress bar placed at the bottom to indicate the user's progress in the evaluation. To select a value on the 7-point scale for each statement pair, the users had used the arrow buttons on the TV remote and click OK to confirm their choice. The option to return to any previous question at any point of time was also added, allowing for easy mistake correction during the testing. Users were also able to easily switch between simultaneous and sequential signing, with the sequential signing being the default option. This decision was based on the results and comments from the preliminary testing. During the preliminary study the UEQ comprised of both signed and written expressions. However, due to the fact that one of the main goals was to assess the understandability and appropriateness of the signed UEQ variant, the separate UEQ in sign language and standard UEQ evaluation versions were created. During testing the UEQ with statements in sign language has been used first, followed by the standard UEQ with textual statements.

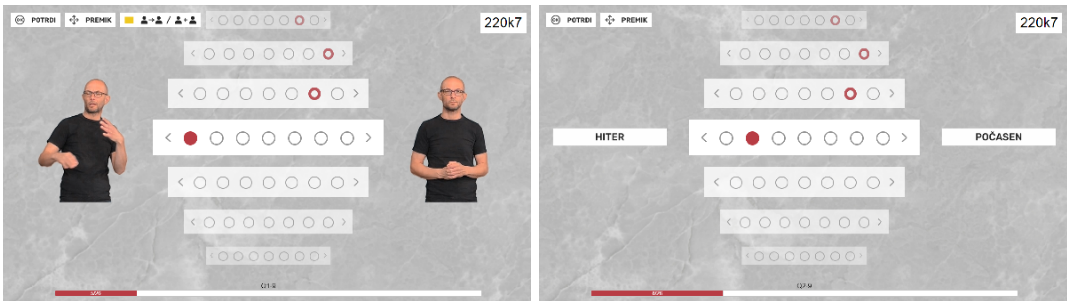

Figure . UEQ in sign language (left) and textual (right) for the 'fast – slow' statement

An example for the statements fast – slow is presented in Figure 1, with the sign language version on the left side and the standard (textual) version on the right side.

The users were able to adjust the position of the interpreter on the screen by pressing the left-up-right-down arrow keys. This allowed the users to move the interpreter out of the way in case it was covering an important part of the screen, without sacrificing the size of the original video. 6 predefined positions were available, as presented in Figure 2.

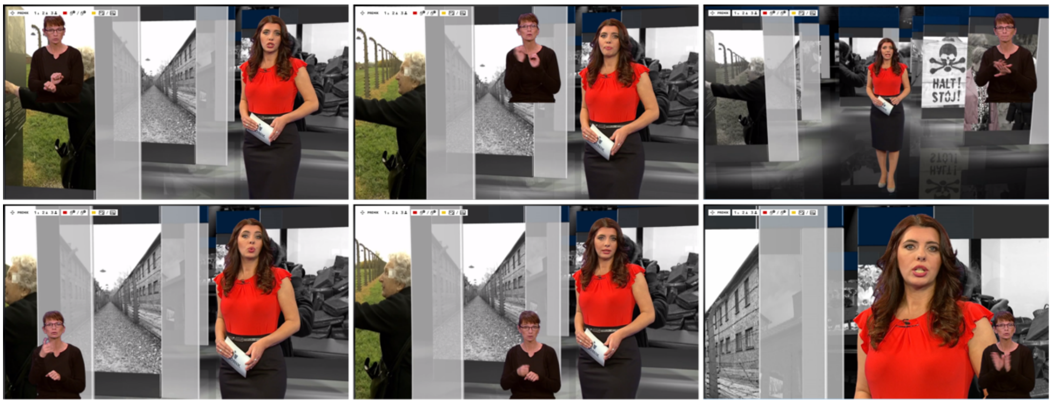

Figure . Sign interpreter positions

By pressing the numbers 1, 2 and 3, the users were able to adjust the size of the overlayed interpreter, with 1 being the smallest and 3 being the largest option, as presented in Figure 3.

Figure . Sign interpreter sizes (from left to right: small, medium, big)

The users also had the option to switch off or turn on the sign interpreter video by pressing the red button on the TV remote control. Last but not least, the layout mode could be changed from "overlay" to "side-by-side" mode, by pressing the yellow button on the remote, as presented in Figure 4.

Figure . Video and sign interpreter layout (on the left overlay example and on the right a side-by-side example)

The second interactive application was a virtual 3D sign language interpreter for a limited vocabulary and for selected domains of use, like emergency notifications, weather forecasts, traffic announcements, schedule changes, delays etc. The virtual interpreter was developed in the scope of another project and is in this experiment used only as an example of an interactive service, intended for user experience evaluation. This application allowed for the same functionalities and options as the first one, allowing the participants to adjust the virtual interpreter's position, size, and layout. Additionally, the option to toggle subtitles was added by pressing the blue button on the TV remote control. The participants were able to evaluate the virtual interpreter's understandability, attractiveness, novelty, etc. using the UEQ method. Figure shows examples of different layouts of a virtual interpreter for the weather forecast domain.

Figure . Example of two different layouts for a virtual sign language interpreter

Furthermore, the introduction, brief experiment explanation and UEQ instructions have also been prepared in written and sign language and were included as videos, creating a seamless user experience throughout the evaluation. An example of the instructions is shown in Figure 6.

Figure . UEQ instructions in sign language

2.3. *Experiment procedure*

User testing was conducted in the DemoLab of the Laboratory for Multimedia (LMMFE), at the Faculty of Electrical Engineering, University of Ljubljana.

Figure 7. User testing room setup (captured during the preliminary study)

The room provided a comfortable environment for the participants, which were sat on a sofa in front of a 55″ OLED TV (LG OLED55C6V) and user interaction with interactive application and UEQ has been limited to the use of a TV remote control from the beginning to the end of the experiment, as shown in .

At the beginning of each evaluation, the participant would fill in a simple questionnaire with the purpose of gathering demographic data, information about their technical experience and other relevant information and sign a consent form for the purpose of evaluation. Each test started with a simple presentation of the experiment idea and an explanation of all user interaction possibilities. At the beginning of each of the two experiments (with adjustable and virtual interpreter), the participant was given a couple of tasks and was asked to:

change the position of the interpreter,

change the size of the interpreter,

change the layout,

turn the interpreter off and then back on (since the interpreter was already present at the beginning of the experiment),

adjust the interpreter size, position, and layout according to their personal preference.

After the participant finished testing the interactive application, a video with an explanation of the UEQ structure and answering procedure was presented. Then the participant first answered the UEQ in sign language, followed by the written UEQ. This procedure was repeated for both experiments. After each step in the experiments the application waited for the participant's confirmation before continuation, and the users always had the option to repeat the last step. At the end of the second experiment the participants filled a post-test questionnaire, followed by a discussion, which was conducted in a manner of a relaxed conversation. All comments and suggestions about both experiments were gathered, along with the opinion about the involvement of deaf users in user testing. The whole step-by-step process with all the steps, average completion timings, and all available user interactions is presented in Figure 8.

