Link to publisher's version: *http://dx.doi.org/10.1177/0748175613513808*

Citation: Ayre C and Scally A J (2014) Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. Measurement and Evaluation in Counseling and Development. 47(1): 79-86.

Title: Critical values for Lawshe's content validity ratio. Revisiting the original methods of calculation.

Abstract:

The content validity ratio originally proposed by Lawshe is widely used to quantify content validity and yet methods used to calculate the original critical values were never reported. Methods for original calculation of critical values are suggested along with tables of exact binomial probabilities.

Introduction:

Content validation refers to a process which aims to provide assurance that an instrument (checklist, questionnaire or scale etc.) measures the content area it is expected to measure (Frank-Stromberg and Olsen, 2004). One way of achieving content validity involves a panel of subject matter experts considering the importance of individual items within an instrument. Lawshe's method, initially proposed in a seminal paper in 1975 (Lawshe, 1975), has been widely used to establish and quantify content validity in diverse fields including health care, education, organizational development, personnel psychology and market research (Wilson et al, 2012). It involves a panel of subject matter 'experts' rating items into one of three categories: 'essential', 'useful, but not essential' or 'not necessary'. Items deemed 'essential' by a critical number of panel members are then included within the final instrument, with items failing to achieve this critical level discarded. Lawshe (1975) suggested that based on 'established psychophysical principles' a level of 50% agreement gives some assurance of content validity.

The CVR (content validity ratio) proposed by Lawshe (1975) is a linear transformation of a proportional level of agreement on how many 'experts' within a panel rate an item 'essential' calculated in the following way:

$$CVR = \frac{n_e - (N/2)}{N/2}$$

CVR =content validity ratio; n_e = Number of panel members indicating an item 'essential'; N = Number of panel members

Lawshe suggested the transformation (from proportion to CVR) was of worth as it could readily be seen whether the level of agreement amongst panel members was greater than 50%. CVR values range between -1 (perfect disagreement) and +1 (perfect agreement) with CVR values above zero indicating that over half of panel members agree an item essential. However, when interpreting a CVR for any given item it may be important to consider whether the level of agreement is also above that which may have occurred by chance. As a result Lawshe reported a table of critical CVR ($CVR_{critical}$) values computed by his colleague Lowell Schipper, where $CVR_{critical}$ is the lowest level of CVR such that the level of agreement exceeds that of chance for a given item, for a given alpha (Type I error probability, suggested to be 0.05 using a one-tailed test). $CVR_{critical}$ values can be used to determine how many panel members need to agree an item essential and thus which items should be included or discarded from the final instrument. In order to include or discard items from a given instrument appropriately it is imperative that the $CVR_{critical}$ values are accurate. Recently, concern has been raised that the original methods used for calculating $CVR_{critical}$ were not reported in Lawshe's paper on content validity and, as both Lawshe and Schipper have since passed away, it is now not possible to gain clarification (Wilson et al, 2012). Furthermore, an apparent anomaly exists in the table of critical values between panel sizes of 8 and 9, where $CVR_{critical}$ unexpectedly rises to 0.78 from 0.75 before monotonically decreasing with increasing panel size up to the calculated maximum panel size of 40. This led Wilson et al (2012) to try and identify the method used by Schipper to calculate the original $CVR_{critical}$ values in Lawshe (1975) in the hope of providing corrected values.

Despite their attempts, Wilson et al (2012) fell short of their aims. They suggested that Schipper had used the normal approximation to the binomial distribution for panel sizes of 10 or more yet these claims were theoretical as they were unable to reproduce the values of $CVR_{critical}$ reported in Lawshe (1975). As values for $CVR_{critical}$ calculated by Wilson and colleagues differed significantly from those reported in Lawshe (1975) it was suggested that instead of the one-tailed test reported Schipper had, in fact, used a two-tailed test as this more closely resembled their results. In addition, for panel sizes below 10 no satisfactory explanation was provided of how $CVR_{critical}$ may have originally been calculated. Furthermore, they were unable to provide a satisfactory explanation for the apparent anomaly between panel sizes of 8 and 9.

In their paper, Wilson et al (2012) produced a new table of $CVR_{critical}$ values using the normal approximation to the binomial distribution. This method we believe to be inferior to calculation of exact binomial probabilities as, by definition, it is ultimately just an approximation and may not be valid for small sample sizes and for proportions approaching 0 or 1 (Armitage et al, 2002). It is understandable that a normal approximation was used for larger panel sizes when Schipper calculated the original $CVR_{critical}$ values in 1975, but as statistical programmes can now readily calculate exact binomial probabilities it would seem more appropriate to do so in the present day. We had further concerns regarding the methods used by Wilson et al (2012) to calculate the normal approximation, as it appeared a continuity correction had not been employed. In cases where the continuous normal distribution has been used to approximate the discrete binomial distribution more accurate results are obtained through use of a continuity correction (Rumsey, 2006; Gallin and Ognibene, 2007).

