Prithvi Journal of Research and Innovation
[A Peer-Reviewed, Open Access Multidisciplinary Bilingual Journal; Indexed in NepJOL]
ISSN 2705-4888 [Print]; ISSN 2705-4896 [Online]; JPPS Star-Rated Journal
Volume 7; 15 December 2025; pp. 56-76
eJournal Site: http://ejournals.pncampus.edu.np/ejournals/pjri/

Research Article

Developing the Content Validity Index of a SERVQUAL Instruments in the Public Sector of Nepal

Amin Palikhe
Department of Marketing, Prithvi Narayan Campus, Tribhuvan University

Article History:
Submitted 07 October 2025
Reviewed 10 November 2025
Revised 13 November 2025
Accepted 01 December 2025

Corresponding Author:
Amin Palikhe
Email: amin@pncampus.edu.np

Article DOI:
https://doi.org/10.3126/pjri.v7i1.87675

Publisher:
Centre for Research and Innovation
Prithvi Narayan Campus
Tribhuvan University, Pokhara, Nepal
[Accredited by UGC, Nepal]

Tel.: +977-61-576837
Email: research@pncampus.edu.np
URL: www.pncampus.edu.np

ABSTRACT

The content validity index is one of the methods of the instrument development, well designed and relevant tool of scale development. There are rarely used scale development instruments in the Nepali public sector and particularly used to gather the reliable data in the medical and health field. This is the foremost calculation of value of the content validation of questionnaire items to assess the perceived service quality delivery instrument in the perspective of Nepali public sector, which uses the modified Delphi method for the development of scale instrument. Therefore, this study aims to develop and validate the questionnaire items to assess the perceived service quality by using the content validation index (CVI) method. The value of content validation of developed questionnaire items has more than 0.70 cut-off point. Among the 87 questionnaire items, 33 items got a sufficient agreement from the experts, which ~~is~~ sufficiently validates in terms of a number of experts in the panel, in which there were eleven experts with an individual difference. The 24 items are more than 60% of agreement, as the refinement list and rest of the items (less than 0.60) are eliminated. In conclusion, it is more transparent, accurate, consistent and reliable ~~of~~ assessment questionnaire items based on the evidence from the calculation of content validation index. Therefore, the calculation of CVI is a theoretically and methodologically proven method, practically to apply for scale development and significantly aids researchers and practitioners. It is recommended ~~and~~ that further suggestions to confirm the assessment questionnaire items are made by using the advance statistical tools like the exploratory factor analysis and confirmatory factor analysis ~~and~~ in order to undergo a pilot study.

KEYWORDS: Content validation index, Delphi method, scale development, SERVQUAL

INTRODUCTION

The scale development is an indicator of latent construct which is a measured range of attitude, behavior, psychological, social, experience and perception. Theoretically, the world cannot assess directly therefore, scales are used to capture that behavior, an actor or feeling as well as it cannot be covered by a single item or variable. The development of multiple items to measure the latent construct that leads to accurate finding of research (De Villis, 2012). There are many scales that have been used to measure perceived service quality of different fields and large amount of literature on scale theory and development. When novel design and question is emerged, even though research becomes necessary to develop new scale (Thapa, 2022; Boateng et al., 2018).

The SERVQUAL model is proven, popular, reliable and great accuracy and used in various industries including private and public sectors. For the optimum performance of service quality delivery, there is no single solution and way out. The different measurements of service quality delivery dimension are required, and the improvement is necessary to provide better services to the citizens. Ater a thorough search of literature, most commonly used simple quality measurement questionnaire was chosen and modified from SERVQUAL model to suit the local situation. All available literature contains questionnaire in English; it may not be understood by many people and questionnaire needs in local language (Prakash & Pallepati, 2016). The instrument proposed by Parasuraman et al. (1988) was criticized and not appropriate and supported by majority of scientific research therefore the dimension of SERVQUAL model should be modified. The question is whether the 22-itmes SERVQUAL model can measure the service quality of Nepalese public sectors. (Babakus & Boller, 1992; Carman, 1990; Engdaw, 2019; Zeithmal et al., 1990).

The Content Validation Index (CVI) is the process of representing and relevance of elements for assessment of instruments to the targeted construct for assessment purpose. It is an assessment tool, like a questionnaire, is essential to proving its validity, particularly when used for research (Yusuff, 2019), with an emphasis to applied in psychological assessment in clinical situation (Haynes et al., 1995). The content validation involves refinement and validation of targeted construct (Smith & McCarthy, 1995). It is primarily focused on content validation of questionnaire items (Haynes et al., 1995). The content adequacy is vital for each questionnaire items that measures what it is to be measured (De Villis, 2012). Nevertheless, there has been no research performed to evaluate the validity of questionnaire items to assessment of perceived service quality in Nepalese perspective by using CVI. Therefore, this study developed appropriate questions or item generation by using two ways, deductive and inductive method (Hinkin, 1995). It is the best practice to define the domain and identify the questionnaire to assess it. When developing the questionnaire, the expert evaluation of questionnaire items is established as a content validity. It is important part for scale development and validation of questionnaire items (Boteng et al., 2018; De Villis, 2012; Morgado et al., 2017). After questionnaire items are generated and experts evaluate each generated items on the basis of relevance and representative to the domain are assessed to the content validity using modified Delphi method. It is the method of scale development and validation of instrument gathering the consensus (Iqbal & Pipon-Young, 2009).

Therefore, content validity needs evidence of content relevance and

representativeness. By adopting modified Delphi method, qualitative and quantitative data, the study aims to calculate the content validation index of questionnaire items by taking the expert opinion to assess the perceived SERVQUAL instrument in the perspective of Nepalese public sectors. This study will provide concrete evidence for content validity of questionnaire items and validated questionnaire items can be used to assess the perceived service quality delivery in Nepalese public sectors for quality measurement. Therefore, the modified Delphi technique along with the CVI significantly contributes and validates to the future researcher in developing the scale instrument before going to research process.

LITERATURE REVIEW

Measuring the service quality is important. The SERVQUAL framework aimed at measuring the service quality in the service sectors. Therefore, the valid and reliable instrument is necessary for quality research and most commonly used valid questionnaire was SERVQUAL questionnaire, chosen and modified to suit and apply in the local situation ((Prakash & Pallepati, 2016) but available literature showed that number of studies validating the questionnaire by using the content validity index method for measuring service quality in different sectors.

Allahyari et al. (2011) have studied to develop a measurement tool in order to evaluate the cognitive-based human errors in the occupational setting. The Occupational Cognitive Failure Questionnaire (OCFQ) was developed to test the content validity. A total 35 items were prepared. A total of 16 panelist has agreed to participate, having a PhD or equivalent. The new instrument with 30 items was developed. The content validity rate for OCFQ was 0.70. The final OCFQ was internally consistent (α=96) and the intraclass correlation coefficient was 0.99. The results show that OCFQ would be a useful instrument for the measurement of performance and safety in the working environment.

Prakash and Pallepati (2016) had studied to content validate the Kannada translated modified service quality (SERVQUAL) questionnaire. The questionnaire consists of 15 closed-ended items that are translated into Kannada language. It was distributed to five subject matter experts and measures the relevance, clarity, simplicity and ambiguity. The results showed that the values for relevance, clarity, simplicity, and ambiguity are 0.97, 0.97,0.97, and 0.94, respectively, for the translated and modified SERVQUAL questionnaire. Overall, the Content Validity Index (CVI) was 0.96. In the conclusion, it can be used in the future for the measurement of the SERVQUAL of a hospital in Kannada.

Ozair et al. (2017) had studied to develop a tool to assess the knowledge and clinical reasoning of healthcare providers on acute asthmatic management in emergency situations. The Content Validity Index (CVI), Face Validity Index (FVI), and Intraclass Correlation Coefficient (ICC) have been used. The ten pannel was recruited from lectures of the emergency department. The result showed that the CVI for three domains, namely diagnosis, treatment, and disposition, was more than 0.83. FVI values were at least 0.83. The ICC was 0.89 with a p-value <0.001. The newly developed tool is named K-CRAMED, a valid tool to assess the knowledge and clinical reasoning of healthcare providers, and further validation is suggested to verify in other settings.

