Technical Appendix for:
"Results on the Standard Error of the Coefficient Alpha Index of Reliability"
by Adam Duhachek, Anne T. Coughlan, and Dawn Iacobucci[1]

April 2005

[1]Duhachek is a professor of marketing at Indiana University, Coughlan is a professor of marketing at Northwestern University, and Iacobucci is a professor of marketing at Wharton. The authors are grateful to Professor Steve Shugan, the AE and Reviewers, participants at the European ACR meeting, and seminar attendees from the marketing departments at the University of Colorado; University of Michigan; Northwestern University; Notre Dame and the University of Southern California for their helpful feedback on this research. Please address correspondence to the first author at: aduhache@indiana.edu.

Technical Appendix for:
"Results on the Standard Error of the Coefficient Alpha Index of Reliability"

Derivation of baseline alpha and standard error (equations (1), (2), and (4))):

First, regarding alpha; the standard equation for alpha is:

$$\alpha = \frac{p}{p-1}\left[1 - \frac{\sum \sigma_i^2}{\sigma_T^2}\right],$$

which simplifies under the assumption of interitem correlation homogeneity to

$$\alpha = \frac{p}{p-1}\left[1 - \frac{p}{p + p(p-1)r}\right]$$

given that there are p 1's on the diagonal of the correlation matrix, $\sum_{i=1}^{p} \sigma_i^2 = p$ and $p(p-1)$ of the r's on the off-diagonals, hence $p+p(p-1)r$.

Then note that the standard error (equation (3)) is the square root of the ratio of the variance divided by the number of observations. Then, using the variance equation (2), note that:

p= number of items, and $trV=p$ for this simple correlation matrix form (hence, $tr^2V=p^2$). Further evaluating the other terms of equation (2), we have first that:

$j'Vj = \sum_{i=1}^{p}\sum_{j=1}^{p} v_{ij}$. This summation includes p entries of "1" (on the diagonal), and $p(p-1)$ entries equal to r (on the off-diagonals). The summation can therefore be restated as:

$$j'Vj = \sum_{i=1}^{p}\sum_{j=1}^{p} v_{ij} = p + p(p-1)r = A_1 \ .$$

Next, we have that:

$$trV^2 = \sum_{i=1}^{p}\sum_{k=1}^{p} v_{ik}v_{ki} = B_1 .$$

In this expression, note that

$$\sum_{k=1}^{p} v_{ik} v_{ki} = 1 + (p-1)r^2 \text{, and thus,}$$

$$trV^2 = \sum_{i=1}^{p} \sum_{k=1}^{p} v_{ik} v_{ki} = p\left[1 + (p-1)r^2\right] .$$

Further, we have that:

$$j'V^2 j = trV^2 + \sum_{i=1}^{p} \sum_{j \neq i}^{p} \sum_{k=1}^{p} v_{ik} v_{kj} \text{ , where}$$

$$\sum_{i=1}^{p} \sum_{j \neq i}^{p} \sum_{k=1}^{p} v_{ik} v_{kj} = p(p-1)\left[2\bar{r} + (p-2)\bar{r}^2\right] = C_1 .$$

We can then restate Q_1 as:

$$Q_1 = \left[\frac{2p^2}{(p-1)^2 A_1^3}\right]\left[A_1(B_1 + p^2) - 2p(B_1 + C_1)\right],$$

which, after substitution and simplification, yields equation (6) for Q_1.

<u>Proposition 1.</u> The comparative-static sign effects of *n*, *p*, and *r* on alpha and its standard error are:

TABLE 1: Comparative-Static Effects on Alpha and Its Standard Error:

	Sign of Comparative-Static Effect On:	
Of:	α_1	SE_1
n	$\frac{\partial \alpha_1}{\partial n} = 0$	$\frac{\partial SE_1}{\partial n} < 0$
p	$\frac{\partial \alpha_1}{\partial p} > 0$	$\frac{\partial SE_1}{\partial p} < 0$
r	$\frac{\partial \alpha_1}{\partial r} > 0$	$\frac{\partial SE_1}{\partial r} < 0$

Proof:

The entries in Table 1 can be verified by simply taking the derivatives of α_1 and Q_1 with respect to *n*, *p*, and *r*. Clearly, α_1 is invariant with respect to changes in *n*. The other two derivatives of α_1 are:

$$\frac{\partial \alpha_1}{\partial p} = -\frac{r(r-1)}{[1+r(p-1)]^2} > 0 \text{ and}$$

$$\frac{\partial \alpha_1}{\partial r} = \frac{p}{[1+r(p-1)]^2} > 0 \ .$$

The derivatives of SE_1 with respect to n, p, and r are:

$$\frac{\partial SE_1}{\partial n} = -\frac{\sqrt{Q/n}}{2n} < 0 \ ;$$

$$\frac{\partial SE_1}{\partial p} = -\frac{(r-1)^2\left[1+r\left(2p^2-p-1\right)\right]}{\sqrt{n}(p-1)^2[1+r(p-1)]^3} < 0 \text{ (recall p} \geq 2\text{); and}$$

$$\frac{\partial SE_1}{\partial r} = -\frac{2p^2(r-1)}{\sqrt{n}(p-1)[1+r(p-1)]^3} < 0.$$

Q.E.D.

<u>Proposition 2.</u> *p* and *r* are complements in increasing the estimate of alpha for $r < \frac{1}{1+p}$, and substitutes in increasing alpha for $r > \frac{1}{1+p}$. Further, *n, p*, and *r* are substitutes in decreasing the standard error of alpha. Formally:

$$\frac{\partial^2 \alpha_1}{\partial p \partial r} > 0 \ \text{for} \ r < \frac{1}{1+p}, \ \text{and} \ < 0 \ \text{for} \ r > \frac{1}{1+p};$$

$$\frac{\partial^2 SE_1}{\partial n \partial p} > 0 \ ; \quad \frac{\partial^2 SE_1}{\partial n \partial r} > 0 \ ; \quad \frac{\partial^2 SE_1}{\partial p \partial r} > 0 \ .$$

Proof: The Proposition can be proven by taking the relevant second derivatives. They are:

$$\frac{\partial^2 \alpha_1}{\partial p \partial r} = \frac{1-r-pr}{[1+(p-1)r]^3}, \ \text{ which is } >0 \text{ for } r < \frac{1}{1+p}, \text{ and } < 0 \text{ for } r > \frac{1}{1+p} \ .$$

This means that, given a $\{p, r\}$ pair for which $r < \frac{1}{1+p}$, any increase in scale length *increases* the marginal benefit of improving *r*, and conversely, any increase in *r increases* the alpha benefit of adding to scale length. This means that in this range, *r* and *p* act as complements in increasing alpha. The converse set of statements is true when $r > \frac{1}{1+p}$, so that in this range, *r* and *p* act as substitutes. Below is a table showing for what values of *r* we have the complementary relationship between *r* and *p*, given any *p* value from 2 to 20:

Value of *p*	**Value of *r* below which** $\frac{\partial^2 \alpha_1}{\partial p \partial r} > 0$:
2	0.333
3	0.25
4	0.2
5	0.167
6	0.143
7	0.125
8	0.111
9	0.1
10	0.091
11	0.083
12	0.077
13	0.071
14	0.067
15	0.063
16	0.059
17	0.056
18	0.053
19	0.05
20	0.048

Note that as *p* increases, the range of *r* over which *r* and *p* are complements becomes smaller and smaller; intuitively, when scale length increases, the initial inter-item correlation must be lower and lower to guarantee that increasing one parameter increases the "productivity" (in terms of alpha increases) of the other parameter.

The second derivatives of the standard error of alpha reported in the Proposition are:

$$\frac{\partial^2 SE_1}{\partial n \partial p} = \frac{(r-1)^2\left[1+r\left(2p^2-p-1\right)\right]}{2n\sqrt{n}(p-1)^2[1+r(p-1)]^3} > 0\ ;$$

$$\frac{\partial^2 SE_1}{\partial n \partial r} = \frac{p^2(r-1)}{n\sqrt{n}(p-1)[1+r(p-1)]^3} > 0\ ;$$

$$\frac{\partial^2 SE_1}{\partial p \partial r} = \frac{4p(r-1)\left(2rp^2-p+2(1-r)\right)}{2\sqrt{n}(p-1)^2[1+r(p-1)]^4} > 0\ .$$

Q.E.D.