Based on the wide discrepancy between $CVR_{critical}$ reported by Wilson et al (2012) and Lawshe (1975) we intended to answer the following questions:

1) Did Wilson et al (2012) correctly employ a method for calculating binomial probabilities?
2) What method was employed by Schipper to calculate $CVR_{critical}$ in Lawshe (1975) for all panel sizes?
3) Are there anomalies in Schipper's table of critical values in Lawshe (1975)?
4) Did Lawshe report $CVR_{critical}$ for a one-tailed test or a two-tailed test?

As a result of our belief that exact binomial probabilities were more appropriate than normal approximations we also intended to calculate exact binomial probabilities for all panel sizes between 5 and 40.

Methods:

We calculated the minimum number of experts required to agree an item 'essential' for a given panel size, such that the level of agreement exceed that of chance. In keeping with previous work we assumed the outcome as dichotomous (i.e. ‘essential’ or ‘not essential’) although we acknowledge it could be considered trichotomous as there are 3 possible outcomes when rating any given item (‘essential’, ’important, but not essential’ and ‘not necessary’). As the CVR is designed to show a level of agreement above that of chance we are only concerned with testing in one direction. Thus, in this case a one-tailed hypothesis test is appropriate.

Hypothesis:

$H_{0:}$ $n_e = N/2$

Significance (α) was set at 0.05.

Using a one-tailed test we would reject H_0 if $P(n_e \geq n_{critical}) \leq 0.05$; where $n_{critical}$ = the lowest number of experts required to agree an item ‘essential’ for agreement to be above that of chance, n_e = the number of experts rating an item as ‘essential’.

We calculated exact $CVR_{critical}$ values for panel sizes between 5 and 40, based on the discrete binomial distribution, computed using Stata Statistical Software: Release 12 (StataCorp (2011), College Station, TX: StataCorp LP). The following command was used:

bitesti N n_e p

Where N= total number of panel members, n_e = number of experts agreeing ‘essential’, p = the hypothesised probability of success (agreeing the item as essential) = ½

Using this method we produced a table of the minimum number of experts (n_e) required to agree an item essential such that we could reject H_0 (i.e. the minimum number of experts such that p≤0 .05). Values for $CVR_{critical}$ were then calculated on the basis of the minimum number of experts required using the formula for calculating CVR given previously in the paper. Exact one-sided p-values are reported.

In order to allow direct comparison, we calculated the exact binomial probabilities according to the method used by Wilson et al (2012), described in their paper, using the Microsoft Excel function:

$n_{critical}$ = CRITBINOM (n,p,1-α)

Where $n_{critical}$ is the minimum number of experts required to agree an item essential, n is the panel size, p is the probability of success = ½ and α = 0 .05

Normal approximation to the binomial was calculated using the following formula incorporating a continuity correction (Armitage et al, 2002). This subtracts 0.5 from the number of panel experts required to agree an item essential to account for using the continuous normal distribution for approximation of the discrete binomial distribution.

$$z = \frac{(n_e - Np - 0.5)}{\surd[Np(1-p)]} \sim N(0,1)$$

Therefore as p= ½ :

$$n_e = z\left(\frac{\sqrt{N}}{2}\right) + \left(\frac{N}{2}\right) + 0.5$$

Where z= normal approximation of the binomial, N= total number of panel members, n_e = number of experts agreeing 'essential', p= probability of agreeing each item essential = ½, 0.5 is the continuity correction

CVR based on the normal approximation was calculated in the following way:

$$CVR = \frac{\left[z\left(\frac{\sqrt{N}}{2}\right)\right]+0.5}{\left(\frac{N}{2}\right)}$$

Therefore:

$$CVR = \frac{\left[\left(z\sqrt{N}\right) + 1\right]}{N}$$

Normal approximations for $CVR_{critical}$ were calculated using this method for all panel sizes to allow comparison with previous work.

