

# Dataset on the Validation and Standardization of the Questionnaire for the Self-Assessment of Service-Learning Experiences in Higher Education (QaSLu)

Roberto Sánchez-Cabrero , Elena López-de-Arana Prado , Pilar Aramburuzabala  and Rosario Cerrillo 

Faculty of Education and Teacher Training, Autonomous University of Madrid, 28049 Madrid, Spain; elena.lopezdearana@uam.es (E.L.-d.-A.P.); pilar.aramburuzabala@uam.es (P.A.); charo.cerrillo@uam.es (R.C.)

\* Correspondence: roberto.sanchez@uam.es

**Abstract:** This dataset shows the original validation and standardization of the Questionnaire for the Self-Assessment of Service-Learning Experiences in Higher Education (QaSLu). The QaSLu is the first instrument to measure university service-learning (USL), validated following a strict qualitative and quantitative process by a sample of experts in USL and generating rating scales for different profiles of professors. The Delphi method was used for the qualitative validation by 16 academic experts, who evaluated the relevance and clarity of the items. After two consultation rounds, 45 items were qualitatively validated, generating the QaSLu-45. Then, 118 instructors from 43 universities took part as the sample in the quantitative validation procedure. Quantitative validation was carried out through goodness-of-fit measures using confirmatory factor analysis and the final configuration optimized using one-factor robust exploratory factor analysis, determining the most optimal version of the questionnaire under the law of parsimony, the QaSLu-27, with only 27 items and better psychometric properties. Finally, rating scales were calculated to compare different profiles of USL professors. These findings offer a valid, strong, and trustworthy instrument. The QaSLu-27 may be helpful for the design of USL experiences, in addition to facilitating the assessment of such programs to enhance teaching and learning processes.

**Dataset License:** License under CC0 format.

**Keywords:** service-learning; higher education; educational evaluation; validity; internal consistency; standardization



**Citation:** Sánchez-Cabrero, R.; López-de-Arana Prado, E.; Aramburuzabala, P.; Cerrillo, R. Dataset on the Validation and Standardization of the Questionnaire for the Self-Assessment of Service-Learning Experiences in Higher Education (QaSLu). *Data* **2024**, *9*, 108. <https://doi.org/10.3390/data9090108>

Academic Editor: Kesheng (John) Wu

Received: 9 July 2024

Revised: 23 August 2024

Accepted: 12 September 2024

Published: 19 September 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Summary

Service-learning (SL) is an instructional method that enhances students' learning of academic subjects through real-world applications, strengthening their commitment to their studies [1,2]. Moreover, it serves as a tool for social change, advancing social justice and sustainable development [3]. Grounded in experiential learning, SL facilitates students in linking classroom knowledge with societal needs, thus promoting civic engagement through intentional social action [4–6]. SL transcends the traditional classroom setting by involving students in community activities, resulting in mutually beneficial outcomes for students and the community [4,7,8].

SL has gained popularity in the higher education sector in recent years [4,7,8], and it may have influenced the current educational legislation. For instance, within the European context, the Renewed Agenda for Higher Education by the European Commission emphasizes the importance of community collaboration and encourages institutions to broaden their civic engagement [9].

This framework makes it clear that there is a need for verified tools to assess the caliber of SL-based university educational initiatives [10]. Even if they do not adhere to the guidelines of this paradigm, certain educational experiences are nonetheless called “SL”;

in these instances, “SL” is more closely associated with volunteer work or fieldwork [11]. Because of this, tools that adhere to the pedagogical foundations of USL are required for the planning, execution, and assessment of university programs [12]. Since most of these data are currently collected from students [13], it would be interesting to hear from instructors.

The Questionnaire for the Self-Assessment of Service-Learning Experiences in Higher Education (QaSLu) [14] is the first instrument to measure university service-learning (USL), validated qualitatively by means of the Delphi method [10,15,16] and quantitatively by means of robust factor analysis [17]. This dataset shows anonymized attributive data on 118 experts in USL, considering their age, gender, academic studies, and level of expertise in USL experiences; the type of the collaborating institution or organization; the intervention on which the service is based; the recipients of the service activity; whether they are involved in online or face-to-face USL; and their answers to the 45 items of the QaSLu.

This dataset on the initial validation and standardization of the QaSLu includes the complete dataset of the original 118 participants’ responses to the 45 items and their overall scores. However, it also includes the total score for the reduced version (QaSLu-27), with only 27 items and better psychometric properties, obtained by one-factor robust exploratory factor analysis.

These data offer valid, strong, and trustworthy initial validation for the instrument, that may be helpful for the design of USL experiences, in addition to facilitating the assessment of such programs for the purpose of enhancing teaching and learning processes.

Additionally, this dataset provides scales, a feature that was previously unavailable that allows teachers to assess their USL experiences based on both their own and the context’s characteristics.

### *1.1. Theoretical Background: The Assessment of SL in Higher Education*

While practitioners and researchers can agree on several quality components of SL, there is currently no standardized quantitative tool available to evaluate how much a course combines these essential components of high-quality practice. Most of the instruments currently in use are qualitative and/or narrowly focused on a select few essential elements or specific fields.

For example, Shumer [18] reported on a three-year project developing the Quintessential Elements of SL, a self-assessment tool with 23 statements in 5 domains for SL practitioners in K–12 settings. Nevertheless, the primary purpose of this instrument was program improvement, as it only allowed for self-rating each item as “strong”, “weak”, or “needs work”. A staged “checklist for planning, implementing, and evaluating SL” was created by Jenkins and Sheehey [19]; it is meant to be used in course design and does not include any ratings.

A rubric for evaluating course syllabi for quality and evidence-based indicators of SL components as found in the literature was created by Kieran and Haack [20]. With scoring options of “excellent”, “satisfactory”, and “developing”, their PRELOAD rubric covers partnership, reflection, engagement, logistics, objectives, assessment, and the definition of SL as important. Despite this, the rubric is still focused on syllabus design rather than actual implementation.

A team lead by Stokamer [21] at her institution developed a set of 10 Principles of Quality Academic Civic Engagement (PQACE) that were tailored to their particular university setting and based on “the SL Civic Engagement literature, best practices, and personal experience”. Botelho et al. [22] identified eight elements of SL quality in STEM courses across the California State University system using surveys of students and faculty, as well as course curricula. These included single-item components (“SL preparation” and “linked to learning objectives”), as well as composite measures (“reflections”, “values focus”, “collaboration with community”, “addressing community need”, “linked to academic content”, and “communication with community”). Based on an analysis of STEM curricula and post-participation student surveys, each measure could be rated on a scale of 1 to 4 (or 5).

The Service-Learning Quality Assessment Tool (SLQAT), which was created to offer a way to assess the quality of the planning and execution of academic SL experiences that provide credit, was explained recently by Furco et al. [23]. This tool considers 28 components that the SL literature considers necessary for high-quality SL that promote beneficial academic and other outcomes for students and organizes them into 5 dimensions. Every component also has an underlying numerical value or weight that indicates how much it is thought to contribute to the development and execution of high-quality SL experiences. The SLQAT can be utilized for a variety of tasks, including faculty development, course design, teacher self-study, and research with either independent or dependent variables.

In Matthews [24], the development of the SLQAT is described as a five-year multi-institutional process. The selection and operational definitions for these elements are discussed, as well as the assumptions and decisions that went into the instrument's creation, the use of expert feedback to create baseline weights that represent the relative importance of each element's contribution, the establishment of rating levels that represent element quality, and the development of protocols for the scoring and use of the instrument all being covered.

In conclusion, the literature shows that there are tools for assessing SL in higher education, but many of them have not been validated. Few reliable and accurate tools are available, even though numerous tools are used to evaluate different aspects of USL. Among the validated tools that assess the quality of the SL experiences created and offered in higher education, some of them are particularly noteworthy.

Escofet-Roig et al. [25]'s questionnaire, for instance, has sixteen items and evaluates university students' experiences with SL in three areas: participation, service, and competencies. Transversally, a fourth dimension is introduced, referring to their overall satisfaction with engagement in the USL experience. These 16 items all adhere to the "funnel" process, starting with the broadest and ending with the most specific, and they all meet the condition of having the shortest form feasible.

A team of experts in the field checked the content validity (validation through expert judgment). The participants were eight teachers who worked in USL and came from various fields. A second version of the instrument was created using the judges' suggestions.

Its reliability was tested using an empirical validation approach for the second version. In a pilot study, 116 university students—84.5% female and 15.5% male—participated in USL experiences, and the questionnaire was administered to them all. Cronbach's alpha coefficients were calculated after the internal consistency of the items was analyzed to assess the instrument's reliability ( $\alpha = 0.9$ ). Acceptable reliability was suggested by the high correlation shown by the index obtained for all the items. Every item's contribution to its corresponding scale (the corrected homogeneity index, which offers the ability to discriminate) was always positive.

Another validated tool is Rodríguez-Izquierdo's validated scale for assessing how USL events affect the growth of student teachers' professional competence [26]. The five components of this scale are as follows: (1) ethical commitment, (2) cooperation with other professionals, (3) the design and development of experiences, (4) readiness for diversity, and (5) readiness for professional development. In total, 366 social studies students took part. A simple, stratified, multistage random sampling procedure was used to choose the sample.

The content was sent to nine expert judges for assessment to increase its validity. The number of elements was reduced and/or they were rearranged in accordance with the suggestions received, considering placement, intelligibility, univocity, and validity requirements. Consequently, a two-part instrument was designed, comprising a section for demographic data and a second section with a Likert-type scale consisting of thirty items and five points, in which students evaluated the degree to which USL supported the development of their professional competency.

Using the Varimax rotation principal component extraction approach, an exploratory factorial study was performed to confirm the construct's validity. Furthermore, for parametric

ter estimation, various confirmatory factorial analyses (CFAs) were employed in accordance with the maximum likelihood criterion.