Ghazali et al. (2018) had measured the content validity of the students' self-efficacy and meaningful learning in the context of Massive Open Online Courses (MOOC) by using the content validity ratio. The research was conducted through the evaluation of 20 expert panels using a purposive sampling method. The panelists were divided into two groups: professional (university experts

in psychology) and field experts (PhD candidates). The four constructs with 35 items for students' self-efficacy and five constructs with 50 itmes for meaningful learning have been validated. The results show that all the sclale has a good content validity and are promoted as a good scale of students' self-efficacy and meaningful learning in the context of MOOC. It items obtained positive Content Validity Rate (CVR) between 0.00 to 0.80. It is suggested to apply further the Structural Equation Modeling for exploring quality items and model development.

Mohamad Marzuki et al. (2018) have studied to translate and validate the English System Usability Scale questionnaire into Malay the main language in Malasysia. This study aims to assess the usability of mobile apps to assess in Malay by translating with cross-cultural adaptation. The 10 item was developed in English and needs to be adapted into local languages to assess the usability of a mobile app. Therefore, forward and backward translation of the questionnaire was conducted in the local people of Malay. The contents of the questionnaire for mobile apps were validated by 10 experts in mobile app development. The face validity of the efficacy of the questionnaire was further probed by testing 10 mobile phone users, involving 54 mobile phone users. The content validity index was 0.91. it indicates good relevance of the 10 items used to assess the usability of a mobile app. The face validity was 0.94; therefore, the study shows the clarity and comprehensiveness of the questionnaires. The reliability showed 0.85, indicating the scale questionnaire is a valid and reliable tool to assess the usability of mobile apps in Malaysia.

De Oliveira et al. (2024) studied the expectations and needs of the users in determining the quality of services. They tried to understand the elements involved in the care process of perceived quality of service offered by health services to satisfy the needs of the users and improve the assistance of service provider. The content validation of an instrument for evaluating the quality of services of reception and risk classification with evaluation of factors for users, managers and health professionals. The Reception Service with Risk Classification in Obstetrics (ACCRO) with evaluation dimensions module for users, managers and health professionals has been developed. The validation for the content committee of experts from specialized areas of the instruments was carried out based on agreement rate, Content Validity Index (CVI), Kappa index for evaluation and achievement of agreement, and content validity of the tools among the experts. All together 201 items with eleven experts were selected. The cut-off point of the validation tool is equal or greater than 0.80 was considered, otherwise eliminated. The results showed that the tool is good and suggested to pilot test to be continue to complete the validation process.

Hamid and Habidin (2024) have studied assessing the performance of manufacturing companies for the success of manufacturing business to validate an instrument through content validity calculation method. The questionnaire was distributed to managers and executives of the manufacturing companies in Malaysia. The assessment was conducted by seven panels of experts. The face validity, Kappa index and content validity have been used to validate the instrument. The instrument had 80 items categorized into three constructs and 14 components. The results showed that the content validity index varies between 0.86 to 1 where scale content validity index an average varies from 0.86 to 1. Therefore, the instrument is high level of validity. It indicated that the instrument is novel for manufacturing performance. This study would be valuable for researchers and practitioners to deeper understanding of the factors of manufacturing performance and for future research agendas.

The content validation index is the

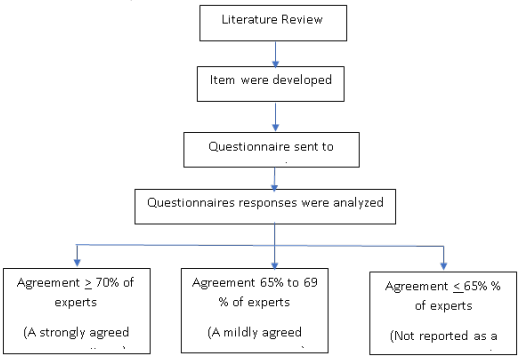
unique method of development of the perceived service quality delivery construct, items and tool that ensures the content validity in detail, commonly used in medical and health, manufacturing and education sectors rather than other sectors to validate the questionnaire for measuring service quality. The previous studies showed that the service quality delivery was measured through extracting and modifying the variables items to suit the local condition. This study is used to develop the content validation index of questionnaire items for measuring the perceived SERVQUAL in public service sector of Nepal.

RESEARCH METHODS

A modified Delphi techniqie, mixed method research (quantitative and qualitative) design with content validitation index employed to achive experts concensus from purposively elevent experts in the individual different were convoked across the single iterative round in Pokhara Metropolitan city to evaluate the content validation of questionnaire to assess the perceived SERVQUAL in Nepalese public sectors. Which follows the following research process for calculation of CVI.

Figure 1

Flow Diagram of Consensus Steps and Modified Delphi Method (*Yilmaz Yalcmkaya et al., 2021)*



In the validation process, researchers have a phone call to the expert for appointment in their suitable time, meeting with panel of experts, who are provided with detailed instruction and given a content validity form with QR code agreement letter in the questionnaire and as well asked to provide feedback for improvement of research instruments. Finally, researchers perform the follow up procedure and receive the form.

The following are the six detail steps of calculation of content validation index.

(a) Preparing Content Validation Form

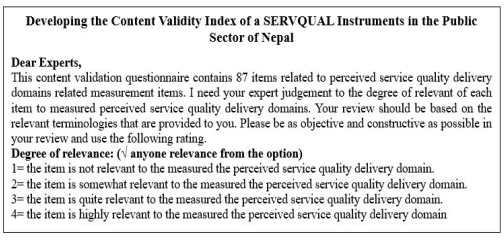
Content Validity (CVI) is widely used and accepted method of measuring content of study construct based on expert rating from not relevance to highly relevance by developing content validity form. It measures the construct of interest as relevance and representativeness to the targeted construct (Yusuff, 2019). Therefore, it is important to ensure the items are appropriate to the construct domain (Davis, 1992; Polit & Beck, 2006; Roebianto et al., 2023). The calculation of content validation questionnaire is done through relevance and representativeness of questionnaire items for model development in the Nepalese context, by adopting 4-point scale (Grand & Davis.1997; Yusoff, 2019). The content validity evidence can be shown by the content validity index (CVI) assessment tools, by utilizing the CVI to support their findings (Yusoff, 2019). It has been applied in two stages (development and judgement) process (Lynn, 1985) and item generation can be inductive and deductive method (Hinkin, 1995; Kapuscinski & Masters 2010).

The literature helps to identify the appropriate content domain for model development. To create reliability and validity in this research, therefore, more than 50 service quality related articles from the national and international journal of various streams are analyzed. In addition, to make validity, the researcher has used accurate information received from various sources to set up the questions and to confirm findings. This study drew the questionnaire and modified the instrument

items on the basis of research objectives and answer the research questions. The previous studies increase the validity of the research instrument (Messick, 1987), and significant items and the content validity being undertaken for validation process. So, a set of questionnaires has been developed by adopting of previously validated measures or rephrasing the items extracted from various literature regarding to service quality related literature review (Adetunji et al., 2013; Barabino et al., 2012; Kant, & Jaiswal, 2017; Santoyo-Sánchez et al., 2022; Tefera & Govender, 2016). The questionnaire items were taken from various service quality literature (Adetunji et al. 2013, Alanezi et al. 2010; Aljukhadar et al., 2022; Avkiran, 1994; Caro, 2007; Parasuraman et al., 1988; Pokharel et al., 2017, Pokharel et al., 2018, Wong, 2012). It ensures reliability or internal consistency (Hamid & Habidin, 2024). The items under the questionnaires are flexible. It has been selected considerable service quality delivery variable that has been used for Nepalese context and finally converted into the content validation form (*Figure 2 and 3*). The content validation form includes questionnaire items, scales, instructions and comment section as well. The developed instrument has three sections, an explanatory part, a demographic background part and content validation questionnaires. This stage of instrument development comprises three stages: selecting the content domain, creating items, and building the instrument (Zamanzadeh et al., 2015). Altogether, the questionnaire contains 87 items.

