Computing Alpha and Omega Reliability Estimates

Larry R Nelson[1]
Curtin University, Western Australia

A working paper dated 8 August 2024[2]. Click here for a possible update.

This study has to do with calculating and comparing estimates of an instrument's reliability, looking at coefficients alpha and omega.

Of the these, alpha has been by far the most popular, a veteran with *many* decades of application behind it. McDonald (1999) reviewed the long-recognized shortcomings of alpha, suggesting "omega" as a fix for alpha's main limitation: the "tau-equivalent" assumption, seldom met in practice. Since then an increasing number of apps have emerged which support the calculation of omega.

The present paper builds on alpha/omega comparisons seen in Nelson (2016). The majority of the datasets included in Table 1 below have been carried over from that work, with the addition of six new ones[3], and two additional statistics: standardized alpha in column 7, and omega hierarchical in column 11.

The new datasets serve in part to address an imbalance in the original paper – now there are more affective scales to counter the former predominance of cognitive instruments.

Nelson used two R packages, MBESS and psych, and the CALIS procedure in SAS University to calculate omega total values. These tools, powerful as they are, lean on software skills which might be seen by some as being in the somewhat advanced realm, tending to involve installation steps often substantially more demanding than those found in many of the (usually) straightforward installers for other apps.

The present paper has included two free alternatives to MBESS: JASP and *Jamovi*[4]. SAS is carried forward here as it has been used in the past to calculate omega.

SPSS also receives mention in an appendix – it came to include support for computing omega in a system update released in the latter half of the year 2020. An omega capability was added to Lertap5 early in 2021 and is discussed in another appendix.

---

[1] l.nelson@curtin.edu.au
[2] Original paper released in February 2021.
[3] Numbers 10, 11, 12, 14, 15, and 16 in Table 1.
[4] *jamovi* is covered in more detail in a related paper.

Table 1

1	2	3	4	5	6	7	8	9	10	11
No.	Source	Type	Items	N	Alpha	Std. Alpha	Omega CALIS	Omega JASP	Omega total	Omega hier.
1	MathsQuiz	Test	15	999	0.80	0.78	0.80	0.81	0.80	0.69
2	UniAA	Test	30	127	0.74	0.74	0.75	0.74	0.76	0.41
3	UniBB	Test	34	132	0.82	0.81	0.82	0.82	0.83	0.49
4	Zmed	Test	100	2,470	0.96	0.96	0.95	0.96	0.96	0.76
5	HalfTime	Test	100	424	0.94	0.93	0.92	0.93	0.94	0.70
6	N Rivers	Test	50	689	0.92	0.92	0.92	0.92	0.92	0.74
7	LenguaBlg	Test	50	5,504	0.81	0.81	0.82	0.82	0.83	0.72
8	LenguaBlg	Test	47	5,504	0.83	0.83	0.84	0.83	0.84	0.72
9	Negocios	Test	60	500	0.86	0.86	0.85	0.86	0.87	0.57
10	FIMS	Test	14	6,371	0.69	0.70	0.70	0.70	0.73	0.53
11	FIMS AUS	Test	14	4,320	0.60	0.60	0.61	0.62	0.63	0.46
12	FIMS JPN	Test	14	2,051	0.73	0.73	0.74	0.74	0.75	0.45
13	DunnSES	Scale	7	201	0.94	0.94	0.94	0.94	0.96	0.92
14	BFIneuro	Scale	5	2,663	0.81	0.81	0.80	0.81	0.85	0.73
15	Blirt8	Scale	8	211	0.78	0.78	0.79	0.79	0.84	0.54
16	Hammer	Scale	29	758	0.93	0.93	0.93	0.94	0.94	0.75

Table 1 readily provides the setting for a variety of questions.

Q1: what is standardized alpha (column 7)?

Alpha and standardized alpha are reliability estimates with corresponding models.

Alpha is often frequently referred to as "Cronbach's alpha". It is based on the "tau-equivalent" model and, as such, has a basic underlying assumption of equal inter-item covariances.

Standardized alpha is not used nearly as often but is certainly supported in numerous apps/programs (three of them being the psych package in R, the reliability module in JASP, and the Scale option in **SPSS**). Its underlying model adds another assumption: items are "parallel", they meet the tau-equivalent assumption, and, in addition, it is assumed that in the population all items have the same variance.

According to Falk and Savailie (2010, p.445), "The decision of which version of alpha to use depends on whether researchers decide to standardize a test's items before adding the items to form a composite score." They elaborate on the importance of recognizing that alpha and standardized alpha are based on different models.

Cho (2018) suggests that standardized alpha is a misnomer as it has nothing to do with alpha, saying that "… there is no reason to include the word alpha in the name of the coefficient". He found that the use of this statistic increased after SPSS included it many years ago[5]. Cho suggested that uninformed users may think standardized alpha to be superior to alpha just because "standardized" sounds better. To clarify matters, he

[5] Original reference is the SPSS User's Guide (Sprecht, 1975).

suggested that alpha be renamed to "tau-equivalent reliability". In an earlier paper, Cho (2016) suggested renaming "standardized alpha" to "parallel reliability".

Looking at Table 1, a comparison of Columns 6 and 7 indicates only a few very minor differences between the two alphas for the instruments included in this study. That is to say, in this study little difference was found in the estimates of tau-equivalent reliability and parallel reliability (columns 6 and 7 respectively)[6].

