Educational Quest: An Int. J. of Education and Applied Social Sciences: Vol. 17, No. 01, pp. 21-27, April 2026
DOI: 10.30954/2230-7311.1.2026.3
Peer Reviewed Journal

Mastering Content Validation: A Guide to Calculating the Content Validity Index (CVI)

Iram Sleem Malik* and Mohammad Hasan

Maulana Azad National Urdu University, College of Teacher Education, Bhopal, India

*Corresponding author: iramm8842@gmail.com

Received: 09-01-2026 **Revised:** 30-03-2026 **Accepted:** 07-04-2026

ABSTRACT

A crucial component of developing research tools is content validity, which guarantees that the tool measures the target components as intended. When evaluating content validity, especially by expert assessment, the Content Validity Index (CVI) is a widely utilised technique. The procedures for constructing the tool, the Selection of the expert panel, the Comprehensive guide for calculating CVI, and the statistical calculations of both item-level (I-CVI) and scale-level (S-CVI) indices are all covered in this paper. A case study illustrates how these techniques are used in practice and provides researchers with an easy-to-follow, sequential procedure.

Keywords: Content Validity, Tool Development, Content Validity Index, CVI, I-CVI, S-CVI and Research Methods

The five sources of validity proof and the consequences are the content, response procedure, core structure, relationship to additional factors, and response structure (Cook and Beckman, 2006). Content validity is demarcated as the extent to which components of an evaluation tool are pertinent and as a representation of the intended construct for a certain evaluation objective (Cook and Beckman, 2006; Haynes *et al.* 1995). A tool for evaluation relates to the specific way of obtaining information from psychological tests, such as questionnaires. The components of an assessment tool cover every facet of the measuring procedure that may have an impact on the information gathered, including questionnaire items, formats for responses, and specifications. The construct refers to a specific concept, The characteristic, dimension, or variable that is being measured. The goal of the evaluation is to determine the intended functions of the tool for measurement. For example, the medical student's Stressor Questionnaire (MSSQ) was created to determine the medical students' sources of stress (Yusoff, 2017), and the Anatomy Education Environment Measurement Inventory (AEEMI) was created to gauge the anatomy educational environment at medical schools (Hadie *et al.* 2017). The applicability of a measurement tool refers to the suitability of its components for the intended structures and functions of evaluation, but the representativeness of an evaluation instrument relates to the extent to which its components are proportionate to the specified construct's aspects (Haynes *et al.* 1995). Despite content validity's two facets (that is, the relevance and representation of a tool for evaluation), the significance of the assessment tool that Davis promoted has been widely used to assess the content validity (Davis, 1992; Polit and Beck, 2006; Polit *et al.* 2007). Establishing content validity is crucial for ensuring the validity of assessment tools, such as questionnaires, particularly in research contexts. scientific purposes. Haynes *et al.* highlighted that "Inferences from assessment tools with unsatisfactory content validity will be suspect, even when other indices of validity are satisfactory." The proof of content validity is often

How to cite this article: Malik, I.S. and Hasan, M. (2026). Mastering Content Validation: A Guide to Calculating the Content Validity Index (CVI). *Educational Quest: An Int. J. Edu. Appl. Soc. Sci.,* **17**(01): 21-27.

Source of Support: None; **Conflict of Interest:** None

represented by the content validity index (CVI) (Davis, 1992; Lynn, 1986). For example, a few recent studies (Hadie *et al.* 2017; Ozair *et al.* 2017; Marzuki *et al.* 2018) determined the information authenticity by employing CVI to increase the legitimacy of a tool for evaluation.

Objectives of the Paper

This study aims to give researchers a useful framework for ensuring the content validity of their instruments or research tools by providing a comprehensive guide on computing the CVI. The full process, from building the tool and choosing an expert panel to computing the I-CVI and S-CVI, will be covered in the article. The guide will be accessible and adaptable to a broad range of research situations, as a case study will demonstrate the implementation of these methodologies.

An Overview of Content Validity Index (CVI)

One of the most effective ways to verify that a measurement tool is valid is to have professionals review the questionnaire's content. The test's content is analysed to ensure it accurately reflects the structure it intends to evaluate. This is known as content validity. Content validity measures how well the measuring instrument's content was sampled (Ling Zhang *et al.* 2022). The Content validation addresses whether the features being measured are represented in the content of this measuring tool. In essence, it involves judgment on the part of the individual or people determining the identification of the items (Almanasreh *et al.* 2019; Taherdoost, 2016).