<u>Derivation of alpha and standard error in the case of "one bad item" (equations (8) and (9)):</u>

In this scenario, the alpha simplification is computed as (equation (8)):

$$\alpha_2 = \frac{p}{p-1}\left[1 - \frac{p}{p + 2(p-1)(cr) + (p-1)(p-2)r}\right];$$

the numerator in brackets again is easy, $\sum_{i=1}^{p} \sigma_i^2$, given the p 1's on the diagonal. Regarding the *cr* terms, there are *p-1 cr*'s in its row, and *p-1 cr*'s in its column, hence *2(p-1) cr*.

The remaining reduced matrix is *(p-1)x(p-1)* with *(p-1)(p-2)* off-diagonals, with element, *r*, hence, *p+2(p-1)cr+(p-1)(p-2)r*.

The formula for *Q* in equation (2) still characterizes the variance of alpha, but in this case, the A, B, and C terms noted in the derivation of Q_1 above are now different. Using the same techniques as in that baseline case, first note that:

p= number of items, and $trV=p$ for this correlation matrix form (hence, $tr^2V=p^2$). Further evaluating the other terms of equation (3), we have first that:

$j'Vj = \sum_{i=1}^{p}\sum_{j=1}^{p} v_{ij}$. This summation includes p entries of "1" (on the diagonal), and *2(p-1)* entries equal to *cr* (the off-diagonal terms pertaining to the one bad item). There remain *(p-1)(p-2)* terms equal to *r*. The summation can therefore be restated as:

$$j'Vj = \sum_{i=1}^{p}\sum_{j=1}^{p} v_{ij} = p + 2(p-1)cr + (p-1)(p-2)r = A_2 \quad .$$

Next, we have that:

$$trV^2 = \sum_{i=1}^{p}\sum_{k=1}^{p} v_{ik}v_{ki} = B_2 \,.$$

In this expression, note that for the first row of the matrix,

$$\sum_{k=1}^{p} v_{1k} v_{k1} = 1 + (p-1)(cr)^2 \ .$$

For the next *(p-1)* rows, we have:

$$\sum_{k=1}^{p} v_{ik} v_{ki} \Big|_{i \neq 1} = (p-1)\left[(cr)^2 + 1 + (p-2)r^2\right].$$

The sum can therefore be written as:

$$trV^2 = \sum_{i=1}^{p}\sum_{k=1}^{p} v_{ik} v_{ki} = B_2 = 1 + (p-1)(cr)^2 + (p-1)\left[(cr)^2 + 1 + (p-2)r^2\right].$$

Further, we have that:

$j'V^2 j = trV^2 + \sum_{i=1}^{p}\sum_{j \neq i}^{p}\sum_{k=1}^{p} v_{ik} v_{kj}$, where through the same sort of calculation,

$$\sum_{i=1}^{p}\sum_{j \neq i}^{p}\sum_{k=1}^{p} v_{ik} v_{kj} = 2(p-1)\left[2cr + (p-2)cr^2\right] + (p-1)(p-2)\left[(cr)^2 + 2r + (p-3)r^2\right] = C_2 \ .$$

We can then restate Q_2 as:

$$Q_2 = \left[\frac{2p^2}{(p-1)^2 A_2^3}\right]\left[A_2(B_2 + p^2) - 2p(B_2 + C_2)\right],$$

which, after substitution and simplification, yields equation (7).

<u>Derivation of alpha and standard error in the case of "two underlying factors" (equations (9) and (10)):</u>

The final scenario considers the reduction of alpha to:

$$\alpha = \frac{p}{p-1}\left[1 - \frac{p}{p + \frac{1}{2}p^2(cr) + pr\left(\frac{1}{2}p - 1\right)}\right],$$

via the now standard $\sum_{i=1}^{p}\sigma_i^2 = p$ for the 1's on the diagonal. Then, for each (p) row, half the items load on each factor, resulting in $.5p$ columns of cr, hence $.5p^2(cr)$, and for each (p) row, there exists an r in each location in the upper and lower submatrices, hence $pr(.5p-1)$, altogether, $p+.5p^2cr+pr(.5p-1)$.

The formula for Q in equation (2) still characterizes the variance of alpha, but in this case, the A, B, and C terms noted in the derivations of Q_1 and Q_2 above are different. Using the same techniques as in that baseline case, first note that:

p= number of items, and $trV=p$ for this correlation matrix form (hence, $tr^2V=p^2$). Further evaluating the other terms of equation (2), we have first that:

$j'Vj = \sum_{i=1}^{p}\sum_{j=1}^{p} v_{ij}$. This summation includes p entries of "1" (on the diagonal). Further, in any of the p rows of the matrix, half of the items load on each factor; so there are ($.5p^2$) instances where an entry of (cr) occurs in the matrix. Finally, in each of the p rows of the matrix, there are entries of r in ($.5p$-1) cells, so across the whole matrix, there are [$p(.5p-1)$] instances of an entry of r. Thus, the summation can be written as:

$$j'Vj = \sum_{i=1}^{p}\sum_{j=1}^{p} v_{ij} = p+.5p^2cr+pr(.5p-1) = A_3 \quad .$$

For the next term,

$$trV^2 = \sum_{i=1}^{p}\sum_{k=1}^{p} v_{ik}v_{ki} = B_3 .$$

In this expression, note that for each row of the matrix,

$$\sum_{k=1}^{p} v_{1k}v_{k1} = 1+.5p(cr)^2+(.5p-1)r^2 \ ,$$

so that the entire term becomes

$$trV^2 = \sum_{i=1}^{p}\sum_{k=1}^{p} v_{ik}v_{ki} = p\left[1+.5p(cr)^2+(.5p-1)r\right] = B_3 \ .$$

Further, we have that:

$j'V^2j = trV^2 + \sum_{i=1}^{p}\sum_{j\neq i}^{p}\sum_{k=1}^{p} v_{ik}v_{kj}$, where through the same sort of calculation,

$$\sum_{i=1}^{p}\sum_{j\neq i}^{p}\sum_{k=1}^{p} v_{ik}v_{kj} = p(.5p-1)\left[2r+.5p(cr)^2+(.5p-2)r^2\right]+.5p^2\left[2cr+(p-2)cr^2\right]=C_3 \ .$$

We can then restate Q_3 as:

$$Q_3=\left[\frac{2p^2}{(p-1)^2 A_3^3}\right]\left[A_3(B_3+p^2)-2p(B_3+C_3)\right],$$

which, after substitution and simplification, yields equation (9).

<u>Demonstration that any combination of *c* and *r* that yields the same $\bar{r}$ yields the same alpha estimate (relevant for section on "Adding an Item to a Scale"):</u>

We show here that the same alpha value results from some $\bar{r}$ combined with c=1 (the baseline case), as from some c<1 combined with a higher r value, such that the *average* is still $\bar{r}$. In other words, for any given p, any combination of c and r that yields the same $\bar{r}$ yields the same alpha estimate.

The formula for α_2 is given in equation (8). When there is "one bad item," this means that there are (p-1) items that have intercorrelations equal to r, and one item that has a correlation of cr with each of the other items. In this case, it is straightforward to show that the *average* intercorrelation among the p items is given by:

$$\bar{r}=\frac{r(p+2c-2)}{p} \ .$$

Note that in the special case where the starting point is our baseline case, c=1 and the above formula reduces to simply $\bar{r}=r$, which is clearly true. In any other general starting situation, just as in the situation where the starting point is the baseline case, $\bar{r}$ is given and known, and hence can be treated as a parameter for the following discussion.

Imagine now changing the value of c; how would r have to change as well, in such a way as to preserve $\bar{r}$ as the average inter-item correlation? This would imply that r, the underlying correlation among the first (p-1) items, has to adjust according to the following rule:

$$r=\frac{p\bar{r}}{p+2c-2}=\tilde{r}.$$

To see what this adjustment rule implies for the value of alpha, we next substitute in $\tilde{r}$ for r in the formula for α_2, to get:

$$\alpha_2(\tilde{r}) = \frac{p\bar{r}}{1+\bar{r}(p-1)} \ ,$$

which is indeed not a function of c. In other words, *any* feasible value of c will produce the same value of α_2, if r adjusts to changes in c in order to preserve the inter-item correlation at $\bar{r}$; alternatively said, adjusting r to preserve $\bar{r}$ also results in holding the value of α_2 constant with respect to changes in c.