Results:

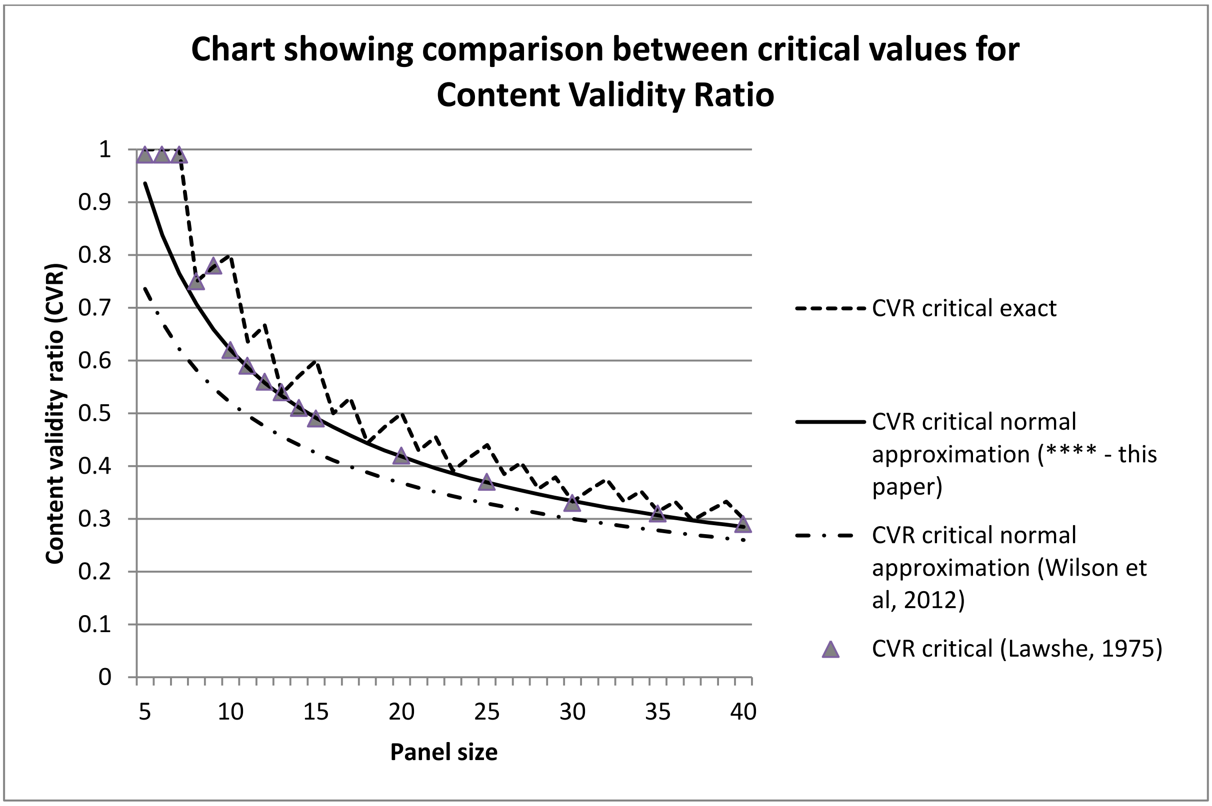
The calculations for $CVR_{critical}$ based on exact binomial probabilities for panel sizes of 5 to 40 are shown in table 1. Calculations using the CRITBINOM function returned values for the critical number of experts 1 fewer for all panel sizes compared with our calculations (table 1).

Table 1: Table showing $CVR_{critical}$ one-tailed test (α= 0.05) based on exact binomial probabilities

N (panel size)	Proportion agreeing essential	$CVR_{critical}$ exact values	One-sided p-value	$N_{critical}$ (minimum number of experts required to agree item essential)- ****and ****, this paper	$N_{critical}$ calculated from CRITBINOM function- Wilson et al (2012)
5	1	1.00	.031	5	4
6	1	1.00	.016	6	5
7	1	1.00	.008	7	6
8	.875	.750	.035	7	6
9	.889	.778	.020	8	7
10	.900	.800	.011	9	8
11	.818	.636	.033	9	8
12	.833	.667	.019	10	9
13	.769	.538	.046	10	9
14	.786	.571	.029	11	10
15	.800	.600	.018	12	11
16	.750	.500	.038	12	11
17	.765	.529	.025	13	12
18	.722	.444	.048	13	12
19	.737	.474	.032	14	13
20	.750	.500	.021	15	14
21	.714	.429	.039	15	14
22	.727	.455	.026	16	15
23	.696	.391	.047	16	15
24	.708	.417	.032	17	16
25	.720	.440	.022	18	17
26	.692	.385	.038	18	17
27	.704	.407	.026	19	18
28	.679	.357	.044	19	18
29	.690	.379	.031	20	19
30	.667	.333	.049	20	19
31	.677	.355	.035	21	20
32	.688	.375	.025	22	21
33	.667	.333	.040	22	21
34	.676	.353	.029	23	22
35	.657	.314	.045	23	22
36	.667	.333	.033	24	23
37	.649	.297	.049	24	23
38	.658	.316	.036	25	24
39	.667	.333	.027	26	25
40	.650	.300	.040	26	25

Figure 1 shows a comparison of $CVR_{critical}$ values from our exact binomial and normal approximation to the binomial calculations and those reported by Lawshe (1975) and Wilson et al (2012). Normal approximation using the continuity correction returned values equal to those reported in Lawshe (1975) for all given panel sizes of 10 and above other than a minor difference of 0.01 for a panel size of 13.