Figure 2

Example of Explanatory, Instruction and Rating Scales in the Calculation for Content Validation Form

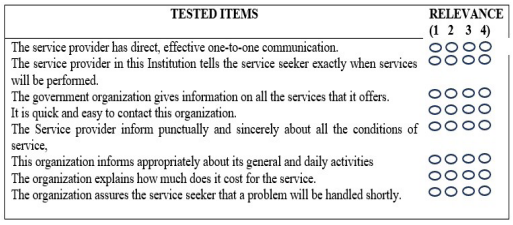


Developing the Content Validity Index of a SERVQUAL Instruments in the Public Sector of Nepal

Dear Experts,
This content validation questionnaire contains 87 items related to perceived service quality delivery domains related measurement items. I need your expert judgement to the degree of relevant of each item to measured perceived service quality delivery domains. Your review should be based on the relevant terminologies that are provided to you. Please be as objective and constructive as possible in your review and use the following rating.
Degree of relevance: (√ anyone relevance from the option)
1= the item is not relevant to the measured the perceived service quality delivery domain.
2= the item is somewhat relevant to the measured the perceived service quality delivery domain.
3= the item is quite relevant to the measured the perceived service quality delivery domain.
4= the item is highly relevant to the measured the perceived service quality delivery domain

Figure 3

Example of Layout for Content Valuation Questionnaire and Items Represents for Measurement for Perceived SERVQUAL



TESTED ITEMS	RELEVANCE (1 2 3 4)
The service provider has direct, effective one-to-one communication.	○○○○
The service provider in this Institution tells the service seeker exactly when services will be performed.	○○○○
The government organization gives information on all the services that it offers.	○○○○
It is quick and easy to contact this organization.	○○○○
The Service provider inform punctually and sincerely about all the conditions of service,	○○○○
This organization informs appropriately about its general and daily activities	○○○○
The organization explains how much does it cost for the service.	○○○○
The organization assures the service seeker that a problem will be handled shortly.	○○○○

(b) Selecting a Review Panel of Experts

After initially development of scale, the expert judgement panel of experts were appointed to assess the instrument's content validity. A specialist group of experts are selected based on the experience, education, engagement in service and service providers in the study area. The previous studies and recommended literature show that at least three experts, from individual differences in establishing content validity and optimum number of experts have not been decided but it is obviously doubtful that more than ten experts participated in the process, due to the decrease in consensus (Hamid & Habidin, 2024; Polit & Back, 2006). Although five to ten experts are considered adequately sufficient for content validation (Lynn, 1986). There are different cut-off points on the basis of number of panelists used for content validation calculation (Allahyari et al., 2011; Davis, 1992; Lawshe, 1975; Lynn, 1996; Polit & Beck, 2006; Polit et al., 2007). This study used 11 experts from individual differences to validate the instrument. All the 11 experts are provided with consent and agreed to participate. Therefore, the exploratory research design with purposive sampling was used to validate the questionnaires items. It has covered eight Professor with PhD, one ward representative, one ward secretary on the behalf of ward president and one public service experts as district administrative executive officer of Pokhara, Nepal, which was purposefully selected from the Pokhara Metropolitan city for measuring the content validity, through

face to face and non-face to face approach (Yusoff, 2019). The selected sample was rational to understand all the situation and environment and as well they can provide real problem solutions for the service delivery system for the improvement and reform of the quality-of-service delivery therefore those samples have been taken. The expert list is presented in Table 1.

Table 1

Panels of Expert's Selected for Content Validity Assessment is Presented

Designation	Experience in year
Professor, PhD	27
Professor, PhD	31
Professor, PhD	29
Professor, PhD	36
Professor, PhD	37
Professor, PhD	35
Professor, PhD	32
Professor, PhD	28
District Administrative Officer	11
Ward Representative	8
Ward Secretary	15
Average experience	24.08 years

Sources: Survey, 2025

The selected expert panels were requested to provide their professional judgment on each questionnaire items regarding various serviced service quality delivery. Their assessments on the item are collated to confirm their content validity.

(c) Conducting Content Validation

For the assessment of the content validity calculation, the modified Delphi techniques are applied. The modified Delphi method provides the experts to review, consolidate and opinion regarding content validation of constructs to be used in research instruments. This technique provides special guidelines and principles for development of models in research. This technique focuses on the consensus of experts and uses iteration process with a set of predictable steps (Garson, 2014). Based on the literature review, there were numbers of iterations in a Delphi technique ranged from two to eight rounds. However, three rounds were optimum number of iterations in Delphi method (Boulkedid et al., 2011; Grime & Wright, 2016). There were numbers of reviews that used single iteration in Delphi method (Beamish et al., 2020; Lamb et al., 2020; Tang et al., 2020; WSRI, 2020; Yilmaz Yalcinka et al., 2021).

There is no specific standard for achieving consensus in Modified Delphi technique but simple and frequently used criteria in the literature were considered for achieving consensus in Modified Delphi technique (Fackrell et al., 2017; Williamson et al., 2012). The consensus in Modified Delphi technique is applied with help of content validity index. Even though, the modified Delphi method includes a standard set of predictable steps. Simply it follows following steps for iteration of Delphi method in survey (DeOnna, 2006; Stines, 2003). Initially, questionnaires are developed and sent to the experts. Second, experts will complete and return the questionnaire to the researcher. Third, the researcher analyzes the data that is gathered from the expert panels and further looks for iteration in modified Delphi technique.

If the researcher thinks it is necessary then the second round of process will be initiated in the survey.

Accordingly, the list of item statements containing 87 items including content validation form was distributed to the experts via face-to-face meeting and non-face to face way. The eleven experts are used and clearly explain objectives of the study. Each expert was asked to rate items in terms of relevance and irrelevance (not relevance, somewhat relevant, quite relevant and highly relevant) of each statement and return back within one week besides each statement, experts were asked to provide comments and suggestion for any additional items that

may or may not have been included when developing the initial list of statement. While any statement does not achieve appropriate cut-off point that statement was removed or modified according to feedback provided by expert panel and supervisor on the basis of necessity for survey.

(d) Reviewing the Items

The content validation form is provided to the experts and requested to critically review and analysis the questionnaire items before providing the final score on the basis of relevance and irrelevance rating scales. Besides, the experts are encouraged to provide valuable comments on the item's modification, refine and addition of possible items for CVI calculation.

(e) Providing Score on Each Item

However, in the content validation, it is important to select the correct content in an instrument item. Therefore, the experts are requested to specify whether the questionnaire items are necessary or not in terms of rating in relevance and irrelevance way. The response criteria are presented in Figure 1 and 2.

(f) Calculating CVI

All the quantitative data were analyzed in MS Excel. The major outcomes of the study are the expert consensus on the relevance of questionnaire items. The definition and formula of the CVI indices are summarized (Polit & Beck, 2006).

I-CVI (item-level content validity index): The proportion of content experts giving item a relevance rating of 3 or 4 for 1 as an agreed item. Then the formula for

$$\text{I-CVI} = \frac{Agreed\ Item}{Number\ of\ expert}$$

The study ensured that ethical issues would be taken care of in terms of citation, confidentiality, not violation of respondent and consent for response.

RESULTS AND DISCUSSION

The collected data has been edited, coded and analyzed. The content validation form enables the respondent's profile and relevance of the content validity of items. These two categories of data are separately analyzed in results and discussion.

Results

Expert Group Profile

The content validation calculation of questionnaire for assessing the perceived service quality experts are Professor with PhD, experience in government service, politician. In the survey, Male experts are 90.91 percent. The average age of the experts is 51.58 years whereas 81. 82 percent of experts are over the age of 50 years. It indicated that the experts are selected from experience. The average experience of experts in public service 24.08 years. Those with over 20 years of experience experts are 72.73 percent in the calculation of content validation of questionnaire. The maximum experts are taken from Professors with PhD. It is 72.73 percent. The overall presented data can be observed in Table 2. Which provides information regarding the sociodemographic profile of the participants in the expert's judgement.