{Bandalos 2018 has a summary of model assumptions in her "Reliability and Validity" chapter (Chapter 10); Meyer 2010 has a similar discussion in his Chapter 3, "Assumptions".}

Q2: what are the omega total values in Table 1's columns 8, 9, and 10?

McDonald (1999) is generally cited as the originator of the omega statistic. Omega is a reliability coefficient which, unlike alpha, does not depend on items being tau-equivalent, a dependence difficult to meet in practice, resulting in alpha generally being an underestimate of reliability. When items are not tau-equivalent (the usual case), omega will be a more accurate estimate of reliability, expected to generally exceed alpha in value.

Here, in this paper, the term "omega total" is used in order to distinguish it from "omega hierarchical" (discussed below). In the current literature this distinction is often not made, with omega total often being simply referred to as just "omega" or as "McDonald's omega".

Column 8 in Table 1 has omega total values computed using the CALIS module in SAS (as in Nelson (2016) and Geldhof et al. (2013)). CALIS works from a general structural equation model (**SEM**).

Column 9 omega total values were obtained using JASP version 0.14.1. JASP uses a confirmatory factor analysis model (**CFA**). (*jamovi* could also be used.)

Column 10 draws on the use of the CRAN psych package, using exploratory factor analysis (**EFA**) and a Schmid-Leiman transformation (Schmidt & Leiman, 1957) to obtain omega statistics. {The S-L transformation is also covered in Thompson 2004.}

Q3: why are the three omega total values different?

This question arises when comparing Columns 8, 9, and 10 in Table 1.

The values in Column 8 (CALIS) and 9 (JASP CFA) should be similar. Their underlying models use somewhat different terminology but are essentially identical. Differences in Columns 8 and 9 omega total values should therefore be slight. And they are.

Column 10's values are based on EFA and a bifactor model. In this case, omega total involves loadings on a general factor plus cross loadings on secondary factors. Differences are now noted in a few of the Table 1 Rows – compare Columns 8, 9, and 10 for each of the 16 datasets. (McNeish (2017) is a worthy reference here.)

[6] JASP 0.14.1 was used to get the values seen in columns 6 and 7.

In theory, alpha will never be greater than omega. Does this hold up in this study, are there any alpha values greater than omega?

We would do well to ask which alpha and which omega.

There is an alpha value in Column 6 that is slightly greater than the corresponding omega value computed by JASP, as seen in Column 9. It is in Row 5, where alpha is 0.94 and JASP omega 0.93; when carried to three decimal places, the values are 0.935 and 0.934 .

This trivial difference would disappear if we were instead to use Row 5's standardized alpha, seen in Column 7.

Now, omega total values from the psych program, Column 10, <u>are</u> all above corresponding alpha <u>and</u> standardized alpha values, although, in the case of Row 5, three decimal places are needed to show it: the Row 5, Column 10 omega total was 0.939. The psych program uses EFA and a bifactor model, as mentioned above.

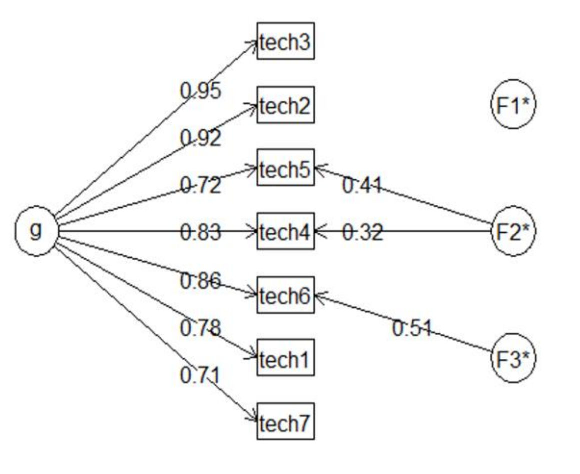
Q4: what is omega hierarchical in Table 1 Column 11?

The figure in Column 11, "Omega hier.", is derived from the same Schmid-Leiman bifactor solution used to find the omega total seen in Column 10, but now <u>only</u> loadings on the general factor are employed in the calculation.

Two examples will help indicate the meaning of omega hierarchical. Figures 1 and 2 were produced by the psych program (omega (h) below is omega hierarchical):

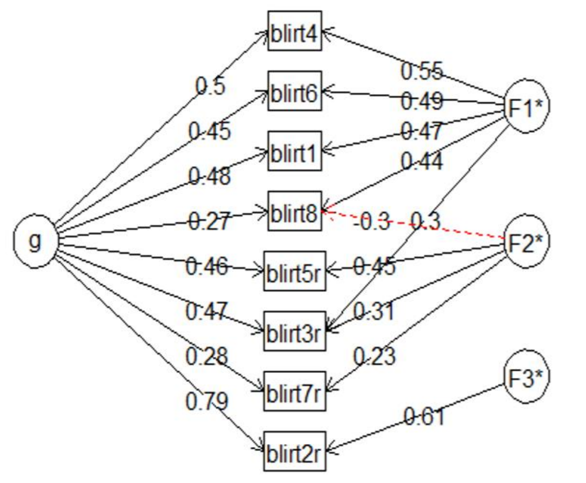
DunnSES: omega (h) = 0.92

Figure 1



Blirt8: omega (h) = 0.54

Figure 2



DunnSES results are seen on the left in *Figure 1*. The general factor, "g", has a strong influence on the measure's seven items, with relatively high loadings: from 0.71 to 0.95. Such is not the case with Blirt8, *Figure 2*, where the influence from "g", the general factor, is less, having loadings ranging from 0.27 to 0.79. The three second-order Blirt8 factors, F1, F2, and F3 have had much more impact than the second-order factors in DunnSES.