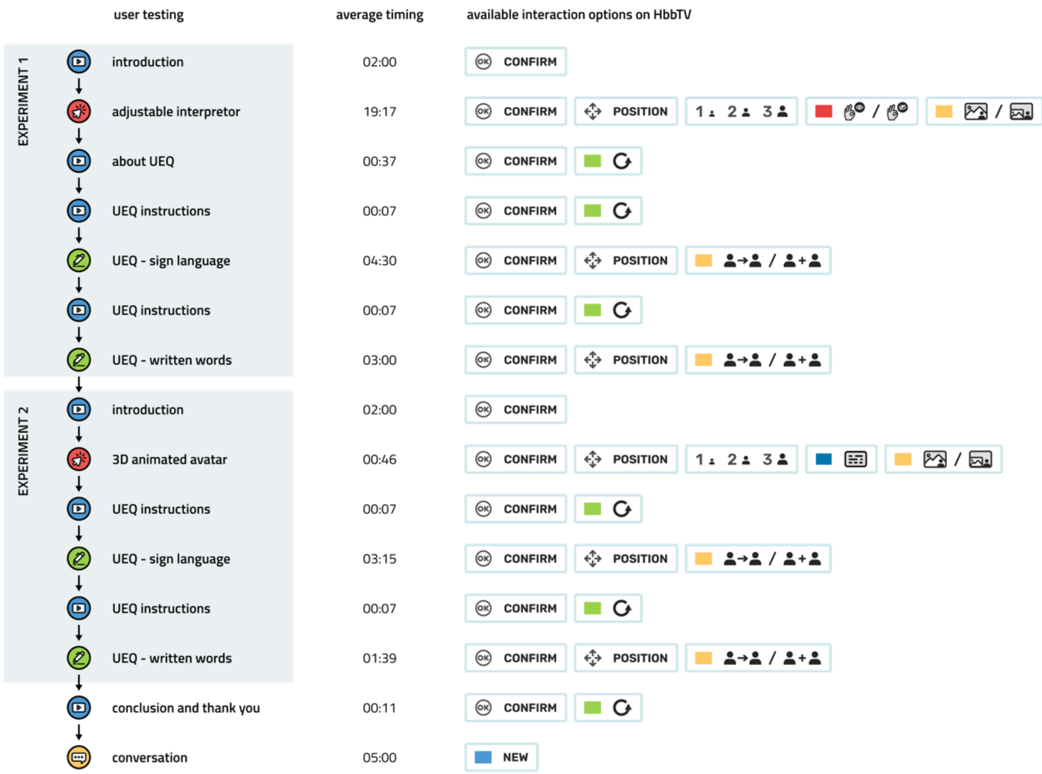

Figure 8. User testing workflow as implemented in the HbbTV application

Despite the pre-recorded instructions and explanations, a sign language interpreter was present during the testing especially with deaf participants to clarify any additional questions and to translate all given comments and suggestions to and from the sign language during the discussion. Additionally, two user experience experts were present at each test session. One expert was responsible for leading the evaluation procedure, answering potential questions, and providing additional explanations, while the other was taking notes about observed difficulties and challenges that one may have, and noted all relevant comments during the discussion phase. The collection of notes in conjunction with user action logs have allowed a comprehensive analysis of our solutions. Additionally, they have given a lot of useful information on the issues and dilemmas related to the chosen sign language expressions, their understandability, and appropriateness.

At the beginning of each user test, a random user ID was generated and displayed in the upper right corner of the screen. The ID was used to associate each participant's UEQ answers, questionnaires, and notes, without the need to identify them by their names, thus allowing anonymisation. The data for all four UEQ questionnaires was automatically stored to a SQL database immediately after each keypress, recording both the answer and the time of completion, enabling further analysis. All responses were exported to a .csv format and transferred into the UEQ Data Analysis Tool [9] as well as to an additional Excel spreadsheet for further analysis.

2.4. *Demographic data*

Most participants' responses were gathered during the Deaf History International 2022 conference, organised by the Association of Deaf Teachers in Slovenia, Association "Theater, Audiovisual Arts and Culture of the Deaf - DLAN", and Deaf History International in Ljubljana and Zagreb [21].

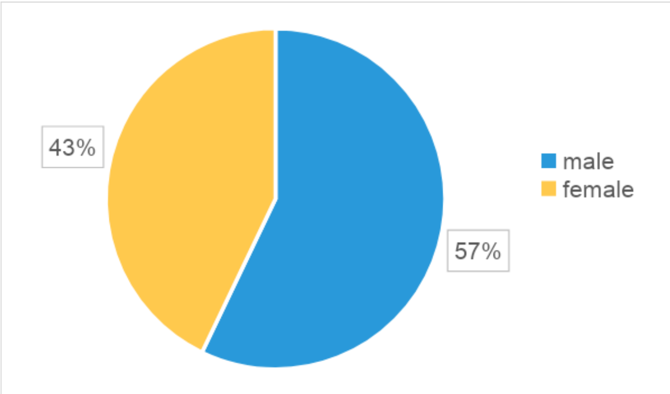

Figure . Gender distribution of the participants

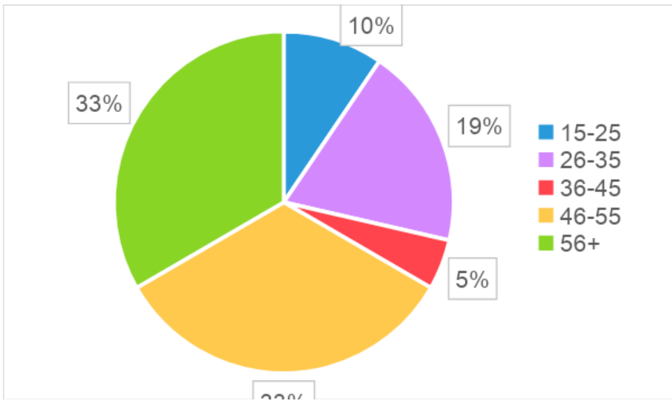

Figure . Age structure of the participants

21 participants took part in the experiment, with 43% being female and 57% male participants, as shown in Figure 9. Age structure was gathered in age ranges with 10% of participants aged 15-25 years, 19% 26-35 years, 5% 36-45 years, 33% 46-55 years, and 33% aged 56 or more, as shown in Figure 10.

Out of all participants 62% were deaf from birth, 19% lost hearing later in life, 19% with no hearing impairment, and no participants were hard of hearing, as shown in

Figure 11. Since the educational level could have an influence on the understanding of the written language, information about the participants' education was also collected. 5% of all participants finished only primary school, 48% secondary school, 14% Short-cycle higher vocational education (SCHVE), 0% obtained a bachelor's degree, 24% a master's degree, and 10% a PhD, as shown in Figure 12.