Table 2

Sociodemographic Profile of the Expert's Participating in the Content Validation of Questionnaire

Profile of experts		Frequency	Percentage
Gender	Male	10	90.91
	Female	1	9.0
Age	31 to 40 years	1	9.09
	41 to 50 years	1	9.09
	Over 50 years	9	81.82

	Average age		51.58 years
Experience	Less than 10 years	1	9.09
	11 to 20 years	2	18.18
	Over 20 years	8	72.73
Organization	University (Professor, PhD)	8	72.73
	District Administration Office (DAO)	1	9.09
	Ward Representative	1	9.09
	Ward Secretary	1	9.09
Positions	Professor, PhD	8	72.73
	DAO officer	1	9.09
	Ward Secretary	1	9.09
	Ward Member	1	9.09
Total		11	100

Source: Survey, 2025

Procedures and Judges Analysis

To identify the relevant items of perceived service quality delivery, during the item generation stage, items were taken from the literature review of service quality in different context were identified and contextualizing by rephrasing of items. The list of existing items was collected (items =87). while applying modified Delphi method, the iteration process, 11 experts with domain of individual differences independently ranked the 87 items rating each according to one of the four acceptance categories: not relevant, somewhat relevant, quite relevant and high relevant and the numeric values are given 1 to 4 respectively for not relevant, somewhat relevant, quite relevant and high relevant. After evaluating the numeric number, the rank 1 and 2 is given the item score equal to zero (item score =0) and rank 3 and 4 is given the items score equal to one (item score = 1). It means, there are irrelevance and relevance items which are further used for calculation of content validity index (CVI). The qualitative data was collected using open-ended questions for item revision or modification, or requirement of additional items in the scale development. The feedback on qualitative responses made by experts, items were refined and grammatical changes have been made. As a result of judges, the panel of experts evaluated and ranked the items.

The content validation calculation for questionnaire aimed to prepare scale development for assessing perceived service quality delivery. Therefore, the statistical analysis was considered. The collected data and calculation of content validation index (CVI) can be seen in the Table 3.

Table 3

Number of Agreement of Item Relevant, Irrelevant, I-CVI, Mean (SD) and Median (Range)

Items Tested	(A)	(DA)	I-CVI	Mean (SD)	Median (Range)
1. The service provider has direct, effective one-to-one communication.	10	1	0.91	3.64(0.67)	4(2-4)
2. The service provider in this Institution tells the service seeker exactly when services will be performed.	9	2	0.82	3.27(1.19)	4(1-4)
3. The government organization gives information on all the services that it offers.	9	2	0.82	3.64(0.81)	4(2-4)
4. It is quick and easy to contact this organization.	8	3	0.73	3.18(1.08)	4(1-4)

5. The Service provider inform punctually and sincerely about all the conditions of service,	10	1	0.91	3.36(0.92)	4(1-4)
6. This organization informs appropriately about its general and daily activities	7	4	0.64	2.82(1.33)	3(1-4)
7. The organization explains how much it costs for the service.	9	2	0.82	3.27(1.19)	4(1-4)
8. The organization assures the service seeker that a problem will be handled shortly.*	6	5	0.55	2.82(1.08)	3(1-4)
9. The digital web site provides information that is easy to understand.	8	3	0.73	3.27(1.10)	4(1-4)
10. Service providers are able to provide instant feedback about the progress and situation of the service they provide.	7	4	0.64	2.82(1.17)	3(1-4)
11. Clarity of correspondence I receive from this organization.	7	4	0.64	3.00(1.26)	4(1-4)
12. It is easy to get the required information.	7	4	0.64	2.82(1.17)	3(1-4)
13. The information is offered timely but time to time it seems obstacle in accessing services due to requirement of extra documentation.	10	1	0.91	3.55(0.69)	4(2-4)
14. The communication is based on the attitude of the service seeker /client.	8	3	0.73	3.18(1.08)	4(1-4)
15. There is discrimination of information on the basis of gender and relationship *(Afano Manche).* *	6	5	0.55	2.55(1.37)	3(1-4)
16.Overall, the communication of this organization is good.	8	3	0.73	3.00(1.00)	3(1-4)
17.The appearance of the physical facilities of this organization represents the type of services that it provides.	9	2	0.82	3.09(0.94)	3(1-4)
18. Cleanliness and token system is symbol of good accessibility.	8	3	0.73	3.27(1.27)	4(1-4)
19. There is enough waiting area for service seeker.	7	4	0.64	2.82(1.17)	3(1-4)
20. This organization has used modern technology for accessing service.	10	1	0.91	3.55(0.69)	4(2-4)
21. There are sufficient service providers.*	4	7	0.36	2.45(1.29)	2(1-4)
22. There is convenient location regarding service facility. *	6	5	0.55	2.73(1.19)	3(1-4)
23. There is always associated a risk of uncertainty even though it cost high for the services.*	5	6	0.45	2.55(1.13)	2(1-4)
24. The physical facilities of the organization are alluring.	8	3	0.73	3.09(1.04)	3(1-4)
25. Materials associated with the service (such as pamphlets and notices) seems good at the organization.*	6	5	0.55	2.82(1.25)	3(1-4)
26. The organization has working hours convenient for all their client/service seeker.	7	4	0.64	2.91(1.04)	3(1-4)
27. I am able to access this service at a time that is convenient for me.*	6	5	0.55	2.45(1.29)	3(1-4)
28. The steps needed to make any change were easy to carry out.	7	4	0.64	3.00(1.10)	3(1-4)
29. This premises of the organization is easy to visit by service seeker. *	5	6	0.45	2.55(1.13)	2(1-4)
30. It is necessary to pay for intermediary to access service.*	5	6	0.45	2.36(1.29)	2(1-4)
31. When service seeker required a service from Service provider, it is easily accessible by any means.	9	2	0.82	3.18(0.98)	3(1-4)
32. Overall, I found this service was easily accessible.	8	3	0.73	2.91(1.14)	3(1-4)

33. Service providers provide service in time.	9	3	0.82	3.27(1.01)	4(1-4)
34. organization is trustworthy.	8	3	0.73	2.82(1.08)	3(1-4)
35. Service delivery takes comparably long time.	7	4	0.64	2.91(1.22)	3(1-4)
36. The organization keeps their records accurately and systematically.	7	4	0.64	2.73(1.27)	3(1-4)
37. The working hours of this organization is convenient to me.	7	4	0.64	2.82(1.33)	3(1-4)
38. It takes long time to receive service from this organization.	7	4	0.64	3.00(1.10)	3(1-4)
39. the organization fulfills the expectation of service seeker timely.*	6	5	0.55	2.82(1.25)	3(1-4)
40. organization provide the service at the single step and through one door policy.*	5	6	0.45	2.64(1.36)	2(1-4)
41. It takes more than usual time to access for the service.	7	4	0.64	2.91(1.38)	4(1-4)
42. Service provider in this organization gives prompt services to service seekers.*	6	5	0.55	2.55(1.37)	3(1-4)
43. The organization enables me to complete the task quickly. *	3	8	0.27	2.09(1.14)	2(1-4)
44. The organization informs me about the time to complete the service process.*	5	6	0.45	2.27(1.19)	2(1-4)
45. The service provider understands that delay in the service hampers the long run reputation of their organization. *	5	6	0.45	2.36(1.29)	2(1-4)
46. The schedule of working hours fits my needs.*	4	7	0.36	2.09(1.22)	2(1-4)
47. Overall, time of service is good.*	6	5	0.55	2.55(1.21)	3(1-4)
48. Customers of this organization feel safe in their interaction with these firms' Service provider.*	5	6	0.45	2.64(1.21)	2(1-4)
49. My personal information is secure.	7	4	0.64	2.91(1.04)	3(1-4)
50. I feel secure in providing sensitive information. *	5	6	0.45	2.82(1.17)	2(1-4)
51. I have confidence that the individuals responsible for this organization will not misuse my personal information. *	6	5	0.55	2.73(1.19)	3(1-4)
52. The privacy of the service seeker is protected on this organization.*	6	5	0.55	2.64(1.29)	3(1-4)
53. The e-government web site assures me of the security it provides.	8	3	0.73	3.00(1.00)	3(1-4)
54. It does not share my personal information with other without my permission. *	6	5	0.55	2.82(1.25)	3(1-4)
55. This organization is equipped with adequate security features.	8	3	0.73	3.27(0.90)	4(2-4)
56. I am confident towards the activities performed in this organization.	8	3	0.73	3.27(1.10)	4(1-4)
57. In every situation, this organization protects from the probable financial losses for me.	8	3	0.73	3.27(0.90)	4(2-4)
58. Overall, I trust that this organization is secured.	7	4	0.64	3.00(1.10)	3(1-4)
59. Service providers follow the rules.	9	2	0.82	3.27(0.79)	3(2-4)
60. Service providers properly listen to and understand concerns of service recipients.	7	4	0.64	2.82(1.17)	3(1-4)
61. Service providers provide clear information about the service.	7	4	0.64	2.64(1.21)	3(1-4)
62. Service providers behave in a friendly/cordial manner.	7	4	0.64	3.00(1.10)	3(1-4)
63. Service providers do not create any hassles/ problems.*	5	6	0.45	2.82(1.17)	2(1-4)