Revelle, the author of the psych package, shies away from using the term “unidimensional”, as does Cho (2016), but other authors (McDonald 1999, Thompson 2004, Meyer 2010, Bandalos 2018) might look at *Figure 1* and possibly interpret it as at least suggesting unidimensionality – there’s a strong dominant general factor with little influence from the second-order factors. Some would say that the DunnSES result displayed in *Figure 1* is evidence of a “purer” measure when compared to the Blirt8 results seen in *Figure 2* – by and large, responses to the DunnSES items are attributable, or “linked”, to a single general factor. Revelle and Zinbarg (2009) suggest that omega(h) is “… an index of how much the test measures one common factor”.

Flora (2020) has a discussion of CFA/EFA, the bifactor model, and omega hierarchical, including mention of unidimensionality. His article has examples and tutorials for using R packages to estimate omega. Goodboy & Martin (2020) also have extensive coverage of CFA and dimensionality, with tutorials and mention of related software.

Note that a measure can have good alpha and omega total reliability figures while having low omega hierarchical – the “Negocios” test (Row 9 in Table 1) is an example – it’s not very “pure”, with omega hierarchical at 0.57, but its reliability is acceptable: both alpha and omega total are at 0.86. This serves to underline something known for decades, but not always acknowledged: alpha is <u>not</u> an index of a measure’s dimensionality. For that matter, neither is omega total.

Q5: should I report alpha, standardized alpha, or omega total in my report?

If theory holds, the indicated correct answer would be: *omega total for sure*. It is widely recognised as a more accurate estimate of reliability. It makes fewer assumptions; it does not assume that items are tau-equivalent or parallel. In addition, omega total is in theory an upper bound for alpha – alpha is expected to be less than omega total unless items are tau-equivalent, in which case alpha will equal omega total.

However, theory has not held for all of the results in this empirical study. An exception to the theory is seen in Row 5 of Table 1 where alpha edges out omega total when it (omega total) has been calculated by SEM modelling in CALIS, and CFA modelling in JASP. This small apparent discrepancy to theory is corrected when omega total has been calculated using EFA and a bifactor model; see Column 10 where the omega total value does exceed alpha.

A similar result was found in another empirical study, Hayes & Coutts (2020). They looked at 17 scales and found two cases where alpha was ever so slightly greater than omega total (CFA). This is further discussed below, under Q6.

The apparent violations of theory mentioned here correct themselves when omega total is derived by applying EFA and a bifactor model, a matter also further discussed below, under Q6.

There could be an argument for also including alpha. Omega remains relatively new. Readers of a technical report may be unfamiliar with omega and uncertain how to interpret it; it could be that the lack of an alpha figure might cause some to fidget.

Standardized alpha is rarely encountered as its assumptions are difficult to meet.

Q6: if I don't have an app for omega, and have to report only alpha, am I really bad off?

From a practical point of view, very often the difference between alpha and omega total may not amount to much. In this study, a comparison of alpha (Column 6 in Table 1, computed with JASP and cross-checked with Lertap5), and omega total (CFA), as computed by either CALIS or JASP (Columns 8 and 9[7]), shows nearly identical values for the 16 instruments involved in this study.

Other studies have noted the great similarity between alpha and omega total values. For example, in the Hayes & Coutts (2020) study mentioned above, when rounded to two decimal places, 11 of 17 scales showed no difference between alpha and omega total (CFA); the six other scales had differences of 0.01.

A study by Savalie & Reise (2019) stated that "Researchers working with well-established unidimensional scales who continue reporting alpha are rarely more than trivially misrepresenting their measures' reliability."

In a simulation study comparing alpha and omega values under varied conditions, Zhang and Yuan (2016) concluded: "… for normally distributed data, there is no substantial difference between alpha and omega estimates regardless of whether the items are tau-equivalent or not ….".

Geldhof et al. (2013) used CALIS to compute omega total, compared results to alpha, and found no compelling reason to advocate against the continued use of alpha.

Even McDonald (1999, Chapter 6), wrote that "… it is difficult to find … empirical examples in which the estimate of alpha is very much lower than that of omega ….".

These literature citations are largely founded on the use of omega total values resulting from the application of CFA. As mentioned, in the present study omega total was also estimated by applying an EFA bifactor model.

Table 1 shows that EFA omega values (Column 10) were occasionally more than trivially greater than corresponding alpha values (Column 6). McNeish (2017) noted the same outcome when processing results from the Big Five Inventory – after observing CFA omega values equal to alpha on all five BFI subscales, or just .01 points above alpha, corresponding EFA omega values were stronger. For example, on the BFIneuro subscale, he found alpha at .81, omega CFA at .82, and omega EFA at 0.88. (Compare with Row 14 in Table 1 where BFIneuro data are also given, and a similar trend observed.)

At the end of the day, and in spite of often highly similar alpha and omega total figures, one could perhaps be "bad off" reporting only alpha as editors, reviewers, and publishers might increasingly come to expect omega total. An alpha-only report *might* now be seen as reflective of dated research, below par, not up to snuff. With this in mind, however, empirical

[7] Technically, CALIS is a general SEM app; CFA is a specific model within SEM.

studies, such as the present one, suggest that alpha and omega total seem unlikely to have meaningful differences unless, *perhaps*, omega total is estimated using EFA.