Q.E.D.

<u>Proposition 3.</u> Heterogeneity due to the presence of a single poor item in a scale has an adverse (negative) effect on alpha, and an adverse (positive) effect on the standard error of alpha, holding *p* and *r* constant:

$$\frac{\partial\alpha_2}{\partial c} > 0, \quad and \quad \frac{\partial SE_2}{\partial c} < 0 \quad .$$

Proof: The expression for α_2 is given in equation (8) in the paper, and that for Q_2 is given in equation (7). The first derivative is given by:

$$\frac{\partial\alpha_2}{\partial c} = \frac{2p^2r}{\left[p+2cr(p-1)+(p-1)(p-2)r\right]^2} > 0 \ .$$

Note that the second derivative is equal to:

$$\frac{\partial SE_2}{\partial c} = \frac{1}{2\sqrt{nQ_2}} \frac{\partial Q_2}{\partial c} \; ,$$

whose sign is the sign of $\frac{\partial Q_2}{\partial c}$. This can be expressed as:

$$\frac{\partial Q_2}{\partial c} = \frac{A \cdot B}{C \cdot D} \; , \quad \textit{where}$$

$$\begin{aligned}
A &= 8p^2 r > 0 \\
B &= \quad -4(1-c)^2 r^3 + 2pr(1-c)\left(3-(5-c)r+(6-5c)r^2\right) \\
&\qquad - p^2 r(1-c)\left[9-(19-c)r + (13-8c)r^2\right] \\
&\qquad + p^3\left[-1+(5-3c)r-\left(11-9c-c^2\right)r^2+2\left(3-4c+c^2\right)r^3\right] \\
&\qquad - p^4 r(1-r)\left[1-(1-c)r\right] \\
C &= p-1 > 0 \\
D &= \left[p^2 r + p(1+2cr-3r)+2r(1-c)\right]^4 > 0 \; .
\end{aligned}$$

The sign of $\frac{\partial Q_2}{\partial c}$ is therefore the sign of term B above. We can note immediately that B is a quartic in *p*. The constant term of this quartic is clearly negative, as is the coefficient of p^4. A numerical analysis of the coefficient of p^3 shows that it is also always negative in the permissible range of *r* and *c* (i.e. *r* and *c* less than 1). A similar analysis demonstrates that the coefficient of p^2 is negative over most of the $\{r, c\}$ range, but for high enough values of *r* and *c* together, this coefficient can be positive. And finally, the coefficient of *p* is always positive.

Together, these imply that for large enough *p*, $\frac{\partial Q_2}{\partial c}$ is always negative and hence $\frac{\partial SE_2}{\partial c}$ is also negative. The question is whether this is true for all $p \geq 2$. To verify that this is so, note that a *sufficient* condition for $\frac{\partial Q_2}{\partial c}$ to be negative would be for it to be negative when $p=1$ (so that we have the weakest case possible for the negative terms in p^3 and p^4 to swamp the [sometimes] positive terms in the lower-order terms in *p*). If we substitute $p=1$ into the "B" term above, we get:

$$B\Big|_{p=1} = (-1+r) < 0 \; .$$

Hence, even in the weakest case, it is clear that B<0. We can reinforce this insight by showing the values of "B" for *p* equal to integers from 2 through 5, which we do below; all are unambiguously negative in the feasible ranges of *r* and *c*, and the increasing negativity is obvious from the series.

$$B\big|_{p=2} = -8\left[1-(cr)^2\right]$$

$$B\big|_{p=3} = -27-9r(1+2c)+6r^2\left(1+3c+5c^2\right)-4r^3\left(1+c-2c^2\right)$$

$$B\big|_{p=4} = 4\left[-16-2r(7+9c)+6r^2\left(3+2c+3c^2\right)-9r^3\left(1-c^2\right)\right]$$

$$B\big|_{p=5} = -125-15r(12+12c)+20r^2\left(15+3c+7c^2\right)-48r^3\left(3-c-2c^2\right)\ .$$

Thus, as we sought to show, indeed $\frac{\partial SE_2}{\partial c} < 0$.

Q.E.D.

<u>**Proposition 4.**</u> **Heterogeneity due to multiple underlying factors also adversely (negatively) affects alpha and adversely (positively) affects the standard error of alpha, holding *p* and *r* constant:**

$$\frac{\partial \alpha_3}{\partial c} > 0, \quad and \quad \frac{\partial SE_3}{\partial c} < 0 \quad .$$

Proof: The expression for α_3 is given in equation (10) in the paper, and that for Q_3 is given in equation (9). The relevant derivatives are given by:

$$\frac{\partial \alpha_3}{\partial c} = \frac{2p^2 r}{(p-1)\left[2+(p+cp-2)r\right]^2} > 0\text{; and}$$

$$\frac{\partial SE_3}{\partial c} = \frac{E \cdot F}{G \cdot H}\ ,\ where$$

$$E = 8p^2 r > 0$$
$$F = -2r(1-c)(1-r) - p\left[2-r(1+c)+r^2(1-c)\right] < 0$$
$$G = (p-1)^2 > 0$$
$$H = \left[2+r(p+cp-2)\right]^3 > 0\ .$$

Thus, $\frac{\partial SE_3}{\partial c}$ is unambiguously negative, as the Proposition claims.

Q.E.D.

Derivation of c_1^* in equation (11):

c_1^* is defined as the value of c that just preserves the value of alpha when moving from the baseline case (with p items) to the "one bad item" case (with (p+1) items). The baseline alpha is given in equation (4), while the "one bad item" alpha is given in equation (8). However, we need to equate the expression in (4) not just to the expression in (8), but to the expression in (8) *where p becomes (p+1)*, to control for the fact that the researcher makes this change by literally adding another item to the scale. Thus, the relevant calculation is to solve the equality below for c to discover the value of c that keeps these two alphas equal:

$$\alpha_1(p) = \alpha_2(p+1) \Rightarrow$$
$$\frac{pr}{1+r(p-1)} = \frac{p+1}{p}\left[1 - \frac{p+1}{p+1+2pcr+p(p-1)r}\right].$$

Solving this equality for c directly yields equation (11):

$$c_1^* = \frac{1+p+r(p-1)}{2(1+p-r)} = \frac{1}{2} + \frac{rp}{2(1+p-r)}\;;\quad \frac{1}{2} < c_1^* \leq 1.$$

Demonstration that c_1^* is increasing in *p* and *r*:

The derivatives of c_1^* with respect to p and r are, respectively:

$$\frac{\partial c_1^*}{\partial p} = \frac{r(1-r)}{2(1+p-r)^2} > 0$$
$$\frac{\partial c_1^*}{\partial r} = \frac{p(1+p)}{2(1+p-r)^2} > 0\quad .$$

Demonstration that the 95% confidence interval around alpha narrows (becomes more precise) when adding "one bad item" that either keeps alpha constant or increases alpha:

The standard error of alpha in the baseline case is given by equation (5), while that in the "one bad item" case is given by $\sqrt{\frac{Q_2}{n}}$, where Q_2 is given by equation (7). However, as in the demonstration immediately above, here too we have to adjust the "one bad item" standard error to account for the fact that (p+1) items are included in that case, not p items. We therefore make this substitution and then ask the following question: if c takes on its alpha-preserving value of c_1^*, then if the researcher does indeed add this "one bad item," does the standard error of alpha shrink, relative to the baseline case?