64. Service providers are willingness to accept complaints or criticisms at the same time.*	6	5	0.55	2.73(1.35)	3(1-4)
65. Service providers are efficient.*	6	5	0.55	2.91(1.14)	3(1-4)
66. Service providers are always readiness to provide service.	7	4	0.64	2.91(1.22)	3(1-4)
67. Service provider are not busy to respond to service seeker 's requests.*	6	5	0.55	2.64(1.12)	3(1-4)
68. Service providers have the capability to answer service seeker's questions.	9	2	0.82	3.18(0.98)	3(1-4)
69. Service providers give explanations and instructions to their service seekers in a friendly way.	7	4	0.64	3.00(1.10)	3(1-4)
70. Service provider respect for service seekers. *	6	5	0.55	2.91(1.14	3(1-4)
71. Service providers are polite and friendly. *	4	7	0.36	2.64(1.12)	2(1-4)
72. Service providers in this organization are always willing to help service seekers.	9	2	0.82	3.27(0.79)	3(2-4)
73. Service provider asked for bribe for receiving service.	7	4	0.64	3.09(1.14)	4(1-4)
74. The behavior of Service providers in this organization is confidential with regard to service seekers.	7	4	0.64	2.82(1.17)	3(1-4)
75. Service providers in this organization are consistently courteous with service seekers.*	6	5	0.55	3.00(1.00)	3(2-4)
76. Service provider in this organization give individual attention to service seekers.	8	3	0.73	3.18(1.08)	4(1-4)
77. Service provider of this organization understand the specific needs of their service seekers.	8	3	0.73	3.00(1.18)	3(1-4)
78. Service providers in the organization try to get rid of the service seeker.*	5	6	0.45	2.45(1.37)	2(1-4)
79. Service providers of the organization are only concerned about their monetary benefits.	8	3	0.73	3.09(1.22)	4(1-4)
80. All the Service providers are closely linked to perform the task in this organization.	7	4	0.64	2.82(1.17)	3(1-4)
81. Service provider takes bribery indirectly for the smooth operation of organization.	7	4	0.64	2.73(1.27)	3(1-4)
82. Service provider perceives that relationship matters in receiving public service.	10	1	0.91	3.45(0.69)	4(2-4)
83. Change in behavior of Service provider due to political interventions are aimed at personal benefits.	8	3	0.73	3.00(1.34)	4(1-4)
84. Service providers are well dressed and appear neat.	7	4	0.64	2.91(1.38)	4(1-4)
85. The Service provider of the organization practices the service culture.	8	3	0.73	3.09(1.04)	3(1-4)
86. Service providers are capable to answer service seeker's questions.	10	1	0.91	3.45(0.69)	4(2-4)
87. Overall, the behavior of Service provider is good.	9	2	0.82	3.09(0.94)	3(1-4)

Sources: Field Survey 2025

(A) is Expert in Agreement and (DA) is Expert in Disagreement

***** *Indicated items completely deleted*

In the iteration round, 87 items were asked to the expert. Out of 87 items, 33 items reached consensus according to the above standard criteria. All 33 items were relevant by at least equal or more than 70% of experts have agreed consensus in the items. The remaining items 54 items that did not reach consensus. Out of 87 items, 24 items are more than 60% of consensus agreement of experts and rest of items were eliminated. In the iteration process, the consensus items of 64% is borderline of mildly agreed

consensus category. Even though, these items need to be further analysis and shall be represented in the retained items. The mean values are ranged from 2.09 to 3.45 and the SD standard deviation (SD) are varied from 0. 69 to 1.38. this shows I-CVI score more than 0.70 but notably exhibited there is high SD values.

Table 4 represents the overall comments given by the experts. Based on the qualitative feedback,

Table 4

Overall Comments of Panel Expert

Overall comments and suggestions
All covering items should be minizine and find out about more sub construct
More simple language should be written
Focus on items regarding to relationship and accountability related
Sentence clarity is required
Questions regarding understanding of the laws and regulation of government, does citizen has knowledge about it ?
Repeated and similar questions should be removed, validity
Write the items regarding to perception of grievances handling and procedure
Items' problem (revised/modified in clarity) by experts
Items: 1, 2, 4,8.10.11,12,14, 15,16, 17,18, 20, 21,22,23,24,26, 30,32, 34, 35, 36, 37, 38, 39, 46, 48, 56, 67,78,

Table 5 exhibits the item modified, rephrase and retained according to the feedback given by experts.

Table 5

Item Modified and Refined According to Qualitative Feedback

Original wording	Refined/ Modified	Reason for changing items
Overall, the communication of this organization is good.	the communication of this organization is good.	The overall cannot measure perfectly in organization
Overall, I found this service was easily accessible.	I found this service was easily accessible.	The overall cannot measure perfectly in organization
Overall, the behavior of Service provider is good.	the behavior of Service provider is good.	The overall cannot measure perfectly in organization
The service provider has direct, effective one-to-one communication	The service provider has effective one-to-one communication	Original items were double meaning questions

Table 6

List of Final 33 Items Retained and Modified (33 Items) with Perceived Factors

Items	Perceived factor
1.The service provider has effective one-to-one communication	Communication
2. The service provider in this Institution tells the service seeker exactly when services will be performed.	Responsive
3. The government organization gives information on all the services that it offers.	Responsive
4. It is quick and easy to contact this organization.	Access
5. The Service provider inform punctually and sincerely about all the conditions of service,	Time
6. The organization explains how much does it cost for the service.	Cost
7. The digital web site provides information that is easy to understand.	Technology
8. The information is offer timely but time to time it seems obstacle in accessing services due to requirement of extra documentation.	Reliable
9. The communication is based on the attitude of the service seeker /client.	Behavior

10. The appearance of the physical facilities of this organization represents the type of services that it provides.	Tangible
11. Cleanliness and token system is symbol of good accessibility.	Accessibility
12. This organization has used modern technology for accessing service.	Technology
13. The physical facilities of the organization are alluring.	Tangible
14. When service seeker required a service from Service provider, it is easily accessible by any means.	Access
15. Service providers provide service in time.	Time
16. Organization is trustworthy.	Reliable
17. The e-government web site assures me of the security it provides.	Security
18. This organization is equipped with adequate security features.	Security
19. I am confident towards the activities perform in this organization.	Reliable
20. In every situation, this organization protects from the probable financial losses for me.	Security
21. Service providers follow the rules.	Accountable
22. Service providers have the capability to answer service seeker's questions.	Capacity
23. Service providers in this organization are always willing to help service seekers.	Willingness
24. Service provider in this organization give individual attention to service seekers.	Customization
25. Service provider of this organization understands the specific needs of their service seekers.	Responsive
26. Service providers of the organization are only concerned about their monetary benefits.	Behavior
27. Service provider perceives that relationship matters in receiving public service.	Relation
28. Change in behavior of Service provider due to political interventions are aimed at personal benefits.	Behavior
28. The Service provider of the organization practices the service culture.	Responsive
30. Service providers are capable to answer service seeker's questions.	Capacity
31. the communication of this organization is good.	Communication
32. I found this service was easily accessible.	Accessibility
33. the behavior of Service provider is good.	Behavior