Q7: Should I look at omega hierarchical?

Doing so could be insightful. If there are a relatively small number of items, the path diagrams (*Figure 1* and *Figure 2* above) can be informative. They may be quite welcome in a report.

But above twenty items or so, the path diagrams will become much too busy, overcrowded, difficult to interpret[8]. Then a simple comparison of omega hierarchical and omega total will carry a message as to the saturation (predominance) of the general factor. If an instrument has a strong, dominant general factor, it will be reflected in a high omega hierarchical value. In such cases, the difference between omega hierarchical and omega total will be low. In Table 1, DunnSES in Row 13 stands out in this regard.

Software used

JASP Version 0.14.1 was used to calculate the alpha and standardized alpha figures seen in Columns 6 and 7 of Table 1, and the CFA-based omega total in Column[9]. JASP will accept files in these formats: .sav, .txt, .csv, and .ods. In this study, the Omega1 macro in Lertap5 was used to prepare csv files having item scores for each of the 16 datasets[10].

Appendix B has more information about JASP and also makes mention of *jamovi*.

The CALIS procedure in the University edition of SAS was used to produce the factor loadings needed to compute the omega total values seen in Table 1 column 9. The SAS script used may be downloaded from here.

The University edition of SAS has now been replaced by "SAS OnDemand for Academics", resulting in a cloud-based app that involves no installation (website is here).

This is a marked improvement in SAS, but CALIS itself does not output an omega estimate of any sort; the factor loadings from CALIS may be copied to a spreadsheet, from which omega may then be calculated[11]. JASP is unquestionably easier to use.

Omega total in Table 1 Column 10, and omega hierarchical in Column 11, were obtained by using the CRAN psych package following a process described in Appendix C.

AMOS and Mplus are two other programs capable of producing omega total estimates. Bandalos (2018, Chapter 13) has an Mplus omega example, Goodboy and Martin (2020) have another.

A version of **SPSS** released late in 2020 provides support for omega users (see Appendix A).

[8] As an example, see these complete psych results for the 34-item UniBB-Test dataset.
[9] JASP 0.15 beta was released in August 2021 with a more robust reliability module.
[10] csv files have "comma-separated values" and are made by a number of apps (Microsoft Excel and Google Sheets to name just two). Lertap5 is not required.
[11] Sum the factor loadings, square the sum and then divide it by the variance of the test scores.

An updated version of an Excel-based test and survey analysis app, **Lertap5**, was released in February 2021 with support for omega users (see Appendix D).

Data sources

Many of the datasets used in Table 1 may be found here. The "Blirt8" dataset is included in this zip file. Hammer data are found here.

Recommendations

At the moment, JASP (or *jamovi*) would be the author's first recommendation for those wanting to obtain an estimate of omega total[12]. It is easy to install, not difficult to master, and uses CFA. Its output is well formatted, and error messages are generally clear. JASP documentation, however, is not yet extensive.

Second recommendation: the CRAN psych package. It may be easy to use for those familiar with R packages, and, perhaps, also for those working from the steps outlined in Appendix C.

An advantage of using CRAN pysch is that its uses EFA to estimate omega total, and it also produces an estimate of omega hierarchical[13]. The chance of omega total exceeding alpha in more than a trivial way has been found, in this study, to result when EFA is used, as in the psych package. McNeish (2017) reported a similar finding.

Not highly recommended, but potentially useful when all item covariances are known to be positive: the closed-form method as used in SPSS and Lertap5.

Note the mention of AMOS and Mplus above – these omega-capable apps were not used in the present study.

[12] Jamovi, unlike JASP, will accept Excel xlsx workbooks as input.
[13] There has been mention of omega hierarchical in JASP GitHub. Perhaps it will be included in a forthcoming release.

References

Bandalos, D.L. (2018). *Measurement theory and applications.* New York, NY: The Guilford Press.

Cho, E. (2016). Making reliability reliable: A systematic approach to reliability coefficients. Organizational Research Methods, 19(4), 651–682.

Cho, E. (2018). Fixing a broken clock: a historical review of the origins of reliability coefficients, including Cronbach's alpha. *Survey Research*, 19(2), 23-54.

Dunn, T.J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105, 399-412.

Falk, C.F. & Savaeli, V. (2011). The Relationship Between Unstandardized and Standardized Alpha, True Reliability, and the Underlying Measurement Model. *Journal of Personality Assessment*, 93(5),445-53.

Flora, D.B. (2020). Your Coefficient Alpha is Probably Wrong, but Which Coefficient Omega is Right? A Tutorial on Using R to Obtain Better Reliability Estimates. *Advances in Methods and Practices in Psychological Science*, 3(4), 484-501.

Geldhof, G.J., Preacher, K.J, & Zyphur, M.J. (2013). Reliability estimation in a multilevel confirmatory factor analysis framework. *Psychological Methods.*

Goodboy, A.K, & Matthew, M.M. (2020) Omega over alpha for reliability estimation of unidimensional communication measures. *Annals of the International Communication Association*, 44(4), 422-439. DOI: 10.1080/23808985.2020.1846135

Hancock, G. R., & An, J. (2020). A closed-form alternative for estimating omega reliability under unidimensionality. *Measurement: Interdisciplinary Research and Perspectives*, 18(1), 1-14.