The scale is an accurate and trustworthy tool for evaluating professional competencies. The whole instrument's Cronbach's alpha was 0.87, and the values for each of the dimensions varied from 0.84 to 0.91. With a  $\chi^2 = 881.22$ ,  $p = 0.00$ , GFI = 0.93, CFI = 0.98, SRMR = 0.067, and RMSEA = 0.064, the CFA's fit was quite excellent. The author listed the following as some of the drawbacks of Cronbach's alpha: the study used a sample of students from a single educational program, and its quasi-experimental approach limited the development of causal links.

The Utrecht Work Engagement Scale for Students (UWES-S-9) was validated by the same author, Rodríguez-Izquierdo [2], a year later. It consists of nine components organized into three dimensions: (a) vigor: level of energy, persistence, and effort in performing academic tasks; (b) dedication: high level of involvement in studies and their career; and (c) absorption: high level of concentration and immersion in what they do when they study. These are all rated on a 7-point Likert scale, where 0 represents not at all/never and 6 represents every day/always.

A total of 342 students from Pablo de Olavide University (UPO) in Seville, Spain, made up the sample, and 183 of them held a bachelor's degree in social studies, while 153 had a double degree in social studies and social work. A multistage, random, stratified sampling process was used, with the most pertinent student characteristics—gender, year, age, and mode of entry—categorized into several strata.

Using two randomly selected subsamples, the internal organization of the assessment tool was verified. An exploratory factor analysis was conducted on the first half of the sample ( $n_1 = 178$ ) to ascertain the number of factors. The instrument corresponded to a one-dimensional structure according to the data (its ECV values ranged from 0.70 to 0.85; its MIREAL values were less than 30). On the second half of the sample ( $n_2 = 164$ ), CFA was carried out. The S-B2 values were greater than 0.01, the NNFI and CFI values were equivalent to or greater than 0.95, and the SRMR and RMSEA values were less than 0.08, indicating very good findings.

Cronbach's alpha and McDonald's omega were calculated for reliability, composite reliability (CR), and maximum reliability (MR), and the instrument produced satisfactory results. Lastly, appropriate results from the calculation of its discriminant validity were obtained.

León-Carrascosa et al. [27] have developed and validated an instrument for assessing the educational benefit of SL in higher education. The three basic dimensions of this tool are training (as an objective), learning (as a means), and service (as a commitment to the community).

A sample of 180 students from 9 Spanish universities in the regions of Madrid, Valencia, Catalonia, and Galicia participated in the study. Incidental non-probabilistic sampling was used. The instrument, which consists of 35 items, is designed to assess the respondents' approach to service-learning. Students answer on a Likert-type scale from 1 to 5.

Expert judges in educational research and specialists in USL in higher education assessed the instrument's content validity. With a Cronbach's alpha of 0.95, the instrument's internal consistency was excellent. Following the CFA, its reliability was examined, yielding satisfactory results across all dimensions and an excellent final value ( $\alpha$  Global = 0.95,  $\alpha$  Formative Dimension = 0.88,  $\alpha$  Learning Dimension = 0.90,  $\alpha$  Service Dimension = 0.91).

The scale developed and validated by Santos-Pastor et al. [28] for evaluating college students' perceptions of the influence SL initiatives have had on their learning experience, as well as their social and personal growth, is another significant contribution. With 41 items and 1 open-ended question, this tool makes it possible to validate and objectify the tangible consequences that this pedagogical approach has according to the formative, professional, personal, and community dimensions. A total of 200 students from 5 Spanish institutions who had taken part in various USL programs, including physical activity with underprivileged populations, were included in its evaluation. Convenience sampling that

conformed to Nunnally's guidelines was applied. Six USL experts, including two from the pedagogical field and one from the physical activity field, evaluated the content critically in order to validate it.

Each of the global scale's dimensions was confirmed, and a confirmatory factor analysis was conducted to evaluate the theoretical dimensions' internal consistency and structure. The outcomes validated the suitability of the selected structural model (RMSEA = 0.08) and the reliability of the scale ( $\alpha = 0.95$ ), as well as that of every dimension ( $\alpha$  ranging from 0.68 to 0.86). The internal consistency of the scale was examined using Cronbach's alpha, which was also determined for each of its dimensions. For the global E-ASAF scale, Cronbach's alpha  $\alpha = 0.95$ .

The selection of the participants, who were chosen from a convenience sample of students who had taken part in physical-education-related USL activities, was the study's main weakness because there are not many of this kind of participant in a university setting.

Also assessed was the psychometric validation of Ruiz-Ordoñez et al. [29]'s VAL-U instrument. The necessity to create tools for evaluating how the USL approach affects college students' civic attitudes and values led to the creation of this particular tool. It has twenty questions with five-point Likert-scale responses. Using information from 162 university students, the authors examined the instrument's internal consistency and factorial structure. The tool was validated as having appropriate psychometric characteristics. After the factor analysis, three factors were identified, and Cronbach's alpha was 0.67.

Furthermore, it is important to take into account the work of Gul et al. [13], who designed and validated a tool for assessing USL management. There were 315 teachers in their validation sample. Its items were chosen using the deductive method. The factor structure of the scale was investigated using exploratory factor analysis (EFA). A total of 21 items in a 4-factor structure were found. This scale had very high reliability, as evidenced by the Cronbach's alpha of 0.96 obtained.

Although there are instruments intended to evaluate SL experiences in higher education, these are rare. Moreover, there are few that have undergone qualitative validation before psychometric validation and with their quantitative and reliability validation have optimal properties.

About qualitative validation, it is usually unknown whether the groups of experts who validated the designs were part of the research teams or not, and these have often been small groups made up of 6–8 people. Another frequent deficiency is related to the sample since most of the validations have been made based on university students with little experience in USL, while studies that have included teachers are quite infrequent.

Regarding quantitative validation, most of these analyses explored the factorial structure of the designed instruments without optimizing their final structure. Many lack a robustness analysis of the items and refine their final configuration towards a single factor; hence, the final scores obtained do not adequately distinguish the quality of the USL experiences and include noise, distractions, or impurities from the other latent factors involved.

Another advantage of optimizing the design of the instrument towards a single factor is the possibility of creating scales that measure the implementation of SL experiences in higher education, a rare feature in this field.

Among the existing questionnaires, the QaSLu-27 is particularly short. There are a few even shorter instruments, with 8, 16, 20, and 21 items, respectively, created by Rodríguez-Izquierdo [2], Escofet-Roig et al. [25], Ruiz-Ordoñez et al. [29], and Gul et al. [13], but these tools have not undergone an item optimization process without affecting the psychometric properties of their design. They were simply designed to be short without assessing their possible conceptual shortcomings and without assessing whether their reliability and validity would be compromised compared to those of longer instruments. The QaSLu-27 includes 27 items, which is considered a balance when compared to other validated instruments, such as the 30-item questionnaire by Rodríguez-Izquierdo [26], the 35-item instrument by León-Carrascosa et al. [27], and the 41-item questionnaire with an open question by Santos-Pastor et al. [28].

Notably, the QaSLu-27 was validated using the Delphi technique, just like previous instruments like those created by Rodríguez-Izquierdo [26], Escofet-Roig et al. [25], León-Carrascosa et al. [27], and López-de-Arana et al. [10]. But the QaSLu-27 has also been exposed to a factorial analysis, in accordance with several researchers' approaches [2,10,25–27]. While authors usually employ CFA to validate their questionnaires, this work has attempted to optimize the QaSLu-27 to ensure its maximum validity, robustness, and reliability. This approach aligns with Rodríguez-Izquierdo's work [2,25].

Regarding reliability, the QaSLu-27's limited number of items did not affect its level of reliability. Given that Cronbach's alpha is particularly susceptible to this problem, the QaSLu-27's Cronbach's alpha coefficient of 0.92 is exceptionally high. The tool created by Rodríguez-Izquierdo [26], despite having more items, had a slightly lower reliability (0.87); in contrast, the tool validated by Rodríguez-Izquierdo [26] with fewer items had a much lower reliability (0.70). This is evident when comparing the reliability of the QaSLu-27 with other instruments. Despite them having fewer components, the instruments created by Ruiz-Ordóñez et al. [29] and Escofet-Roig et al. [25] showed values of about 0.90 and 0.67, respectively.

Although it should be mentioned that they include more items, the proposals of León-Carrascosa et al. [20] and Santos-Pastor et al. [28] obtained a higher reliability (0.95). With a reliability of 0.96, only the instrument developed by Gul et al. [13] outscored the QaSLu-27; this is likely because of their study's much larger sample size.

Therefore, analysis of the previous literature reflects the need to develop new instruments, optimizing their design qualitatively and quantitatively to solve most of the deficiencies observed.

## 1.2. Methods for Validation and Standardisation

This dataset was collected from a representative sample of 118 participants selected according to its appropriateness for the target population. Two main methods were used in the data validation process, and another method used was standardization. In the validation process, we used the Delphi method, due to its scientific guarantee when dealing with qualitative validation, and factor analysis, as this is the most accepted method in the scientific field for carrying out quantitative validation. Finally, for standardization, we proceeded by creating different rating scales considering the attributive characteristics of the participating sample.

### 1.2.1. The Delphi Method

The Delphi method is a qualitative technique proposed by Dalkey and Helmer [30]. Dalkey and Helmer [30] provide a collective judgment that emerges from a group of geographically dispersed experts [31–34]. Participation in the validation process should be anonymous [35,36] to decrease the impact of social desirability and avoid responses being influenced by other people's contributions or postulates, allowing alternative thoughts to emerge [37–39]. The iterative process that is established starts with the application of a questionnaire, through which the experts' assessments are collected, which are then analyzed statistically to provide feedback to the group by sharing the results, and then decisions are made, on which their opinions will be requested again [36].