The confidence interval in the baseline case is given by:

$$CI_{base} = 1.96\sqrt{\frac{Q_1}{n}} = 2.77186\sqrt{\frac{p(1-r)^2}{n(p-1)[1+r(p-1)]^2}} \ ,$$

while the confidence interval in the "one bad item" case (with p converted to $(p+1)$) is given by:

$$CI_{1bad} = 1.96\sqrt{\frac{Q_2}{n}} = 2.77186\sqrt{\frac{(1+p)^2 J}{np[K]^3}} \ , \quad where$$

$$\begin{aligned} J \ = \ & -4(1-c)^2(1+c)r^3 + 2(1-c)r\left[3-(5+c)r+\left(4+c-2c^2\right)r^2\right](1+p) \\ & +(1-r)\left[1-2(2-c)r+\left(5-2c-2c^2\right)r^2\right](1+p)^2 \\ & +r(1-r)^2(1+p)^3 \end{aligned}$$

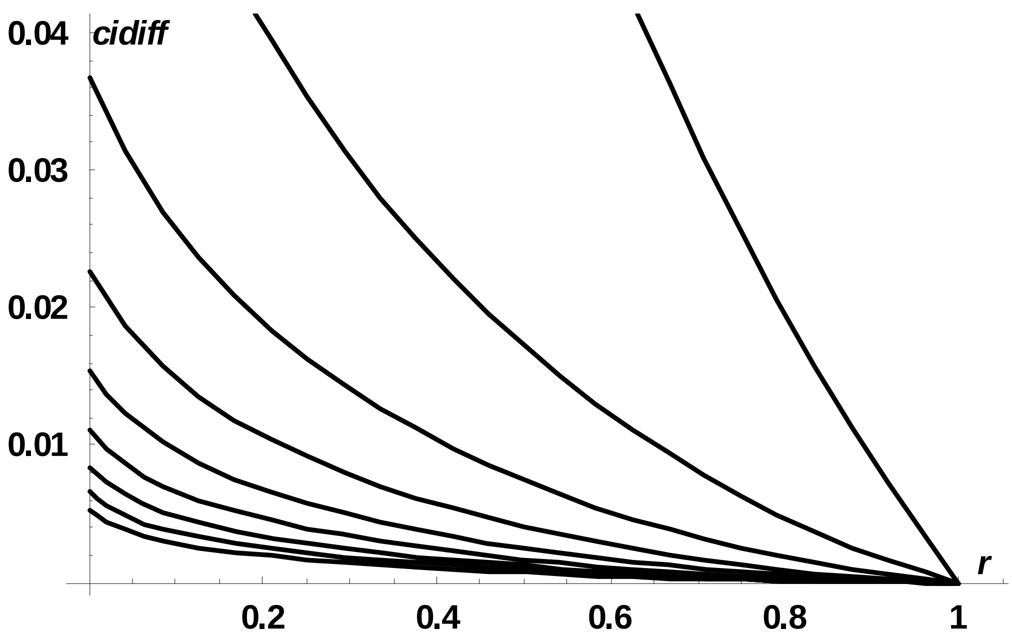
$$K = \left[1+2r-2cr+p-r(3-2c)(1+p)+r(1+p)^2\right]^3 .$$

We next take the difference $[CI_{base} - CI_{1bad}]$; factor out common terms (such as the "2.77186" in the numerator and the "n" in the denominator); and substitute in c_1^* for c. If this difference is then *positive*, it means that adding the one bad item in fact decreases the size of the 95% confidence interval for alpha, when c is such that this addition just preserves the value of alpha. After tedious calculation, we find that if the following expression is positive, the 95% confidence interval for alpha shrinks when adding an alpha-preserving item to the scale:

$$CIdiffsign = \sqrt{\frac{p(1-r)^2}{(p-1)[1+r(p-1)]^2}} - \sqrt{\frac{(1-r)^2\left[2(1+p)^3+2r(p-1)(p+1)^3+r^2\left(p-2+2p^2-p^3\right)+r^3\left(2-3p+p^3\right)\right]}{2p(1+p)^2[1+r(p-1)]^3}}.$$

While this expression is clearly not analytically very tractable, it is in fact easy to numerically verify that *CIdiffsign* is positive for all feasible *r* values and for integer *p* values throughout the whole likely range. Specifically, see the graph below, which plots this expression for *p* taking on integer values from 2 to 10 and *r* ranging from 0 to 1 (the lowest curve in the graph is for *p*=10, and the highest curve is for *p*=2).

Figure TA-1: Demonstration of Sign of Difference Between the Confidence Interval for the Baseline Case and the Confidence Interval for the "One Bad Item" Case



Notes:

The curves plotted here are of the relevant terms of the confidence intervals for these two cases that determine whether the difference between the confidence intervals is positive or negative. They do not represent the confidence interval itself (which, among other things, depends on the sample size, *n* – a factor that does not affect the positivity or negativity of these curves, and which is therefore omitted). Thus, the important insight to draw from this graph is the *sign*, rather than the value, of the points on the curves. Positivity of the points on the curves indicates that the confidence interval for the "one bad item" case is smaller (hence, the estimate of alpha is more precise) than the confidence interval for the baseline case. The curves represent integer values of *p* from 2 to 10. The highest curve is for *p*=2, and the loweste curve is for *p*=10.

As *r* approaches 1 (its upper bound), the difference between the confidence intervals approaches zero mathematically; as *r* approaches zero, the difference approaches a strictly positive number:

$\frac{\sqrt{p^2}-\sqrt{p^2-1}}{\sqrt{p(p-1)}}$. As *p* approaches infinity, the difference between the confidence intervals mathematically approaches zero; but as the graph above indicates, finite values of *p* imply that the confidence interval for the baseline case is strictly greater in size than for the "one bad item" case (and therefore, adding one bad item that just preserves the value of alpha also makes that estimate more precise). By extension, adding "one bad item" that actually increases the alpha estimate will also tighten the 95% confidence interval.

The remainder of these results supplement the article:

Derivation of c_2^*:

c_2^* is defined as the value of *c* that just preserves the value of alpha when moving from the baseline case (with *p* items) to the "two underlying factors" case (with (*p*+2) items). The baseline alpha is given in equation (5), while the "two underlying factors" alpha is given in equation (10). However, we need to equate the expression in (5) not just to the expression in (10), but to the expression in (10) *where p becomes (p+2)*, to control for the fact that the researcher makes this change by literally adding two more items to the scale. Thus, the relevant calculation is to solve the equality below for *c* to discover the value of *c* that keeps these two alphas equal:

$$\alpha_1(p)=\alpha_3(p+2) \Rightarrow$$
$$\frac{pr}{1+r(p-1)}=\frac{r(p+2)[p+c(2+p)]}{(p+1)[2+pr+cr(p+2)]} \quad .$$

Solving this equality for *c* directly yields:

$$c_2^*=\frac{p(2r+p)}{(2+p-2r)(2+p)} \quad ; \quad c_2^*\leq 1 \, .$$

Demonstration that c_2^* is increasing in *p* and *r*:

The derivatives of c_2^* with respect to *p* and *r* are, respectively:

$$\frac{\partial c_2^*}{\partial p}=\frac{4(1-r)(2p+p^2+2r)}{(2+p)^2(2+p-2r)^2}>0$$
$$\frac{\partial c_1^*}{\partial r}=\frac{4p(1+p)}{(2+p)(2+p-2r)^2}>0 \quad .$$

Calculations Showing When Moving From the Baseline Case to the "Two Underlying Factors" Case Decreases the Size of the 95% Confidence Interval:

Moving from the baseline case to the "two underlying factors" case may or may not improve the precision of the alpha estimate. To establish this, we first subtract the expression for the 95% confidence interval in the "two underlying factors" case from that in the baseline case, in the same fashion as we did above when comparing the baseline case to the "one bad item" case, assuming that $c = c_2^*$.

