

To appear in *Journal of Occupational and Organizational Psychology* 2005

Big Five validity: Aggregation method matters

Peter Warr, Institute of Work Psychology, University of Sheffield,
Dave Bartram and Anna Brown, SHL Group plc, Thames Ditton

Abstract

Correlations between Big Five personality factors and other variables have been examined in three different ways: direct scoring of items within a factor, application of a composite-score formula, and taking the average of single-scale correlations. Those methods were shown to yield consistently different outcomes in four sets of data from sales-people and managers. Factor correlations with job performance were greatest for direct scoring, and were reduced by half when scale correlations were averaged. The insertion of previously-suggested estimates into the composite-score formula yielded intermediate correlations with performance. It is necessary to interpret summary accounts of correlations with a compound construct in the light of the aggregation method employed.

Representing multiple aspects of personality through the Big Five framework has been important in job and other settings. Drawing on previous findings, meta-analytic reviews have examined, for example, the associations of each personality factor with overall job performance. Those meta-analyses have varied in the procedure adopted to combine research findings across studies, and it is important to examine the implications of that methodological inconsistency. This paper will consider the principal methods for aggregation, and illustrate their application in two sets of analyses.

Average correlations with Big Five factor scores (for example, across several studies) have been computed in three different ways. A first approach (“Method 1”) is based on the averaging of item scores. In the case of inventories designed explicitly to provide scores on the five factors (here referred to as “FFM” measures, for “five-factor model”), a simple averaging of within-scale (i.e., within-factor) items yields a factor score in a straightforward manner. Correlations of factor scores with, say, job performance can then be computed, and an average correlation for each factor can be calculated across different studies. This procedure was used in the meta-analysis by Hurtz and Donovan (2000) to summarize findings about the Big Five correlates of overall job performance; all inventories in their review had been designed to index the Big Five factors.

Average factor scores can also be directly scored from inventories not constructed in Big Five terms (“non-FFM” measures), by designating particular scales as representative of a Big Five factor. Averaging responses to the scales allocated to a factor can yield an overall factor score, and correlations between those factor scores and a criterion can be averaged across studies. This method depends on the appropriate allocation of scales to a factor, but otherwise is the same as for FFM measures, above. For both FFM and non-FFM inventories, item-averaging may be termed “Method 1a”.

When one does not have raw data, the basic formula for Pearson correlation yields a derivation that allows the computation of a composite-score correlation. This is based on the correlations between each scale and the criterion, scale standard deviations, and between-scale correlations (e.g., Guilford & Fruchter, 1978, p. 394). A composite-score correlation

computed in that way (“Method 1b”) is identical to the directly-scored value (Method 1a) for the same set of data. The formula is presented in the Appendix.

If standard deviations for single scales within a Big Five factor are not known (for example, not having been reported), then we might assume that they are all the same, as is the case when working with standardized scores. The formula for Method 1b can be simplified, giving rise to another established formula (“Method 2”). This formula (also in the Appendix) draws on the correlations between single scales and the criterion and on the average correlation between predictor scales themselves (e.g., Hunter & Schmidt, 1990, pp. 455-463). A composite-score correlation computed using Method 2 thus produces (for a given set of data) the same result as Method 1b, if that method examines standardized scores.

Furthermore, between-predictor correlations (r_{xx}) (required for both formulae) are not usually published, so that it often becomes necessary in secondary analyses to import an average intercorrelation value from other research reports. In examinations of the Big Five factor of Conscientiousness, Mount and Barrick (1995) inserted as the average between-predictor correlation (mean r_{xx}) an estimated value of 0.47, whereas Salgado (1997, 1998) assumed an average of 0.34. Using those estimates, the average across-study (uncorrected) composite-score correlation of Conscientiousness with overall job performance was reported (from different data-sets) as 0.18 by Mount and Barrick (1995) and 0.10 by Salgado (1997).

A third form of scale-aggregation to yield Big Five correlations with a criterion (“Method 3”) merely averages the correlations between the specified scales and the criterion. (This contrasts with Method 1, which averages items rather than correlations.) Correlations between a criterion and those scales defined as components of a factor can be averaged to yield an overall factor correlation value, and within-study average factor correlations can then be averaged across studies. The procedure was used by Barrick and Mount (1991), in their examination of Big Five factors’ average correlation with overall job performance.