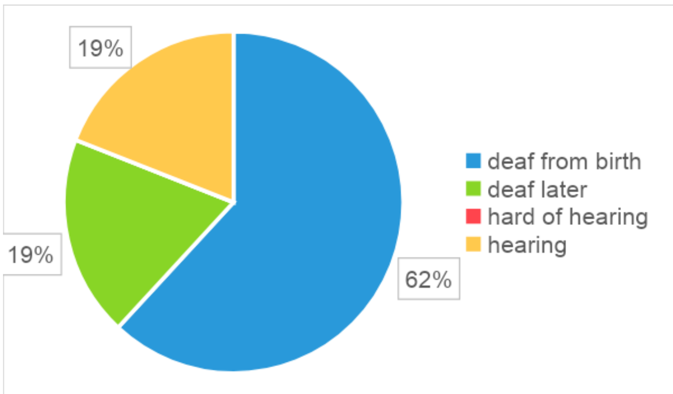

Figure 11. Type of deafness

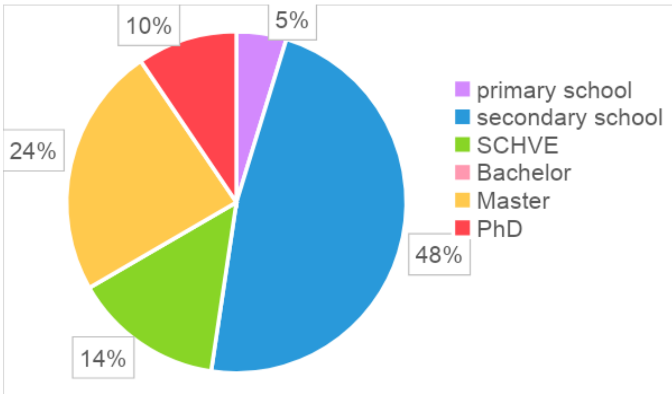

Figure 12. Participants' education levels

3. **Results**

In all standard user experience evaluations with the UEQ questionnaire, the pragmatic and hedonistic UEQ values represent the measure of success or failure. In this case, however, the UEQ values themselves are of secondary importance, as the appropriateness of the UEQ questionnaire, translated into Slovenian sign language, is evaluated. Therefore, the results are presented in the form of a comparison between the standard UEQ questionnaire values and the UEQ in sign language questionnaire values.

For this purpose, the following Theorems were defined:

Theorem 1. *The responses obtained with the standard UEQ questionnaire and the UEQ questionnaire in sign language will be similar, without significant statistical differences.*

Theorem 2. *The results of the evaluation using standard UEQ questionnaire and the sign language translation of the UEQ questionnaire will give comparable results in all UEQ scales (attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty).*

Theorem 3. *The deaf users will prefer the user experience evaluation procedure when using sign language translation of the UEQ questionnaire or a combination of both as compared to the evaluation when using a standard UEQ questionnaire in written form only.*

3.1. Statistical analysis of the gathered responses

First, we evaluated all gathered data from the experiments and performed the statistical analysis. The raw data was then further evaluated using the official UEQ Data Analysis Tools. In some cases, minor anomalies were spotted, which are presented and explained with additional information in Discussion.

The subtraction difference between the values of the answers from UEQ in sign language and standard UEQ was calculated, with the maximal possible difference between standard UEQ and UEQ in sign language being 6 or -6, and the target value being 0. The analysis focused on the answers to each UEQ question shows, that the best match between standard UEQ and UEQ in sign language was obtained for UEQ question 11 with 19 exact matches and the worst match was obtained for question 25 with 11 exact matches. When focused on each participant, the analysis shows that the best result was 26 exact matches, meaning that one participant answered both UEQs equally, and the worst result was 11 exact matches for a participant. The distribution of all differences is presented in Figure 13.

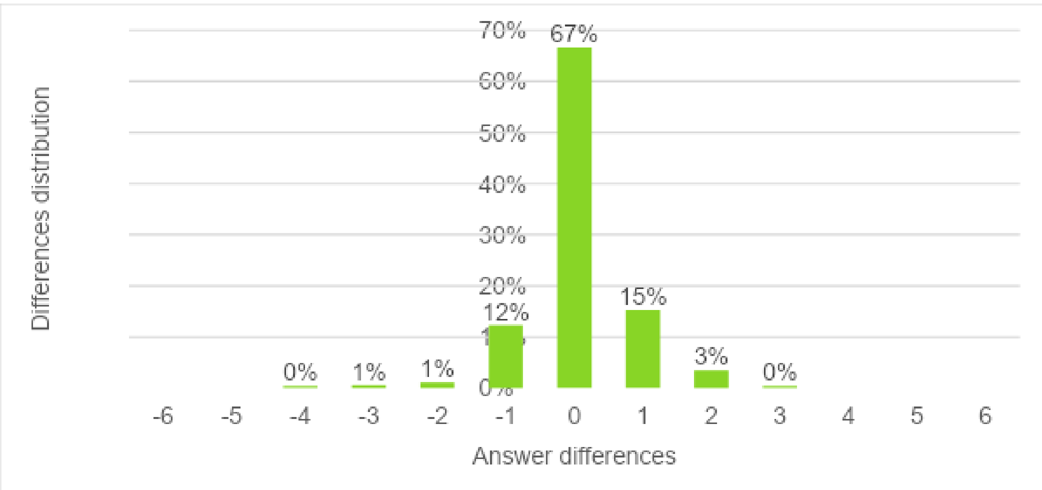

Figure . Differences between standard UEQ and UEQ in sign language distribution

However, the UEQ results are always presented using UEQ scales and calculated accordingly, thus the distribution of differences within each UEQ scale has been analysed and, as presented in Figure 14, the distributions of differences for each UEQ scale are very similar.

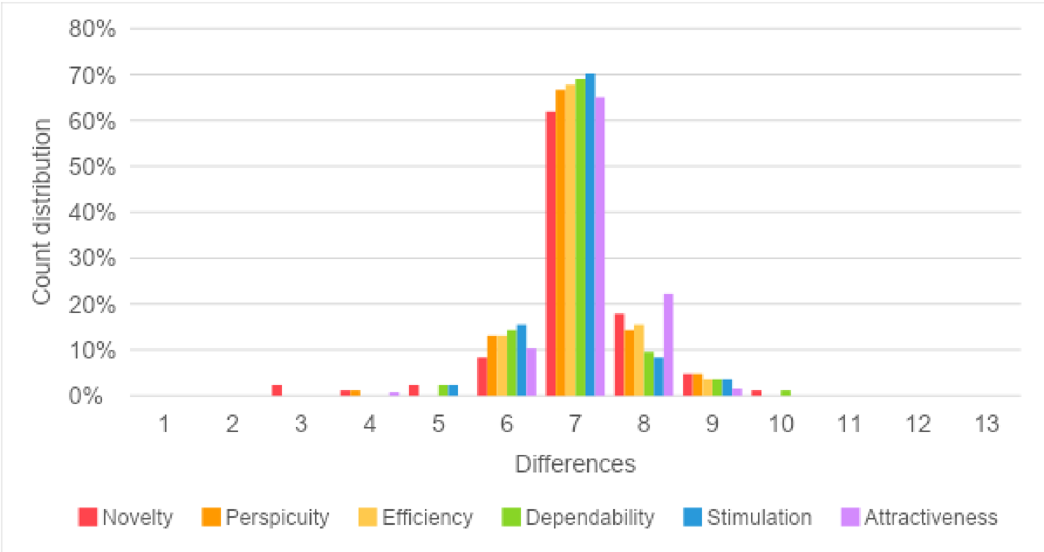

Figure . Differences distribution between standard UEQ and UEQ in sign language per scales