The confidence interval in the baseline case is given by:

$$CI_{base} = 1.96\sqrt{\frac{Q_1}{n}} = 2.77186\sqrt{\frac{p(1-r)^2}{n(p-1)[1+r(p-1)]^2}} \ ,$$

while the confidence interval in the "two underlying factors" case (with p converted to (p+2)) is given by:

$$CI_{2\,factor} = 3.92\sqrt{\frac{(p+2)\left\{2(1-cr)^2 + p\left[2-2r(1+c)+r^2\left(1+c^2\right)\right]\right\}}{n(p+1)^2\left[2+pr+cr(p+2)\right]^2}} \ .$$

Now suppose that $c = c_2^*$; under what conditions then does adding two items to a baseline scale to create a scale with two underlying factors nevertheless tighten up the 95% confidence interval? To answer this, we substitute in c_2^* for c above and set the two confidence interval expressions equal to each other. After accounting for common terms, a sufficient condition for the 95% confidence interval to shrink in this situation is that the expression *CIdiffsign2* below be positive:

$$CIdiffsign2 = \sqrt{\frac{p}{p-1}} - \sqrt{\frac{8+4p\left(3+r^2\right)+6p^2+p^3}{(p+1)(p+2)^2}} \ .$$

This can be straightforwardly solved for r to show that, when $c = c_2^*$, adding the two items shrinks the 95% confidence interval as long as:

$$r < \frac{(2+p)}{\sqrt{2p(p-1)}} \ .$$

The Table below reports values of r (given values of p) above which the 95% confidence interval around alpha actually gets larger (i.e., for which the estimate of alpha is less precise) when the researcher adds two items to a baseline scale to create a "two underlying factors" scale:

Table TA-1: Values of *r* Above Which Creating a Second Underlying Factor Increases the 95% Confidence Interval Around Alpha

Value of *p*	Value of *r* above which adding 2 items (creating a 2nd underlying factor) enlarges rather than shrinks the 95% confidence interval around alpha
2 through 6	No values of *r* in the feasible range ($0 < r \le 1$)
7	$r > 0.982$
8	$r > 0.945$
9	$r > 0.917$
10	$r > 0.894$
11	$r > 0.876$
12	$r > 0.862$
13	$r > 0.849$
14	$r > 0.839$
15	$r > 0.830$
16	$r > 0.822$
17	$r > 0.815$
18	$r > 0.808$
19	$r > 0.803$
20	$r > 0.798$

Specifically, for p=2, r would have to be greater than 2 in value for the 95% confidence interval to increase when making this change – an infeasible r value. The corresponding critical values of r for other p values up to 6 are: r=1.443 (for p=3); r=1.225 (for p=4); r=1.107 (for p=5); and r=1.033 (for p=6).

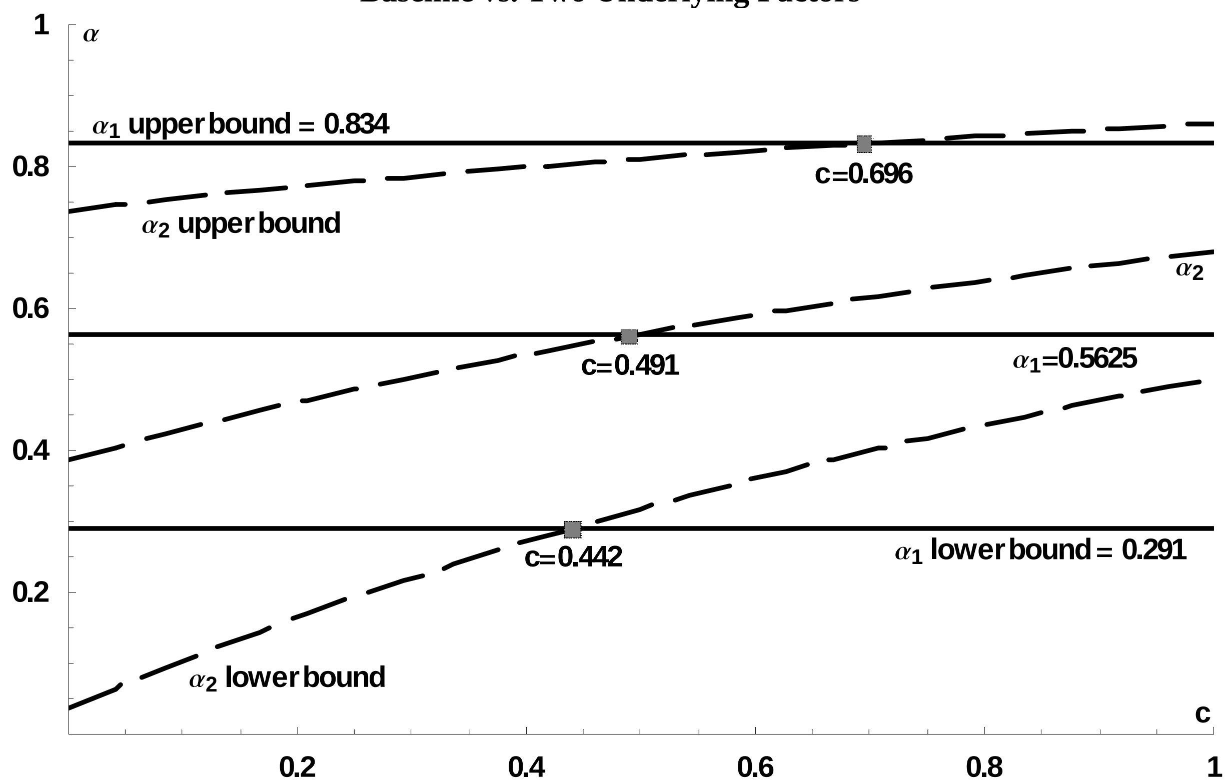
For the intuition regarding this finding, consider that the two incremental items have a correlation with the other p items of only $c \cdot r$, not r. The greater the number of items in the baseline scale, the greater will be the number of new, lower inter-correlations in the expanded correlation matrix – and hence, the more extreme will be the lowering of the average item inter-correlation. Further, the incremental value of adding another pair of items to a scale has already been shown to diminish with initial scale length; eventually, the diminution in average item inter-correlation simply swamps the beneficial effect of adding the extra two items.
As the Table above shows, the range over which this occurs is not very large, nor is it likely that a researcher is considering adding two more items to a scale that already has 7 or more items. In this spirit, we conclude that "most of the time," adding two items to a scale and thereby creating a second underlying factor indeed shrinks the 95% confidence interval around alpha, providing a more precise estimate (with the caveat that it is possible to reverse this general conclusion for high enough baseline values of p and r).

For comparison purposes with the baseline-to-one-bad-item illustration in the text of the paper, we next present here an example showing what can happen when the change is from the baseline case to the "two underlying factors" case. Consider once again the example of p=3 and r=0.3, with a baseline alpha value of 0.563. Comparing that baseline case with the two-underlying-factors case, the critical value of c is 0.491. With this baseline scenario and assuming only a sample size of n=30, we have that:

- For $0 < c < 0.442$, the lower and upper bounds of the confidence interval around alpha with two underlying factors (p=5), and alpha itself, are *lower* than with the baseline alpha (p=3);
- For $0.442 < c < 0.491$, the upper bound with two underlying factors (and alpha) are *lower* than in the baseline, but the lower bound is now *larger* in value than in the baseline;
- For $0.491 < c < 0.696$, the upper bound with two underlying factors is *lower* than in the baseline, but the lower bound (and alpha) are now *larger* in value than in the baseline case;
- For $0.696 < c \leq 1$, the upper and lower bounds, and alpha itself, are all *larger* in value in the case of two underlying factors than in the baseline case.

In short, for $0.442 < c < 0.491$, the alpha estimate *falls*, but causes the lower bound of the 95% confidence interval to *rise* (that is, to improve), a result attributable to the beneficial effect of adding enhancing p. Clearly, however, a "bad enough" pair of items (one for which $c < 0.442$) will cause both alpha to decrease, and the lower bound of the 95% confidence interval to decrease. We show this graphically in Figure TA-2:

FIGURE TA-2: Values of Alpha and 95% Confidence Interval Upper and Lower Bounds: Baseline vs. Two Underlying Factors



Notes: All curves in this graph arise from the following parameter values: p=3, r=0.3, n=30. In this situation, the baseline value of alpha (denoted as "α_1" in the graph above) is 0.5625, and the 95% confidence interval is bounded (below) by 0.291 and (above) by 0.834 (denoted as "α_1 lower bound" and "α_1 upper bound," respectively). If we add two items that create a second underlying factor, whose correlation with the other items in the scale is $c \cdot r$, producing a 5-item scale, the resulting alpha value and boundaries of the 95% confidence interval are represented by dotted lines, and are labeled "α_3," "α_3 lower bound," and "α_3 upper bound." The lower bound for the baseline α intersects the lower bound for the "two underlying factors" α at

c = 0.442. The α estimate for the baseline α intersects that for "two underlying factors" at c = 0.491. The upper bound for the baseline α intersects the upper bound for the "two underlying factors" α at c = 0.696.