Inferences set out in the Appendix show that the average of the correlations between specified scales and a criterion is necessarily smaller than the estimate from the composite score correlation formula. This becomes intuitively clear if we think of the proportion of variance accounted for by each scale separately. When predictor scales have an intercorrelation of zero, the contribution of each to the prediction of a complex criterion is unique, and the proportion of variance predicted by their combination is equal to the sum of the proportions of variance predicted by each separately. Conversely, when the predictor scales are intercorrelated perfectly (all r 's = 1.00), a single “average” scale will explain all the variance in the prediction. However, correlations between predictors are usually between 0.00 and 1.00, so that each scale contributes unique variance over and above the others, enhancing the combined prediction. Even scales with relatively low validity coefficients can become important within a composite, provided their correlations with other predictors are not too high.

In summary, one may examine the across-study average correlation between Big Five personality factors and job performance (or some other variable) by combining within and across studies in one of three ways:

Method 1. By direct scoring: within each study, combine items across scales within a factor (or use a single factor score if the inventory is so designed) and correlate the combined factor score with a criterion; then compute the across-study average of those within-study factor correlations. Raw data are needed for the Method-1a application of this procedure.

Alternatively, the formula for Method 1b in the Appendix may be applied to descriptive statistics to obtain the composite-score correlation.

Method 2. By applying a formula for the correlation with a derived composite score: within each study, compute the average correlation between scales within each Big Five factor; relate that average between-scale correlation to the average scale-criterion correlation as in the formula for Method 2 in the Appendix; then compute the across-study average of those within-study composite-score correlations. When scale standard deviations are all the same (e.g., standardized), Methods 2 and 1 yield identical results.

Method 3. By averaging the criterion correlations of scales making up a factor: within each study, average for each factor the scale-criterion correlations; then compute the across-study average of those within-study average factor correlations. This procedure yields an estimate of the factor validity that is biased downwards.

Empirical Implications

It is to be expected that these different aggregation procedures will yield different associations between Big Five factors and variables such as job performance. However, direct empirical comparisons appear not to have been carried out.

Mount and Barrick (1995, pp. 174-175) compared the associations of Conscientiousness with overall job performance in respect of the composite-score formula and the average-correlation method. The composite-score procedure (Method 2, above) yielded an uncorrected validity coefficient of 0.18, whereas the average uncorrected correlation for Method 3 (from Barrick & Mount, 1991) was 0.13. However, the two datasets overlapped only partly, with 173 and 117 studies reviewed in the two analyses respectively.

The meta-analysis by Hurtz and Donovan (2000) was restricted to inventories designed explicitly to assess the Big Five factors, and thus permitted application of the direct scoring method (1, above). (However, in a minority of cases two sub-scales had to be combined into a Big Five factor, and Method 2 was then used to obtain a composite within-study correlation.) With those FFM inventories, the average uncorrected correlation of Conscientiousness with overall job performance was found to be 0.14. However, the personality measures examined were (intentionally) distinct from those in other meta-analyses.

Salgado (2003) applied Methods 1 and 3, but again in relation to different inventories. He raised the possibility that average factor validities of inventories designed explicitly to assess Big Five factors ("FFM" inventories) may be greater than the factor validity coefficients for non-FFM measures (those for which allocation of scales to a Big Five factor is made subsequently). In separate analyses of studies with the two kinds of inventory, he reported that no differences exist, relative to the criterion of rated overall performance, for three of the five personality factors, but suggested that non-FFM validity was lower than FFM validity in the case of Emotional Stability and Conscientiousness.

However, that conclusion was based on differential application of aggregation methods. For the FFM inventories, Salgado used in each study the reported validity coefficients for a factor, based directly on individuals' scale responses (referred to here as Method 1). But from published reports about non-FFM inventories it was not possible to

create an overall factor score for each individual by averaging responses to the items in scales allotted to a factor (as in Method 1), so that the criterion-related validity of factors in non-FFM inventories was computed as an average of the single scale correlations (Method 3). Salgado's comparison of the validity of FFM and non-FFM inventories was thus based on aggregation procedures that differed from each other.

Method

In order to compare correlations with Big Five factors (e.g., in FFM versus non-FFM inventories) methods of aggregation need to be the same in all cases. Similarly, in order to learn about the consequences of different methods of combination, it is essential to compare the same instruments and the same respondents. These conditions have not yet been simultaneously met in the literature, and that deficiency has prompted the analyses presented here.