The Delphi method attempts to obtain a view on a specific topic that is as consensual as possible among different experts by conducting repeated rounds of questions [31]. This procedure responds to the rational assumption that a collective judgment is more reliable than that of a single individual [34]. The agreement or consensus that different experts reach on a topic implies that this method is based on collective intelligence, which is built on mediatized, controlled, and centralized collaboration [10].

The literature confirms the effectiveness of the Delphi method for the construction and validation of instruments, mainly in the field of educational sciences [31,35,40–43], which positions it as an ideal method for the study presented here. Moreover, as there is coherence between the principles of the Delphi method and the conceptualization of USL,

it is considered a suitable method for the validation of an instrument whose aim is the assessing the quality of SL experiences in higher education [15].

### 1.2.2. Factor Analysis

The most widely accepted method for the quantitative validation of the collection of data using a scientific instrument is known as factor analysis (FA) [44,45]. FA, in simple terms, involves analyzing the variation matrix of the results on the items of an instrument after it has been administered to a representative sample of a population. In this result matrix, there is covariation among the results of the items that follows structured patterns, known as factors. That is, the results of some items covary based on a factor or a reason that links them, which is shown through their covariance. Therefore, FA is the analysis of the existence and composition of these inherent factors in the result matrix [44,46].

There are two basic types of FA: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). EFA explores the factor structure that exists in each result matrix, while CFA aims to compare this result matrix with a predetermined structure, accepting this composition as valid if the differences are not significant. The use of EFA or CFA is determined by the research purpose. CFA is more appropriate when intending to validate a factorial structure based on the dimensions of an instrument, previously designed with that structure in mind, whereas EFA is usually employed to identify an unknown factorial structure, seeking to determine that the instrument measures what it was generally intended to measure when it was designed [46].

Both types of FA use a result matrix as the source of their analysis. However, there is a third, much more complex approach which does not base its results on a single matrix but on the combination of all possible matrices resulting from the elimination of some of the items. This is known as robust exploratory factor analysis (REFA) [47]. Under REFA, sometimes millions of matrices are analyzed, aiming to identify the items of the instrument that are least linked to the main factors and whose elimination from the instrument's composition does not significantly negatively affect the validity and goodness of fit of the instrument or whose elimination would even positively impact its psychometric properties. The use of REFA allows researchers to design reduced versions of an instrument by eliminating undesired items or factors, facilitating a robust design based on the law of parsimony.

In the creation of this database, all three types of FA were employed, demonstrating the possibilities and psychometric properties of the optimized version of the QaSLu with 27 items.

### 1.2.3. The Development of Rating Scales

The creation of rating scales for a previously validated assessment instrument allows for the comparison of a participant's results based on their most prominent characteristics and situates them on a scale that establishes their position within the continuum of their reference population [48,49]. The development of rating scales is especially useful in the clinical field for determining the severity of symptoms or the effectiveness of a treatment and in the educational field for establishing a student's standing relative to their reference group [50,51].

For rating scales to be useful, they must include various attributive variables relevant to the field of study and allow the conversion of the raw scores (RSs) into standard scores. The most used types are percentiles or Z-scores.

This process involves creating double-entry tables where the intersection of the scores using the instrument, along with the various attributive variables considered, results in different distributions of the scores within the instrument's scoring range.

For the creation of this dataset, the variables that defined the profile of a USL experience manager the best and conversion from RSs into percentiles were used.

### 1.3. List of Publications Based on the Dataset

- López-de-Arana Prado, E.; Aramburuzabala Higuera, P.; y Opazo Carvajal, H. Diseño y validación de un cuestionario para la autoevaluación de experiencias de aprendizaje-servicio universitario. *Educ. XX1* **2020**, *23*, 319–347. <http://doi.org/10.5944/educXX1.23834> [10].
- López-de-Arana Prado, E.; Aramburuzabala Higuera, P.; Cerrillo, R. Respondiendo a los Objetivos de Desarrollo Sostenible a través de la implementación del Aprendizaje-Servicio Universitario. In *El Aprendizaje-Servicio universitario ante los retos de la Agenda 2030*; Gutiérrez, J.G., Morera, F.J.A., Ramírez, A.C., Eds.; UNED-Universidad Nacional de Educación a Distancia: Madrid, Spain, 2023; pp. 568–578 Available online: <https://dialnet.unirioja.es/servlet/articulo?codigo=9332822>, accessed on 12 September 2024 [52].
- López-de-Arana Prado, E.; Aramburuzabala, P.; Cerrillo, R.; Sánchez-Cabrero, R. Validation and Standardization of a Questionnaire for the Self-Assessment of SL Experiences in Higher Education (QaSLu-27). *Educ. Sci.* **2024**, *14*, 615. <https://doi.org/10.3390/educsci14060615> [17].

### 1.4. Future Potential Benefits and Related Research Projects Based on the Dataset

This dataset can be utilized by any researcher interested in USL in various ways. Some of these are described below. However, these are not the only possibilities, as the creativity and initiative of researchers can generate new opportunities for its use.

Firstly, this dataset allows for the development of new studies on USL by using meta-analysis methodology and combining this dataset with other open-access studies that openly share their datasets. Nevertheless, it is also possible to combine it with new original results, enriching new studies on USL and expanding the samples on which conclusions are based.

The 45 items and the results from the 118 participants for each item could also be utilized to design new instruments on USL. Both reduced versions of the QaSLu and versions adapted to new populations would be feasible to create with ease, or even new instruments using some of the QaSLu items, potentially identifying better items to replace some of the existing ones.

Another possibility would be comparison with other samples with different characteristics. Whether from other countries or with different profiles, it would be possible to clearly observe their differences.

Finally, the existence of rating scales allows for various uses—for example, evaluating new experiences in USL, evaluating the work of USL managers, or creating new rating scales or modifying and expanding existing ones with larger and more representative samples.

## 2. Data Description

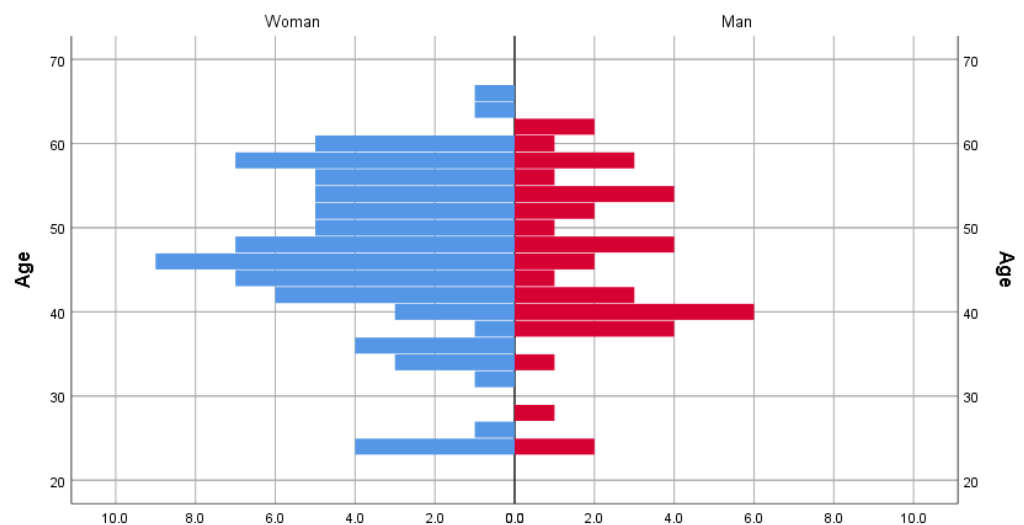
The data for this article include two different files that constitute valuable resources for researching USL. The first is a table with the raw and anonymized results from 118 university experts who participated in the original validation of the QaSLu. The second is the standardization of the QaSLu-27 scale. Both files can be downloaded from Supplementary Materials of this publication or Harvard Dataverse [14]. Both files are described in detail below:

### 2.1. Validation of the QaSLu Scale File

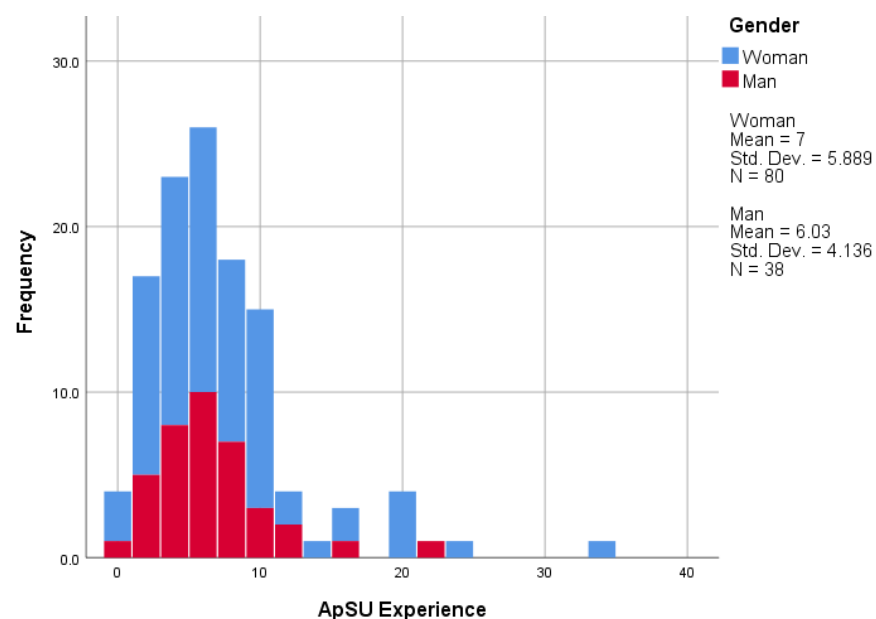
This file is presented in multiple formats to ensure its accessibility with different types of software. Specifically, it is available in .sav format for use with IBM SPSS Statistics software (the file was created with version 25 but is compatible with all current versions of the software), in .xlsx format for use with versions old Microsoft Excel post-2007, and in .csv format to facilitate its conversion for various statistical and computational software.