Derivation of c_3^*:

c_3^* is defined as the value of *c* that just preserves the value of alpha when moving from the "one bad item" case (with (*p*+1) items) to the "two underlying factors" case (with (*p*+2) items). The "one bad item" alpha is given in equation (8), while the "two underlying factors" alpha is given in equation (10). However, we need to adjust both alpha expressions to reflect the fact that the number of items is (*p*+1) in the "one bad item" case, and (*p*+2) in the "two underlying factors" case. Thus, the relevant calculation is to solve the equality below for *c* to discover the value of *c* that keeps these two alphas equal:

$$\alpha_2(p+1) = \alpha_3(p+2) \Rightarrow$$
$$\frac{p+1}{p}\left[1 - \frac{p+1}{p+1+2pcr+p(p-1)r}\right] = \frac{r(p+2)[p+c(2+p)]}{(p+1)[2+pr+cr(p+2)]} \ .$$

After algebraic manipulation, the two alpha expressions are equal if the expression *Z* below is zero (and the "two underlying factors" alpha is larger than the "one bad item" alpha if *Z* is positive):

$$Z = -2rc^2(p+2) + c\left[p^2(p+1) - r\left(p^2+3p-2\right)\right] + \left[2 + p(4+r) + p^2(1-r) - p^3\right]$$

Z is a quadratic in *c*. This quadratic approaches negative infinity as *c* approaches either negative infinity or positive infinity. Further, *Z* is clearly positive when *c*=1:

$$Z\big|_{c=1} = 2(1-r)(p+1)^2 \ .$$

Thus, the function *Z does* have a positive range, and this range does include at least one feasible *c* value. But *Z* is not always positive for all *c* values. Consider *c*=0, the other limit of the range of *c*. We have:

$$Z\big|_{c=0} = 2 + p(4+r) + p^2(1-r) - p^3 \ .$$

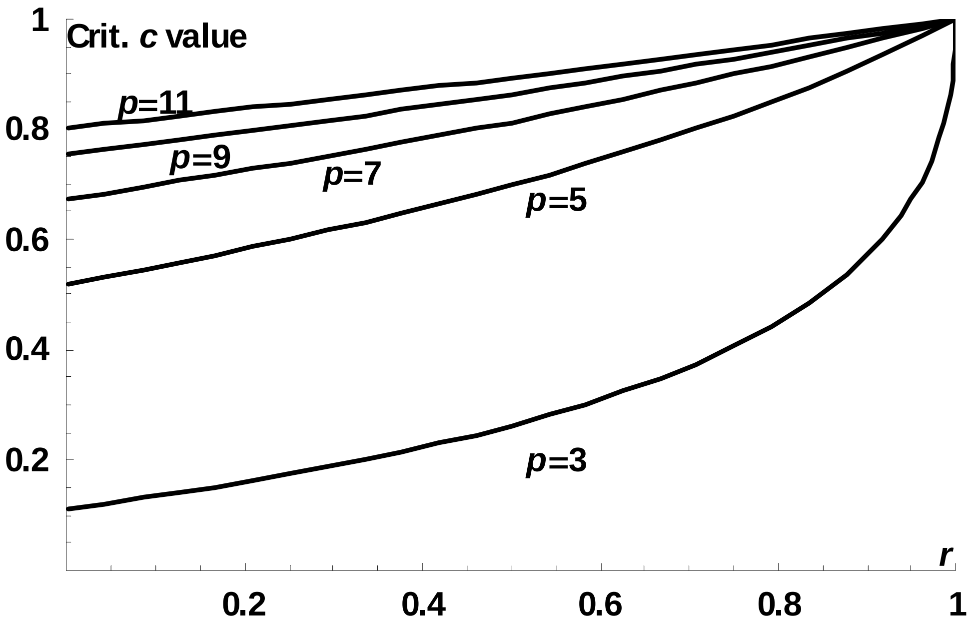
Note that when *c*=0, *Z* is positive for *p*=2 (it equals (6-2*r*)), but is negative for *p*=3 (it then equals (-4-6*r*)). This suggests that when *c* is low, the alpha of an already-long scale is not enhanced by the addition of one more item that creates a second underlying factor.

We can solve for the two *c* roots of the *Z* expression above. One of the roots is always greater than 1 in value, and this (larger) root is the critical value of *c* above which *Z* is always negative. But the smaller *c* root is sometimes between 0 and 1 in value, and it is this smaller *c* root that is given by c_3^*:

$$c_3^* = \frac{-r\left(p^2+3p-2\right)+(1+p)\left[p^2-\sqrt{r^2(p-2)^2+p^4-2r\left(p^3+6p^2-8p-16\right)}\right]}{4r(2+p)}.$$

We show these critical values of *c*, as functions of *r*, for several salient values of *p* in Figure TA-3 below. Note that c_3^* is increasing in *r* for any value of (*p*+1) greater than 3, and is similarly increasing in *p* for any value of *r* in these parameter ranges. Thus, it exhibits the same comparative-static properties as did c_1^* and c_2^*. Further, note that when the researcher starts with 3 items in the "one bad item" scale [i.e., (*p*+1)=3], it is always alpha-improving to add the incremental item (we therefore omit this case from the Figure). However, for any higher value of *p*, there exist {*r*, *c*} pairs for which it does not increase alpha to add the incremental item. As *p* increases, the minimum hurdle that *c* must surpass in order to increase alpha increases, just as it did in the first two illustrations. In the limit, as *p* approaches infinity, c_3^* approaches 1. Thus in the limit, it is impossible to add one more item to a scale that increases the estimate of alpha.

FIGURE TA-3: Critical Values of *c* Above Which Alpha Increases When Moving from "One Bad Item" to "Two Underlying Factors" Cases



Notes: When p=2, alpha increases when moving from "one bad item" (p=3) to "two underlying factors" (p=4) case, for all feasible values of c and r ($c \in (0,1]$, $r \in (0,1]$). Each curve represents the value of c, for any value of r, that just holds alpha constant when moving from the "one bad item" case to the "two underlying factors" case, for each value of p.

Demonstrating that Moving From the "One Bad Item" Case to the "Two Underlying Factors" Case Does Not Always Improve the 95% Confidence Interval, Even When Alpha Itself Remains Constant:

Adding an incremental item in this third illustration does not always tighten the confidence interval around alpha, even when the estimate of alpha itself remains constant. A sufficiently low value of r (given p) guarantees that adding the incremental item decreases the variance around alpha, however. It is in these relatively low-r cases that adding another item, even one that creates a second underlying factor, has a chance of improving the precision of the estimate of alpha. The formula governing this relationship is not simple, so we present Table TA-2 here to illustrate the critical values of r for each value of p, under the assumption that the value of c is such as to keep the estimate of alpha itself constant:

TABLE TA-2: Values of *r* and *c* That Preserve Both Alpha and Its 95% Confidence Interval, Moving from "One Bad Item" Case to "Two Underlying Factors" Case, For Various *p* Values

Value of *p*	Value of *r* that preserves the size of 95% confidence interval (for *c* = alpha-preserving *c* value)	Value of *c* that preserves alpha	Implied (Constant) Alpha Value
2	0.5	N/A (any $c \in (0, 1]$ *increases* alpha)	N/A
3	0.523810	0.272727	0.666667
4	0.590909	0.615385	0.833333
5	0.650000	0.769231	0.900000
6	0.696970	0.847826	0.933333
7	0.733990	0.892617	0.952381
8	0.763514	0.920354	0.964286
9	0.787440	0.938650	0.972222
10	0.807143	0.951327	0.977778
11	0.823609	0.960461	0.981818
12	0.837553	0.967254	0.984848
13	0.849498	0.972441	0.987179
100	0.980008	0.999592	0.999798

To illustrate, imagine p=3; i.e., the researcher is considering switching from a "one bad item" case with 4 items in it [(p+1) items] to a "two underlying factors" case with 5 items in it [(p+2) items]. In order for adding that incremental (p+2) item to make the estimate of alpha more precise, given a c value of 0.272727 (with this change preserving alpha), the underlying correlation among the first three items (*r*) must be less than 0.52381. If instead *r* were, say, 0.6, then adding the extra item (assuming *c*=0.272727) would cause the alpha estimate to remain constant, but be *less* precise. The table shows that the greater the number of underlying items, the higher the hurdle. We illustrate this graphically in Figure TA-4 below, for the case of *p*=3.