While the differences seem to be quite consistent in distribution manner, paired t-test has been used to check if answers in text and sign language are the same. Question pairs were defined as vectors from UEQ in sign language answers and standard UEQ, as shown in (1).

$$X_{SSLi} = \left(x_{SSLi1}\ x_{SSLi2}\ \vdots\ x_{SSLip}\right) \text{ and } X_{TXTi} = \left(x_{TXTi1}\ x_{TXTi2}\ \vdots\ x_{TXTip}\right) \quad (1)$$

The paired t-test has two alternative hypotheses, i.e., the null hypothesis H_0 assumes that the mean of difference of paired answers in text and sign language is equal to zero, while the alternative hypotheses H_a implies that there is a significant difference between the two populations. In equation for t-value (2), $\bar{x}_{TXT}$ represents the observed mean of responses for one question in textual form, $\bar{x}_{SSL}$ represents the observed mean of responses for the same question in sign form (corresponding pairs - e.g. $x_{TXT14} - x_{SSL14}$), $s_{TXT-SSL}$ represents sample standard deviation of the differences of textual and sign pair responses, and n represents sample size.

$$t = \frac{\bar{x}_{TXT} - \bar{x}_{SSL}}{\sqrt{\frac{s^2_{TXT-SSL}}{n}}} \quad (2)$$

All t-values for individual questions were compared with corresponding T-value for $\alpha = 0.05$ with 20 degrees of freedom, which corresponds to value of 2.086 [22], and all P-values with T.DIST.2T function were calculated. We rejected hypothesis H_0 for questions 6 and 9, due to the P-value being 0.0155, and concluded from alternative hypotheses H_a that responses to questions 6 and 9 are significantly different, while for all the other questions that is not the case.

To detect irregularities, user errors, or misunderstood signs, an in-depth analysis of each separate question is useful, however, in our study, the statistical analysis of results of UEQ scales is more important.

As mentioned, six UEQ scales are obtained by calculating the mean of responses to the corresponding questions. Attractiveness is determined by 6 questions, while all other scales (Perspicuity, Efficiency, Dependability, Stimulation, and Novelty) are determined by 4 questions. For this reason, we define vectors

$$Z_{TXTi} = \left(z_{TXTi1}\ z_{TXTi2}\ \vdots\ z_{TXTip}\right) and\ Z_{SSLi} = \left(z_{SSLi1}\ z_{SSLi2}\ \vdots\ z_{SSLip}\right) \tag{3}$$

with

$$Z_{TXTi} = \frac{1}{m_i}\sum_{j\in M_i} X_{TXTj} and\ Z_{SSLi} = \frac{1}{m_i}\sum_{j\in M_i} X_{SSLj}, \tag{4}$$

where each of the vectors Z_{TXTi} and Z_{SSLi}, for $i = 1, \dots, 6$ corresponds to one of the UEO scales with cardinality of M_i, i.e., $m_i = |M_i|$, equal to 6 for Attractiveness, and with cardinality of M_i equal to 4 for all the other UEQ scales. Appropriate statistical test is needed to compare responses for both scales. In this manner both samples were tested one against another by Multivariate Paired Hotelling's T-Square test [23], with the null hypothesis

$$H_0: \mu_{TXT} = \mu_{SSL} \tag{5}$$

and the alternative hypothesis

$$H_a: \mu_{TXT} \neq \mu_{SSL}, \tag{6}$$

where μ_{TXT} *and* μ_{SSL} are mean vectors for vectors Z_{TXTi} and Z_{TXTi}, respectively.

To compare two corresponding samples, vectors Y_i are defined as:

$$Y_i = Z_{TXTi} - Z_{SSLi} \tag{7}$$

and

$$\mu_y = \mu_{TXT} - \mu_{SSL} \tag{8}$$

Subsequently, the null hypothesis that the population mean vector is equal to 0 can be tested, which would mean that all of its elements are equal to 0. This can be tested against the alternative that the vector μ_y is not equal to 0, which would mean that at least one element of the vector μ_y is not equal to 0:

$$H_0: \mu_0 = 0\ against\ H_0: \mu_0 \neq 0 \tag{9}$$

The sample mean vector was defined as:

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{10}$$

and the S_y matrix with sample variance-covariance of the vectors Y_i as:

$$S_y = \frac{1}{n-1}\sum_{i=1}^{n} (Y_i - \overline{Y})(Y_i - \overline{Y})'. \tag{11}$$

Finally, paired Hotelling’s T-Square test is given with the expression

$$T^2 = n\overline{Y}'S_y^{-1}\overline{Y} \tag{12}$$

and the corresponding F-statistic is defined with the expression

$$F = \frac{n-p}{p(n-1)}T^2, \tag{13}$$

where $n - p$ and p represent degrees of freedom and H_0 can be rejected at level α, if the F-value exceeds the value from F-value with $n - p$ and p degrees of freedom, evaluated at level α. Results of the Paired Hotelling’s T-Square test for all UEQ scales are presented in Table and for virtual interpreter in Table 2.

Table . Results of the Paired Hotelling's T-Square test for all UEQ scales for the adjustable interpreter

UEQ Scales	n	p	df1	df2	T2	F-value	p-value
Attractiveness	21	6	6	15	14.676	1.835	0.152
Perspicuity	21	4	4	17	3.364	0.715	0.589
Efficiency	21	4	4	17	7.071	1.503	0.230
Dependability	21	4	4	17	1.196	0.254	0.904
Stimulation	21	4	4	17	8.301	1.764	0.166
Novelty	21	4	4	17	5.148	1.094	0.380

Table . Results of the Paired Hotelling's T-Square test for all UEQ scales for the virtual interpreter

UEQ Scales	n	p	df1	df2	T2	F-value	p-value
Attractiveness	19	6	6	13	8.771	1.056	0.398
Perspicuity	19	4	4	15	2.318	0.483	0.748
Efficiency	19	4	4	15	8.688	1.810	0.157
Dependability	19	4	4	15	11.565	2.409	0.075
Stimulation	19	4	4	15	10.287	2.143	0.104
Novelty	19	4	4	15	1.351	0.281	0.887

Proof of Theorem 1. The results of the Paired Hotelling's T-Square test show, that we cannot reject our hypothesis H_0 for each of the UEQ scales since all p-values are larger than 0.05. This indicates that the responses in standard UEQ and UEQ in sign language are not significantly different. □

3.2. *Analysis and comparison of the UEQ results for both experiments*

Firstly, the calculated results of the UEQ from UEQ_Data_Analysis_Tool were obtained for the UEQ in sign language along with calculated results from the standard UEQ in text. In order to determine the differences of mean values, variances, standard deviations, and confidence intervals between both UEQ versions, differences of mean values for each UEQ scale were compared, with the results presented in Table , with the UEQ in sign language results in the 2nd column, standard UEQ results in the 3rd column and the differences in the 4th column.