Prior research has demonstrated two methods for assessing content validity: the qualitative technique and the quantitative method. The need, clarity, and applicability of the questionnaire items that are verified by expert opinion are all part of the qualitative content evaluation process (Patrick *et al.* 2011; Eriksson *et al.* 2015). In terms of quantitative assessment, the experts have provided qualitative input, and the items' content has been corrected. A quantitative method for assessing an instrument's content validity is the Content Validity Index (CVI). It includes a panel of experts systematically evaluating both the scale as a whole and every individual item. The item-level Content Validity Index (I-CVI) and the scale-level Content Validity Index I (S-CVI) are the two levels at which the CVI is normally computed. The S-CVI offers a broad indication of the scale's validity, whereas the I-CVI quantifies the percentage or proportion of experts who concur on the relevance of each item. Because of its simple interface, precision, and capacity to measure expert consensus, the CVI is often employed.

Tool Construction

Identifying the components

Defining the components of the measurement tool must be done. It is the initial step towards developing a valid tool. Constructs are abstract ideas that a research tool is designed to evaluate. Examples include knowledge, motivation and satisfaction, etc. To ensure that the tool captures all key aspects of the idea, the identification of the components or dimensions is crucial. It requires a thorough examination of existing literature. To move to the next step in the creation process, it is the responsibility of researchers to ensure that every component has a clear definition.

Items Generation

The subsequent stage calls for the creation of items. These items represent dimensions after identification. Expert interviews, existing literature, various published works, and different exploratory research can guide this process and can be of assistance when generating new items. In addition, each component should cover a specific construct feature. Illustratively, one concept might be "student engagement." In such cases, the questions should assess cognitive, emotional, and behavioural involvement. Furthermore, redundancy needs to be avoided. Each item must be precise and concise. Additionally, it should be pertinent to the construct.

For example, we might create Items like:

- ❒ • "The student participates keenly in classroom discussions."
- ❒ • "The student shows interest in the subject matter."
- ❒ • "The student regularly completes assignments on time."

These Items must be relevant. They should cover each component of the construct. This ensures the validity of content, which is vital for the intended demographic.

Content Validity Procedure

Several essential steps are involved in content validation to guarantee the reliability of a measuring instrument. The steps involved in computing CVI are as follows:

VALIDATION OF STUDENT ENGAGEMENT IN CLASSROOM SCALE:

Content Validation

Dear Experts,

I'd like your expert assessment of the item's degree of relevance and clarity concerning the measured domains/ dimensions. Please refer to the rating scale below for guidance.

Degree of Relevance:

1. The Item lacks relevance to the determined dimension.
2. The item has some relevance to the determined dimension.
3. The item is very relevant to the determined dimension.
4. The item is extremely relevant to the determined dimension.

The Content Validation Performa

Fig. 1: An illustration of instructions and rating scale in CVI Performa

The first step is creating a content validation performa that outlines the domains and items to be evaluated. Experts should be given guidelines on how to rate the relevance of each item on the form. Fig. 1 shows a sample content validation form complete with a rating scale that allows experts to rank Items according to their relevance.

Dimension 1: STUDENTS INVOLVEMENT

Operational Definition: The extent to which students actively take part in and participate in class discussions, assignments, and activities. This comprises their behavioral participation (attending courses, participating in discussions, finishing assignments), emotional involvement (e.g., demonstrating interest and motivation), and cognitive involvement (e.g., Thinking analytically about the subject matter).

Tested Items	Relevance
1. The student actively thinks about the course material during class discussions.	1 2 3 4 ○○○○
2. The student asks questions demonstrating critical thinking about the subject matter.	1 2 3 4 ○○○○
3. The student is strongly interested in learning the course material.	1 2 3 4 ○○○○
4. The student shows enthusiasm during classroom activities.	1 2 3 4 ○○○○
5. The student regularly attends and participates in class.	1 2 3 4 ○○○○
6. The student completes assignments on time and with effort.	1 2 3 4 ○○○○

Fig. 2: An example of the layout of CVI Performa with dimension, its operational definition, and items measuring the dimension

1. Expert Panel Selection

The panel's competence in evaluating the items has major effects on the validity of the CVI computation. The selection of experts needs to be predicated on their previous experience, competence and knowledge of the components under evaluation. To maintain a balance between reliability and practicality, the panel should ideally consist of five to ten experts. Since it offers a variety of opinions on the items, the backgrounds of the experts can be diverse, which can improve the credibility of their conclusions. The selection procedure must be open and transparent, with specific expert selection criteria. An individual's level of competence in the subject matter being researched is typically considered when choosing someone to provide feedback on a tool for evaluation (such as a questionnaire).