This file contains descriptive variables and the results of the QaSLu questionnaire from 118 participants who were experts in USL. The data are organized into rows (participants)

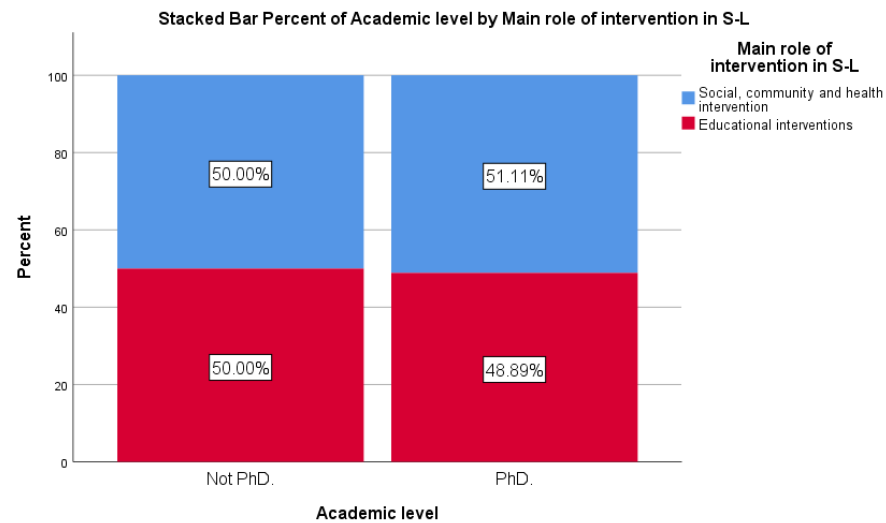
and columns (descriptive variables and QaSLu items). The variables included and their levels are age (continuous discrete variable); age distributed into groups (Under 40 years old; From 40 to 49 years old; Over 50 years old); gender (Female; Male; Other); academic level (Not PhD.; PhD.); experience in years (continuous discrete variable); level of expertise in USL experiences (Novice or beginner; Experienced); type of collaborating institution or organization (Social centers, NGOs and foundations; Schools; Alliances with Higher Schools or Universities; Government public administrations; Government public administrations); intervention on which the service is based (Social, community and health intervention; Educational interventions); the recipients of the service activity (Disadvantaged, vulnerable and disabled groups; Students); and online or face-to-face USL (SL is exclusively face to face; SL includes online activity in whole or in part). Additionally, the scores of the 45 QaSLu items using a 5-point Likert scale (Never; Seldom; Sometimes; Frequently; Always) are included, along with the total scores for the sum of the 45 items and the reduced 27-item version of the QaSLu. Figures 1–5 visually display the final data configuration according to the attributive variables considered.



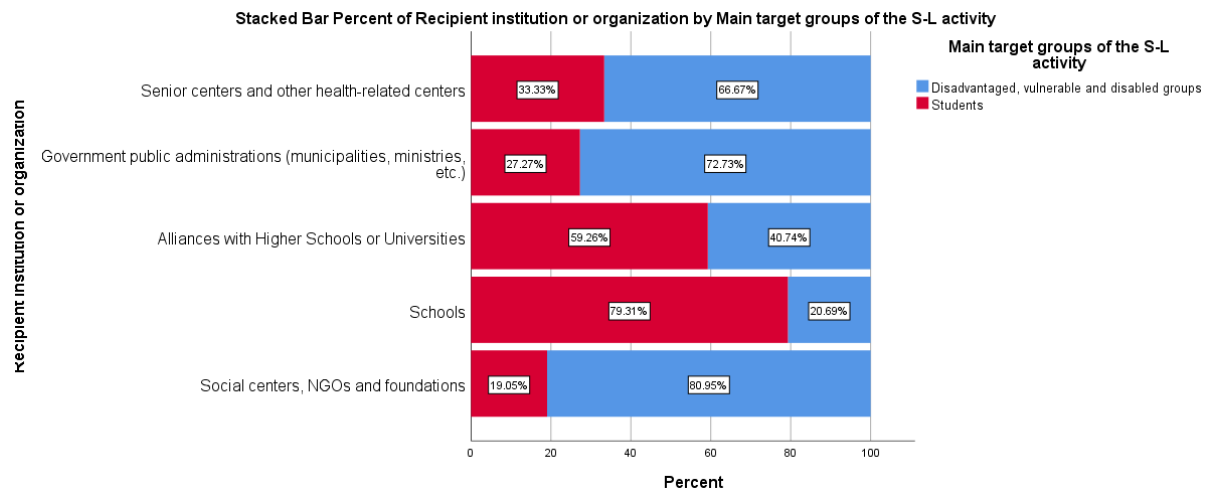
**Figure 1.** Population pyramid. Extracted from Lopez-de-Arana Prado et al. (2024) [17].



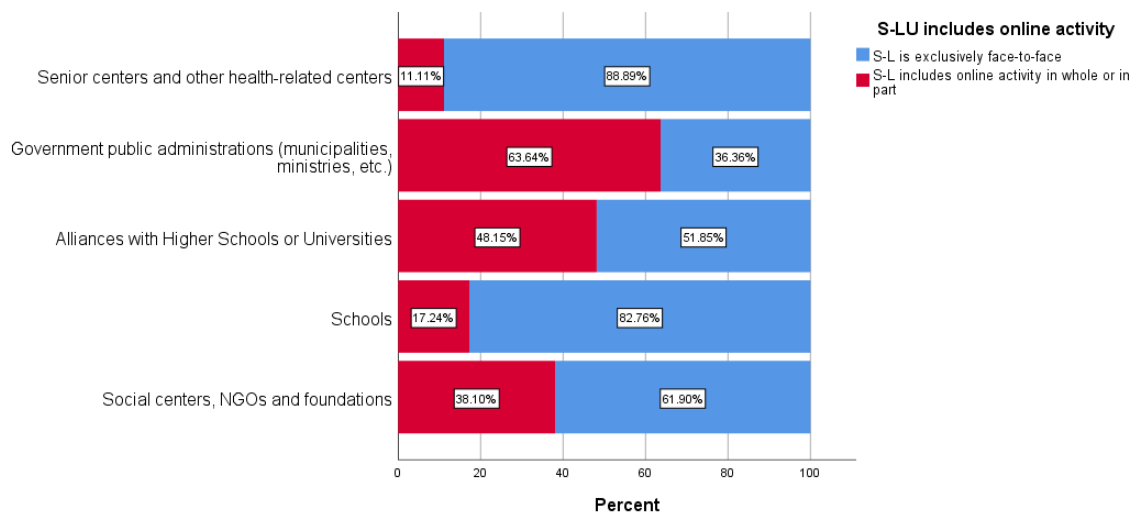
**Figure 2.** Distribution of the participant sample according to experience in SL in higher education and gender. Extracted from Lopez-de-Arana Prado et al. (2024) [17].



**Figure 3.** Distribution of the participant sample according to academic studies by intervention on which the USL service is based.



**Figure 4.** Distribution of the participant sample according to the type of the collaborating institution or organization by recipients of the service activity in USL.



**Figure 5.** Distribution of the participant sample according to type of the collaborating institution or organization by online or face-to-face USL.

## 2.2. Standardization of the QASLU-27 Rating Scale File

This file is presented in .xlsx format for use with versions of Microsoft Excel post-2007 and shows the standardization of the QaSLu-27 according to the profile of the participants to whom the test was applied, considering the different attributive variables measured and outlined in Section 2.1.

The rating scales allow for the comparison of any participant's score with the scores previously obtained by other participants with the same characteristics, placing them on a percentile-based scale and into different categories according to their level (Low, 1–25th percentile; Medium/Low, 26–44th percentile; Medium, 45–55th percentile; Medium/High, 56–74th percentile; High, 76–99th percentile).

The file includes four rating scales distributed across four Excel sheets. The four norms are (1) a gender- and age-based scale, (2) a gender- and USL-experience-based scale, (3) a scale based on gender and the type of collaborating institution in which USL is performed, and (4) a scale based on gender and the virtuality of the USL practice. The scales developed for the QaSLu-27 are listed below (Tables 1–4).

**Table 1.** Gender- and age-based scale.

| Level       | Percentile | Female    |             |           | Male      |             |           |
|-------------|------------|-----------|-------------|-----------|-----------|-------------|-----------|
|             |            | <40 Years | 40–49 Years | >50 Years | <40 Years | 40–49 Years | >50 Years |
| Low         | 1          | 0–48      | 0–68        | 0–61      | 0–48      | 0–57        | 0–54      |
|             | 10         | 49–69     | 69–71       | 62–67     | 49–52     | 58–73       | 55–59     |
|             | 20         | 70–73     | 72–78       | 68–75     | 53–64     | 74–79       | 60–70     |
| Medium/Low  | 30         | 74–76     | 79–83       | 76–80     | 65–74     | 80–85       | 71–75     |
|             | 40         | 77–80     | 84–85       | 81–86     | 75–77     | 86–89       | 76–82     |
| Medium      | 50         | 81–85     | 86–93       | 87–89     | 78–82     | 90–94       | 83–85     |
| Medium/High | 60         | 86–93     | 94–97       | 90–92     | 83–93     | 95–96       | 86–88     |
|             | 70         | 94–99     | 98–99       | 93–98     | 94–95     | 97–98       | 89–95     |
| High        | 80         | 100–103   | 100–102     | 99–100    | 96–98     | 99–104      | 96–101    |
|             | 90         | 104–107   | 103–107     | 101–107   | 99–107    | 105–107     | 102–107   |
|             | 99         | 108       | 108         | 108       | 108       | 108         | 108       |
| N = 118     |            | 16        | 32          | 32        | 10        | 14          | 14        |

Extracted from Lopez-de-Arana Prado et al. (2024) [17].