UEQ results of both approaches can be identified as significantly close as the absolute differences in values for individual UEQ scales range between 0.012 (Dependability) and 0.127 (Attractiveness), with the highest relative difference around 5% (Attractiveness). This seems to be a good result, especially considering that the UEQ scale values are in practice interpreted as excellent, good, above average, below average, bad, which further reduces the significance of such minor differences in result values.

Table . Comparison of UEQ Mean values for the adjustable interpreter

UEQ Scales	Mean SSL	Mean TXT	Δ Mean
Attractiveness	2.429	2.556	0.127
Perspicuity	2.440	2.512	0.071
Efficiency	2.190	2.286	0.095
Dependability	2.190	2.202	0.012
Stimulation	2.512	2.464	-0.048
Novelty	2.262	2.310	0.048

Similarly to the procedure of the adjustable interpreter, mean values were calculated and compared for the virtual interpreter which are presented in Table 4.

Table : UEQ Mean values for the virtual interpreter

UEQ Scales	Mean SSL	Mean TXT	Δ Mean
Attractiveness	2.026	2.105	0.079
Perspicuity	1.947	2.026	0.079
Efficiency	1.868	2.013	0.145
Dependability	1.592	1.895	0.303
Stimulation	2.171	2.197	0.026
Novelty	2.289	2.329	0.039

Again, the results of both UEQ versions were comparable, with the difference in values for individual UEQ scales being between 0.26 (Stimulation) and 0.303 (Dependability). This also seems to be a good result, especially considering that UEQ scale results are in practice interpreted as excellent, good, above average, below average, bad, as presented in Figure .

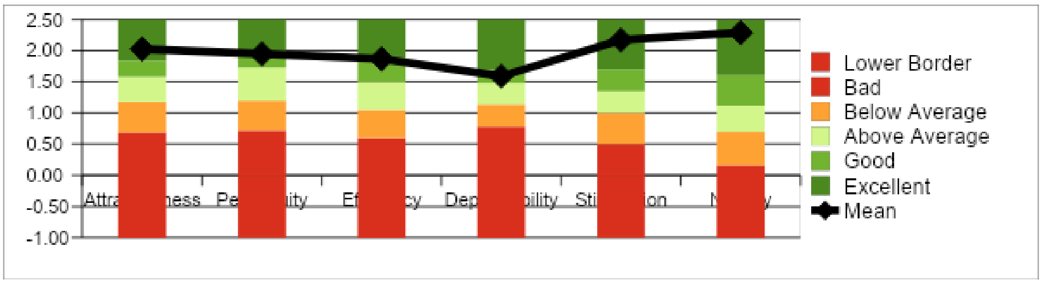

Figure . UEQ results from UEQ in sign language for virtual interpreter

The UEQ_Data_Analysis_Tool also includes embedded tools for detecting data inconsistencies and calculating scale consistencies. Data inconsistencies are reported when differences between the best and the worst answer values on the same scale exceed 3 points. If such a difference would appear in three or more scales, that participant's data is marked as critically suspicious and should be eliminated. For both experiments no critically suspicious data was detected, however some minor inconsistencies were identified, with a maximum of one inconsistency per user. Most inconsistencies were identified in Dependability scale.

The scale consistency is checked against the whole scale by calculating an average of correlation of all answers in UEQ scale with Cronbach's Alpha-Coefficient and Gutman's Lambda2 Coefficient. Cronbach's Alpha-Coefficient is calculated as:

$$\propto \; = \frac{n\bar{r}}{1+(n-1)\bar{r}}, \qquad (14)$$

where n represents the sample size (6 for Attractiveness and 4 for other scales) and $\bar{r}$ represents an average of all correlations between sample items. As it can be noticed from the equation, the sample size is fixed (6 or 4) and the actual correlation has an influence on the coefficient value.

Guttman's Lambda2 coefficient, for one of the UEQ scales, is calculated as [24]:

$$\lambda_2 = 1 - \sum_{j \in M_j} \frac{\sigma_j^2}{\sigma_x^2} + \frac{\sqrt{\frac{n}{(n-1)} \sum_{i \in M_i} \sum_{j \in M_j} \sigma_{ij}^2}}{\sigma_x^2} \; i \neq j, \qquad (15)$$

with cardinality of M_i M_j, i.e., $m_i = |M_i|$ *and* $m_j = |M_j|$, equal to 6 for Attractiveness, and with cardinality of M_i M_j, equal to 4 for all the other UEQ scales. Variable n represents the number of questions for corresponding scale (6 for Attractiveness and 4 for other scales), σ_j^2 represents the variance of the j^{th} question, σ_x^2 represents the variance of the observed scores on the composite measures for corresponding question, and σ_{ij}^2 represents the covariances between items *i* and *j*.

Using the rule of thumb, it is recommended that Cronbach's Alpha-Coefficient and Gutman's lambda2 coefficient are higher than 0.7. In case of the adjustable interpreter these coefficient values are above the recommended threshold value for all the scales except in the case of Efficiency ($\alpha = 0.67$) and Dependability ($\alpha = 0.46$) for UEQ in sign language, and in the case of Dependability for standard UEQ with the value of 0.69. This result could either be a consequence of having a small sample size, since only 21 people participated in the study, or a consequence of corresponding statement understandability. Namely, 14% of the participants had issues with understanding the statements for question 8, and more than 50% of participants did not exactly understand the user experience context of statements 17 and 19. The results of scale consistency for Cronbach's Alpha and Gutman's Lambda 2 are presented in Table 5.

Table . UEQ Scale consistency of the adjustable interpreter

UEQ Scales	SSL α	TXT α	SSL λ_2	TXT λ_2
Attractiveness	0.79	0.92	0.85	0.92
Perspicuity	0.75	0.90	0.84	0.91
Efficiency	0.67	0.81	0.63	0.78
Dependability	0.46	0.69	0.47	0.65
Stimulation	0.84	0.88	0.84	0.87
Novelty	0.76	0.79	0.75	0.76

In the experiment with the virtual interpreter, the scale consistency results were analogous to the first experiment, as the calculated coefficient values are above the recommended threshold value of 0.7 for all scales, except in the case of Efficiency and Dependability for the UEQ in sign language. In this case the Cronbach's Alpha-Coefficient for Efficiency was 0.55 and for Dependability 0.65, which again could be a consequence of poor understanding of signed statements, as well as a relatively small number of participants (19). The results are presented in Table 6.