Table 1 summarises the suggested number of professionals and how it affects the standard CVI cut-off score. While most recommendations recommend a minimum of six

experts, it may be agreed that two experts are the least satisfactory for content validation. The total number of experts for content validity should range from 6 to 10, considering suggestions (Davis, 1992; Lynn, 1986) and the author's experience.

2. Content Validity Index (CVI) Acceptable Values Determining Expert Panel Size

The number of experts engaged in the evaluation process determines the allowable Content Validity Index (CVI) values. According to Davis (1992, a CVI of at least 0.80 has been suggested for a panel of two experts. As recommended by Polit & Beck (2006) and Polit *et al.* (2007), the CVI should be 1.00 if the expert panel consists of three to five members. A CVI of at least 0.83 has been proposed for panels with six or more experts, as suggested by Polit & Beck (2006), Polit *et al.* (2007), and Lynn (1986). Similarly, Lynn (1986) recommends a CVI of not less than 0.83 when the panel consists of six to eight experts. According to Lynn (1986), panels including 9 or more experts should have a slightly decreased acceptable CVI of 0.78.

Table 1: The Impact of Expert Numbers on the Acceptable CVI Cut-Off Range

Panel Size (Number of Experts)	Minimum Acceptable CVI Score	Cited References
2 Experts	0.80 or higher	Davis (1992)
3 to 5 Experts	1.00	Polit & Beck (2006), Polit *et al.* (2007)
6 or more Experts	0.83 or higher	Polit & Beck (2006), Polit *et al.* (2007)
6 to 8 experts	0.83 or higher	Lynn (1986)
9 or more Experts	0.78 or higher	Lynn (1986)

3. Conducting Content Validation

Online platforms or in-person interactions can be used to validate content. When selecting the approach, factors like response rate, time, and cost should be taken into account. Whereas the online technique is efficient and economical when appropriate follow-up processes are followed, the in-person approach may be more expensive but offers real-time feedback.

4. Evaluating Dimensions and Items

Experts examine the items and dimensions listed on the validation form. They evaluate whether every item accurately presents the intended concept and provide suggestions for improvement. Fig. 2 shows a hypothetical layout for item and dimension definitions.

5. Rating/Scoring of Each Item

Experts evaluate each item after examining it to see how relevant it is to the construct (Figs. 1 and 2). The ratings must be gathered and ready for further evaluation.

6. Calculating the Content Validity Index (CVI)

CVI has dual forms: The Item-level CVI (I-CVI) and the Scale-level CVI (S-CVI).

1. Item-Level Content Validity Index (I-CVI)

The proportion of experts who score an item as either three or four on a 4-point or 5-point Likert scale, where 1 represents lack of relevance, 2 indicates somewhat relevant, 3 indicates rather relevant, and 4 indicates extremely significant, is used to compute the I-CVI. An item's I-CVI would be 0.80, for instance, if eight experts out of ten score it a three or a four. However, this criterion may change based on the scenario. According to Polit & Beck (2006) and Polit *et al.* (2007), the CVI score shouldn't be less than 0.83 even when at least six experts are assessing the research tool (Yusoff, M.S.B. 2019).

2. Scale-Level Content Validity Index (S-CVI)

There are two ways to calculate the S-CVI: the first computes the I-CVI scores for entire items on the scale and averages them (S-CVI/AVE); the second computes the percentage of items that all experts evaluated as 3 or 4 in relevance (S-CVI/UA). Fig. 3 provides an overview of the definitions and formulas used for calculating the CVI indices. Before calculating the CVI, as indicated in Table 2, the relevance ratings must first be categorised as either 1 (for a rating of 3 or 4) or 0 (for a rating of 1 or 2) to compute the CVI.

To determine how different CVI indices are calculated, Table 2 displays the relevance ratings given by 10 experts for an item scale.

To demonstrate the CVI indices calculation (as shown in Fig. 3), the examples are based on the data presented in Table 2.

(i) **Experts in agreement**: Sum the relevance ratings all experts give each item. For instance, for Item no. 2, the experts in the agreement are calculated as (1+1+1+1+1+1=1+1+1+1) =10.