**Table 2.** Gender- and USL-experience-based scale.

| Level       | Percentile | Female              |                        | Male                |                        |
|-------------|------------|---------------------|------------------------|---------------------|------------------------|
|             |            | Beginner (<5 Years) | Experienced (>5 Years) | Beginner (<5 Years) | Experienced (>5 Years) |
| Low         | 1          | 0–48                | 0–61                   | 0–48                | 0–54                   |
|             | 10         | 49–68               | 62–75                  | 49–55               | 55–74                  |
|             | 20         | 69–70               | 76–80                  | 56–57               | 75–77                  |
| Medium/Low  | 30         | 71–73               | 81–85                  | 58–66               | 78–83                  |
|             | 40         | 74–78               | 86–88                  | 67–73               | 84–85                  |
| Medium      | 50         | 79–82               | 89–93                  | 74–82               | 86–88                  |
| Medium/High | 60         | 83–89               | 94–97                  | 83–93               | 89–93                  |
|             | 70         | 90–98               | 98–99                  | 94–95               | 94–97                  |

Table 2. Cont.

| Level   | Percentile | Female              |                        | Male                |                        |
|---------|------------|---------------------|------------------------|---------------------|------------------------|
|         |            | Beginner (<5 Years) | Experienced (>5 Years) | Beginner (<5 Years) | Experienced (>5 Years) |
| High    | 80         | 99–100              | 100–102                | 96–100              | 98–99                  |
|         | 90         | 101–107             | 103–107                | 101–107             | 100–107                |
|         | 99         | 108                 | 108                    | 108                 | 108                    |
| N = 118 |            | 31                  | 49                     | 14                  | 24                     |

Extracted from Lopez-de-Arana Prado et al. (2024) [17].

Table 3. Scale based on gender and the type of collaborating institution in which USL is performed.

| Level       | Percentile | Female                                     |                          | Male                                       |                          |
|-------------|------------|--------------------------------------------|--------------------------|--------------------------------------------|--------------------------|
|             |            | Social, Community, and Health Intervention | Educational Intervention | Social, Community, and Health Intervention | Educational Intervention |
| Low         | 1          | 0–65                                       | 0–48                     | 0–54                                       | 0–48                     |
|             | 10         | 66–70                                      | 49–68                    | 55–61                                      | 49–63                    |
|             | 20         | 71–77                                      | 69–74                    | 62–73                                      | 64–72                    |
| Medium/Low  | 30         | 78–80                                      | 75–79                    | 74–78                                      | 73–76                    |
|             | 40         | 81–85                                      | 80–84                    | 79–84                                      | 77–83                    |
| Medium      | 50         | 86–92                                      | 85–88                    | 85–88                                      | 84–93                    |
| Medium/High | 60         | 93–97                                      | 89–93                    | 89                                         | 94                       |
|             | 70         | 98–99                                      | 94–97                    | 90–95                                      | 95–96                    |
| High        | 80         | 100–102                                    | 98–100                   | 96–98                                      | 97–101                   |
|             | 90         | 103–107                                    | 101–107                  | 99–107                                     | 102–107                  |
|             | 99         | 108                                        | 108                      | 108                                        | 108                      |
| N = 118     |            | 38                                         | 42                       | 22                                         | 16                       |

Extracted from Lopez-de-Arana Prado et al. (2024) [17].

Table 4. Scale based on gender and virtuality of the USL experience.

| Level       | Percentile | Female                         |                                                  | Male                           |                                                  |
|-------------|------------|--------------------------------|--------------------------------------------------|--------------------------------|--------------------------------------------------|
|             |            | SL Is Exclusively Face to Face | SL Includes Totally or Partially Online Activity | SL Is Exclusively Face to Face | SL Includes Totally or Partially Online Activity |
| Low         | 1          | 0–48                           | 0–61                                             | 0–54                           | 0–48                                             |
|             | 10         | 49–70                          | 62–68                                            | 55–68                          | 49–54                                            |
|             | 20         | 71–75                          | 69–76                                            | 69–75                          | 55–64                                            |
| Medium/Low  | 30         | 76–78                          | 77–82                                            | 76–77                          | 65–77                                            |
|             | 40         | 79–83                          | 83–88                                            | 78–84                          | 78–84                                            |
| Medium      | 50         | 84–87                          | 89–91                                            | 85–88                          | 85–93                                            |
| Medium/High | 60         | 88–94                          | 92–97                                            | 89–92                          | 94–95                                            |
|             | 70         | 95–97                          | 98–99                                            | 93–95                          | 96–97                                            |
| High        | 80         | 98–102                         | 100                                              | 96–101                         | 98–99                                            |
|             | 90         | 103–107                        | 101–107                                          | 102–107                        | 100–107                                          |
|             | 99         | 108                            | 108                                              | 108                            | 108                                              |
| N = 118     |            | 48                             | 32                                               | 28                             | 10                                               |

Extracted from Lopez-de-Arana Prado et al. (2024) [17].

### 3. Methods

#### 3.1. Design of the Instrument for Collecting Information

QaSLu is the acronym for Questionnaire for Self-Assessment of SL Experiences in Higher Education (“Cuestionario para la Autoevaluación de Experiencias de Aprendizaje-Servicio Universitario” in Spanish) and previously underwent strict qualitative and quantitative

validation. This instrument was designed to provide a single final score so that any user could rate their USL experience and compare it to their reference population, which is why they are not considered dimensions, as the existence of dimensions would have prevented users from being able to self-assess and score themselves on a reference scale.

The design of the questionnaire was based on existing tools for the self-assessment of USL experiences. Some of the sources specified criteria for quality assurance of USL activities or practices [53,54]. Others were rubrics [55–57].

Content analysis [58] was carried out with the item as the units of the analysis. Although it is assumed and accepted that the trajectory of each USL experience is unique, there are orientative stages for its design. Following CLAYSS' proposal [59], the content of the items was categorized as belonging to the following stages: (1) the pre-experience stage (in this stage, the focus is to motivate university students and to diagnose the social needs in the community); (2) planning (this stage is for planning all the necessary steps for developing the USL experience); (3) execution (in this stage, the people involved develop the USL experience); and (4) closure and multiplication (in this stage, the USL experience is finished with a celebration, in which all the people involved participate, the last evaluation is carried out, and the university teacher and community partners reflect on the USL experience's continuity and multiplication). These four commonly accepted stages in the USL process served as theoretical categories for the design of the questionnaire. However, the QaSLu was designed to give a single final score to enable university teachers to evaluate themselves in the exercise of their USL experiences.

Once this task was completed, new items were drafted, trying to encompass all the nuances. The initial Questionnaire for the Self-Assessment of SL Experiences in Higher Education was constituted as follows: (1) the objective of the questionnaire and instructions, (2) identification of the person filling it in, (3) data on the SL experience, (4) self-evaluation of the experience through 54 items that are assessed through a Likert-type scale with 5 response options, and (5) acknowledgments.

In the qualitative validation by means of the Delphi method, the reliability of the results depends on the suitability of the people chosen as experts [60], which makes it necessary to define the profile of the members of this group. This study determined that the experts had to meet the following characteristics: (1) had at least 4 years of experience in research [61] and the design and implementation [62,63] of USL experiences; (2) were part of the university academic staff; and (3) were a doctor or a doctoral candidate.

Ludwig [64], in a review of the Delphi method, concluded that in most studies, the group consists of between 15 and 20 people. Therefore, 20 people who matched the pre-established characteristics were contacted. Ultimately, 16 completed the process in its entirety, with 75% being female experts and 25% male experts in USL.

The study followed the four characteristic phases of the Delphi method [33,34]. In the definition phase, it was established that two types of validity would be analyzed: on the one hand, construct validity, following the process explained in the study by Bakieva et al. [41], and on the other, content validity, for which the work of Gil and Pascual-Ezama [65] was taken as a reference.

The phase of forming the group of experts served to determine the characteristics that the people who were to be consulted had to meet. They were then identified and selected to establish contact and request their participation.

The modified online Delphi method was used for the implementation phase of the consultation rounds. The term “modified” refers to the adaptation of the original method proposed by Cabero-Almenara [66], who limited the iterative process to 2 rounds to avoid the desertion of the people who made up the group of experts. The label “online” used by Cruz-Ramírez and Rúa-Vásquez [67] is due to the fact that the consultations were conducted via email.

While the aim of the first round of consultation was to analyze the relevance and clarity of all the items, the aim of the second round was to analyze the clarity of the items that had been reformulated. The experts rated the items using a four-choice Likert scale.

By not introducing any intermediate response alternatives, the respondent is forced to take a position in favor or against [68], thus facilitating decision-making on whether to keep the items that make up the questionnaire.

The previous phase was combined with the result phase, in which the assessments of the group of experts were analyzed. SPSS version 25 was used for the data analysis. As the results were extracted, decisions were made according to the degree of consensus found [69].

On the one hand, Kendall's coefficient was calculated to determine the level of agreement between the answers given by the experts (construct validity—relevance = 0.11; content validity—clarity = 0.17). On the other hand, a descriptive analysis was carried out to provide the mean, standard deviation, and percentiles of the responses collected. All the results and decisions were shared with the group of experts.

The QaSLu-45 was qualitatively validated through the Delphi method, and it constituted 45 items with a 5-point Likert scale based on frequency (Never; Seldom; Sometimes; Frequently; Always) [10].

Subsequently, quantitative validation through the goodness-of-fit measures of confirmatory factor analysis followed, and the final configuration was optimized for one factor by means of robust exploratory factor analysis, determining the most optimal version of the questionnaire under the law of parsimony, the QaSLu-27, with only 27 items and better psychometric properties. These factor analyses were developed for a single-factor configuration, as the aim was for the QaSLu to provide a single final score with a model optimized for a single factor/dimension as far as possible.

The 45-item QaSLu (QaSLu-45) was given to the participants to complete [14]. The sampling process was haphazard and opportunity-driven [70]. All conference proceedings published up until 2018 that had been arranged by the Spanish Network of USL were examined to create the sample. Every author received an email encouraging them to take part in the study. The email requested that the writers pass it on to any contacts who may have been willing to take part. As a result, snowball sampling was added to accidental sampling [71].

