and universities are also often those with fewer resources supporting accreditation support at the program and unit levels. As such, faculty members at these universities and colleges are more likely to serve on unit and program-level assessment teams and be responsible for testing and interpreting any self-developed instruments. This is a time-consuming process, and faculty members themselves may be uninformed about how to do this.

In the field of Teacher Education, the major accreditation bodies, AAQEP and CAEP, require that validity of instruments be reported to justify data quality (AAQEP, 2021; CAEP, 2022), and, in many cases, the minimum acceptable validity type that must be reported is content validity (Chepko, 2016). For faculty members unfamiliar with statistics or the method itself, testing for content validity can seem daunting. The calculations can be tedious and time-consuming, and there are few resources available to make it easier. Manually entering numbers into the formulas on a calculator, and starting over in the case of small errors, is not a good use of time. If there was a program available for doing this, justifying the funds for purchasing it could be difficult in the current American higher education budgeting struggles.

Although there is an ethical expectation that faculty members do work that assists their programs and universities (Dressel, 1971), work on assessment teams often does not fall neatly into service, teaching, or scholarship (Lewis, 2016) and, thus, may be undervalued for promotion and tenure reviews. Manually computing values for establishing content validity is an undue burden for researchers and evaluators, and particularly for those volunteering their time on assessment and accreditation teams. For researchers and assessment teams who wish to use the Lawshe approach to establish content validity, ws1 340.49 211.97 Tm0 g0 G[(or those)4(volunt)-3(e)-5(e)4(rin)

t

item, the less valid the item may be (Lawshe, 1975). As an aside, I highly recommend including additional space for written comments or subsequent interviews with the Content Evaluation Panel members to gain a deeper understanding of their ratings.

To examine the content validity of each item on an instrument, a formula was proposed by Lawshe (1975) and confirmed by later researchers as a generally accurate method of calculating and interpreting content validity results (Ayre & Scally, 2014). The value is called the Content Validity Ratio (CVR), which is a direct linear transformation of the number of raters agreeing an item is *Essential.* The results of the CVR can help determine which items should be revised or removed from the instrument. The formula is:

$$CVR = \frac{\bar{} \bar{2}}{\bar{2}}$$

Here, n_e is the number of panel members indicating an item is *Essential*, and *N* is the total number of panel members responding to an item (regardless of their rating). This formula may be computed using a calculator, but for instruments with a large number of items, and/or Content Evaluation Panels with a large number of members, it is a time-consuming process with opportunities for calculation errors. In examples provided by La oaG g0mula maIh

Here, n_e is the number of panel members indicating an item is essential, and N is the total number of panel members responding to an item, regardless of their rating.

validity. However, more sophisticated statistical analyses has concluded that the CVR value is a more accurate measurement of content validity.

The CVR values range from 1, which represents a perfect disagreement (or, no panel members mark an item as *Essential*) to +1, which is a perfect agreement (all panel members mark an item as *Essential*). Ayre and Scally (2014) constructed a table of acceptable CVR values, similar to that originally proposed by Lawshe (1975), which may be used as a comparison for acceptable item CVR according to the number of experts on the Content Evaluation Panel (see their publications for the complete tables).

The number of Content Evaluation Panel members determines the level of acceptable CVR values. If using up to seven Content Evaluation Panel members, all members must agree that an item is essential for it to demonstrate acceptable content validity (PAE = 1, CVR = 1.00; Ayres & Scally, 2014). However, if using 20 panel members, only 15 must agree that an item is essential in order to meet the content validity criteria (PAE = .75, CVR = .500; Ayres & Scally, 2014). This is simply explained because it is far more difficult to get larger groups of experts to agree on a single concept than it is a smaller group.

be close to 1.00 to be considered valid (since any individual item with a CVI less than that should have been deleted), while a panel with 20 members would demonstrate a lower CVI, but still be considered valid, due to the lower CVR threshold requirement with larger panels.