We will examine average validity coefficients derived through the three methods, analyzing raw data to provide a (Method 1) bench-mark correlation against which to compare other estimates. The magnitude of correlations will be compared between the bench-mark (Method 1) and both Method 3 (mean of single-scale correlations) and Method 2 (application of a composite-score formula). Method 2a assumes that scale standard deviations are all the same, and in addition Methods 2b and 2c have employed estimated rather than actual mean scale intercorrelations. As indicated above, those estimates have varied between Big Five investigations (Mount & Barrick, 1995; Salgado, 1997, 1998), and the implications of that variation will be studied here.

Analysis 1

Three sets of data were examined in which sales performance was considered as a function of scores on the Customer Contact Sales Questionnaire (CCSQ) (SHL, 1997). Respondents were 119 car sales executives (Study 1), 78 retail sales employees (Study 2), and 90 door-to-door book salespeople (Study 3) (Warr, Bartram, & Martin, 2005). The factor of Conscientiousness has been shown to be the important Big Five characteristic in sales jobs (Vinchur, Shippmann, Switzer, & Roth, 1998), and that factor was examined here.

Table 1 about here

Table 1 presents the correlations of sales performance with each of six CCSQ scales conventionally viewed as making up the Conscientiousness factor. The lower part of the table shows overall correlations for each of the aggregation methods described above. Method 1 (direct scoring) generated an average validity coefficient for Conscientiousness of 0.22 (in the right-hand column), whereas Method 3 (mean scale correlation) indicated an average validity of 0.11, half the value from Method 1. Application of the formula to create a composite-score correlation (Method 2a) resulted in an average across-study correlation of 0.21.

Insertion of the different estimates of mean r_{xx} used by Mount and Barrick (1995) (Method 2b in Table 1) and Salgado (1997, 1998) (Method 2c) gave rise to average correlations with sales performance of 0.15 and 0.16 respectively. Composite-score correlations based on assumptions from previous analyses thus fell mid-way between the bench-mark (Method 1) value and the average-correlation figure (Method 3). In practice, the mean r_{xx} across the present three studies was 0.14; insertion of that value rather than the

previous estimates yielded the composite-score correlations shown as Method 2d. The average correlation between Conscientiousness and overall job performance was in that case 0.21, as found after application of each study's own r_{xx} (Method 2a).

Analysis 2

The generality of this pattern has been examined in a study to predict specific competencies from designated sets of personality scales. This second analysis is formally identical to Analysis 1 (a single criterion value with a composite predictor), but does not involve Big Five constructs. Table 2 summarizes findings from an investigation of 413 managers' competencies as rated by their immediate superiors. For illustrative work behaviours within the SHL Universal Competency Framework™ (Bartram, in press), correlations are shown with scales of the Occupational Personality Questionnaire (OPQ; SHL, 1993, 1999) that have been defined as relevant predictors in respect of that particular competency. (The scales, labelled as A, B, C, etc. in Table 2, thus vary between the competencies. Bartram, in press, presents details of relationships between scales and competencies.)

Table 2 about here

The bottom rows of the table contain correlations between each competency and the different forms of compound predictor. It can be seen that differences in predictor-criterion correlations between Methods 1, 2 and 3 are similar to those shown in Table 1 for the Customer Contact Styles Questionnaire. Validity coefficients for these four competencies are considerably higher when aggregate scores are computed from each individual's responses across the designated OPQ scales (Method 1) than when an average of the scale-criterion correlations is taken (Method 3). To conserve space, the table contains only illustrations from the full data-set. Average correlations across all 14 competencies in this set (including the four in the table) were found to be 0.19 and 0.10 for Methods 1 and 3 respectively. As in the first analysis, directly calculated aggregate scores produced validity coefficients that were about twice the magnitude of average scale-criterion correlations.

Discussion

These analyses have illustrated that validity coefficients for personality factors within the Big Five and similar frameworks are systematically different between different kinds of scale combination. Correlations were substantially larger when individuals' responses were first aggregated into a factor score (Method 1, here treated as a bench-mark), in comparison with a computed average of separate scales' coefficients (Method 3). It is essential to interpret correlations with Big Five and other compounds in the light of the aggregation method that has been applied.