FIGURE TA-4: From One Bad Item to Two Underlying Factors (*p*=3): {*r, c*} Values' Effect on Alpha and Its 95% Confidence Interval

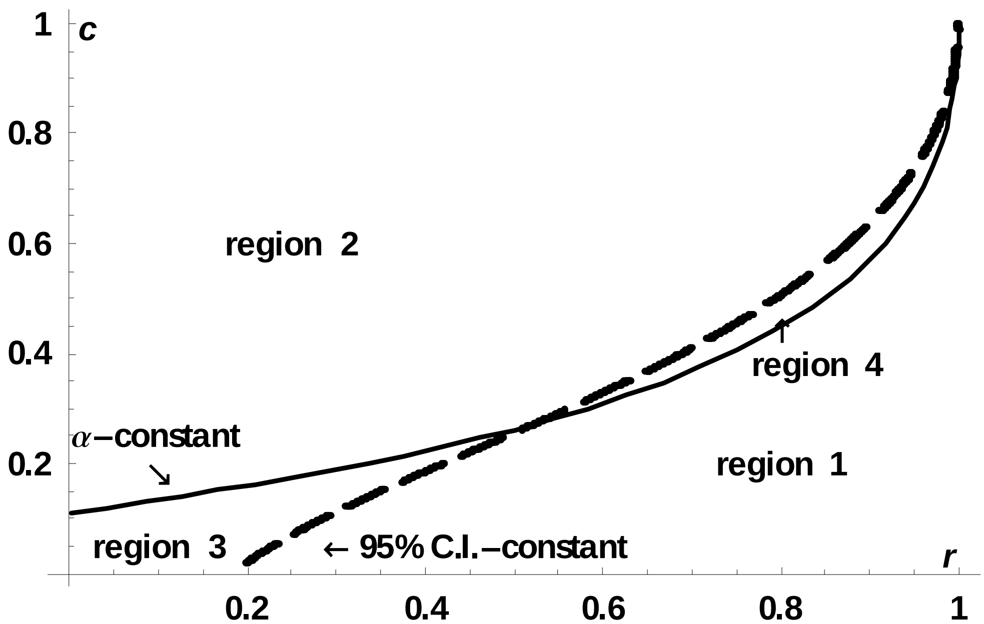


Figure TA-4 (for *p*=3) shows two curves. The solid curve is the set of {*r, c*} values for which the estimate of alpha remains constant when *p*=3 and the researcher adds an item that moves the scale from the "one bad item" case to the "two underlying factors" case (reproduced from Figure TA-3). The dashed curve presents the {*r, c*} values for which the confidence interval is constant and the researcher is performing the same transformation on the scale. The two curves cross at {*r*=0.52381, *c*=0. 272727}, as the table above reports. Points on the "α-constant" curve to the right of this intersection are {*r, c*} pairs where adding the extra item preserves the value of alpha, but decreases its precision (i.e., increases the size of the confidence interval).

In "region 1" (where *r* is high but *c* is low), adding the extra item causes both the value of alpha *and* its precision to fall. Conversely, in "region 2" (where *r* is low relative to *c*), adding the extra item causes both the value of alpha *and* its precision to increase. These insights are intuitively plausible: when *r* is low relative to *c*, adding the extra item that creates two underlying factors has a strong positive impact on the average *r* in the scale, which counteracts the negative impact of creating a second factor within one scale. But when *r* is high relative to *c*, the effect is reversed. In either of these regions, it is fairly clear what the researcher's best action is—that is, add the extra item if in region 2; don't add the extra item if in region 1—given that both alpha and its precision move in the same direction.

But in “region 3” and “region 4,” the effect on alpha and its precision are not of the same directionality. In “region 3” (low *r* and low *c*), adding the extra item lowers alpha, but makes the estimate more precise. In “region 4” (high *r* and intermediate *c*), adding the item increases alpha, but makes the estimate less precise. Thus, the impact of adding an item to a scale is not fully revealed by merely assessing its effect on the estimate of alpha itself. The decision made by the researcher of course also rests upon the actual value of alpha derived in these cases, an issue we consider in the next section.

In sum, in this illustration we consider the effect of moving from (p+1) to (p+2) items in a scale. In the (p+1)-item case, the scale contains one bad item. In the (p+2)-item case, the scale is made up of two underlying factors. The question considered is when and whether adding the (p+2) item to the scale improves the reliability of the scale. We show here that doing so *can* improve the estimate of alpha, if the underlying correlation among the first *p* items is low enough relative to the *c* factor characterizing the last two items added. Intuitively, the last item must not drag down the average correlation too much; but on the other hand, lengthening a scale always has a positive impact of its own. It is the counterbalancing of these effects that determines the net effect on alpha. However, in this illustration we also show that an increase in alpha is not necessarily accompanied by an increase in its precision. When the underlying correlation among the original *p* items is high, adding the incremental item may increase alpha but also increase the size of the confidence interval – that is, decrease its precision. In a situation like this, the researcher has to decide whether adding the extra item is “worth it,” which is a choice between the improvement in the mean estimate of alpha and the possibility (revealed through the confidence interval) that alpha could be lower.

<u>Demonstration of Result 1:</u> **When the researcher’s goal is to maximize the size of alpha and the initial scale has 3 items, the equilibrium number of items is:**

- **the baseline 3 items, if the {*r*, *c*} pair lies in either “region 1” or “region 2” of Figure 2;**
- **the “two underlying factors” 5-item scale if the {*r*, *c*} pair lies in either “region 3” or “region 4” of Figure 2.**

In “region 1” of Figure 2, with very low *c* values and a range of low to high *r* values, the researcher is best off staying with the baseline set of 3 items in the scale. Adding one “bad” item does not improve alpha (the points in region 1 lie below the top curve); adding two items to move from the baseline case to the “two underlying factors” case also does not improve the value of alpha (the points in region 1 lie below the middle curve). And, if the researcher were in the “one bad item” case, it would not improve alpha to add another item that creates a “two underlying factors” case; conversely, it would improve alpha to drop the one bad item. The unique equilibrium number of items in this situation is therefore three in region 1.

In “region 2” of Figure 2, if the researcher already has “one bad item” in a 4-item scale, alpha will improve with the fifth item that transforms the scale into a “two underlying factors” scale. However, it is not an equilibrium to have these five items in the scale, because these points lie below the “Illust. 2” curve in Figure 2. Alpha is higher for the baseline case of 3 items than it is for the 4-item scale with “one bad item” as well (since the points in region 2 lie below the “Illust.

1" curve in Figure 2). Hence, in "region 2" we also find that the alpha-maximizing scale length is three.

In "region 3" of Figure 2, moving from the baseline case to "one bad item" decreases alpha, but moving from the baseline case to "two underlying factors" *increases* alpha. Further, if the researcher starts with the "one bad item" case, alpha improves with the addition of one more item that transforms the scale into a "two underlying items" scale. Thus, the unique equilibrium number of items in "region 3" is five, with two underlying factors.

Finally, in "region 4" of Figure 2, we have $\{r, c\}$ pairs for which moving from the baseline case to the "one bad item" case improves alpha; moving from the baseline case to the "two underlying factors" case improves alpha; and moving from the "one bad item" case to the "two underlying factors" case also improves alpha. The unique equilibrium in this portion of the space is therefore also to have a 5-item scale, with two underlying factors.

In sum, for the case of p=3, we see that the equilibrium number of factors is either the baseline 3 items (if the $\{r, c\}$ pair lies in either "region 1" or "region 2"), or the "two underlying factors" 5-item scale (if the $\{r, c\}$ pair lies in either "region 3" or "region 4"). It is never an equilibrium in this situation to add just one bad item to the baseline scale, given the opportunity to add another item that creates a second underlying factor. The intuition for the results rests on the relative size of c and r, given the low number of original items in the scale (p=3). When c is high enough relative to r, the beneficial effect of adding items swamps the negative effect of creating a scale with two underlying factors. But if c is too low relative to r, the researcher is best off sticking with the original 3-item scale. More generally, the implication is that when starting with a small number of items, adding as many items as possible is likely to be a good strategy for increasing the estimate of alpha.