DISCUSSION

The content validation of questionnaire to assess the perceived service quality delivery items measure by using modified Delphi method for developing of new reliable scale development and particularly establish a new tool in the perspective of Nepalese public sectors. The use of Delphi technique gives valuable feedback of field experts to item selection, modification and refinement as well as strong recommendation to develop a reliable scale. The iteration process in modified Delphi method helps to eliminate unnecessary items. Therefore, the modified Delphi technique along with content validation index was used. The similar iteration process has been seen in the study of Beamish et al. (2020). Lamb et al. (2020). Tang et al. (2020), WSRI, (2020) and Yilmaz Yalcinka et al. (2021).

There were no validated tools found in past studies regarding assessment of questionnaire items to perceived service quality in Nepalese perspective. alike the same situation, the new scale instrument needs to be developed. The same practices

used by De Oliveira et al. (2024) and different methods used by Adhikari et al. (2023).

This study is based on the quantitative indicator of content validity of modified scale for perceived service quality delivery in the perspective of Nepalese public sectors. It was revealed that the items scale with I-CVI more than 0.70 was found acceptable and good in terms of relevance. The CVI was calculated, should be at least 0.70, to be considered higher content validity but 0.80 should be for a new scale instrument (Dragostinov et al., 2022; Hamid & Habidin, 2024). The agreement in between 0.65 and 0.70 were considered moderate (Yilmaz Yalcmkaya et al., 2021). Where, the I-CVI≥ 0.78 is considered as excellent and in between I-CVI >0.78 and >0.50 need to be revised and < 0.50 is unacceptable (Kipli & Khairani, 2022; Polit et al., 2007; Pursio et al., 2024; Yamad et al., 2010). According to size of expert equal to eleven, or more than the consensus of items drop down to less than 60%, indicated the I-CVI value is 0.59 (Lawshe, 1975; Takom et al., 2025). There were numbers of previous studies that uses less than ten experts and more than two rounds of iteration in modified Delphi method of content validation for scale development in different field but there was no similar study to compare with this type of content validation of questionnaire of the study. However, using the similar method on other studies that calculate content validation for scale development by using the I-CVI studies had similar results that retained the validation items that has equal or more than 0.70 and moderately retained the items for scale (Abbady et al., 2021; Dragostinov et al., 2022; Hamid & Habidin, 2024; ; Hassanein et al., 2013; Shomah et al., 2025; Yilmaz Yalcmkaya et al., 2021; Yamada et al., 2010). Those studies assessed more than ten experts and I-CVI range 0.60 to 1.0 at items varying in terms of relevance for scale development. There are no ideal assessment criteria therefore the cut-off point is 0.70 but various literature explained their retaining criteria with reason. The qualitative feedback helps to include supplement items (Dragostinov et al., 2022). Besides, Table 3 also explains the mean value of each questionnaire items. The overall mean rating was 2.09 to 3.55. The mean value more than 3 is sufficient to be a good valid item in scale instrument. The study done by Allahyari et al. (2011) and Pearl, et al. (2017) indicated that the mean values and the values of higher than 1.5 or not less than fifty percent of the panels' decision indicated that the mean judgement is much closer to the five-point scale value of the strongly agree, therefore the judgement is good to use for scale development. The higher the SD indicating that the majority of the experts rated the item relevant but a minority experts refused the inclusion of items in domain therefore, the variations was appeared. It can be possible to individual difference of experts, strong claim for revision and enhancing clarity and consensus. The similar situation appeared in the study of Dragostinov et al., (2022). This is the one of the best method of items evaluation under unweighted approach of computing mean of raw items score for standardizing the scale (Armor, 1973). Generally, it does not make much difference for scale performance, either using unweighted (mean and sum score) and weighted (factor analysis) items score is computed of scales (Boateng et al., 2018). The expert panels agreed on the 33 items and rest of the items need to be modified or revised because some of the items are in the borderline of the validation. Some of the items need to be rephrased to make it clear and easily understand. A few expert comments on the items that are not relevant to ask for the citizen. Items need to be revised by considering all the comments from expert panels.

The researcher limits the research area therefore, further recommendation for analysis, to be carried out to test construct validity, discriminant validity, and convergent validity by using exploratory

factor analysis and confirmatory factor analysis using as statistical tool, as suggested by Ghazali et al. (2018) and Wang and Shahid (2024).

CONCLUSION

In conclusion, the instrument development, the content validity form is crucial tool to gather and undergoes the systematic process to develop the content validity on scale development. For this, modified Delphi method frequently deployed to scale development for the reliable and valid scale. Using statistically inclusion and exclusion criteria, expert ranking, expert opinion and feedback in the modified Delphi method in iteration process the list of items reduced further. Utilizing the Content Validity Index (I-CVI) and the considering the validation of cut-off point ($\geq$0.70), the questionnaire item instruments claims that the items capture the content of perceived service quality delivery domain and can be utilized in future by researchers for assessing the perceived SERVQUAL of Nepalese public services. The items are more valid, reliable, clear and understable. Therefore, the scale questionnaire can assess to measure the perceived service quality delivery in the perspective of Nepalese public sector. So that, the study embraces the theoretically significant implication and contribution of assessed questionnaire items to the further researcher and partitioners, policy maker and firm as well but this study has certain limitation. This study is calculation of just content validation index of questionnaire to assess the perceived service quality delivery, limited to Pokhara Metropolitan city may create problem in generalization and mostly rely on few self-administered survey and single iteration process that leads to study biases. Therefore, further recommendation to confirm the validity of the scale on the large scale of sample size of population by using exploratory factor analysis and confirmatory factor analysis is necessary. It is suggested to undergo a pilot study.

ACKNOWLEDGEMENTS

First of all, the author thanks to all the experts for their valuable responses for making the paper relevant for future researcher and the Centre for Research and Innovation (CRI) for their funding to the support of paper publication. The author also thanks to all the voluntarily participator in this paper publication. The special thanks to Professor, Dr. Vikash Kumar K.C. Department of Statistics, Prithvi Narayan Campus, Pokhara.

REFERENCES

Abbady, A. S., El-Gilany, A. H., El-Dabee, F. A., Elsadek, A. M., ElWasify, M., & Elwasify, M. (2021). Psychometric characteristics of the of COVID stress scales-Arabic version (CSS-Arabic) in Egyptian and Saudi university students. *Middle East Current Psychiatry, 28*(14), 1-9. https://doi.org/10.1186/s43045-021-00095-8

Adetunji1, O., Yadavalli, V.S.S., & Malada, A. (2013). Assessment of the quality of service provided by a national regulatory institution. *South African Journal of Industrial Engineering, 24*(1), 29-49.

Adhikari, Y., Wagley, M. P., & Dahal, N**.** **(**2023). Developing the contextualized SERVQUAL instrument for measuring the service quality of Nepali resorts: An application of the modified Delphi technique. *Sage Open, 13*(3), 1-16. https://doi.org/10.1177/21582440231195952

Aljukhadar, M., Belisle, J.-F., Dantas, D. C., Sénécal, S., & Titah, R. (2022). Measuring the service quality of governmental sites: Development and validation of the e-Government service quality (EGSQUAL) scale. *Electronic Commerce Research and Applications, 55*, 1-12. https://doi.org/10.1016/j.elerap.2022.101182

Allahyari T., Hassanzadeh Rangi, N., Khosravi Y., & Zayeri, F. (2011). Development and evaluation of a new questionnaire for rating of cognitive failures at work. *International Journal of Occupational Hygiene, 3*(1), 6-11.