Judge, Thoresen, Bono, and Patton (2001) drew attention to this issue in respect of the average correlation between job satisfaction and performance. They pointed out that the correlation of performance with overall satisfaction as a single composite (which they examined) is likely to be larger than the average of correlations with specific facet satisfactions (as studied in some previous research). Similarly, Salgado's (2003) suggestion that personality inventories designed to assess Big Five factors ("FFM" inventories) might be more valid than non-FFM inventories for certain factors (see above) has yet to be supported by comparisons that employ the same method of aggregation.

Average correlations within each Big Five compound were computed by DeNeve and Cooper (1998) in relation to respondents' subjective well-being. Their detailed presentation allows consideration of an additional issue – the varying breadth of predictors within an average correlation. For example, 23 indicators of Extraversion were considered. Although most of the cited scales could be defined as facets of that Big Five factor (Activity, Dominance, Sociability, etc.), DeNeve and Cooper also included several that tapped the broad compound of Extraversion itself. The cited average correlations thus derived not only from separate facets but also from the overall compound. That is likely to be a common procedure. It is desirable to present those two sets of results separately, rather than merging them into a single average correlation.

The size of the discrepancy between outcomes from Methods 1 and 3 is in part a function of the variation in magnitude between correlations for the single scales (shown in earlier rows of the tables). In cases where scale-criterion correlations include some particularly low individual values, a mean of those (Method 3) will of course be markedly reduced. For example, Detail Conscious is a poor predictor of sales performance in Table 1, despite being part of the conventional construct definition of Conscientiousness; the low values for Detail Conscious naturally reduce the mean correlation value. On the other hand, the inclusion of some scales with low validity does not necessarily reduce the value of a computed composite-score correlation (Methods 1 and 2). Such scales can leave unaffected a moderate or large aggregate correlation rather than decreasing it, as occurs for the mean in Method 3.

Conversely, when a Big Five factor is represented by a set of single scales that are all similarly associated with a criterion, the discrepancy between their average correlation and the correlation with their composite is likely to be smaller than those illustrated here. For example, the first three scales in Table 1 (Competitive, Results Oriented, and Energetic) are similar in representing the achievement-orientation form of Conscientiousness, whereas the other three scales (Structured, Detail Conscious, and Conscientiousness as measured here) more concern dependability. Supplementary analyses of the former scales yielded the same divergence as shown for all six scales in Table 1, but the difference between Methods 1 and 3 was reduced; mean values were 0.27 and 0.19.

Method 2 (deriving a composite-score correlation) is of interest here in part because of the need in secondary analyses of published data to assume an average between-predictor correlation. Table 1 showed that different assumptions by Mount and Barrick (1995) and Salgado (1997, 1998) ($r_{xx} = 0.47$ and 0.34 respectively) created similar mean composite-score correlations of Conscientiousness with sales performance (0.15 and 0.16). Outcomes from both estimates were well below the bench-mark factor-criterion correlation from Method 1 applied to actual data (averaging 0.22), but did not differ greatly between themselves. Use of the actual mean value of r_{xx} from these three sets of data (0.14) yielded an average composite-score correlation of 0.21, identical to the average value for Method 2a, which used each study's own r_{xx} value. It is clearly important in this aggregation process that estimates of mean r_{xx} are open to justification.

The present results are consistent across three studies in Analysis 1 and across 14 competencies in Analysis 2. Two different inventories (CCSQ and OPQ) were examined, and both sales-people and managers were included in the research. It is clear, as expected from their statistical basis, that the three aggregation methods yielded different results, and that composite-score correlations using Method 2 were sensitive to the assumption made

about the mean level of between-predictor correlations. Previous research findings need to be interpreted with that in mind.

In the future, composite-score correlations (as in Method 2) could be calculated more effectively if (in addition to reporting predictor-criterion correlations, as is usual) primary research publications about the Big Five were required also to present a study's average between-predictor correlation for each factor. That additional information would permit the incorporation in subsequent investigations of assumptions about mean r_{xx} that are clearly justifiable through specified previous evidence. For example, it would then be possible to use an estimated mean r_{xx} value that applies to the particular Big Five scales that are under examination, rather having to resort to different or unspecified Big Five indicators. More generally, it is clear from the present analyses, and from a consideration of statistical principles, that aggregation Method 3 is not appropriate for the examination of correlations with Big Five factors. That procedure produces misleadingly low estimates of the validity of a composite.