Table 6. UEQ Scale consistency of virtual interpreter

UEQ Scales	SSL α	TXT α	SSL λ_2	TXT λ_2
Attractiveness	0.90	0.96	0.89	0.96
Perspicuity	0.88	0.92	0.88	0.92
Efficiency	0.55	0.84	0.59	0.83
Dependability	0.65	0.78	0.70	0.77
Stimulation	0.93	0.87	0.90	0.87
Novelty	0.82	0.91	0.77	0.91

In addition, the variances and the standard deviations for both UEQ versions, as well as standard deviations of their differences were calculated for both experiments. For the adjustable interpreter the standard deviation of the answer value differences for each individual UEQ question varies between 0.31 and 1.28, while the standard deviations of answer value differences belonging to each UEQ scale are presented in Table . The 2nd column represents the standard deviation of UEQ scales in sign language, the 3rd column a standard deviation of standard UEQ and the 4th column a standard deviation of differences between standard and UEQ in sign language evaluation.

Table . Standard deviation of differences between standard UEQ and UEQ in sign language per scale for the adjustable interpreter

UEQ Scales	Std. dev. SSL	Std. dev. TXT	Std. dev. TXT-SSL
Attractiveness	0.569	0.659	0.217
Perspicuity	0.680	0.625	0.318
Efficiency	0.612	0.624	0.340
Dependability	0.602	0.692	0.422
Stimulation	0.673	0.653	0.269
Novelty	0.772	0.851	0.640

As the UEQ evaluation results are usually interpreted on the level of each scale, these results are more significant than each individual answer value. While the standard deviation of the UEQ scale results in sign language varies from 0.569 (Attractiveness) to 0.772 (Novelty) and of the standard UEQ from 0.624 (Efficiency) to 0.851 (Novelty), it is a little bit lower for the differences and it varies from 0.217 (Attractiveness) to 0.650 (Novelty). It can be noticed that the highest values of standard deviation are for novelty scale in all three cases.

In the same manner as for the first experiment variance and standard deviations were calculated for virtual interpreter, which are presented in Table .

Table . Standard deviation of differences between standard UEQ and UEQ in sign language per scale for adjustable interpreter

UEQ Scales	Std. dev. SSL	Std. dev. TXT	Std. dev. TXT-SSL
Attractiveness	0.778	0.932	0.285
Perspicuity	0.739	0.794	0.354
Efficiency	0.704	0.664	0.481
Dependability	0.778	0.699	0.483
Stimulation	0.825	0.715	0.456
Novelty	0.809	0.722	0.419

While the standard deviation of the UEQ in sign language scale results varies from 0.704 (Efficiency) to 0.825 (Stimulation) and of the standard UEQ from 0.664 (Efficiency) to 0.932 (Attractiveness), it is a little bit lower for the differences and it varies from 0.285 (Attractiveness) to 0.483 (Dependability).

Finally, to address the Theorem2, a comparison of means and confidence intervals of the UEQ scales for the adjustable interpreter are presented in graphical form in Figure .

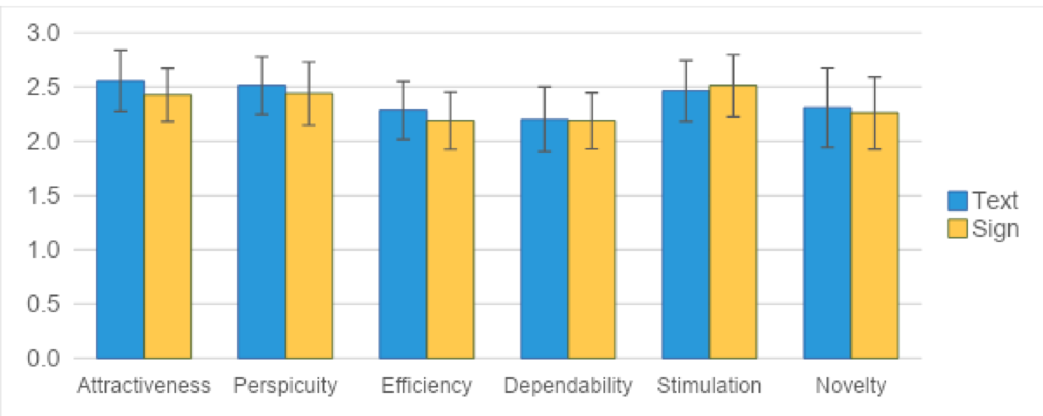

Figure . Comparison of UEQ results between standard UEQ and UEQ in sign language for the adjustable interpreter

The same comparison of the standard UEQ and UEQ in sign language for the virtual interpreter can be observed in Figure .

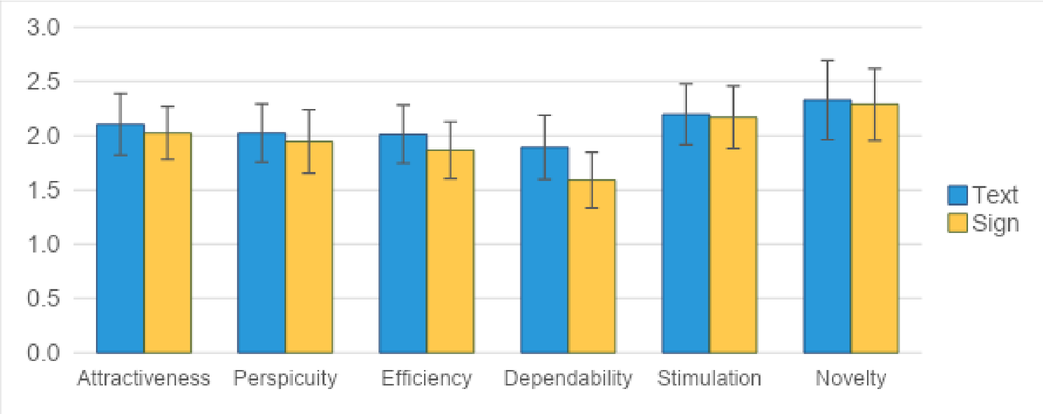

Figure . Comparison of UEQ results between standard UEQ and UEQ in sign language for virtual interpreter

The values of means and their corresponding confidence intervals are in both cases coequal, except for the Dependability scale, where, as already mentioned, participants had issues with understanding some of the statements.

Proof of Theorem 2. The comparison of UEQ scale results between the standard UEQ and UEQ in sign language have shown that they were significantly close. Due to the representation and interpretation of UEQ results it can be concluded that the UEQ in sign language would return comparable results as the textual one. □

3.3. *Evaluation of participants' preference of the UEQ version*

Other results and information, obtained through a structured interview after the evaluations, have shown that all of the deaf participants, involved in this study, would like to use the evaluated solutions daily and that all of them believe that early inclusion of deaf users is beneficial to the development process of any solution. The analysis of demographic data relating to the UEQ results, it can be concluded that UEQ in sign language would mean a lot to the participants with lower education, especially for those who could not read. Although only 2 participants could not read, only one of all the participants did not find UEQ in sign language beneficial, while all other participants stated that they would prefer the UEQ in sign language or a combination of both, to the standard UEQ in written form only. They also prefer to be independent during such activities and consequently sign language evaluations are needed and welcome.