A total of 118 teachers—67.80% female and 32.20% male—from 43 higher education institutions took part in the event, with a minimum age of 24, a maximum age of 65, and a mean age was 46.16 years ( $SD = 9.60$ ).

Regarding the participants' educational background, the vast majority (76.3%) had a doctorate. The average length of time for which they had used SL was 6.69 years ( $DT = 5.39$ ), with 0.5 years being the shortest duration and 34 years being the longest.

Two types of institutions were partnered with to produce SL: formal education centers, which made up 50.85% of the sample, and centers unrelated to formal education, which made up the remaining 49.15%.

Regarding the mode of the SL experiences assessed, 64.41% used in-person interactions, whereas 35.59% involved virtual activities of some kind.

The reliability of the original 45 items' configuration was excellent as measured through Cronbach's alpha ( $\alpha = 0.90$ ). However, the reliability of the 27-item configuration (QaSLu-27) was even better ( $\alpha = 0.92$ ), despite it having 40% fewer items.

### 3.2. Procedure for Collecting the Data

There were three phases of collecting the data. Starting in September 2019, the first phase's goal was to locate and set up the sample. Sampling was incidental and based on opportunity [70]. To configure the sample, all the proceedings of conferences organized by the Spanish Network of USL that had been published up to 2018 were reviewed to collect the emails of the authors.

In the second phase, an email was sent to all the chosen authors inviting them to participate in the study and asking them to complete the QaSLu-45. Therefore, incidental sampling was complemented by snowball sampling [71]. This phase was prolonged to January 2021 due to the SARS-CoV-2 virus-related global health emergency.

To validate the questionnaire, a descriptive analysis of the responses collected was conducted during the third phase, which took place in 2021.

Finally, as the initial step towards validation, we used the statistical program IBM SPSS (version 25.0) to perform an exploratory factor analysis of the major components. The principal component factor analysis findings revealed 13 components with eigenvalues greater than 1.0, accounting for 69.95% of the total variance. In order to determine whether this analysis was ideal, López-Aguado and Gutiérrez-Provecho's recommendations were followed [44], and two tests were run—the Kaiser–Meyer–Olkin Sampling Adequacy Test (KMO = 0.793) and Bartlett's Sphericity Test ( $\chi^2 = 2689.72$ ;  $p = 0.00$ )—demonstrating that this factorial analysis was optimum.

In terms of the QaSLu-45's dependability, these 45 items together demonstrated outstanding reliability as determined by Cronbach's alpha ( $\alpha = 0.90$ ) [72] and an excellent robust goodness of fit as determined by the CFA model's fit measures. According to Kelley's suggested criterion, its Comparative Fit Index was  $>0.95$  (CFI = 0.973) since the minimum–maximum difference in the CFA/degrees of freedom was between 1 and 3 (CMIN/DF = 1.053) [73], its root mean square error of approximation was less than 0.06 (RMSEA = 0.042), and its weighted root mean square residual was less than 1.0 (WRMR = 0.0978) [74].

Performing a robust unweighted least squares (RULS) factor exploratory analysis for a configuration with a single principal component was the next step [47], conducted through the FACTOR software (version 12) designed by Ferrando and Lorenzo-Seva [75]. The parsimony principle was adhered to in this analysis, enabling the development of an optimal and sturdy design that yielded a distinct, legitimate, and trustworthy score. This new analysis recommended removing the following 18 items from the design: 1, 4, 6, 7, 8, 9, 11, 16, 18, 20, 23, 25, 28, 31, 35, 38, 41, and 42. This left a final configuration of 27 items (QaSLu-27) that is highly correlated, psychometrically resilient and optimized for a single principal component [14,17]. Furthermore, the statistical program IBM SPSS (version 25.0) was used to determine the final questionnaire's reliability using Cronbach's alpha and the model fit measures of CFA.

Lastly, a new RULS factor exploratory analysis of the key components was performed for the final configuration of a single component to ascertain the increase in the validity, reliability, and robustness of the final QaSLu-27 configuration. RULS showed that it was not necessary to eliminate any additional items from the QaSLu-27. However, a significant improvement was noted in the Kaiser–Meyer–Olkin Sampling Adequacy Test (KMO = 0.863), as well as Bartlett's Sphericity Test ( $\chi^2 = 1447.76$ ;  $p = 0.000$ ). The principal component factor analysis findings with the QaSLu-27 revealed seven components with eigenvalues greater than 1.0, accounting for 65.59% of the total variance. These findings support the idea that the final configuration's validity and robustness are significantly higher than those of the 45-item design. As can be observed in Table 5, the robust goodness-of-fit statistics determined using CFA were somewhat better, and the reliability measured by Cronbach's alpha increased ( $\alpha = 0.92$ ), despite the QaSLu-27 containing 27 items instead of 45.

Finally, the gender and age of the participants, their experience in SL, the type of institution at which the SL was developed, and the nature of their SL experience (virtual or face to face) were taken into consideration when establishing the rating scales for the standardization of the new and reduced QaSLu-27. The percentile distribution of each item as reported by the research participants was taken into consideration when calculating these scales.

**Table 5.** QaSLu model fit measures.

| Measure             | QaSLu-45 |                   |                | QaSLu-27 |                   |                |
|---------------------|----------|-------------------|----------------|----------|-------------------|----------------|
|                     | Estimate | Threshold [66,67] | Interpretation | Estimate | Threshold [66,67] | Interpretation |
| CMIN                | 994.830  | —                 | —              | 332.746  | —                 | —              |
| DF                  | 945.000  | —                 | —              | 324.000  | —                 | —              |
| CMIN/DF             | 1.053    | Between 1 and 3   | Excellent      | 1027     | Between 1 and 3   | Excellent      |
| CFI                 | 0.973    | >0.95             | Excellent      | 0.979    | >0.95             | Excellent      |
| WRMR                | 0.0978   | <1.0              | Excellent      | 0.0985   | <1.0              | Excellent      |
| RMSEA               | 0.042    | <0.06             | Excellent      | 0.046    | <0.06             | Excellent      |
| Cronbach's $\alpha$ | 0.90     | >0.90             | Excellent      | 0.92     | >0.90             | Excellent      |

Extracted from Lopez-de-Arana Prado et al. (2024) [17].

### 3.3. Limitations

This dataset has some limitations and aspects for improvement. The most important one is related to the size of the sample, since although it is within the range of scientifically acceptable sizes according to Thompson [76] and Mundfrom et al. [77], with the number of items in the sample being 4.37x in its reduced version, it is advisable to enlarge the sample in order to improve the reliability of the indices calculated with the factorial analyses. On the other hand, it should also be considered that the high degree of specialization of the sample reduces its heterogeneity and could compromise the measurement of reliability indicators such as Cronbach's alpha.

However, the purpose of this dataset is for it to be shared and to form part of larger future research that expands the sample and its variability, so these limitations are not too serious.

### 3.4. Ethical Considerations

This questionnaire guarantees the participants' anonymity because, as stated in the instrument description, the personal information collected through it does not allow for participant identification. This is in accordance with the Spanish Organic Law 3/2018 of 5 December 2018 on Data Protection and Guarantee of Digital Rights.

## 4. User Notes

Users of this dataset can make use of either the QaSLu-27 or the QaSLu45 and can even create and validate a new version to suit their needs.

It is possible to create new rating scales in combination with the existing ones or even extend the sample to make the scale more appropriate for the target population.

The .sav version of the dataset is recommended, as it includes the categories for each variable.

In case of any doubts about its use, it is possible to review previous linked articles on validation following the Delphi method [10] or quantitative validation using RULS [17].

## 5. Patents

In accordance with the provisions of the Law on Intellectual Property (Royal Legislative Decree 1/1996 of 12 April 1996), the intellectual property rights are registered in this register in the manner set out below:

REGISTRY ENTRY NUMBER 16/2024/4083

Title: Questionnaire for the Self-Assessment of SL Experiences in Higher Education (QaSLu-45)/Cuestionario para la autoevaluación de experiencias de aprendizaje-servicio universitario (CApSU-45)

Intellectual property object: Selection and arrangement of data

Kind of work: Data collection

Place of publication: Madrid

Dissemination date: 25 October 2019.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/data9090108/s1>, The QaSLu-27, the QaSLu-45, the raw dataset, and the standardization of the QaSLu-27 can be downloaded via this link.

**Author Contributions:** Conceptualization, R.S.-C., E.L.-d.-A.P., P.A., and R.C.; methodology, R.S.-C., E.L.-d.-A.P., and R.S.-C.; validation, R.S.-C.; formal analysis, R.S.-C.; investigation, E.L.-d.-A.P., P.A., and R.C.; resources, E.L.-d.-A.P.; data curation, R.S.-C.; writing—original draft preparation, E.L.-d.-A.P., P.A., R.C., and R.S.-C.; writing—review and editing, E.L.-d.-A.P., P.A., R.C., and R.S.-C.; visualization, E.L.-d.-A.P.; supervision, E.L.-d.-A.P. and P.A.; project administration, R.S.-C. and E.L.-d.-A.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** This scientific research and all the data collected were ethically approved in 2020 by the National Commission for Scientific and Technological Research of Chile (CONICYT Fondecyt/Initiation) with approval number 11170623.