References

- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. Personnel Psychology, 44, 1-26.
- Bartram, D. (in press). The Great Eight competencies: A criterion-centric approach to validation. Journal of Applied Psychology.
- DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. Psychological Bulletin, 124, 197-229.
- Guilford, J. P., & Fruchter, B. (1978). Fundamental statistics in psychology and education, sixth edition. New York: McGraw-Hill.
- Hunter, J. E., & Schmidt, F. L. (1990). Methods of meta-analysis: Correcting error and bias in research findings. London: Sage.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. Journal of Applied Psychology, 85, 869-879.
- Judge, T. A., Thoresen, C. J., Bono, J. E., & Patton, G. K. (2001). The job satisfaction-job performance relationship: A qualitative and quantitative review. Psychological Bulletin, 127, 376-407.
- Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. Research in Personnel and Human Resources Management, 13, 153-200.
- Salgado, J. F. (1997). The five-factor model of personality and job performance in the European Community. Journal of Applied Psychology, 82, 30-43.
- Salgado, J. F. (1998). Big Five personality dimensions and job performance in army and civil occupations: A European perspective. Human Performance, 11, 271-288.
- Salgado, J. F. (2003). Predicting job performance using FFM and non-FFM personality measures. Journal of Occupational and Organizational Psychology, 76, 323-346.

- SHL (1993). OPQ Concept Model: Manual and user's guide. Thames Ditton, England: SHL Group plc.
- SHL (1997). Customer contact manual and user's guide. Thames Ditton, England: SHL Group plc.
- SHL (1999). OPQ32: Manual and user's guide. Thames Ditton, England: SHL Group plc.
- Vinchur, A. J., Shippmann, J. S., Switzer, F. S., & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. Journal of Applied Psychology, 83, 586-597.
- Warr, P. B., Bartram, D., & Martin, T. (2005). Personality and sales performance: Situational variation and interactions between traits. International Journal of Selection and Assessment, 13, 87-91.

Appendix

Formulae for calculating composite scores by Methods 1b and 2 are given below. Details of their application are illustrated in the text.

Method 1b

The formula for a composite-score correlation based on raw scores is as follows. Let the predictor variables (here Big Five scales) be denoted x_1, x_2, \dots, x_n ; let the composite score be denoted X ; and let the criterion variable (here job performance) be denoted y . Let $r_{x_1y}, r_{x_2y}, \dots, r_{x_ny}$ be correlations between the individual predictors (here, scales) and y . Let $r_{x_1x_2}, r_{x_1x_3}, \dots, r_{x_{n-1}x_n}$ be correlations between the individual predictors (here, scales). A basic formula for the Pearson correlation between two variables is their covariance divided by the product of their standard deviations:

$$r_{Xy} = \frac{\text{cov}(X, y)}{SD_X SD_y} = \frac{\text{cov}(x_1 + \dots + x_n, y)}{SD_X SD_y} = \frac{\text{cov}(x_1, y) + \dots + \text{cov}(x_n, y)}{SD_X SD_y} =$$

$$= \frac{r_{x_1y} SD_{x_1} SD_y + \dots + r_{x_ny} SD_{x_n} SD_y}{SD_X SD_y} = \frac{r_{x_1y} SD_{x_1} + \dots + r_{x_ny} SD_{x_n}}{SD_X}.$$

Where SD_X (standard deviation of a composite score) is:

$SD_X = \sqrt{SD_{x_1}^2 + \dots + SD_{x_n}^2 + 2(r_{x_1x_2} SD_{x_1} SD_{x_2} + \dots + r_{x_{n-1}x_n} SD_{n-1} SD_n)}$, then the composite-score correlation is:

$$r_{Xy} = \frac{r_{x_1y}SD_{x_1} + \dots + r_{x_ny}SD_{x_n}}{\sqrt{SD_{x_1}^2 + \dots + SD_{x_n}^2 + 2(r_{x_1x_2}SD_{x_1}SD_{x_2} + \dots + r_{x_{n-1}x_n}SD_{n-1}SD_n)}} \quad (1)$$

Because predictor scale intercorrelations cannot be greater than 1.00, the denominator in the above formula cannot be greater than the sum of predictor-scales standard deviations:

$$\begin{aligned} & \sqrt{SD_{x_1}^2 + \dots + SD_{x_n}^2 + 2(r_{x_1x_2}SD_{x_1}SD_{x_2} + \dots + r_{x_{n-1}x_n}SD_{n-1}SD_n)} \leq \\ & \leq \sqrt{SD_{x_1}^2 + \dots + SD_{x_n}^2 + 2(SD_{x_1}SD_{x_2} + \dots + SD_{n-1}SD_n)} = \\ & = \sqrt{(SD_{x_1} + \dots + SD_{x_n})^2} = SD_{x_1} + \dots + SD_{x_n} \end{aligned}$$

Therefore the composite score correlation cannot be less than the weighted average of the individual scales' correlations:

$$r_{Xy} \geq \frac{r_{x_1y}SD_{x_1} + \dots + r_{x_ny}SD_{x_n}}{SD_{x_1} + \dots + SD_{x_n}}$$

Method 2

The formula for a composite-score correlation for standardized scores by Hunter and Schmidt (1990, p. 455) is as follows.