Hayes, A. F. & Coutts, J. J. (2020). Use Omega Rather Than Cronbach's Alpha for Estimating Reliability. But … *Communication Methods and Measures*, 14(1), 1-24.

McDonald, R. P. (1999). Test theory: A unified approach. Mahwah, NJ: Erlbaum.

McNeish, D. (2017). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, 23(3) 412-433. (An Epub.)

Meyer, J. Patrick (2010). *Reliability.* New York: Oxford University Press, Inc.

Nelson, L. R. (2016) Coefficients alpha and omega, an empirical comparison. Accessed November 25, 2020 from ResearchGate. Also available here.

Revelle, W. (2016, May). Using R and the psych package to find ω. Accessed October 19, 2016 from http://personality-project.org/r/psych/HowTo/omega.pdf.

Revelle, W. and Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological Assessment* 31 (12) p 1395-1411.

Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, 74(1), 145–154.

Revelle, W. (2008). psych: Procedures for personality and psychological research (an R package).

Specht, D. A. (1975). *SPSS: Statistical package for the social sciences, version 6: Users guide to subprogram reliability and repeated measures analysis of variance*. Ames, IA: Iowa State University.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: understanding concepts and applications.* Washington, DC: The American Psychological Association.

Zhang, Z., & Yuan K. (2016). Robust Coefficients Alpha and Omega and Confidence Intervals with Outlying Observations and Missing Data: Methods and Software, *Educational and Psychological Measurement*, Vol. 76(3) 387–411

The references found on this Wikipedia page are also useful (as at January 29, 2021).

Appendix A: SPSS

Version 27.0.1.0 of **SPSS** was released in the latter half of the year 2020. This was the first version to support the calculation of omega total[14].

An introduction to this version is here. To be noted is that SPSS does not support CFA[15] (confirmatory factor analysis). Instead, SPSS omega relies on a "closed form" solution, as described in Hancock & An (2020), and in Hayes & Coutts (2020)[16].

SPSS was used to obtain estimates of alpha, standardized alpha, and omega total for each of the 16 datasets in Table 1. Table 2 below has results. Most of the JASP and SPSS omega values agree.

When asked to compute standardized alpha, SPSS gave a warning, reporting that "the determinant of the covariance matrix is zero (or approx.)" for all 16 datasets. However, it nonetheless reported both alpha and standardized alpha values.

Table 2 allows a ready comparison of JASP and SPSS results. Column 6a has been copied from Table 1 Column 6. Column 6b has SPSS results. Similarly, Table 2 Column 7 has been copied from Table 1, while Column 7b has SPSS results. The alpha values are the same. The standardized alpha values are very similar but not all are identical.

Table 2

1	2	3	4	5	6a	6b	7	7b	9	9b
No.	Source	Type	Items	N	Alpha	SPSS Alpha	Std. Alpha	SPSS Std. Alpha	Omega JASP	SPSS Omega
1	MathsQuiz	Test	15	999	0.80	0.80	0.78	0.80	0.81	0.81
2	UniAA	Test	30	127	0.74	0.74	0.74	0.74	0.74	0.71
3	UniBB	Test	34	132	0.82	0.82	0.81	0.82	0.82	0.80
4	Zmed	Test	100	2,470	0.96	0.96	0.96	0.96	0.96	0.96
5	HalfTime	Test	100	424	0.94	0.94	0.93	0.94	0.93	0.93
6	N Rivers	Test	50	689	0.92	0.92	0.92	0.92	0.92	0.92
7	LenguaBlg	Test	50	5,504	0.81	0.81	0.81	0.81	0.82	0.88
8	LenguaBlg	Test	47	5,504	0.83	0.83	0.83	0.83	0.83	0.83
9	Negocios	Test	60	500	0.86	0.86	0.86	0.86	0.86	0.85
10	FIMS	Test	14	6,371	0.69	0.69	0.70	0.69	0.70	0.69
11	FIMS AUS	Test	14	4,320	0.60	0.60	0.60	0.60	0.62	NA
12	FIMS JPN	Test	14	2,051	0.73	0.73	0.73	0.73	0.74	0.74
13	DunnSES	Scale	7	201	0.94	0.94	0.94	0.94	0.94	0.94
14	BFIneuro	Scale	5	2,663	0.81	0.81	0.82	0.81	0.81	0.81
15	Blirt8	Scale	8	211	0.78	0.78	0.78	0.78	0.79	0.78
16	Hammer	Scale	29	758	0.93	0.93	0.93	0.93	0.94	0.93

[14] IBM SPSS Version 27.0.1.0. Omega statistic discussed here. Parallel reliability discussed here.
[15] According to Hayes & Coutts (2020), "SPSS has no built in procedures for confirmatory factor analysis."
[16] Both papers suggest CFA is preferred; the closed-form method is a possible option when CFA is not available.

When computing omega, SPSS again reported problems with dataset covariance matrices; it gave warnings for all but one of the tests (the exception was the FIMS test in Row 10).

For the FIMS AUS dataset (Row 11), SPSS found that "Omega could not be estimated due to negative or zero item covariances". SPSS found no problems with the four scales (Rows 13 – 16).

A few (4) SPSS alpha values exceeded SPSS omega values. This is not expected as, in theory, alpha will not be greater than omega.

Of special note are the results for the data in Row 7 of Table 2. In this case, the SPSS estimate of omega total showed a less than trivial increase over the omega total (CFA) from JASP. Here again negative covariances were probably a factor – three items in the LenguaBlg dataset had negative covariance values. When these items were omitted (Row 8 in Table 2), the SPSS estimate of omega total dropped to equal the outcome from JASP.