<u>Discussion of the Equilibrium Number of Scale Items when the Baseline is *p*=5:</u>

The insights shown above for the p=3 case are altered when we increase the number of items in the baseline scale to 5, as shown in Figures TA-5 and TA-6 below:

FIGURE TA-5: $\{r, c\}$ Values That Keep Alpha Constant for Illustrations 1, 2, and 3 (p=5)

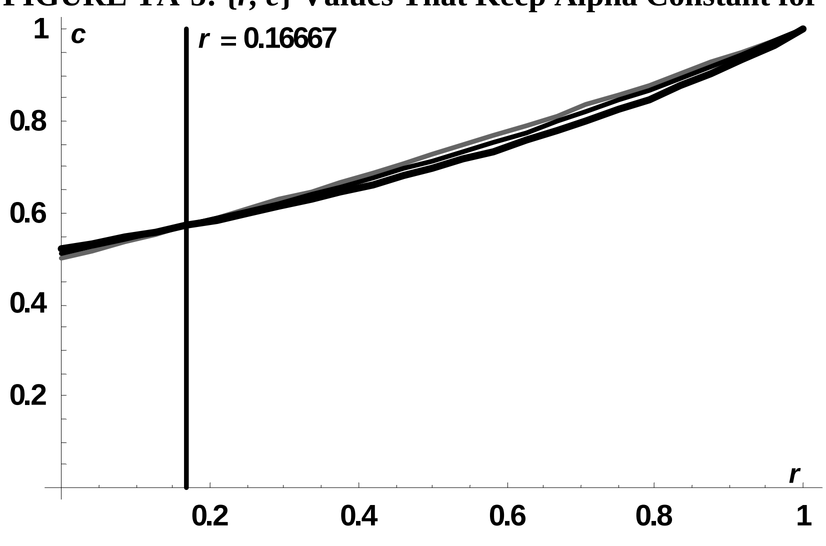
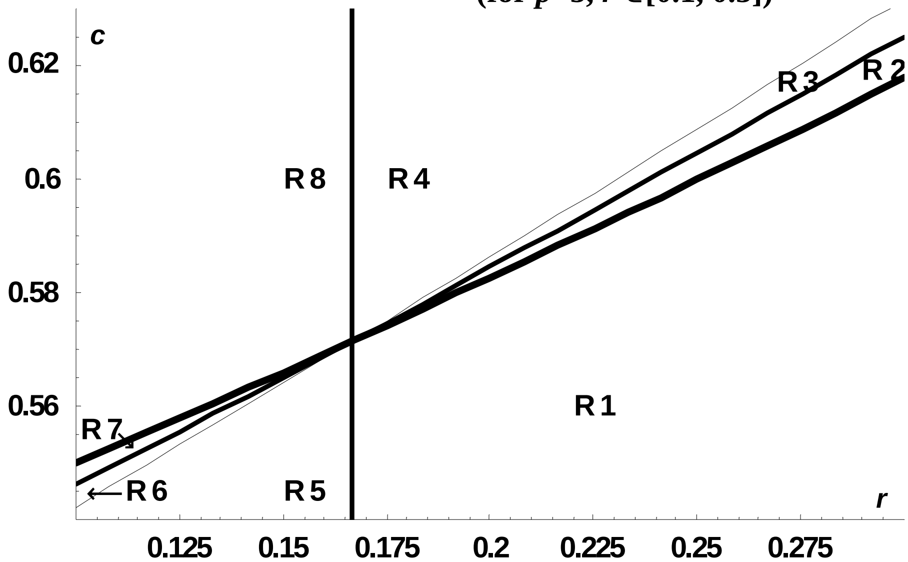


FIGURE TA-6: {*r, c*} Values That Keep Alpha Constant for Illustrations 1, 2, and 3 (for *p*=5, *r* ∈[0.1, 0.3])



Notes to Figures TA-5 and TA-6: The lightest curve is the locus of {*r, c*} pairs that keep alpha constant when moving from baseline case to "one bad item" case ("Illustration 1"). The medium-weight curve is the locus of {*r, c*} pairs that keep alpha constant when moving from baseline case to "two underlying factors" case ("Illustration 2"). The heaviest curve is the locus of {*r, c*} pairs that keep alpha constant when moving from "one bad item" case to "one bad item" case ("Illustration 3"). {*r, c*} pairs above a given curve imply that alpha *increases* with the indicated change. {*r, c*} pairs below a given curve imply that alpha *decreases* with the indicated change. "R1," "R2," etc. refer to "region 1," "region 2," etc. as in Figures 2 and 3. All three curves cross at {*r* = 0.16667, *c* = 0.571429}.

Figure TA-6 presents the same data as Figure TA-5, but for $r \in [.1, .3]$ in order to clearly depict the cross-over points of the curves. We indicate eight regions of the space in Figure TA-6 (these regions also exist in Figure TA-5, but we label them here for clarity), four of which lie to the left of the point of intersection of the three curves, and four of which lies to the right of the point of intersection. The common intersection point of all three curves is at {r=0.167, c=0.571}, and corresponds to an alpha value of exactly 0.5.

Regions 1 through 4 have the same characterization as regions 1 through 4 in Figure 2, since in this portion of the space (r>0.167), the curves are arranged identically to those in Figure 2. Thus, we conclude that in regions 1 and 2, the equilibrium number of items in the scale is five, the baseline number; while in regions 3 and 4, the equilibrium number of items is the "two underlying factors" total of 7 items.

But interestingly, there is a complete reversal in the order of the curves for r<0.167. In region 5, we have {r, c} pairs for which no addition of extra items ever increases alpha; and in region 8, we have {r, c} pairs for which adding two items (ending up with the "two underlying factors" case and 7 items in the scale) is the equilibrium. These two regions (characterized by the lowest and highest values of c) have the same implications for equilibrium number of items in the scale as in regions 1 and 4 (respectively) of Figure TA-6. It is in regions 6 and 7 that we see a difference. In region 6, alpha increases when the researcher adds “one bad item” to a baseline 5-item scale, but alpha does *not* increase if a seventh item is added, creating “two underlying factors” – either from the baseline starting point or from the “one bad item” starting point. Thus, in region 6, the equilibrium action is to add “one bad item,” but to stop there and not to add any further items.

Meanwhile, in region 7, alpha increases when moving from the baseline case to the “one bad item” case and when moving from the baseline case to the “two underlying factors” case. One might be tempted to conclude that increasing the length of the scale from 5 to 7 items (creating the “two underlying factors” scenario) would be optimal. But these $\{r, c\}$ pairs lie below the top curve, and hence moving from the “one bad item” case to the “two underlying factors” case *decreases* alpha. Thus, in region 7 of Figure TA-6, the equilibrium number of items in the scale is again 6, with “one bad item.”

In sum, when p=5 and r<0.167, three equilibria exist: in region 5, the equilibrium number of items in the scale is the baseline 5; in regions 6 and 7, the equilibrium number of items is 6, with the “one bad item” case; and in region 8, the equilibrium number of items is 7, with the “two underlying factors” case. Together with the discussion of the r>0.167 range of the Figure, we again see that in general, when c is low relative to r, the researcher is best off sticking with the original baseline scale of five items; and when c is very high relative to r, the researcher maximizes alpha by making the scale as long as possible. Now, however, there is also a portion of the space where the researcher optimally stops at adding just one more item, rather than maximizing the length of the scale to maximize the value of alpha. The optimal number of items (when seeking to maximize alpha) is thus not always an “all or nothing” proposition; it is possible for an intermediate number of items to be best, even when the possibility exists to add more items to the scale.

We should add that although the portion of the space in Figure TA-6 where “one bad item” is the equilibrium appears to be small, it is not insignificant. We use these three illustrative scenarios merely to depict a range of possible situations a researcher might face. In a real research situation, the set of options open to the researcher might result in an intermediate number of items being optimal over a much larger portion of the space than we see here.