Let the predictor variables (here Big Five scales) be denoted x_1, x_2, \dots, x_n ; let the composite score be denoted X ; and let the criterion variable (here job performance) be denoted y . Let \bar{r}_{xy} be the average correlation between the individual predictors (here, scales) and y ; that is, let \bar{r}_{xy} be the average of r_{x_1y}, r_{x_2y} and so on. Let \bar{r}_{xx} be the average correlation between the individual predictors (here, scales); that is, let \bar{r}_{xx} be the average of $r_{x_1x_2}, r_{x_1x_3}, r_{x_2x_3}$ and so on.

Let \bar{c}_{xx} be the average covariance between the indicators; that is, let \bar{c}_{xx} be defined by

$$\bar{c}_{xx} = \frac{1 + (n-1)\bar{r}_{xx}}{n}.$$

Then the relationship between the average correlation \bar{r}_{xy} and the composite-score correlation r_{Xy} is given by

$$r_{Xy} = \frac{\bar{r}_{xy}}{\sqrt{\bar{c}_{xx}}} \quad (2)$$

Because predictor scale intercorrelations cannot be greater than 1.00, the denominator in the above formula cannot be greater than 1.00:

$$\sqrt{\bar{c}_{xx}} = \sqrt{\frac{1 + (n-1)\bar{r}_{xx}}{n}} \leq \sqrt{\frac{1 + (n-1)}{n}} = \sqrt{\frac{n}{n}} = 1.$$

Therefore the composite score correlation cannot be less than the average of the individual scales' correlations:

$$r_{Xy} \geq \bar{r}_{xy}$$

Table 1. Achieved sales as a function of Conscientiousness: Factor correlations from three methods of aggregation across predictor scales.

CCSQ scale within the Conscientiousness factor	Correlation with sales criterion			
	Study 1	Study 2	Study 3	Mean value across 3 studies
Competitive	.27	.30	.13	.23
Results Orientated	.18	.37	.22	.26
Energetic	.13	-.10	.23	.09
Structured	.15	-.06	.07	.05
Detail Conscious	.01	.03	-.05	.00
Conscientiousness	.10	-.14	.14	.03
Methods 1a and 1b: Composite-score correlation, using the mean of raw scores	.26	.20	.21	.22
Method 2a: Composite-score correlation, using standardized scores and actual mean r_{xx} for each study	.25	.16	.21	.21
Method 2b: Composite-score correlation, using Mount and Barrick's estimate of mean r_{xx} (0.47)	.19	.09	.17	.15
Method 2c: Composite-score correlation, using Salgado's estimate of mean r_{xx} (0.34)	.21	.10	.18	.16
Method 2d: Composite-score correlation, using mean r_{xx} from these three studies (0.14)	.26	.13	.23	.21
Method 3: Mean of correlations with single scales	.14	.07	.12	.11

Table 2. Manager-rated competencies as a function of OPQ compounds within a theoretical framework: Correlations from three methods of aggregation across predictor scales.

OPQ scale linked with the specific competency (different in each case)	Correlation with competency criterion			
	Competency 1.1. Deciding and initiating action	Competency 2.1. Working with people	Competency 5.2. Creating and innovating	Competency 8.2. Entrepreneurial and commercial thinking
A	.17	.00	.02	.19
B	-.01	.17	.24	.11
C	.07	.14	.12	.07
D	.16	.13	.21	.05
E	.20	.09	.02	.19
F	.16	.11		.17
G		.11		.08
Methods 1a and 1b: Composite-score correlation, using the mean of raw scores	.27	.20	.23	.27
Method 2a: Composite-score correlation using standardized scores and actual mean r_{xx} for each competency	.27	.20	.22	.26
Method 3: Mean of correlations with single scales	.13	.11	.12	.12