Figure 1:



Discussion:

We have produced a table of exact values for $CVR_{critical}$ including the minimum number of panel members required such that agreement is above that of chance. We believe we are the first to produce a table of values for $CVR_{critical}$ from exact binomial probabilities. In contrast to previous work, all of the values for $CVR_{critical}$ are calculated based on an *achievable* CVR, given the discrete nature of the variables under investigation.

Comparison with previous work:

Comparison to Lawshe (1975):

The exact critical values for CVR we have produced are equal to those given in Lawshe (1975) for panel sizes below 10, allowing for adjustments and rounding (see figure 1). We therefore believe that Lawshe (1975) calculated exact binomial probabilities for panel sizes below 10. This approach is reasonable as the use of a normal approximation for a binomial distribution is only justifiable when:

$Np > 5$ and $N(1-p)>5$ (Rumsey, 2006). Where N = the number of panel members and p = the probability of success in any trial

This would be satisfied for panel sizes above 10 assuming p= ½.

We do not believe that there is an anomaly for panel sizes between 8 and 9 in $CVR_{critical}$ reported in Lawshe (1975). It can be seen in figure 1 that $CVR_{critical}$ does increase between panel size of 8 and 9, related to the discrete nature of both the panel size and number of experts who can agree any item is essential. It can be seen from our calculations that, although the overall pattern is for $CVR_{critical}$ to fall with increasing panel size, there are a number of instances where $CVR_{critical}$ increases. This is an important consideration when determining panel size for those using the CVR method to gain content validity.

For panel sizes of 10 and above the normal approximation to the binomial has been calculated and we have been able to reproduce the same values reported in Lawshe (1975) notwithstanding a minor discrepancy for a panel size of 13 (see figure 1). As the normal distribution is based on a continuous distribution and it is being used to approximate a discrete distribution, Schipper has correctly used a continuity correction which will more likely result in more accurate approximations. It would appear that Schipper and Wilson et al used identical methods for calculating $CVR_{critical}$ with the exception of the continuity correction.

As the values we have calculated are the same as those of Lawshe (1975) it is apparent they have also used a one-tailed test at p=0.05 as they originally reported, and, not a two-tailed test as suggested by Wilson et al (2012).

Can the critical CVR values given by Lawshe (1975) be used to accurately determine panel size?

In general, use of the originally calculated CVR values from Lawshe (1975) yields an equal value for the critical number of experts required as shown in our exact calculations. The only discrepancy occurs for a panel size of 13 where the exact $CVR_{critical}$ is marginally under that reported by Lawshe. Importantly, our findings would suggest that questionnaires and checklists developed using the $CVR_{critical}$ values originally reported by Lawshe (1975) remain valid.

<u>Comparison to Wilson et al (2012):</u>

$CVR_{critical}$ based on exact binomial probabilities:

The exact $CVR_{critical}$ based on binomial probabilities we have calculated using Stata differ from those given by the CRITBINOM function in Microsoft Excel employed by Wilson et al (2012) as a result of the discrepancy in the critical number of experts required to agree an item 'essential' produced by each method (see table 1). We believe that Wilson et al (2012) have incorrectly interpreted the result returned from the CRITBINOM function and therefore the $CVR_{critical}$ based on the exact binomial probabilities shown in figure 1 of their paper are incorrect. The method used by Wilson et al (2012) returns one fewer than the true critical number of experts required to ensure agreement above that of chance for a given value of α (table 1) yet no mention of this can be seen in their paper. This can be illustrated through an example using a panel size of 15.

Example: Considering a panel size (n) of 15, probability of success (p) of 0.5 and α= 0.05

CRITBINOM (15, 0.5, 0.95) = 11

As this utilises the cumulative binomial probability the interpretation of this result is ''there is at least a probability of 0.95 of getting <u>*11 or fewer*</u> successes''. Thus, there is at most a probability of

0.05 of getting *12 or more* successes. This is the critical number we are interested in to assure a level of agreement above that of chance at α set at 0.05.

The error in calculating exact binomial probabilities may explain why Wilson et al (2012) failed to realise that Shipper had calculated exact binomial probabilities up to a panel size of 10.