Proof of Theorem 3. During the interviews and discussions 95% participants preferred the use of UEQ in sign language (or a combination of UEQ in sign language with added text) to the use of a standard UEQ in text only. □

4. **Discussion**

The main goal of the presented research was to confirm the working hypothesis that the results of the evaluations using standard UEQ questionnaire and the sign language translation of the UEQ questionnaire will give comparable results in all scales (Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty). During the preliminary evaluations of two different interactive solutions, intended for the deaf and hard of hearing population, a sufficiently small UEQ scale result differences were obtained between the sign language translation of the UEQ questionnaire and the standard UEQ questionnaire. Furthermore, the standard deviations of corresponding answers' value differences on the level of UEQ scales are also relatively small. Therefore, the results of UEQ in sign language and standard UEQ can be claimed to be comparable. However, considering the analysis and the origin of the differences, the way how results are calculated, and the sample size obtained, it can be expected that with the sample size of obtained group above 30 it would be sufficient to get reliable results with all scale consistency coefficients above 0.7.

It should also be noted that very high UEQ values given to both evaluated solutions could mostly be contributed to the lack of such solutions in real life and consequently the deaf community is gratefully embracing any solution that would help them in everyday life.

Related to this aspect is other information obtained during the interviews with participants. It includes a high praise for both, the developed applications, which were evaluated, as well as for the evaluation procedure using sign language version of the UEQ questionnaire. Namely, all deaf participants have expressed the wish to participate in the development of such interactive solutions, if possible, without the presence of human sign language interpreter which, in addition, represents a financial cost. Therefore, the need to use sign language-based evaluations seems even greater and more relevant, which is based on the proof of the Theorem3, stating that deaf users will prefer the user experience evaluation procedure when using UEQ in sign language.

Other interview observations included some suggestions for different sign language translations of individual UEQ statements given by seven deaf participants. Despite conducting an extensive study, choosing only official signs, the preliminary study, and testing, there were some additional suggestions to change particular sign interpretations for 'inventive', 'conventional', 'motivating', 'organized', and 'cluttered' UEQ statements. On the other hand, participants understood statements 'secure' and 'not secure' but did not understand their meaning in the context of user experience. For statements 'meets expectations' and 'does not meet expectations' 48% of participants stated that they understood the statements but claimed they did not have any expectations of the evaluated applications. Besides, it was noticed that for the UEQ questions, whose meaning was not clearly understood, the participants answered randomly, which decreased the statistical significance of the data.

For the future use of the UEQ evaluations with deaf participants it would be recommended to use a combined UEQ questionnaire (statements in text and sign language together). Due to the verification process of the UEQ in sign language adequacy, only sign language interpretations were used in the experiment. The analysis of simultaneous vs. sequential statement pair playback has shown, that only 5% of the participants preferred the simultaneous option, while all other participants stated that they prefer the sequential playback of signed pair statements.

The results of this study have been satisfactory and have given some useful directions for future work in this domain. First and foremost, it is planned to extend the study with additional participants with an updated experiment to confirm the existing and obtain statistically more reliable results. Additionally, some of the sign language translations could be modified as suggested by some of the participants and additional evaluations using updated statements could then be conducted. Finally, both evaluated interactive applications show great promise for their usefulness and wider adoption by the Deaf community and should be developed further.

5. Conclusions

The domain of the presented research is becoming increasingly relevant as the Web and digital interactive solutions in general are becoming key elements of communication and everyday life. As an estimated 80 million people in the EU alone live with a disability, it is more necessary than ever to ensure that everyone has equal access to digital products and services. The EU Web Accessibility Directive (Directive (EU) 2016/2102) [25,26] is a step into this direction, which aims to consolidate accessibility standards, making web accessibility a legal requirement. Consequently, evaluations of such solutions are very important and making the evaluations themselves more accessible is another step in the right direction. Furthermore, with a few improvements of the evaluation method presented in this study such evaluations could be adapted for different platforms and use cases and as such, could be used to reach a wider population without the need of a sign interpreter present and would also increase the inclusiveness of the Deaf and Hard of hearing community.

Author Contributions: "Conceptualization, Klemen Pečnik, Matevž Pogačnik and Žana Juvan; methodology, Klemen Pečnik and Matevž Pogačnik; software, Klemen Pečnik and Žana Juvan; validation, Klemen Pečnik and Žana Juvan; statistical analysis, Gregor Dolinar and Klemen Pečnik; resources, Klemen Pečnik and Žana Juvan; data curation, Klemen Pečnik and Gregor Dolinar; writing—original draft preparation, Klemen Pečnik, Matevž Pogačnik and Žana Juvan; writing—review and editing, Klemen Pečnik, Matevž Pogačnik and Žana Juvan; visualization, Žana Juvan; supervision, Matevž Pogačnik; project administration, Klemen Pečnik; All authors have read and agreed to the published version of the manuscript."

Funding: "This research received no external funding"

Institutional Review Board Statement: "The study was conducted in accordance with the Declaration of Helsinki, and in accordance with the Guidelines for ethical conduct in research involving people, defined by the ethics committee of the University of Ljubljana."

Informed Consent Statement: "Informed consent was obtained from all subjects involved in the study."

Data Availability Statement: Raw experiment data from SQL database and demographic data (scanned, and only in Slovenian language) available upon request.

Acknowledgments: The authors would like to sincerely thank the Association of Deaf Teachers in Slovenia for their help in promoting this research and gathering of study participants.

Conflicts of Interest: "The authors declare no conflict of interest."

6. References

1. Initiative (WAI), W.W.A. WCAG 2 Overview Available online: https://www.w3.org/WAI/standards-guidelines/wcag/ (accessed on 16 August 2022).
2. World Health Organization Deafness and Hearing Loss Available online: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss (accessed on 16 August 2022).
3. ZDGNS Gluhost Available online: http://zveza-gns.si/o-zvezi/o-gluhoti/ (accessed on 16 August 2022).
4. Godoi, T.; Silva Junior, D.; Valentim, N. A Case Study About Usability, User Experience and Accessibility Problems of Deaf Users with Assistive Technologies. In; 2020; pp. 73–91 ISBN 978-3-030-49107-9.
5. Varuh človekovih pravic RS Varuh Svetina v luči mednarodnega dneva o uresničevanju pravic gluhih in naglušnih Available online: https://www.varuh-rs.si/index.php?id=1965&L=tseexcwihoggpxrj&tx_news_pi1%5Bnews%5D=6252&tx_news_pi1%5Bcontroller%5D=News&tx_news_pi1%5Baction%5D=detail&cHash=143ccdd6c11be6b04b2d635baa64545b (accessed on 22 September 2022).
6. Ramovš, F. Slovar slovenskega knjižnega jezika 2021.
7. Schlenker, P. Visible Meaning: Sign Language and the Foundations of Semantics. *Theor. Linguist.* **2018**, *44*, 123–208, doi:10.1515/tl-2018-0012.
8. International Sign Available online: https://www.eud.eu/about-us/eud-position-paper/international-sign-guidelines/ (accessed on 25 July 2018).
9. RTV Slovenija Slovenski znakovni jezik je del jezikovne dediščine Slovenije Available online: https://www.rtvslo.si/dostopno/clanki/slovenski-znakovni-jezik-je-del-jezikovne-dediscine-slovenije/583831 (accessed on 21 September 2022).
10. ZZTSJZ Poklicna kvalifikacija tolmač Available online: https://www.tolmaci.si/poklicna-kvalifikacija-tolmac/ (accessed on 22 September 2022).