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

**Data Availability Statement:** The raw data from this study can be found in Supplementary Materials to this document and in Harvard Dataverse [14] for non-commercial use.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Ribeiro, Á.; Aramburuzabala, P.; Paz-Lourido, B. *Research Report on the Institutionalisation of Service-Learning in European Higher Education*; Research procedures and main findings; European Association of Service-Learning in Higher Education: Madrid, Spain, 2021.
2. Rodríguez-Izquierdo, R.M. Aprendizaje Servicio y Compromiso Académico En Educación Superior. *Rev. De Psicodidáctica* **2020**, *25*, 45–51. [\[CrossRef\]](#)
3. Aramburuzabala, P.; Cerrillo, R.; Tello, I. Aprendizaje-servicio: Una propuesta metodológica para la introducción de la sostenibilidad curricular en la universidad. *Profesorado* **2015**, *19*, 78–95.
4. Aramburuzabala, P.; McIlrath, H.; Opazo, H. *Embedding Service-Learning in European Higher Education*; Routledge: Oxford, UK, 2019.
5. Bringle, R.G.; Clayton, P.H. Civic Learning: A Sine Qua Non of Service Learning. *Front. Educ.* **2021**, *6*, 606443. [\[CrossRef\]](#)
6. Maravé-Vivas, M.; Gil-Gómez, J.; Valverde-Esteve, T.; Salvador-Garcia, C.; Chiva-Bartoll, O. A Longitudinal Study of the Effects of Service-Learning on Physical Education Teacher Education Students. *Front. Psychol.* **2022**, *13*, 787346. [\[CrossRef\]](#)
7. Cerrillo, R.; McIlrath, L. Service Learning as a Community of Practice in Irish Higher Education: Understanding Cultural and Historical Nuances. In *Service Learning at a Glance*; Nova Science Publishers, Inc.: New York, NY, USA, 2022; pp. 19–40.
8. Deeley, S.J. *Assessment and Service-Learning in Higher Education: Critical Reflective Journals as Praxis*; Springer International Publishing: Cham, Switzerland, 2022; ISBN 978-3-030-94439-1.
9. Commission to the European Parliament. *Renewed EU Agenda for Higher Education*; Commission to the European Parliament: Brussels, Belgium, 2017.
10. López-de-Arana, E.; Higuera, P.A.; Carvajal, H.O. Diseño y validación de un cuestionario para la autoevaluación de experiencias de aprendizaje-servicio universitario | Design and validation of a questionnaire for self-assessment of university service-learning experiences. *Educ. XXI* **2020**, *23*, 319–347. [\[CrossRef\]](#)
11. Álvarez-Castillo, J.; Usarralde, M.M.; González, H.; Fernández, M. El Aprendizaje-Servicio En La Formación Del Profesorado de Las Universidades Españolas. *Rev. Española De Pedagog.* **2017**, *75*, 199–217. [\[CrossRef\]](#)
12. Martín-García, X.; Puig Rovira, J.M.; Palos Rodríguez, J.; Rubio Serrano, L. Mejorando La Calidad de Las Prácticas de Aprendizaje-Servicio. *Enseñanza Teach.* **2018**, *36*, 111. [\[CrossRef\]](#)
13. Gul, D.R.; Ahmad, D.I.; Tahir, D.T.; Ishfaq, D.U. Development and Factor Analysis of an Instrument to Measure Service-Learning Management. *Heliyon* **2022**, *8*, e09205. [\[CrossRef\]](#)
14. López de Arana Prado, E.; Sánchez-Cabrero, R.; Aramburuzabala, P.; Cerrillo, R. Dataset of the Initial Validation of Questionnaire for the Self-Assessment of University Service-Learning Experiences (QaSLu). *Harv. Dataverse* **2024**. [\[CrossRef\]](#)

15. López-de-Arana, E.; Aramburuzabala, P.; Opazo, H.; Quintana, A.; Franco, L. Coherencia entre los principios del método Delphi y la conceptualización del Aprendizaje-Servicio Universitario. In *El Papel Del Aprendizaje-Servicio En La Construcción De Una Ciudadanía Global*; Ballesteros, C., Gutiérrez, J.G., Lázaro, P., Higuera, P.A., Eds.; Universidad Nacional de Educación a Distancia: Madrid, Spain, 2020; pp. 689–698.
16. López-de-Arana, E.; Martínez-Muñoz, L.; Calle-Molina, M.; Aguado-Gómez, R.; Santos-Pastor, M.<sup>a</sup>. Construction and validation of an instrument for evaluating the quality of university servicelearning projects using the Delphi method | Construction and validation of an instrument for evaluating the quality of university service-learning projects using the Delphi method. *Rev. Española De Pedagog.* **2023**, *81*, 381–402. [\[CrossRef\]](#)
17. López-de-Arana Prado, E.; Aramburuzabala, P.; Cerrillo, R.; Sánchez-Cabrero, R. Validation and Standardization of a Questionnaire for the Self-Assessment of Service-Learning Experiences in Higher Education (QaSLu-27). *Educ. Sci.* **2024**, *14*, 615. [\[CrossRef\]](#)
18. Shumer, R. Self-assessment for service-learning. In *Studying Service-Learning: Innovations in Education Research Methodology*; Billig, S.H., Waterman, A.S., Eds.; Routledge: Oxford, UK, 2003; pp. 149–171.
19. Jenkins, A.; Sheehy, P. A checklist for implementing service-learning in higher education. *J. Community Engagem. Scholarsh.* **2011**, *4*, 52–60. [\[CrossRef\]](#)
20. Kieran, L.; Haack, S. PRELOAD: A rubric to evaluate course syllabi for quality indicators of community engagement and service-learning components. *J. Community Engagem. High. Educ.* **2018**, *10*, 39–47.
21. Stokamer, S.T. The intersection of institutional contexts and faculty development in service-learning and community engagement. In *Reconceptualizing Faculty Development in Service-Learning/Community Engagement: Exploring Intersections, Frameworks, and Models of Practice*; Berkey, B., Meixner, C., Green, P.M., Rountree, E.E., Eds.; Stylus Publishing: Sterling, VA, USA, 2018; pp. 221–239.
22. Botelho, J.; Eddy, R.M.; Galport, N.; Avila-Linn, C. Uncovering the quality of STEM service-learning course implementation and essential elements across the California State University system. *Mich. J. Community Serv. Learn.* **2020**, *26*, 1–19. [\[CrossRef\]](#)
23. Furco, A.; Brooks, S.O.; Lopez, I.; Matthews, P.H.; Hirt, L.E.; Schultzenberg, A.; Anderson, B.N. Service-Learning Quality Assessment Tool (SLQAT). *J. High. Educ. Outreach Engagem.* **2023**, *27*, 181–200.
24. Matthews, P.H.; Lopez, I.; Hirt, L.E.; Brooks, S.O.; Furco, A. Developing the SLQAT (Service-Learning Quality Assessment Tool), a quantitative instrument to evaluate elements impacting student outcomes in academic service-learning courses. *J. High. Educ. Outreach Engagem.* **2023**, *27*, 161–180.
25. Escofet-Roig, A.; Folgueiras Bertomeu, P.; Luna González, E.; Palou Julián, B. Elaboración y validación de un cuestionario para la valoración de proyectos de aprendizaje-servicio. *Artic. Publ. En Rev. (Mètodes D'investigació I Diagnòstic En Educ.)* **2016**, *21*, 929–949.
26. Rodríguez-Izquierdo, R.M. Validación de una escala de medida del impacto del aprendizaje-servicio en el desarrollo de las competencias profesionales de los estudiantes en formación docente. *Rev. Mex. De Psicol.* **2019**, *36*, 63–73.
27. León-Carrascosa, V.; Sánchez-Serrano, S.; Belando-Montoro, M.-R. Diseño y validación de un cuestionario para evaluar la metodología Aprendizaje-Servicio. *Estud. Sobre Educ.* **2020**, *39*, 247–266. [\[CrossRef\]](#)
28. Santos-Pastor, M.L.; Cañadas, L.; Muñoz, L.F.M.; Rico, L.G. Diseño y validación de una escala para evaluar el aprendizaje-servicio universitario en actividad física y deporte. *Educ. XX1* **2020**, *23*, 67–93. [\[CrossRef\]](#)
29. Ruiz-Ordóñez, Y.; Salcedo-Mateu, A.; Turbi, Á.; Novella, C.; Moret-Tatay, C. VAL-U: Psychometric Properties of a Values and Civic Attitudes Scale for University Students' Service-Learning. *Psicol. Reflexão E Crítica* **2022**, *35*, 3. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Dalkey, N.; Helmer, O. An Experimental Application of the Delphi Method to the Use of Experts. *Manag. Sci.* **1963**, *9*, 458–467. [\[CrossRef\]](#)
31. Cabero-Almenara, J.; Infante-Moro, A. Empleo del método Delphi y su empleo en la investigación en comunicación y educación. *Edutec Rev. Electrónica De Tecnol. Educ.* **2014**, *48*, a272. [\[CrossRef\]](#)
32. Pérez-Escoda, A.; García-Ruiz, R.; Aguaded-Gómez, I. Media Competence in University Teaching Staff. Validation of an Instrument of Evaluation. *@Tic Rev. D'innovació Educ.* **2018**, *21*, 1–9. [\[CrossRef\]](#)
33. Reguant-Álvarez, M.; Torrado-Fonseca, M. El mètode Delphi. *REIRE Rev. D'innovació I Recer. En Educ.* **2016**, *9*, 87–102. [\[CrossRef\]](#)
34. Varela-Ruiz, M.; Díaz-Bravo, L.; García-Durán, R. Descripción y usos del método Delphi en investigaciones del área de la salud. *Investig. En Educ. Médica* **2012**, *1*, 90–95. [\[CrossRef\]](#)
35. López-Meneses, E.J.; Bernal-Bravo, C.; Leiva-Olivencia, J.J.; Martín-Padilla, A.H. Validación del instrumento didáctico de valoración de observatorios digitales sobre MOOC: CUVOMOOC<sup>®</sup> mediante el Método Delphi. *Campus Virtuales* **2018**, *7*, 95–110.
36. Rowe, G.; Wright, G. The Delphi Technique as a Forecasting Tool: Issues and Analysis. *Int. J. Forecast.* **1999**, *15*, 353–375. [\[CrossRef\]](#)
37. Geist, M.R. Using the Delphi Method to Engage Stakeholders: A Comparison of Two Studies. *Eval. Program Plan.* **2010**, *33*, 147–154. [\[CrossRef\]](#)
38. Landeta, J. Current Validity of the Delphi Method in Social Sciences. *Technol. Forecast. Soc. Chang.* **2006**, *73*, 467–482. [\[CrossRef\]](#)
39. Landeta-Rodríguez, J. *El Método Delphi: Una Técnica De Previsión Para La Incertidumbre*; Ariel España: Barcelona, Spain, 1999; ISBN 978-84-344-2836-2.
40. Cabero-Almenara, J.; Barroso-Osuna, J. La utilización del juicio de experto para la evaluación de TIC: El coeficiente de competencia experta. *Bordón. Rev. De Pedagog.* **2013**, *65*, 25–38. [\[CrossRef\]](#)
41. Bakieva, M.; Meliá, J.M.J.; González-Such, J.; Barajas, Y.E.L. Colegialidad docente: Validación lógica del instrumento para autoevaluación docente en España y México. *Estud. Sobre Educ.* **2018**, *34*, 99–127. [\[CrossRef\]](#)

42. George-Reyes, C.E.; Trujillo-Liñán, L. Aplicación del Método Delphi Modificado para la Validación de un Cuestionario de Incorporación de las TIC en la Práctica Docente. *Rev. Iberoam. De Evaluación Educ.* **2018**, *11*, 135. [CrossRef]
43. McCartney, K.; Burchinal, M.R.; Bub, K.L. Best Practices in Quantitative Methods for Developmentalists. *Monogr. Soc. Res. Child Dev.* **2006**, *71*, 1–145. [CrossRef] [PubMed]
44. López-Aguado, M.; Gutiérrez-Provecho, L. Cómo realizar e interpretar un análisis factorial exploratorio utilizando SPSS. *REIRE Rev. D'innovació I Recer. En Educ.* **2019**, *12*, 1–14. [CrossRef]
45. Sánchez-Cabrero, R.; Sandoval-Mena, M.; Saez-Suanes, G.P.; Prado, E.L.-A. Design and Validation of the Classroom Climate for an Inclusive Education Questionnaire (CCIEQ). *Environ. Soc. Psychol.* **2024**, *9*, 1754. [CrossRef]
46. Marsh, H.W.; Morin, A.J.S.; Parker, P.D.; Kaur, G. Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis. *Annu. Rev. Clin. Psychol.* **2014**, *10*, 85–110. [CrossRef]
47. Lorenzo-Seva, U.; Ferrando, P.J. Robust Promin: A Method for Diagonally Weighted Factor Rotation. *LIBERABIT Rev. Peru. De Psicol.* **2019**, *25*, 99–106. [CrossRef]
48. Blais, M.; Baer, L. Understanding Rating Scales and Assessment Instruments. In *Handbook of Clinical Rating Scales and Assessment in Psychiatry and Mental Health*; Baer, L., Blais, M.A., Eds.; Humana Press: Totowa, NJ, USA, 2010; pp. 1–6. ISBN 978-1-59745-387-5.
49. Newman, F.L.; Newman, F.L. Global Scales: Strengths, Uses and Problems of Global Scales as an Evaluation Instrument. *Eval. Program Plan.* **1980**, *3*, 257–268. [CrossRef]
50. Sánchez-Cabrero, R. Mejora de la satisfacción corporal en la madurez a través de un programa específico de imagen corporal. *Univ. Psychol.* **2020**, *19*, 1–15. [CrossRef]
51. Pounder, J.S. A Behaviourally Anchored Rating Scales Approach to Institutional Self-Assessment in Higher Education. *Assess. Eval. High. Educ.* **2000**, *25*, 171–182. [CrossRef]
52. López-de-Arana Prado, E.; Aramburuzabala Higuera, P.; Cerrillo, R. Respondiendo a los Objetivos de Desarrollo Sostenible a través de la implementación del Aprendizaje-Servicio Universitario. In *El Aprendizaje-Servicio universitario ante los retos de la Agenda 2030*; Gutiérrez, J.G., Morera, F.J.A., Ramírez, A.C., Eds.; UNED-Universidad Nacional de Educación a Distancia: Madrid, Spain, 2023; pp. 568–578. Available online: <https://dialnet.unirioja.es/servlet/articulo?codigo=9332822> (accessed on 12 September 2024).
53. RMC Research Corporation. *Service-Learning Policies and Practices: A Research-Based Advocacy Paper*; National Service Learning Clearinghouse: Scotts Valley, CA, USA, 2008.
54. Stark, W. *Europe Engage: Quality Standards for Service Learning Activities*; Europe Engage: Helsinki, Finland, 2017.
55. Campo, L. Una rúbrica para evaluar y mejorar los proyectos de aprendizaje servicio en la universidad. *RIDAS. Rev. Iberoam. Aprendiz. Serv.* **2015**, *1*, 91–111.
56. Rubio-Serrano, L.; Puig-Rovira, J.M.; Martín-García, X.; Palos-Rodríguez, J. Analizar, repensar y mejorar los proyectos: Una rúbrica para la autoevaluación de experiencias de Aprendizaje-Servicio. *Profr. Rev. Currículum Form. Profr.* **2015**, *19*, 111–126.
57. Puig-Rovira, J.M.; Martín-García, X.; Rubio-Serrano, L. Cómo evaluar proyectos de aprendizaje servicio? *Voces La Educ.* **2017**, *2*, 122–132.
58. Krippendorff, K. *Content Analysis: An Introduction to Its Methodology*; SAGE Publications, Inc.: New York, NY, USA, 2019; ISBN 978-1-07-187878-1.
59. CLAYSS. *Cómo Desarrollar Proyectos De Aprendizaje Y Servicios Solidario En La Educación Media (Secundaria Y Enseñanza Técnica)*; CLAYSS: Buenos Aires, Argentina, 2016.
60. Aponte-Figueroa, G.; Cardozo-Montilla, M.; Melo, R. Método Delphi: Aplicaciones y Posibilidades En La Gestión Prospectiva de La Investigación y Desarrollo. *Rev. Venez. Análisis Coyunt.* **2012**, *18*, 41–52.
61. Steurer, J. The Delphi Method: An Efficient Procedure to Generate Knowledge. *Skelet. Radiol.* **2011**, *40*, 959–961. [CrossRef]
62. Kennedy, H.P. Enhancing Delphi Research: Methods and Results. *J. Adv. Nurs.* **2004**, *45*, 504–511. [CrossRef]
63. Price, B. Delphi Survey Research and Older People. *Nurs. Older People* **2005**, *17*, 25–31. [CrossRef]
64. Ludwig, B. Predicting the Future: Have You Considered Using the Delphi Methodology? *J. Ext.* **1997**, *35*, 1–4.
65. Gil, B.; Pascual-Ezama, D. La metodología Delphi como técnica de estudio de la validez de contenido. *An. De Psicol. Ann. Psychol.* **2012**, *28*, 1011–1020. [CrossRef]
66. Cabero-Almenara, J. Formación del profesorado universitario en TIC. Aplicación del método Delphi para la selección de los contenidos formativos. *Educ. XXI* **2014**, *17*, 111–131. [CrossRef]
67. Cruz-Ramírez, M.; Rúa-Vásquez, J.A. Surgimiento y desarrollo del método Delphi: Una perspectiva cuantitativa. *Biblios J. Librariansh. Inf. Sci.* **2018**, *71*, 90–107. [CrossRef]
68. Abal, F.J.P.; Auné, S.E.; Lozzia, G.S.; Attorresi, H.F. Funcionamiento de la Categoría Central en Ítems de Confianza para la Matemática. *Rev. Evaluar* **2017**, *17*, 18–31. [CrossRef]
69. Martínez-Piñeiro, E. La Técnica Delphi como estrategia de consulta a los implicados en la evaluación de programas. *Rev. Investig. Educ.* **2003**, *21*, 449–463.
70. Pérez-Luco Arenas, R.; Lagos Gutiérrez, L.; Mardones Barrera, R.; Sáez Ardura, F. Taxonomía de diseños y muestreo en investigación cualitativa. Un intento de síntesis entre las aproximaciones teórica y emergente. Taxonomy of designs and sampling in qualitative research. An attempt of synthesis between theoretical and emerging approaches. *Ámbitos. Rev. Int. Comun.* **2017**, *39*, 1–18.
71. Atkinson, R.; Flint, J. *The A-Z of Social Research*; SAGE Publications, Ltd.: New York, NY, USA, 2003; ISBN 978-0-85702-002-4.

- 
72. Arigita-García, A.; Sánchez-Cabrero, R.; Barrientos-Fernández, A.; Mañoso-Pacheco, L.; Pericacho-Gómez, F.J. Pre-Eminence of Determining Factors in Second Language Learning: An Educator's Perspective from Spain. *Heliyon* **2021**, *7*, e06282. [[CrossRef](#)]
  73. Martín-Antón, L.J.; Almedia, L.S.; Sáiz-Manzanares, M.-C.; Álvarez-Cañizo, M.; Carbonero, M.A. Psychometric Properties of the Academic Procrastination Scale in Spanish University Students. *Assess. Eval. High. Educ.* **2023**, *48*, 642–656. [[CrossRef](#)]
  74. Hu, L.; Bentler, P.M. Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Struct. Equ. Model. A Multidiscip. J.* **1999**, *6*, 1–55. [[CrossRef](#)]
  75. Ferrando, P.J.; Lorenzo-Seva, U. Program FACTOR at 10: Origins, Development and Future Directions. *Psicothema* **2017**, *29*, 236–240. [[CrossRef](#)]
  76. Thompson, B. *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*; American Psychological Association: Washington, DC, USA, 2004; ISBN 1-59147-093-5.
  77. Mundfrom, D.J.; Shaw, D.G.; Ke, T.L. Minimum Sample Size Recommendations for Conducting Factor Analyses. *Int. J. Test.* **2005**, *5*, 159–168. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.