Alanezi, M. A., Kamil, A., & Basri, S. (2010) A proposed instrument dimensions for measuring e-government service quality. *International Journal of U- and E-Service, Science and Technology, 3,* 1-18.

Armor, D. J. (1973). Theta reliability and factor scaling. *Sociological Methodology, 5,* 17-50. https://doi.org/10.2307/270831

Avkiran, N. K. (1994). Developing an instrument to measure customer service quality in branch banking. *International Journal of Bank Marketing, 12*(6), 10-18. http://dx.doi.org/10.1108/02652329410063223

Babakus, E. & Boller, G. W. (1992). An empirical assessment of the SERVQUAL scale. *Journal of Business Research*, *24*, 253–268.

Bakar, C., Seval Akgün, H., & Al Assaf, A. F. (2008). The role of expectations in patients' hospital assessments: A Turkish university hospital example. *International Journal of Health Care Quality Assurance*, *21*(5), 503-516. https://doi.org/10.1108/09526860810890477

Barabino, B., Deiana, E., & Tilocca, P. (2012). Measuring service quality in urban bus transport: a modified SERVQUAL approach. *International Journal of Quality and Service Sciences, 4(3*), p. 238–252. https://doi.org:10.1108/17566691211269567

Beamish, A. J., Brown, C., Abdelrahman, T., Ryan, H. E., Harries, R.I., Egan, R., Ansell, J., Evans, T. Hopkins, L., James, O., Lewis, S. Lewis, WG, Luton, O., Mellor, K., Robinson. D., Thomas, R., & Williams, A. (2020). *International surgical guidance for COVID-19: Validation using an international Delphi process - Cross-sectional study. International Journal of Surgery 79.* 309-316. https://doi.org/10.1016/j.ijsu.2020.06.015

Boateng, G.O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H.R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research. *Frontiers in Public Health 6(*149). https://doi.org/10.3389/fpubh.2018.00149

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O., & Alberti, C. (2011) Using and reporting the Delphi method for selecting health care quality indicators: A systematic review. *PLoSONE, 6*(6), e20476. https://doi.org/10.1371/journal.pone.0020476

Carman, J. M. (1990). Consumer Perceptions of service quality: An assessment of the SERVQUAL dimensions. *Journal of Retailing*, *66*(1), 33-35.

Caro, L. M., & García, J. A. M. (2007). Measuring perceived service quality in urgent transport service. *Journal of Retailing and Consumer Services, 14, 60-72. https://doi.org/10.1016/j.jretconser.2006.04.001*

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research, 5*(4), 194–197. https://doi.org/10.1016/S0897 1897(05)80008-4

De Oliveira, D. A., Ferreira, Dd. M., Lisboa, L. L., & Nobre, T. T. X. (2024). Preparation and validation of the content of an instrument to assess the quality of services of reception with risk classification in obstetrics. *PLoS ONE 19*(12): e0315816. https://doi.org/10.1371/journal.pone.0315816

DeOnna, J. (2006). *Developing and validating an instrument to measure the perceived job competencies linked to performance and staff retention of first-line nurse managers employed*

in a hospital setting [Unpublished PhD thesis]. College of Education, Pennsylvania State University.

DeVellis, R.F. (2012). *Scale development: Theory and application*. Los Angeles, CA: Sage Publications.

Dragostinov, Y., Harðardóttir, D., McKenna, P. E., Robb, D. A., Nesset, B., Ahmad, M. I., Romeo, M. Lim, M. Y., Yu, C. Jang, Y., Diab, M., Cangelosi, A., Demiris, Y., Hastie, H., & Rajendran, G. (2022). Preliminary psychometric scale development using the mixed methods Delphi technique. *Methods in Psychology, 7*, 1-10. https://doi.org/10.1016/j.metip.2022.100103

Engdaw, B. D. (2019). The impact of quality public service delivery on customer satisfaction in Bahir Dar city administration: The case of Ginbot 20 sub-city. *International Journal of Public Administration.* https://doi.org/10.1080/01900692.2019.1644520

Eubank, B. H., Mohtadi, N. G., Lafave, M. R., Wiley, J. P., Bois, A. J., Boorman, R. S., & Sheps, D. M. (2016). Using the modified Delphi method to establish clinical consensus for the diagnosis and treatment of patients with rotator cuff pathology. *BMC Medical Research Methodology*, *16*, 56. https://doi.org/10.1186/s12874-016-0165-8

Fackrell, K., Potgieter, I., Shekhawat, G. S., Baguley, D. M., Sereda, M., &, Hoare, D. J. (2017). Clinical interventions for hyperacusis in adults: a scoping review to assess the current position and determine priorities for research, *BioMed Research International*. https://doi.org/10.1155/2017/2723715

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational and Psychological Measurement*, *58*(3), 357-381. https://doi.org/10.1177/0013164498058003001

Garson, G. D. (2014). *The Delphi method in quantitative research*. Statistical Associated Publishing.

Ghazali, N., Nordin, M. S., Hashim, S., & Hussein, S. (2018). Measuring content validity: Students' self-efficacy and meaningful learning in Massive Open Online Course (MOOC) scale. *Advances in Social Science, Education and Humanities Research, 115*, 128-133. https://doi.org/10.2991/icems-17.2018.25

Grant, J. S. & Davis, L. L. (1997). Focus on quantitative methods: Selection and use of content experts for instrument development. *Research in Nursing & Health*, *20*, 269–274.

Grime, M. M. & Wright, G. (2016). Delphi method. *Wiley Stat Ref: Statistic Reference Online,* 1–6. https://doi.org/10.1002/9781118445112.stat07879

Hamid, M. A. & Habidin, N. F. (2024). Scale development and content validity assessment for manufacturing performance. *International Journal of Research and Innovation in Social Science (IJRISS), 8*(15),230-241.https://doi.org/https://dx.doi.org/10.47772/IJRISS.2024.ICAME2420

Hassanein, S., El-Sayed, Z. & Raouf, H. A. (2013). Validity and reliability of scales used for objective structured clinical examination: Piloting modified tools. *Journal of Biology, Agriculture and Healthcare, 3,* 55-62.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validation in psychological assessment: A functional method of concepts and methods. *psychological assessment, 7*(3). 238-248.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. Journal of Management, 21(5), 967–988. http://dx.doi.org/10.1177/014920639502100509

Iqbal, S. & Pipon-Young, L. (2009). The Delphi method: The psychologist.

The British psychological Society. bps.org.uk. Available at: https://thepsychologist.bps.org.uk/volume-22/editio n-7/delphi-method.

Kant, R. & Jaiswal, D. (2017). The impact of perceived service quality dimensions on customer satisfaction: an empirical study on public sector banks in India. *International Journal of Bank Marketing, 35* (3), 411-430. http://dx.doi.org/10.1108/IJBM-04-2016-0051

Kapuscinski, A. N., & Masters, K. S. (2010). The current status of measures of spirituality: a critical review of scale development. *Psychology of Religion and Spirituality, 2*(4), 191–205. http://dx.doi.org/10.1037/a0020498.

Kipli, M., & Khairani, A. Z. (2020). Content validity index: An application of validating CIPP instrument for programme evaluation. *IOER International Multidisciplinary Research Journal, 3*(4), 31–40.