$CVR_{critical}$ based on the normal approximation to the binomial:

$CVR_{critical}$ values reported by Wilson et al (2012) based on a normal approximation to the binomial distribution are markedly lower than those we have calculated using a continuity correction for all given panel sizes (see figure 1). It is clear from their formula for calculating the critical value for CVR that a continuity correction was not used by Wilson et al. Conversely, the values given in Lawshe are consistent with the normal approximation using the continuity correction and are therefore closer to the exact binomial probabilities we have reported in this paper. On this basis, we believe that the recalculated values for $CVR_{critical}$ reported in Wilson et al (2012) are inaccurate and therefore should not be used.

Wilson et al (2012) and Lawshe (1975) have both calculated $CVR_{critical}$ values for panel sizes of 10 or more (Wilson et al, 2012 used the normal approximation for all panel sizes) based on a normal approximation of the binomial distribution. We believe this is an inferior method to the exact calculations we have reported for the following reasons:

1) If the normal approximation value for $CVR_{critical}$ is higher than that produced from exact calculations of binomial probability the panel size deemed necessary will be higher than required.
2) If the normal approximation value for $CVR_{critical}$ is lower than that produced from exact calculations of binomial probability the panel size deemed necessary may be lower than required

Presented below is a simplified table of $CVR_{critical}$ values, calculated using exact binomial probabilities, which includes the number of experts required to agree any given item is essential (table 3).

Table 3: Simplified table of $CVR_{critical}$ including the number of experts required to agree an item essential

Panel size	$N_{critical}$ (minimum number of experts required to agree an item essential for inclusion)	Proportion agreeing essential	$CVR_{critical}$
5	5	1	1.00
6	6	1	1.00
7	7	1	1.00
8	7	.875	.750
9	8	.889	.778
10	9	.900	.800
11	9	.818	.636

12	10	.833	.667
13	10	.769	.538
14	11	.786	.571
15	12	.800	.600
16	12	.750	.500
17	13	.765	.529
18	13	.722	.444
19	14	.737	.474
20	15	.750	.500
21	15	.714	.429
22	16	.727	.455
23	16	.696	.391
24	17	.708	.417
25	18	.720	.440
26	18	.692	.385
27	19	.704	.407
28	19	.679	.357
29	20	.690	.379
30	20	.667	.333
31	21	.677	.355
32	22	.688	.375
33	22	.667	.333
34	23	.676	.353
35	23	.657	.314
36	24	.667	.333
37	24	.649	.297
38	25	.658	.316
39	26	.667	.333
40	26	.650	.300

It can be seen from table 3 that preferred panel sizes exist, when the addition of a further panel member leads to a significant reduction in the required proportion level of agreement that an item is ‘essential’ for it to be included (e.g. between panel sizes of 12 and 13). In addition, it is also immediately apparent that increasing the panel size by 1 will actually increase the required proportion level of agreement on occasions (e.g. between panel sizes of 13 and 14). We believe this table is of most use to researchers’ wishing to quantify content validity using the CVR method, both to decide the most appropriate panel size and when determining whether a critical level of agreement has been reached.

<u>Conclusions:</u>

The method used by Schipper to calculate the original critical values reported in Lawshe’s paper has been suggested and we have been able to successfully reproduce the values using discrete calculation for panel sizes below 10 and normal approximation to the binomial for panel sizes of 10 and above. We have identified problems with both the discrete calculations and normal approximation to the binomial suggested by Wilson et al. Consequently, we do not believe that values for $CVR_{critical}$ reported in Wilson et al should be used to determine whether a critical level of agreement has been reached and therefore whether items should be included or excluded from a given instrument. Although, it is safe to use the values for $CVR_{critical}$ proposed by Lawshe to determine whether items should be included on an instrument we believe that exact $CVR_{critical}$ based on discrete binomial calculations is most appropriate.

References:

Armitage, P., Berry, G. and Matthews, J.N.S. (2002). Statistical Methods in Medical Research, Fourth edition. Blackwell: Oxford.

Frank-Stromberg, M., Stromberg, S.J. (2004). Instruments for clinical health-care research. Jones and Bartlett. London

Gallin, J.I., Ognibene F.P. (2007). Principles and practice of clinical research. Second edition. Elsevier. Massachusetts.

Lawshe, C.H. (1975). A quantitative approach to content validity. Personnel psychology, 28, 563-575.

Rumsey, D. (2006). Probability for dummies. Wiley. Indianapolis

Wilson, F. R., Pan, W.& Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. Measurement and evaluation in counselling and development. May. doi: 10.1177/0748175612440286