11. NECC What Is a Sign Language Interpreter? Available online: https://facstaff.necc.mass.edu/faculty-resources/deaf-and-hard-of-hearing-services-resources/what-is-a-sign-language-interpreter/ (accessed on 23 September 2022).
12. Brooke, J. SUS - a Quick and Dirty Usability Scale. In; 1996; pp. 189–194.
13. Lewis, J. Psychometric Evaluation of the Post-Study System Usability Questionnaire: The PSSUQ.; January 1 1992; Vol. 2, pp. 1259–1263.
14. Hart, S.G. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2006**, *50*, 904–908, doi:10.1177/154193120605000909.
15. Kirakowski, J.; Corbett, M. SUMI: The Software Usability Measurement Inventory. *Br. J. Educ. Technol.* **2006**, *24*, 210–212, doi:10.1111/j.1467-8535.1993.tb00076.x.
16. PIADS – Updates, News and Resources on the Use of PIADS to Measure Impact of Assistive Devices Available online: http://piads.at/ (accessed on 21 September 2022).
17. Team UEQ User Experience Questionnaire (UEQ) Available online: https://www.ueq-online.org/ (accessed on 18 August 2022).
18. Schrepp, M.; Hinderks, A.; Thomaschewski, J. Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios.; June 22 2014; pp. 383–392.
19. Oliveira, T.; Escudeiro, P.; Escudeiro, N.; Rocha, E.; Barbosa, F.M. Automatic Sign Language Translation to Improve Communication. In Proceedings of the 2019 IEEE Global Engineering Education Conference (EDUCON); April 2019; pp. 937–942.
20. ZDGNS Zveza društev gluhih in naglušnih Slovenije Available online: http://zveza-gns.si/ (accessed on 24 August 2022).
21. Alves, A.; Ferreira, S.; Veiga, V.; Silveira, D. Evaluation of Potential Communication Breakdowns in the Interaction of the Deaf in Corporate Information Systems on the Web Http://Www.Sciencedirect.Com/Science/Article/Pii/S1877050912007892. *Procedia Comput. Sci.* **2012**, *14*, 234–244, doi:10.1016/j.procs.2012.10.027.
22. Mohamad, N.; Hashim, N.L. UX Testing for Mobile Learning Applications of Deaf Children. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, doi:10.14569/IJACSA.2021.0121134.
23. Schnepp, J.; Shiver, B. Improving Deaf Accessibility in Remote Usability Testing. In Proceedings of the The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility - ASSETS '11; ACM Press: Dundee, Scotland, UK, 2011; p. 255.
24. VINTAR, Š. Lexical Properties of Slovene Sign Language: A Corpus-Based Study. *Sign Lang. Stud.* **2015**, *15*, 182–201.
25. United Nations Convention on the Rights of Persons with Disabilities and Optional Protocol 2006.
26. Oh, J.; Jeon, S.; Kim, B.; Kim, M.; Kang, S.; Kwon, H.; Kim, I.; Song, Y. Avatar-Based Sign Language Interpretations for Weather Forecast and Other TV Programs. *SMPTE Motion Imaging J.* **2017**, *126*, 57–62, doi:10.5594/JMI.2016.2632278.
27. Orero, P.; Martín, C.A.; Zorrilla, M. HBB4ALL: Deployment of HbbTV Services for All. In Proceedings of the 2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting; June 2015; pp. 1–4.
28. Schrepp, D.M. User Experience Questionnaire Handbook. 15.
29. Douglas, S.P.; Craig, C.S. Collaborative and Iterative Translation: An Alternative Approach to Back Translation. *J. Int. Mark.* **2007**, *15*, 30–43, doi:10.1509/jimk.15.1.030.
30. RedOrbit HbbTV Emulator Available online: https://chrome.google.com/webstore/detail/redorbit-hbbtv-emulator/mmgfafehampkahlmoahbjcjcmgmkppab?hl=en (accessed on 15 June 2022).
31. HbbTV Specifications Available online: https://www.hbbtv.org/resource-library/specifications/ (accessed on 20 August 2022).
32. Guidelines for Positioning of Sign Language Interpreters in Conferences Including Web-Streaming.Pdf.
33. HBB4ALL Sign Language Interpreting in HBBTV. 18.
34. Sousa, V.; Rojjanasrirat, W. Translation, Adaptation and Validation of Instruments or Scales for Use in Cross-Cultural Health Care Research: A Clear and User-Friendly Guideline. *J. Eval. Clin. Pract.* **2011**, *17*, 268–274, doi:10.1111/j.1365-2753.2010.01434.x.
35. Equal Dreams Placement of Sign Language Interpreters in Media.
36. HBB4ALL Sign Language Interpreting in HbbTV. 18.
37. Quandt, L.C.; Willis, A.; Schwenk, M.; Weeks, K.; Ferster, R. Attitudes Toward Signing Avatars Vary Depending on Hearing Status, Age of Signed Language Acquisition, and Avatar Type. *Front. Psychol.* **2022**, doi:10.3389/fpsyg.2022.730917.
38. DHI 11th DHI Conference in Ljubljana and Zagreb 2021 Available online: https://www.dhi2022slocro.com (accessed on 25 August 2022).
39. Zach T-Distribution Table. *Statology* 2018.
40. The Pennsylvania State University Multivariate Paired Hotelling's T-Square Available online: https://online.stat.psu.edu/stat505/lesson/7/7.1/7.1.8 (accessed on 23 August 2022).
41. Sign Language Network Guidelines for Positioning of Sign Language Interpreters in Conferences, Including Web-Streaming. **2016**.

42. Orero, P.; Martín, C.A.; Zorrilla, M. HBB4ALL: Deployment of HbbTV Services for All. In Proceedings of the 2015 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting; June 2015; pp. 1–4.
43. EUR-Lex Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the Accessibility of the Websites and Mobile Applications of Public Sector Bodies Available online: https://eur-lex.europa.eu/eli/dir/2016/2102/oj (accessed on 26 August 2022).
44. What Is Directive (EU) 2016/2102? Available online: https://directive2102.eu (accessed on 26 August 2022).