(ii) **Universal agreement (UA)**: A score of "1" is

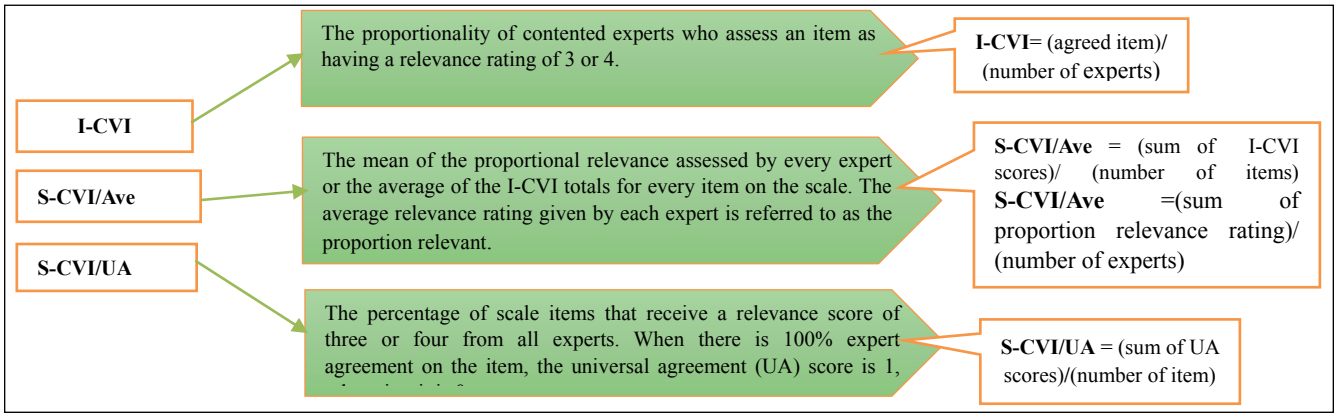

Fig. 3

Note: *The description and formula are derived from the guidelines provided by Lynn, Davis, Polit & Beck, and Polit et al.*

Table 2

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	Expert 7	Expert 8	Expert 9	Expert 10	Expert in Agreement	I-CVI	UA
Items													
Q 1	0	0	0	0	0	0	0	0	0	0	0	0	0
Q 2	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 3	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 4	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 5	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 6	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 7	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 8	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 9	1	1	1	1	1	1	1	1	1	0	9	0.9	0
Q 10	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 11	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 12	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 13	1	1	1	1	1	1	1	1	1	1	10	1	1
Q 14	1	1	1	1	1	1	1	1	1	1	10	1	1
Proportion relevance	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.86	**S-CVI/Ave**	0.93	
											S-CVI/UA		0.86

given to an item where 100% of the experts agree on the relevance. For instance, Item No. 2 received a score of 1 since all experts gave it a relevant rating of 1, whereas Q1 received a score of 0 because some experts did not give it a 1.

(iii) **I-CVI:** Calculate the I-CVI by dividing the number of experts who agree by the total number of experts. For example, the I-CVI for Item no. 2 is calculated as 10 divided by 10 experts, resulting in 1.

(iv) **S-CVI/Ave** (based on I-CVI): which is derived from the I-CVI, is the mean of the I-CVI scores for every item. For instance, the S-CVI/Ave is calculated as [(0+1+1+1+1+1+1+1+0.9+1 +1+1+1+1)/14) = 0.93.

(v) **S-CVI/Ave** (based on proportion relevance): The S-CVI/Ave is [(0.93 + 0.93 + 0.93 + 0.93 + 0.93 + 0.93 + 0.93 + 0.93 + 0.93 + 0.86)/10]. This is the mean of the proportion of relevance rankings obtained from every expert. **The S-CVI/UA** measures the average of the UA values for each item. For instance, the

S-CVI/UA [(0+1+1+1+1+1+1+1+0+1+1+1+1+1) /14] = 0.86.

The computation above demonstrates that the I-CVI, S-CVI/Ave, and S-CVI/UA reach an adequate stage, confirming that the questionnaire scale has attained acceptable content validity. (The works by Hadie *et al.*, Ozair *et al.*, Lau *et al.*, and Marzuki *et al.* for other examples of how to report the content validity index.)

Example Results

I-CVI values and CVI scores are computed following expert review. With S-CVI/Ave and S-CVI/UA values of 0.93 and 0.86, respectively, we can conclude that the CV (content Validity) is satisfactory and acceptable.

Analysis of the Findings

For the tool, the majority of items are relevant and reflect the concept well, and the results indicate good content validity. Items with CVI scores of zero or lower have seen some changes to enhance the total validity of the tool. The revision procedure has a specific purpose. It aims to ensure each item is fitting for relevance and matches the domain being measured. The tool's effectiveness is improved by these changes and modifications of items having low scores. The changes are also made to items with lower CVI scores. The main focus is on maintaining the validity and reliability of the tool. Given all this, it is expected that the estimated outcome is a more precise representation and an improved measurement of the content specified. This enhances the efficacy of the tool. The changes make it more useful, efficient, and accurate.