The previous page makes mention of problems SPSS encountered when processing data from the [FIM AUS dataset](). The error message from SPSS was informative:

Reliability Statistics

McDonald's Omega	N of Items
.[a]	14

a. Omega cannot be estimated due to negative or zero item covariances. This may be due to items needing to be reverse scored, or to violations of model assumptions.

The FIMS AUS dataset is from a short cognitive test of mathematics.

Cognitive items are never reverse scored When, as above, SPSS mentions a possible need for items to be "reverse scored", it is assuming that affective items are in use, such as those found in scales. That was not the case here.

What has happened is that one of the 14 test items had negative correlations with some of the other items. Table 3 below has been copied from a Lertap5 spreadsheet; SPSS, JASP and *jamovi* will readily make similar reports.

Table 3

Lertap5 IStats matrix, created: 31-Oct-20. h

Record No.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14
Q12	-0.01	-0.05	-0.04	0.00	0.02	0.01	0.08	-0.02	0.00	-0.02	0.01	1.00	0.05	-0.07
Q13	0.00	0.04	0.03	0.05	0.11	-0.01	0.03	0.01	0.08	0.02	0.18	0.05	1.00	0.04
Q14	0.15	0.23	0.24	0.23	0.08	0.19	0.05	0.09	0.10	0.19	0.16	-0.07	0.04	1.00
average	*0.10*	*0.14*	*0.11*	*0.15*	*0.08*	*0.13*	*0.07*	*0.05*	*0.07*	*0.12*	*0.14*	*0.00*	*0.05*	*0.13*
SMC	*0.10*	*0.18*	*0.11*	*0.17*	*0.06*	*0.16*	*0.04*	*0.03*	*0.04*	*0.12*	*0.14*	*0.02*	*0.04*	*0.14*

Q12 had negative correlations with Q1, Q2, Q3, Q8, Q10, and Q14.

An item meant to form part of a test or scale should not correlate negatively with other items. When negative correlations are found, coefficient alpha's value will be affected – it'll be lower than what it otherwise would be. Many times, an app will tell us what alpha would be if an item were deleted from the test or scale. In this example, alpha with all 14 items included was 0.602 – omitting Q12 boosted alpha to 0.626.

The closed-form method SPSS uses to calculate omega sputters when it encounters negative item correlations (or covariances). Such incidents will not arise when standard data analysis procedures are followed -- an initial look at item correlations/covariances should always be made to make sure that negative inter-item correlations or covariances are not present. If they are, indicated items should be removed from the test (or scale).

Appendix C mentions the use of an RStudio RMD script for calculating omega estimates using the closed-form method. It also found a fatal error -- no estimate could be made.

JASP processed the FIMS AUS dataset without problem but did note that Q12 “correlated negatively with the scale”.

Appendix B: JASP / *jamovi*[17]

JASP is a project based at the University of Amsterdam. The website is here.

The JASP work undertaken for this paper involved the use of csv data files[18]. A number of apps will open such files. The screen snapshot in Figure 3 displays the "blirt8.csv" file when opened in Microsoft Excel.

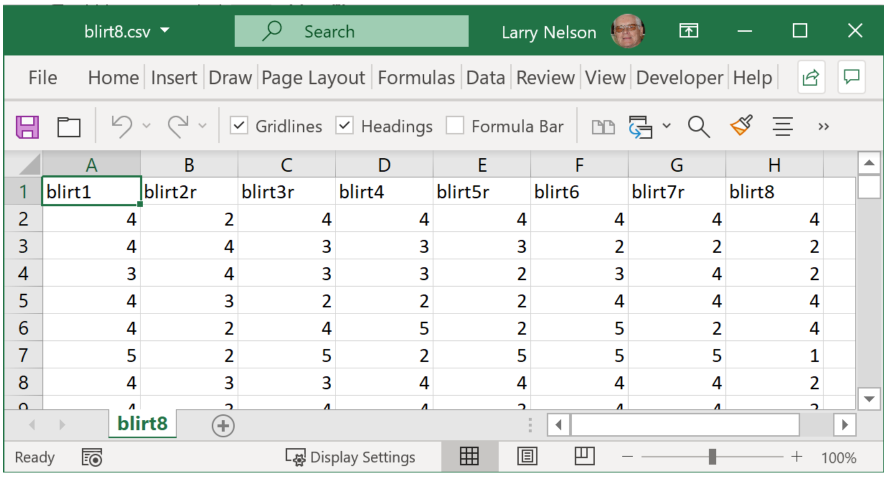


Figure 3

The figure shows the first eight rows of the csv file. The first row contains "headers" corresponding to the names of the eight items used in the BLIRT8 dataset.

Each subsequent row contains item scores for a person. The first person had a response scored as a "4" on the first item, "blirt1". If the eight item scores were summed to get a person's total score, in this case the first person would have a total score of 30.

There are a number of ways to create a csv file similar to this example. The file must have item scores, as exemplified above for Blirt8. The "omega1" special macro in Lertap5 will create csv files with item scores and was often used in this study. Many users may have item scores already at hand in an Excel workbook, or in Google Sheets, in which case saving a copy as a csv file will be straightfor-ward.

The screen snapshot in Figure 4 below shows JASP version 0.14.1 with the first eight blirt8.csv rows displayed.