The use of the Lawshe approach is not without critics. Beckstead (2009), among others, captured the problematic nature of content validity and the Lawshe approach. Although it is beyond the scope of this paper to discuss the statistical and methodological criticisms, they are

validation using the Lawshe app the statistical model underlying interrater agreement, the collapsing of response categories, the correction for chance agreement among raters, and the age-old philosophical problem of

(Beckstead, 2009, p. 1277). If you choose to pursue the Lawshe approach as a method for establishing content validity, the calculations can be done quite easily using Microsoft Excel. Next, I will explain how to set up an Excel worksheet with headings and formulas to compute PAE, CVR, and CVI. This has not been tested with other spreadsheet software programs.

Setting Up the Workbook

I recommend setting up each worksheet in Excel using headings. Each of the formulas presented here works under the assumption of including a first row of headings

should label them with something that allows you to quickly compare the CVR result to the actual item on the instrument, so ensure it has a meaningful label. Second, if you assign them a number or arbitrary code, make sure you create a codebook (either digitally or by hand) that allows you to remember which item is represented by the number or code. For practical reasons, it is almost certainly easiest if you enter the items in the order they are presented on the instrument itself.

Column B: Essential. Under this heading in Column B, you will be typing the total number of *Essential* ratings given by a Content Evaluation Panel for each item. On the first item (Row 2), for example, you will enter 5 if the total number of experts who indicated that item to be *Essential* totaled five. With this column and the subsequent two (Columns C and D), you have two practical options. If you have a small number of raters, you can manually count the number of *Essential* ratings for each item and input that number in this column. This would be relatively easy to do with five or fewer raters. With more than five raters, I recommend the options outlined in Appendix A, which describe the steps to set up a second worksheet in your Excel file, input each item and rater manually, and create sums of the three rating options with a column for rater comments, if they are present. Simply, this option will likely save you time and energy of manually calculating scores, as well as eliminating errors that come from manually counting (or

the total ratings for each item. Additionally, it allows you to organize and store each panel member ratings and comments within your workbook. This can easily be completed by a research assistant, if available, to expedite the analysis.

Column C: Useful. Under this heading in Column C, you will be inputting the number of total *Useful* ratings for each item. The same directions from the *Essential* category apply to this

(left) and the blue box (bottom right) that you should see once you have entered the CVR formula.

Writing PAE Formula

The PAE, or Proportion Agreeing Essential, is the formula for calculating the percentage of respondents rating an item *Essential* out of the total number of raters for the item. This is a relatively easy calculation to do manually since it represents the percentage of raters out of the total who indicated an item was *Essential*. For example, if eight out of ten experts rated an item *Essential*, you would calculate it as: 8/10, or .80 (80%). However, Excel will calculate this automatically, if you enter this formula in Column F, Row 2 (cell F2; under header PAE):

After entering this formula, hit Enter. Similar

Similar to the CVI directions, you may need to adjust the F40 number, depending on how many items you have included on your instrument.

Example

Now that I have explained the headers and formulas that can be entered to calculate CVR, let me give you a short example. In this abbreviated example, five raters responded to a Lawshe rating of three items on an instrument. The raters responded as such:

T

	Item 1	Item 2	Item 3
Rater 1	Essential	Essential	Not Necessary
Rater 2	Essential	Essential	Useful, Not Essential
Rater 3	Essential	Useful, Not Essential	Useful, Not Essential
Rater 4	Essential	Essential	Not Necessary
Rater 5	Essential	Essential	Useful, Not Essential
Total	5 Essential	4 Essential, 1 Useful	3 Useful, 2 Not Necessary

I only want to calculate the three PAE items in cells F2, F3, and F4, so it is modified to calculate those three ( ).

Once the data are entered, the Excel formulas give us the outputs for each item (Item 1 CVR = 1, PAE = 1; Item 2 CVR = .6, PAE = .8, and Item 3 CVR = -1, PAE = 0). In this example, Item 1 has demonstrated content validity using the Lawshe approach. I would likely delete Items 2 and 3, since they did not meet the CVR thresholds for ratings with three experts on the Content Evaluation Panel (Ayre & Scally, 2014). These would require 100% agreement of *Essential*, and they did not meet this. This should be done prior to reporting the final CVI, but for the purpose of providing an example of the formula calculations using the Excel formulas I developed, I left them in. The instrument CVI is .2 and the mean PAE is .6, which would not be close to demonstrating content validity (recall that it is typically acceptable for an instrument to have a CVI between .70 and .80).

It is important to note that most researchers, such as Gilbert and Prion (2016), recommend that all item-level CVR values of 1 should be converted to .99 prior to calculation. To do this, you can manually change the number in each of the row where CVR computes to 1 (such as it did in Item 1 of this example), by typing the number .99 in its place, and the CVI formula will recalculate accordingly (even though the results will be negligible). I also recommend you document this change (1 to .99) when you report the results. For sake, I did not do this in this example, but I would modify those values and report the subsequent CVI in reports or studies.

Reporting Results

The amount of information you include about your instrument depends on the purpose of your documentation. If you are reporting the technical adequacy or development of an

instrument, you may want to include development about the items and constructs and information about the item-level CVR values. Typically, in articles where instrument validity is important to report, but is not the focus of the article, we often see a justification of the Lawshe method, a brief description of the expert raters, their process of rating and returning the scores, and the instrument CVI.

When it is required to report the item-level CVRs (such as white papers, development papers, or technical reports), I place the full item-level results in a table. While this may be a journal preference, typically all PAE, CVR, and CVI values are rounded to the hundredth place, and CVR and CVI values that compute to 1 are often converted to .99. I table results, such as Table 2, including the raw number of expert ratings of each item, the PAE, the CVR, and a source to which I compare the CVR threshold as having met or not met the contentment validity guidelines.

Table 2. *Results of Content Evaluation Panel Rating of Reading Comprehension Assessment*

Item	N_e	N_u	N_n	PAE[a]	CVR[b]	Ayre & Scally
Comprehension 1	10	-	-	1.00	.99	met
Comprehension 2	9	1	-	1.00	.80	met

are on assessment or accreditation teams. Even researchers, with greater resources, can benefit from using the formulas presented here in their calculations. They can be easily entered and reported by research assistants and minimizes nearly any calculation errors that might otherwise occur when done manually.

There are no readily available online calculators, nor are there easy directions for using more complex data analysis software to perform the formula calculations required in the Lawshe approach. Microsoft Excel can be easily transformed into a vital tool for instrument validation. Using the directions provided in this paper, anyone can easily calculate and report essential content validity results. With very little advanced programming, and very little knowledge of Excel formula writing, departments or research teams could use these directions to set up templates available for calculating the essential formulas in the Lawshe approach.

The purpose of this paper was not to dive deeply into the theoretical underpinnings or criticisms of the Lawshe approach, nor was it to compare the Lawshe approach with other forms of content validity testing. Before engaging in this work, I highly recommend exploring the essential readings that were cited throughout this paper to further investigate if the Lawshe approach is appropriate for your needs. To obtain a Microsoft Excel template with a sheet for entering and calculating individual instrument numbers (individual item data entry), and a sheet including all formulas as described in this paper, please contact me at the email provided.

content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development, 47*(1), 79-86.

Beckstead, J. W. (2009). Content validity is naught. *International Journal of Nursing Studies*, *46*(9), 1274-1283.

Boyer, E. L. (1990). Scholarship reconsidered: Priorities of the professoriate. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

CAEP (2022). *CAEP* revised 2022 standards workbook. Retrieved from: https://caepnet.org/~/media/Files/caep/accreditation-resources/caep-2022-standards-workbook-final.pdf?la=en

Chepko, S. (2016). Establishing content validity. Slides presented at the September 2016 CAEP Convention. Retrieved from: https://www.education.ne.gov/educatorprep/Meetings/CAEP-Sep2016/NE-ContentValidityReliability.pdf

Davis, L. L. (1992). Instrument review: Getting the most from a panel of expvre5s

Lewis, S. (2016). *Perceptions of University Faculty Regarding Accreditation in a College of Education* (Doctoral dissertation, University of South Florida).

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing research*, *35*(6), 382-386.

Park, T. (2012). Do faculty members get what they deserve?: A review of literature surrounding the determinants of salary, promotion, and tenure. Journal of the Professoriate, 6(1), 28-47.

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 & health*, *30*(4), 459-467.

Tilden, V. P., Nelson, C. A., & May, B. A. (1990). Use of qualitative methods to enhance content validity. *Nursing Research*, *39*(3), 172-175.

Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for *Measurement and Evaluation in Counseling and Development*, *45*(3), 197-210.

Figure 1. Headings.

Figure 2. Formula.

Figure 3. Example Data Inputs