Lamb, C. R., Desai, N.R., Angel, L., Chaddha, U., Sachdeva, A., Sethi, S., Bencheqroun, H., Mehta, H., Akulian, J., Argento, A.C., Diaz-Mendoza, J., Musani, A., & Murgu, S. (2020). *Use of tracheostomy during the COVID-19 pandemic: CHEST/AABIP/AIPPD: Expert Panel Report. Chest, (158).* https://doi.org/*10.1016/j.chest.2020.05.571*

Lamm, K. W., Lamm, A. J., & Edgar, D. (2020). Scale development and validation: Methodology and recommendations. *Journal of International Agricultural and Extension Education, 27*(2), 24-35. https://doi.org/10.5191/jiaee.2020.27224

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 381–5. https://doi.org/10.1097/00006199-198611000-00017.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50(*9), 741-749. https://doi.org/10.1037/0003-066X.50.9.741

Mohamad Marzuki, M. F., Yaacob, N. A., & Yaacob, N. M. (2018). Translation, cross-cultural adaptation, and validation of the Malay version of the system usability scale questionnaire for the assessment of mobile apps. *JMIR human factors*, *5*(2), e10308. https://doi.org/10.2196/10308

Morgado, F.F., Meireles, J. F., Neves, C.M., Amaral, A., Ferreira, M.E., (2017). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia Reflexao e Crítica 30*(3). https://doi.org/10.1186/s41155-016-0057-1

Ozair M. M., Baharuddin, K. A., Mohamed, S. A., Esa, W., &Yusoff, M.S.B. (2017). Development and validation of the knowledge and clinical reasoning of acute asthma management in emergency department (K-CRAMED). *Education in Medicine Journal. 9*(2), 1–17. https://doi.org/10.21315/ eimj2017.9.2.1

Parasuraman, A., Zeithaml, A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, Spring, 12 – 40.

Pearl, C. E., Vasquez, E., Marino, M. T., Wienke, W., Donehower, C., Gourwitz, J., Rosenberg, M. S., & Duerr, S. R. (2017). Establishing content validity of the quality indicators for classrooms serving students with Autism Spectrum Disorders instrument. *Teacher Education and Special Education*, *41*(1), 58-69. https://

doi.org/10.1177/0888406416687814

Pokharel, T., Dahal, A. & Adhikari, R. (2017). *Public services in Nepal: Citizen's experiences*. Lalitpur, Nepal: Nepal Administrative Staff College.

Pokharel, T., Subedi, BP., Adhikari S. H., Adhikari, R., & Gupta, AK. (2018). *Quality of public service in Nepal: Nepal National Governance Survey 2017/18.* Thematic Paper. Lalitpur: Nepal Administrative Staff College.

Polit, D. F. & Beck C. T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*(5), 489–97. https://doi.org/10.1002/ nur.20147

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). *Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing and Health, 30*(4), 459–467. https://doi.org/10.1002/nur.20199

Prakash S, & Pallepati A. (2016). Cross-cultural content validation of a modified service quality questionnaire in Kannada. *Journal of Indian Association of Public Health Dentistry, 14* (2), 171-174.

Pursio, K., Kankkunen, P., & Kvist, T. (2024). Professional autonomy among registered nurses-validation of the translation of the dempster practice behaviour scale and survey results. *Nursing open*, *11*(5), e2185. https://doi.org/10.1002/nop2.2185

Roebianto, A., Savitri, S. I., Aulia, I., Suciyana, A., & Mubarokah, L. (2023). Content validity: Definition and procedure of content validation in psychological research. *Testing, Psychometrics, Methodology in Applied Psychology, 30*, 5-18. https://doi.org/10.4473/TPM30.1.1

Santoyo-Sánchez, G., Merino-Soto, C, Flores-Hernández, S., Pelcastre-Villafuerte, B. E., Reyes-Morales, H. (2022). Content validity of a scale designed to measure the access of older adults to outpatient health services. *International Journal of Environmental Research and Public Health, 19 (*10102), 1-16. https://doi.org/10.3390/ijerph191610102

Schwab, D. P. (1980). Construct validity in organizational behavior. In Staw, B. & Cummings, L. L. (Eds.), *Research in Organizational Behavior* (2nd ed., pp. 3-43). Greenwich, JAI Press.

Shomah, M. I., El-Gilany, A. H., Mashaly, M., Elshaer, M., Sadeq, M., & El-Mashad, N. (2025). Adaptation and validation of a scale for quality standard measures in hospital blood bank: A pilot study. *Asian Journal of Transfusion Science.* https://doi.org/10.4103/ajts.ajts_21_24

Smith, G. T. & McCarthy, D. M. (1995). Methodological consideration in the refinement of clinical assessment of instruments. *Psychological Assessment, 7*, 300-308.

Stines, A. C. (2003). *Forecasting the competencies that will define best in class business to business market managers: An emergent Delphi-hybrid competency forecasting model* [Unpublished doctoral dissertation]. College of Education, Pennsylvania State University.

Takom, L., Menggat, R., Talib, R., & Edmund, N. (2025). Using content validity index to estimate instrument validity in social sciences research. *Buletin Ilmiah Psikologi and Kaunseling, 7* (1), 61-78.

Tang, Z., Sun, B., & Xu, B. (2020). A quick evaluation method of nosocomial infection risk for cancer hospitals during the COVID-19 pandemic. *Journal of Cancer Research and Clinical Oncology, 146,* 1891-1892. https://doi.org/10.1007/s00432-020-03235-5

Tefera, O. & Govender, K. (2016). From SERVQUAL to HOTSPERF: Towards the development and validation of

an alternate hotel service quality measurement instrument, *African Journal of Hospitality, Tourism and Leisure, 5* (4), 1-17.

Thapa, M. (2022). Impact of service quality on student satisfaction: a case study of public sector higher secondary institution**.** THE BATUK: *A Peer Reviewed Journal of Interdisciplinary Studies,8* (1), 1-15. https://doi.org/10.3126/batuk.v8i1.43500

Wang, F., & Sahid, S. (2024). Content validation and content validity index calculation for entrepreneurial behavior instruments among vocational college students in China. *Multidisciplinary Reviews*, *7*(9), 2024187. https://doi.org/10.31893/multirev.2024187

Welsh Surgical Research Initiative (WSRI) (2020). Collaborative, surgery during the COVID-19 pandemic: Operating room suggestions from an international Delphi process, *British Journal of Surgery*, *107* (11), 1450–1458, https://doi.org/10.1002/bjs.11747

Williamson, P. R., Altman, D. G., Blazeby, J. M., Clarke, M., Devane, D., Gargon, E., &Tugwell, P. (2012). Developing core outcome sets for clinical trials: issues to consider. *Trials 13*(132). https://doi.org/10.1186/1745-6215-13-132

Wong, K. L., Ong, S. F., & Kuek, T. Y. (2012). Constructing a survey questionnaire to collect data on service quality of business academics. *European Journal of Social Sciences ,29* (2), 209-221.

Yamada, J., Stevens, B., Sidani, S., Watt-Watson, J., & de Silva, N. (2010). Content validity of a process evaluation checklist to measure intervention implementation fidelity of the EPIC intervention. Worldviews Evid Based Nurs *7*(3*),* 158–164. https://doi.org/10.1111/j.1741-6787.2010.00182.x .

Yılmaz Yalçınkaya, E., Karadağ Saygı, E., Özyemişci Taşkıran, Ö., Çapan, N., Kutlay, Ş., Sonel Tur, B., El, Ö., Ünlü Akyüz, E., Tekin, S., Ofluoğlu, D., Zİnnuroğlu, M., Akpınar, P., Özekli Mısırlıoğlu, T., Hüner, B., Nur, H., Çağlar, S., Sezgin, M., Tıkız, C., Öneş, K., İçağasıoğlu, A., & Aydın, R. (2021). Consensus recommendations for botulinum toxin injections in the spasticity management of children with cerebral palsy during COVID-19 outbreak. *Turkish Journal of Medical Sciences*, *51*(2), 385–392. https://doi.org/10.3906/sag-2009-5

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal 11*(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A. (2015). Design and implementation content validity study: Development of an instrument for measuring patient-centered communication. *Journal of Caring Sciences, 4*(2), 165–178. https://doi.org/10.15171/jcs.2015.017

Zeithaml, V. A., Parasuraman, A. & Berry, L. L. (1990). *Delivering quality service: Balancing customer perceptions and expectations*. The Free Press, New York.