Figure 5 captures the JASP interface after initiating a "Single-Test Reliability Analysis", and Figure 6 displays the resultant output.

[17] An example of using *jamovi* in a reliability study is here; jamovi will read Excel xlsx workbooks directly.
[18] The same csv files were also used with SPSS.

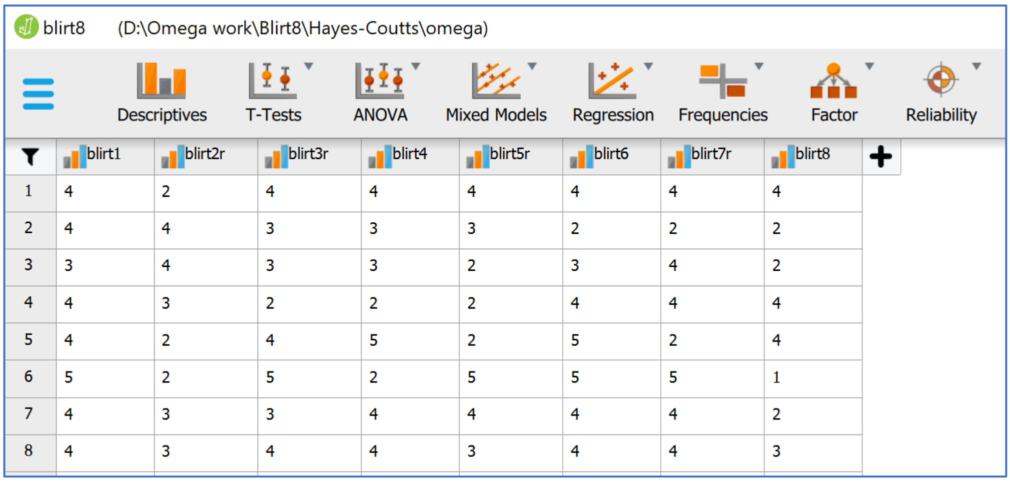


Figure 4

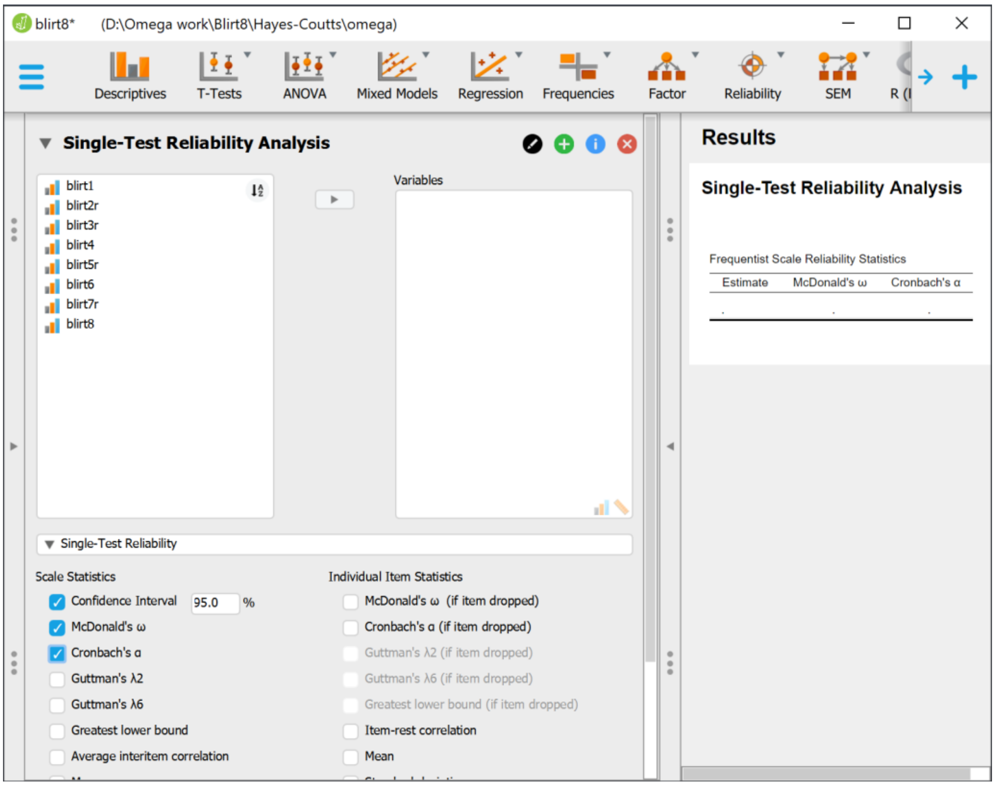


Figure 5

Once the eight items are moved from the left-hand panel (Figure 5) over to the "Variables" panel, action will begin automatically. Results are as seen in Figure 6.

Results

Single-Test Reliability Analysis

Frequentist Scale Reliability Statistics

Estimate	McDonald's ω	Cronbach's α
Point estimate	0.785	0.780
95% CI lower bound	0.741	0.731
95% CI upper bound	0.829	0.821

Figure 6

Author's note: JASP is solidly based on the incorporation of selected CRAN (R) packages, a fact which, for the most part, is not apparent when JASP runs. *jamovi* is similar.

Appendix C: CRAN psych module

CRAN, the Comprehensive R Archive Network, is a repository of application packages and tools contributed by numerous authors, created using the R scripting language. An introduction may be found at this webpage. Discussions of the advantages and disadvantages of R and CRAN are numerous, as an internet search will reveal (try "pros and cons of R").