Research Implications

A high content validity tool indicates that the research tool measures the desired construct precisely, resulting in more valid and reliable findings or results. Researchers can be more confident in the precision and relevance of their findings if the tool measures precisely what it is supposed to.

CVI's Limitations

The CVI method depends on subjective expert judgment. The validity of the results may be affected by the variations in the way experts rate relevance. To collect and interpret experts' ratings, it takes time to conduct a CVI evaluation. However, these challenges are frequently justified by the advantages of a true measure of content validity.

Future Directions

Further studies need to focus on identifying different techniques to improve the objectivity and impartiality of the assessment of content validity. Standardised threshold levels for CVI scores can be developed so that they can be applied in different study situations, to ensure the expert ratings are clearer and easier to compare in different study scenarios. Future studies can also examine the impact of different threshold levels on the validity of research tools and the process of assessment. By addressing these domains, researchers can contribute significantly to more precise and accurate measurement of the content validity index, which will further enhance the applicability and appropriateness of research tools.

CONCLUSION

This paper provides a comprehensive and easy-to-follow set of steps on how to calculate the Content Validity Index (CVI) for a single construct using MS Excel, which aids researchers validate questionnaire items methodically and effectively. It offers a systematic approach to evaluating the relevance of items based on expert ratings. The results show that a satisfactory S-CVI/Average index is an assurance of overall content validity of the scale; it is imperative to carefully evaluate individual I-CVI scores to identify and improve items that are not strong enough. Items that score below the standard cutoff should be modified, re-rated by experts, or further validated to improve the research tool. This paper is a significant resource for researchers who want to advance the rigour, validity, and credibility of their research tools in various fields of study.

REFERENCES

Almanasreh, E., Moles, R. and Chen, T.F. 2019. Evaluation of methods used for estimating content validity. *Res. Social Adm. Pharm.*, **15**(2): 214-221.

Davis, L.L. 1992. Instrument review: getting the most from a panel of experts. *Applied Nursing Research*, **5**(4): 194–7.

Lau, A.S., Yusoff, M.S., Lee, Y-Y., Choi, S-B., Xiao, J-Z. and Liong, M-T. 2018. Development and validation of a Chinese translated questionnaire: a single simultaneous tool for assessing gastrointestinal and upper respiratory tract related illnesses in pre-school children. *Journal of Taibah University Medical Sciences*, **13**(2): 135–41.

Ling Zhang, James D. Basham and Richard Allen Carter, 2022. Measuring personalized learning through the Lens of UDL: Development and content validation of a student self-report instrument, *Studies in Educational Evaluation,* **72**: 101121.

Lynn, M.R. 1986. Determination and quantification of content validity. *Nursing Research*, **35**(6): 381–5.

Marzuki, M.F.M., Yaacob, N.A. and 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.

Eriksson, P. *et al*. 2015. *Qualitative Methods in Business research: A Practical Guide to Social Research.*

Patrick, D.L., Burke, L.B., Gwaltney, C.J., Leidy, N.K., Martin, M.L., Molsen, E. and Ring, L. 2011. Content validity--establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2--assessing respondent understanding. *Value Health*, **14**(8): 978-88.

Polit, D.F., Beck, C.T. and Owen, S.V. 2007. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, **30**(4): 459–67.

Taherdoost, Hamed, 2016. Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research (August 10, 2016). Available at SSRN: https://ssrn.com/abstract=3205040 or http://dx.doi.org/10.2139/ssrn.3205040

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

Cook, D.A. and Beckman, T.J. 2006. Current concepts in validity and reliability for psychometric instruments: theory and application. *The American Journal of Medicine*, **119**(2): 166.e7–16.

Hadie, S.N.H., Hassan, A., Ismail, Z.I.M., Asari, M.A., Khan, A.A., Kasim, F. *et al*. 2017. Anatomy education environment measurement inventory: a valid tool to measure the anatomy learning environment. *Anatomical Sciences Education*, **10**(5): 423–32.

Haynes, S.N., Richard, D. and Kubany, E.S. 1995. Content validity in psychological assessment: a functional approach to concepts and methods. *Psychological Assessment*, **7**(3): 238.

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

Polit, D.F. and 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.

Yusoff, M.S.B. 2017. A systematic review on validity evidence of medical student stressor questionnaire. *Education in Medicine Journal*, **9**(1): 1–16.