Readers not familiar with CRAN and R packages may be best advised to muster a considerable amount of patience before "taking the plunge". There's a great deal to like, a large number of free packages. Many of them are very powerful data analysis tools. However, the learning curve can be rather steep, and at times the documentation may be seen as rather abrupt, less than friendly for beginners. The same might be said for error messages encountered when running some packages.

The following resources cover the steps used in this paper to get the omega reliability estimates seen in columns 10 and 11 of Table 1, and the path diagrams displayed in *Figure 1* and *Figure 2*.

(1) Use this document for instructions on the installation of R and RStudio.

(2) Then, confer with this document for instructions on the installation and use of two R packages, "psych", and "TAM", and note: "TAM" is not required for calculating omega statistics – its installation may be skipped.

The document covers two "test drives". The first uses RStudio and an RMD script, "Omega-From-IScores.Rmd". A Microsoft Word document with sample output produced by this script may be downloaded from this link[19].

The second test drive uses the R programming environment and the "Omega-IscoresProg.R" script. Output is seen in the document itself.

● There can at times be limitations when using the psych package. For example, when processing data collected using the 21-item Beck Depression Inventory with a sample of Australian students, the following error messages appeared:

```
Warning in fa.stats(r = r, f = f, phi = phi, n.obs = n.obs, np.obs = np.obs, :The estimated
weights for the factor scores are probably incorrect.Try a different factor score estimatio
n method.

Warning in fac(r = r, nfactors = nfactors, n.obs = n.obs, rotate = rotate:An ultra-Heywood
case was detected. Examine the results carefully

Warning in cov2cor(t(w) %*% r %*% w): diag(.) had 0 or NA entries; non-finite result is dou
btful
```

[19] A look at the output is recommended.

● An RMD script for calculating omega using the closed-form approach employed in SPSS is available here. It works in the same manner found in the “Omega-From-IScores.Rmd” script mentioned above.

An example of the output produced by the script may be downloaded from here. Note (a): the script assumes that the pysch package has been installed, as mentioned in (2) above. Note (b): the closed-form method is not the preferred method for calculating omega. It’s a resource perhaps useful when the methods based on CFA or EFA are not at hand.

McNeish (2017) has another example of the output produced by the psych package and is a recommended reference.

Appendix D: Lertap5

Lertap5 is a veteran item, test, and survey analysis system packaged as an Excel “app”: Lertap5.xlsm.

Some of its support for omega has been mentioned in main the text above. For example, the “Omega1 macro” in Lertap5 is designed for use with the R psych package seen in Appendix C.

Hancock & An (2020) present a closed-form method for estimating omega total, and extensively document both its strengths and weaknesses. The method is used in SPSS to derive an estimate of omega total.

Hancock & An’s Appendix A has related R code – it has been used in the “RMD script for calculating omega” mentioned towards the end of Appendix C in this document.

Hayes & Coutts (2020) developed and made freely available omega macros for use with SAS and SPSS. They also discussed the closed-form method presented by Hancock & An (2020), referring to it as the “HA algorithm”. Table 1 in their article has a comparison of omega total estimates using CFA, EFA, and HA.

The closed-form method has now been installed in Lertap5 and tested with the 16 datasets used in this paper. Results match the SPSS omega estimates seen in the last column of Table 2 in Appendix A[20] of this document[21].

Figure 7 displays Lertap5’s closed-form results using the Blirt8 dataset[22]

Closed-form lambda estimates:		
blirt1	0.652	
blirt2r	0.676	
blirt3r	0.582	
blirt4	0.709	
blirt5r	0.548	
blirt6	0.634	
blirt7r	0.379	
blirt8	0.343	
Sum of lambda estimates:		4.52
Score variance:		26.28
omega estimate:		**0.779**

Figure 7

Lertap5.xlsm has been programmed using VBA, Visual Basic for Applications. VBA code lines used to calculate the lambda estimate for a given item are shown below:

[20] A match is to be expected as the closed-form method involves a simple non-iterative procedure.
[21] Lertap5 was also unable to process the FIMS AUS results, reporting that “No estimate was possible; too many negative covariance values”.
[22] Compare with Table 1 in Hayes & Coutts (2020). In Figure 7 above, “lambdas” are estimated factor loadings.

```
For ColNo1 = 1 To nits

  If ColNo1 <> rownumber Then

   If ColNo1 <> nits Then

     For ColNo2 = ColNo1 + 1 To nits

       If ColNo2 <> rownumber Then

         If ColNo2 <> nits + 1 Then

           Let lambdaNumerator = lambdaNumerator + (Cov(rownumber, ColNo1) * Cov(rownumber, ColNo2))

           Let lambdaDenominator = lambdaDenominator + Cov(ColNo1, ColNo2)

         End If

       End If

     Next ColNo2

   End If

  End If

Next ColNo1
```

In the VBA code above, “nits” is the number of items; “rownumber” is the sequential number of an item; it will range from 1 to nits as controlled by a For/Next sequence not shown above.

Cov(nits,nits) holds covariance values.   The lambda estimate for an item will be numerator divided by denominator.

To be mentioned is that CFA or EFA are preferred over the closed-form approach. See Hayes & Coutts (2020) and also Hancock & An (2020) for discussions on the limitations of the closed-form method. Lertap5 will present appropriate warnings when some of the limitations are encountered; they are often related to negative covariances.

Examples: