A Method Factor Predictor of Performance Ratings

Michael D. Biderman
University of Tennessee at Chattanooga

Nhung T. Nguyen
Towson University

Billy Mullins
Jason Luna
Vikus Corporation

Authors' Note: Correspondence regarding this article should be sent to Michael Biderman, Department of Psychology / 2803, U.T. Chattanooga, 615 McCallie Ave., Chattanooga, TN 37403. Tel.: (423) 425-4268. Email: Michael-Biderman@utc.edu

Paper presented at the 23rd Annual Conference of The Society for Industrial and Organizational Psychology, San Francisco, CA, 2008.

Poster

TITLE

A method factor predictor of performance ratings

ABSTRACT

A method factor estimated from a Big Five questionnaire was positively correlated with supervisor ratings of performance. Validities of the Big Five variables were negligible or negative in the context of the method factor. The possibility that the method factor represented impression management is discussed.

PRESS PARAGRAPH

This study investigated validity of the Big Five in a model in which a common method factor was estimated. In this model, the common method factor had significant positive validity for prediction of supervisor ratings of customer service, sales, and collections performance of financial services employees while the validities of the Big Five were negligible or negative. The possibility that the method factor found here represented impression management, i.e., faking, is discussed.

The resurgence of the use of personality tests in employee selection has generated a renewed interest in common method variance. Method variance is that associated with individual differences in respondent tendencies, such as acquiescence or faking, when a common source of data (e.g., self-report) is used. When self-report is used for all variables in a study - independent and dependent variables - the presence of such common variance has been found to positively bias substantive relationships among constructs of interest in organizational research (e.g., Doty & Glick, 1998), hence the phrase "common method bias". For the most part, interest in method variance has been on its influence on the validity of conclusions about relationships between substantive variables. For this reason, method variance has often been treated as a nuisance to be partialled out when examining theoretical relationships (e.g., Conger, Kanungo, & Menon, 2000), and much of the research on method effects - the sources of method variance - has been devoted to discovering ways to get rid of them or minimize their influence through construction and wording of test items. For example, Podsakoff, MacKenzie, Lee, & Podsakoff (2003), p. 887-895) devoted a section to techniques for controlling common method effects.

Whereas most researchers have focused on the nuisance aspects of method effects, a few have examined possible substantive aspects of them. For example, Williams, Gavin, & Williams 1996) studied negative affectivity (NA), the tendency of individuals to respond negatively to items independently of item content in self-report questionnares. Williams, et. al. (1996) first examined NA as a nuisance variable, presenting a method for removing the effects of NA from correlations between theoretical constructs based on structural equation models. But they also examined the relationships of NA treated as a substantive construct to other variables. Recently, Motl and DiStefano (2002), Horan, DiStefano, and Motl (2003), and Quilty, Oakman, & Risko, (2006) have focused on the substance of positive and negative affect method factors from the Rosenberg Self Esteem Scale and the relationship of those factors to other variables. In performance management literature, halo error has been found to be related to performance (Sulsky & Balzer, 1988). In other words, variables estimated as method effects could explain relevant criterion variance.

Method variance that is common to only one type of variable in a study may also be problematic. If the scales in a set of predictor variables, such as the Big Five scales, are subject to method effects, then the variance associated with such effects is added to the variance of all the scales in the set. If that variance is correlated with a criterion, then the observed relationships of all scales to the criterion will be augmented. However, if the common method variance is unrelated to the criterion, then the observed relationships of the scales to the criterion would be weaker than would the relationships between pure measures of the dimensions (e.g., Ones & Connelly, 2007).

Studies of method effects among only independent variables have mostly involved examinations of the effects of faking on predictor validities in selection situations. Unfortunately, there is not consensus on the impact of such method effects. On the one hand, Hough, Eaton, Dunnette, Kamp, & McCloy (1990), Ones, Viswesvaran, & Reiss (1996), Ones, & Viswesvaran (1998) found little evidence of the relation of validity to faking or response distortion. . However, other studies have found evidence of moderation. These include Anderson, Warner, & Spencer, 1984; Douglas, et. al., 1996; Mueller-Hanson, Hegfgestad, & Thornton, 2003, and Christiansen, Burns, &

Montgomery, 2005. Thus, it appears that there is need for further research on the impact of method effects such as faking on validity of predictors in selection. This study was designed to provide a look at a new way of measuring method effects - as a method factor extracted from the predictor items - and to examine the effect of estimation of such a factor on validity of the Big Five personality dimensions.

The effect of common method variance on the structure of the Big Five factors was first reported in Schmit and Ryan (1993) who factor analyzed responses to individual items of the NEO FFI (Costa & McCrae, 1989) of applicant and non-applicant samples. In their analysis of the non-applicant sample, they found the expected five-factor solution. However, in the applicant sample, they found that a six-factor solution fit the data best and that the sixth factor shared cross-loading with four of the Big 5 dimensions. They labeled the sixth factor an "ideal employee" factor defined as the tendency to distort responses on a selection test in order to increase the chance of being hired. Later studies (e.g., Frei, 1998; Frei, Griffith, Snell, McDaniel, & Douglas, 1997) comparing variance structures of the Big Five measures between faking good vs. honest groups, showed differences in the number of latent variables, error variances, and correlations among latent variables across group. Recently, Biderman and Nguyen (2004) investigated a model in which a common method factor specifically representing response distortion or faking was included.

The study reported here investigated the utility of estimating method effects as a factor common to all items in a Big Five questionnaire in a concurrent validation study. First, we present results based only on the Big Five dimensions, the type of analysis that is typically conducted with no attempt to estimate or account for method effects. Then, we present an analysis based on a model in which method effects are taken into account.

METHOD

Participants. Participants were 764 employees of a national private personal finance company with job titles of "Sales Associate" or "Sales Manager". Eighty-six percent were female; 59% were White, 24% Black, 9% Hispanic and 8% Other. The essential duties of each job were the same with respect to interacting with customers. Each job required the incumbent to perform duties and tasks in the areas of selling, customer service, and debt collections. Supervisors of the incumbents provided criterion data for those who had worked for at least 60 days. Each supervisor provided criterion data on from 10 to 30 incumbents. Criterion scores were provided in July and August, 2007, the same time that incumbents filled out the Big 5 questionnaire.

Big Five questionnaire. The Big Five personality questionnaire was the 50-item scale (Goldberg, 1999) available on the IPIP web site. The Big Five items were embedded contiguously within a 117 item online questionnaire being evaluated for use as a selection instrument. Items were displayed one-at-a-time using a web-based computer system. A 5-point response scale, with labels "I disagree", "I somewhat disagree", "I neither agree nor disagree", "I somewhat agree", and "I agree" was used. For positively worded items, "I agree" was at the top of the screen under the item. For negatively worded items, "I disagree" was at the top of the screen. Participants responded by clicking on a check box next to the preferred response. Reliability estimates of summated scales for the five dimensions were .82, .70, .71, .83, and .73 for extraversion, agreeableness, conscientiousness, stability, and openness/intellect respectively.

Criteria. Criteria were supervisor ratings of customer service, sales, and collections, each a single rating on a 5-point scale. Ratings were subjective, but followed examination of objective (e.g., sales) data when possible. An overall criterion computed as the average of the above three values was also analyzed. Alpha for the overall criterion was .82.

Measurement and structural models. Two confirmatory factor analysis measurement models were applied to the 50 individual items. The first contained five latent variables labeled E, A, C, S, and O for the Big Five dimensions of extraversion, agreeableness, conscientiousness, stability, and openness/Intellect respectively. Each item loaded on the appropriate latent variable. Correlations among the latent variables were estimated. Thus, this was a standard CFA of the 50-item Big Five questionnaire, presented in Figure 1. The second model was identical to the first with the exception that a sixth latent variable, labeled M, was included on which all 50 items were required to load. M was constrained so that it was orthogonal to all of the Big Five latent variables. Thus, M represented variation that was common to the items in addition to but orthogonal to variation due to the Big Five latent variables. Model 2 is that presented in cell 3A in Table 4 of Podsakoff, et. al. (2003) where the latent variable is called an "unmeasured latent methods factor". It is presented in Figure 2.

For the investigation of the effect of M on the validities of the Big Five, structural models regressing the criteria onto the Big Five latent variables and M were conducted. The form of the structural models is illustrated in Figure 3. All models were applied using Mplus Version 4.2 (Muthen & Muthen, 1998-2006)

RESULTS

Table 1 presents means, standard deviations, and correlations between Big Five summated scale scores and the three criteria. The correlations among the Big Five scales are all positive; most are significantly so. Positive correlations among Big Five scale scores have been reported previously and are the substance of at least one theory of higher order factors of the Big Five (Digman, 1997). Positive correlations among IPIP Big Five dimensions were found previously (Lim & Ployhart,2006) although the mean of the correlations in Table 1, .35, is larger than the mean of .18 from that study.

Examination of the simple validities in Table 1 shows that most are positive and generally small, although with the exception of conscientiousness the validities are about as large as mean correlations for prediction of subjective ratings reported in the Barrick & Mount (1991) meta analysis. The only two significant simple validity coefficients were those that would have been expected based on prior research - agreeableness predicted customer service ratings and extraversion predicted ratings of sales performance (e.g., Vinchur, Schippmann, Switzer, & Roth, 1998). Conscientiousness was not a significant predictor of any of the criteria in contrast to its generally recognized validity (e.g., Schmidt & Hunter, 1998).

Table 2 presents results of the structural model regressions of the four criteria onto the Big Five scale scores. The results of the multiple regression analyses paralleled the simple validities presented in Table 1. Controlling for the other Big Five variables, agreeableness predicted customer service ratings and extraversion predicted sales ratings. Conscientiousness was not significantly related to any of the criteria. Openness/Intellect was negatively related to customer service, collections, and the overall criterion.

In summary, with the exception of the validities of conscientiousness, the results of validation of the Big Five data are not considerably different from previously reported mean Big Five validities (Barrick & Mount, 1991). For the most part the validities are positive but small. Even the negative coefficients associated with Openness/Intellect are not surprising given the small (.02) mean validity for this dimension predicting subjective ratings found by Barrick & Mount (1991).

Before the impact of method effects on the validity coefficients could be evaluated, an assessment of the utility of a method effect measurement model was required. This involved comparing goodness-of-fit of Model 1, in which no method effects were estimated, to Model 2, which included a method factor. Results are presented in Table 3. Inspection of the table shows that Model 2 fit the data better than Model 1 ($\chi^2(50)= 607.791$, $p<.001$)[1]. The impact of including the method factor was evident in aspects of the data other than goodness-of-fit. For example, because the method effect latent variable, M, accounted for some of the covariance between items, loadings of the items on the Big Five latent variables were smaller when the method effect latent variable was included in the model. Moreover, when M was added to the model, the correlations among the Big Five latent variables became less positive, such that the mean of those correlations was reduced from .45 to 0. These results suggest that there were individual differences in participants' tendencies to inflate (or deflate) all of their responses to the Big Five items that were captured by the latent variable, M.

Given the evidence of superior fit of Model 2, regression analyses assessing the validity of the Big Five dimensions with the method latent variable included were conducted. The structural model multiple regressions of the criteria onto the latent variables shown in Figure 3 were conducted. Results are presented in Table 4. Three main differences between these results and the results from the regressions involving Model 1 reported in Table 2 are apparent. First, the method latent variable is the only variable with consistently significant positive correlations with the criteria. Second, the validity coefficients of the Big Five variables are consistently negative, some significantly so. Thus, the validities from a model in which a method factor was estimated are markedly different from those from analyses excluding a method latent variable.

Finally, the multiple Rs in Table 5 from regressions of Model 2 factor scores were larger than the corresponding multiple Rs from Big Five latent variable regressions. These differences occurred even though both analyses used the same 50 items as predictors, albeit in different ways.

DISCUSSION

This study represents the first investigation of Big Five validities within the context of a model in which an item-level method latent variable was estimated along with the Big Five latent variables. The results indicated that addition of a method factor markedly improves goodness-of-fit of a confirmatory factor analysis of Big Five items and that validities of the factors from the method effect model are quite different from validities from traditional analyses of the Big Five dimensions.

The first question that arises is the source of the method factor found here. Our guess is that it represents individual differences in impression management, i.e., faking. In order to provide a limited test of this hypothesis, we drew on data presented previously that utilized the same 50-item scale employed here. If the method factor represents

individual differences in faking, then these individual differences should be similar in some respect to other data in which faking is believed to have occurred and they should be different from other data in which it is believed that participants responded honestly.

To assess such differences, we were able to obtain data from two previous studies utilizing the 50-item IPIP scale (Biderman & Nguyen, 2004 and Wrensen & Biderman, 2005). Both studies employed honest response conditions and also employed instructed faking conditions. First, the honest response data from those two studies were concatenated onto the present data. Then the faking condition data from the same studies were added. The result was a data matrix consisting of 764 rows from the present study, 369 rows of honest condition data from the previous studies, and 369 rows of faking condition data from the previous studies. Model 2 was estimated for the 1502 case concatenated dataset. Factor scores for M were computed and then for each dataset - this dataset, the two previous honest condition datasets, and the two previous faking condition datasets - the mean of those factor scores was computed.[2]

The comparisons of mean M factor scores used the analysis of variance followed by a Tukey *b* post hoc comparison of means. Mean value of M for the present study was .05. For the faking conditions of the previous two studies, means were 0.01 each. For the honest response conditions of the previous two studies, means were -.08 and -.10. Differences between the three groups of means were significant ($p < .05$). Thus the mean M in the present study was larger than the means of the faking conditions of the previous studies which in turn were larger than the means from the honest conditions of the previous studies. These comparisons are thus consistent with the hypothesis that participants in the present study engaged in systematic impression management that was on the average more positive than that found in the two previous studies.

If the M latent variable indeed represents individual differences in a form of impression management, that may explain the positive validity associated with M. We consider the method effect found here to be a general tendency to "add a little" to or "subtract a little" from the response that would otherwise be elicited by an item in the absence of any tendency to present oneself in a positive (or negative) light. We believe that the tendency affects all items and that it is an individual difference variable. If the tendency to present oneself well permeates an employee's other behaviors in the workplace, such positive self presentation may influence supervisors' ratings of performance. On the other hand, the tendency to "subtract a little" might have the opposite influence on supervisor ratings. This agrees with the findings of Piedmont, McCrae, Riemann, and Angleitner (2000) who noted that individual differences in socially desirable responding - differences that may be at least part of what are captured by the M latent variable here - may represent a substantive personality variable, one that may be predictive of performance. The individual differences captured by M here may also be related to influence tactics reviewed by Higgins, Judge, & Ferris (2003). For example, these authors found that *ingratiation* defined as using behaviors designed to increase the target's liking of oneself or to make oneself appear friendly was positively correlated with performance assessments. To the extent that the individual differences reflected in M may be called an influence tactic, then the tactics such as ingratiation reviewed by Higgins, et. al. (2003) may shed light on what was found here.

It is hard to believe that an impression management tendency as strong as the one found here has permeated all other self-report data. It is likely that in some data

collection situations there is little common method variance. On the other hand, it is also likely that in some situations common method variance is quite prevalent. The faking literature mentioned above demonstrates that under instructions or incentives to fake there is very likely substantial common method variance due to faking. The method effects literature has pointed out other sources of common method variance, such as negative affectivity. The upshot of this is that a theory of method effects is needed. Such a theory would explicate the conditions under which method effects might occur and when they might not occur. The data reported here clearly point out the need for such a theory.

Although the above hypothesis for M has some logical bases, the negligible or negative validity coefficients for the Big Five when M is estimated are more difficult to explain. Since M was estimated to be orthogonal to the Big Five latent variables, the Big Five dimensions are uncorrelated with M, eliminating an explanation in terms of classical suppression. The three major findings were the lack of validity of conscientiousness and the negative validity of stability and openness/intellect. One explanation for the low validity of conscientiousness is that the situation presented here is not representative of typical validation studies involving the Big Five. It could be that in the culture of the organization, behaviors associated with high conscientiousness were not viewed favorably by raters. Another possibility is that the estimate of validity of conscientiousness was biased toward zero due to restriction of range on the conscientiousness dimension. The negative validity of stability for the collections criteria may reflect the fact that those who are low on stability are more suited to the onerous task of collecting bad debts (e.g., Robins, 1995). The negative validities of openness/intellect may reflect the fact that that in an attempt to create a structured working environment, those employees whose positions on the openness/intellect dimension led them to engage in behaviors that were outside the established norms were viewed more negatively than those scoring low on openness/intellect.

Finally we note that this is the first study in which individual differences that might be called impression management have been estimated in a single incumbent sample using only the predictor items. Other studies have estimated faking through extrapolation from other scales such as social desirability scales or through the use of difference scores requiring an honest response condition. In this study, only the Big Five items were used to estimate the method factor in a single setting. We believe that such an immediate measure of impression management will ultimately be the most useful measure. Clearly, though, further investigation is required to establish the nomological net for M and to determine the conditions under which individual differences represented by M will occur.

REFERENCES

Anderson, C. D., Warner, J. L., Spencer, C. S. (1984). Enflation bias in fel-assessment examinations: Implications for valid employee selection. Journal of Applied Psychology, 69, 574-580.

Barrick, M., & Mount. (1991). The Big Five personality dimensions and job performance and job performance: A meta analysis. *Personnel Psychology,* 44, 1-26.

Biderman, M. D., & Nguyen, N. T. (2004). Structural equation models of faking ability in repeated measures designs. Paper presented at the 19th Annual Society for Industrial and Organizational Psychology Conference, Chicago, IL.

Christiansen, N. D., Burns, G. an., & Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. Human Performance, 18, 267-307.

Conger, J. A., Kanungo, R. N., & Menon, S. T. (2000). Charismatic leadership and follower effects. *Journal of Organizational Behavior, 21*, 747-767.

Costa, P.T., & McCrae, R.R. (1989). *The NEO PI/FFI manual supplement.* Odessa, FL: Psychological Assessment Resources.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of personality and social psychology*, *73*, 1246-1256.

Doty, D. H., & Glick, W. H. (1998). Common methods bias: Does common methods variance really bias results? *Organizational Research Methods*, *1*, 374–406.

Douglas, E. F., McDaniel, M. A., & Snell, A. F. (1996). The validity of non-cognitive measures decays when applicants fake. Paper published in the Academy of Management Proceedings: August.

Frei, R.L. (1998). *Fake this test! Do you have the ability to raise your score on a service orientation inventory*. University of Akron. Unpublished doctoral dissertation.

Frei, R.L., Griffith, R.L., Snell, A.F., McDaniel, M.A., & Douglas, E.F. (1997). *Faking of non-cognitive measures: Factor invariance using multiple groups LISREL*. Paper presented at the 12th Annual Meeting of the Society for Industrial & Organizational Psychology: St. Louis, MO.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 1-28). Tilburg, The Netherlands: Tilburg University Press.

Higgins, C.A., Judge, T.A., & Ferris, G.R. (2003). Influence tactics and work outcomes: A meta-analysis. Journal of Organizational Behavior, 24, 89-106.

Horan, P. M., DeStefano, C., & Motl, R. W. (2003). Wording effects in self-esteem scales: Methodological artifact or response style? Structural Euqation Modeling, 10, 435-455.

Hough, L.M., Eaton, N.K., Dunnette, M.D., Kamp, J.D., & McCloy, R.A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology, 75*, 581-595.

Lim, B., & Ployhart, R.E. (2006). Assessing the Convergent and Discriminant Validity of Goldberg's International Personality Item Pool: A Multitrait-Multimethod Examination. *Organizational Research Methods*, *9*, 29-54.

Motl, R. W., & DeStefano, C. (2002). Longitudinal invariance of self-esteem and method effects associated with negatively worded items. *Structural Equatioin Modeling, 9*, 562-578.

Mueller-Hanson, R., Heggestad, E. D., & Thornton III, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. Journal of Applied Psychology, 88, 348-355.

Muthén, L.K. and Muthén, B.O. (1998-2006). Mplus User's Guide. Fourth Edition. Los Angeles, CA: Muthén & Muthén.

Ones, D. & Connelly, B. S. (2007). Multiple Measures of a Single Conscientiousness Trait: Validities Beyond .35! Paper presented at the 22nd Annual Conference of the Society for Industrial & Organizational Psychology. New York: NY.

Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology,* 81, 660-679.

Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. Human Performance,11, 245-269.

Piedmont, R. L., McCrae, R. R., Riemann, R., & Angleitner, A. (2000). On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. Journal of Personality and Social ZPdychology, 78, 582-593.

Podsakoff, P.M., MacKenzie, S. B., Lee, J., & Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879-903.

Quilty, L. C., Oakman, J. M., & Risko, E. Correlates of the Rosenberg Self-Esteem Scale method effects. Structural Equation Modeling, 13, 99-117.

Robins, K. W. (1995). Effects of personality and situational judgment on job performance. *Dissertation Abstracts International: Section B: The Sciences and Engineering, Vol 55(9-B*), Mar . pp. 4155

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin, 124, 262--274.

Schmit, M.J., & Ryan, A.M. (1993). The Big Five in Personnel Selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology, 78*, 966-974.

Sulsky, L.M., & Balzer, W.K. (1988). Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. *Journal of Applied Psychology, 73*, 497-506.

Vinchur, A. J., Schippmann, J. S., Switzer, III, F. S., & Roth, P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology, 83*, 586-597.

Williams, L. J., Gavin, M. B., & Williams, M. L. (1996). Measurement and nonmeasurement processes with negative affectivity and employee attitudes. Journal of Applied Psychology, 81, 88-101.

Wrensen, L. B., & Biderman, M. D. (2005). Factors related to faking ability: A structural equation model application. Paper presented at the 20th Annual

Conference of The Society for Industrial and Organizational Psychology, Los Angeles, CA. April.

Footnotes

[1] The model reported in the text utilized maximum likelihood estimation. A model using a more conservative estimation procedure recommended in cases when the data may not be multivariate normal - option MLMV in Mplus - was also applied and yielded results equivalent to those presented here.

[2] Although factor scores were computed so that the mean over all 1502 cases was zero, the means of subsets of the scores may be different from zero, just as the means of subsets of Z-scores may be different from zero.

Table 1. Means, standard deviations, and correlations of summated scales and criteria. Reliability estimates of multiple item scales are on the diagonal.

	1	2	3	4	5	6	7	8	9
1 Extraversion	(.82)								
2 Agreeableness	.295[c]	(.70)							
3 Conscientiousness	.275[c]	.311[c]	(.71)						
4 Stability	.318[c]	.258[c]	.495[c]	(.83)					
5 Openness/Intellect	.483[c]	.287[c]	.408[c]	.345[c]	(.73)				
6 Customer Service	.034	.100[b]	.008	.014	-.039				
7 Sales	.090[a]	.058	.056	.036	.034	.625[c]			
8 Collections	.030	.050	.050	.038	-.045	.601[c]	.557[c]		
9 Overall	.060	.081[a]	.045	.034	-.021	.864[c]	.849[c]	.849[c]	(.81)
Mean	3.88	4.48	4.52	4.12	3.98	3.46	2.93	2.96	3.12
SD	0.71	0.44	0.42	0.66	0.54	0.79	0.79	0.84	0.69

---

[a] $p < .05$ [b] $p < .01$ [c] $p < .001$

Table 2. Results of regression of criteria onto the Big Five latent variables. Entries are standardized regression coefficients from the structural regression model.

	Big Five Latent Variable					
Criterion	E	A	C	S	O	Multiple R
Customer service	0.104	0.113[a]	0.010	0.003	-0.181	0.164
Sales	0.126[a]	0.020	0.035	0.001	-0.068	0.118
Collections	0.108	0.044	0.105	0.002	-0.206[b]	0.152
Overall	0.132[a]	0.069	0.059	0.002	-0.178[a]	0.152

---

[a] $p < .05$ [b] $p < .01$ [c] $p < .001$

Table 3. Results of application of the confirmatory factor analyses.

	Model 1	Model 2
	CFA without M	CFA with M
Chi-square	3468.578[c]	2860.787[c]
df	1165	1115
CFI	0.775	0.830
RMSEA	0.051	0.045
SRMR	0.066	0.047
Chi-square Difference		607.791[c]
Mean E loading	0.57	0.45
Mean A loading	0.45	0.33
Mean C loading	0.46	0.31
Mean S loading	0.60	0.40
Mean O loading	0.48	0.36
Mean M loading	.---	0.35
Mean Big Five intercorrelation	0.45	0.00

---

[a] $p < .05$ [b] $p < .01$ [c] $p < .001$

Table 4. Results of regression of criteria onto the Big Five and the method factor (M). Entries are standardized regression coefficients from the structural regression model.

	Latent Variable						
Criterion	E	A	C	S	O	M	Multiple R
Customer Service	-.017	-.037	-.107	-.107	-.128[a]	.121[a]	.253
Sales	.021	-.090	-.066	-.090	-.044	.141[a]	.200
Collections	-.026	-.134	-.049	-.136[a]	-.144[a]	.146[a]	.277
Overall	-.010	-.105	-.087	-.132[a]	-.120[a]	.161[a]	.277

---

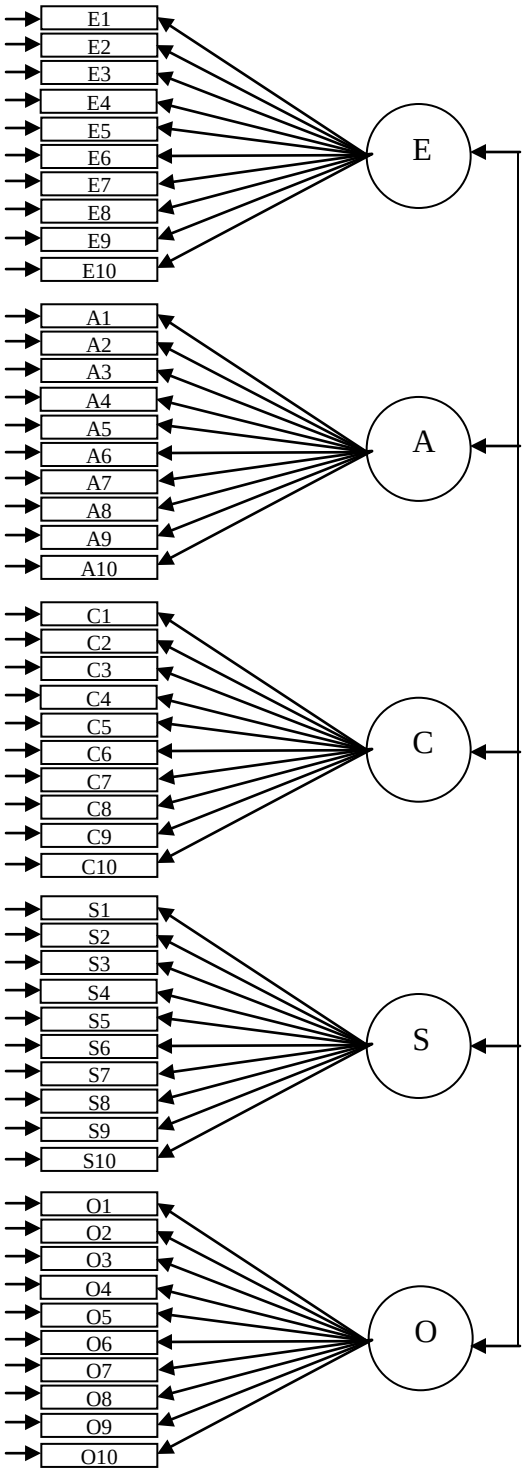
[a] $p < .05$ [b] $p < .01$ [c] $p < .001$

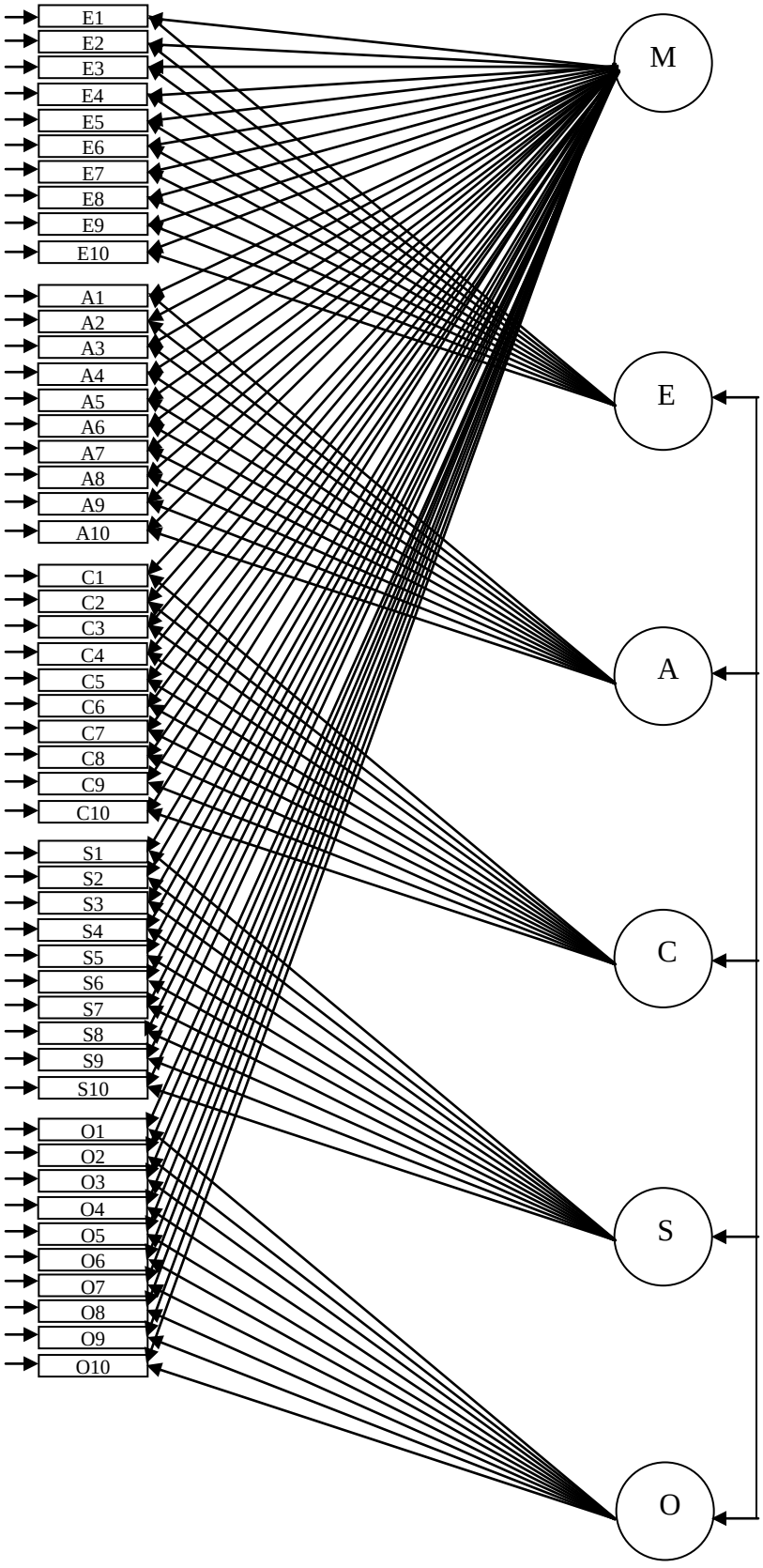
Figure Captions

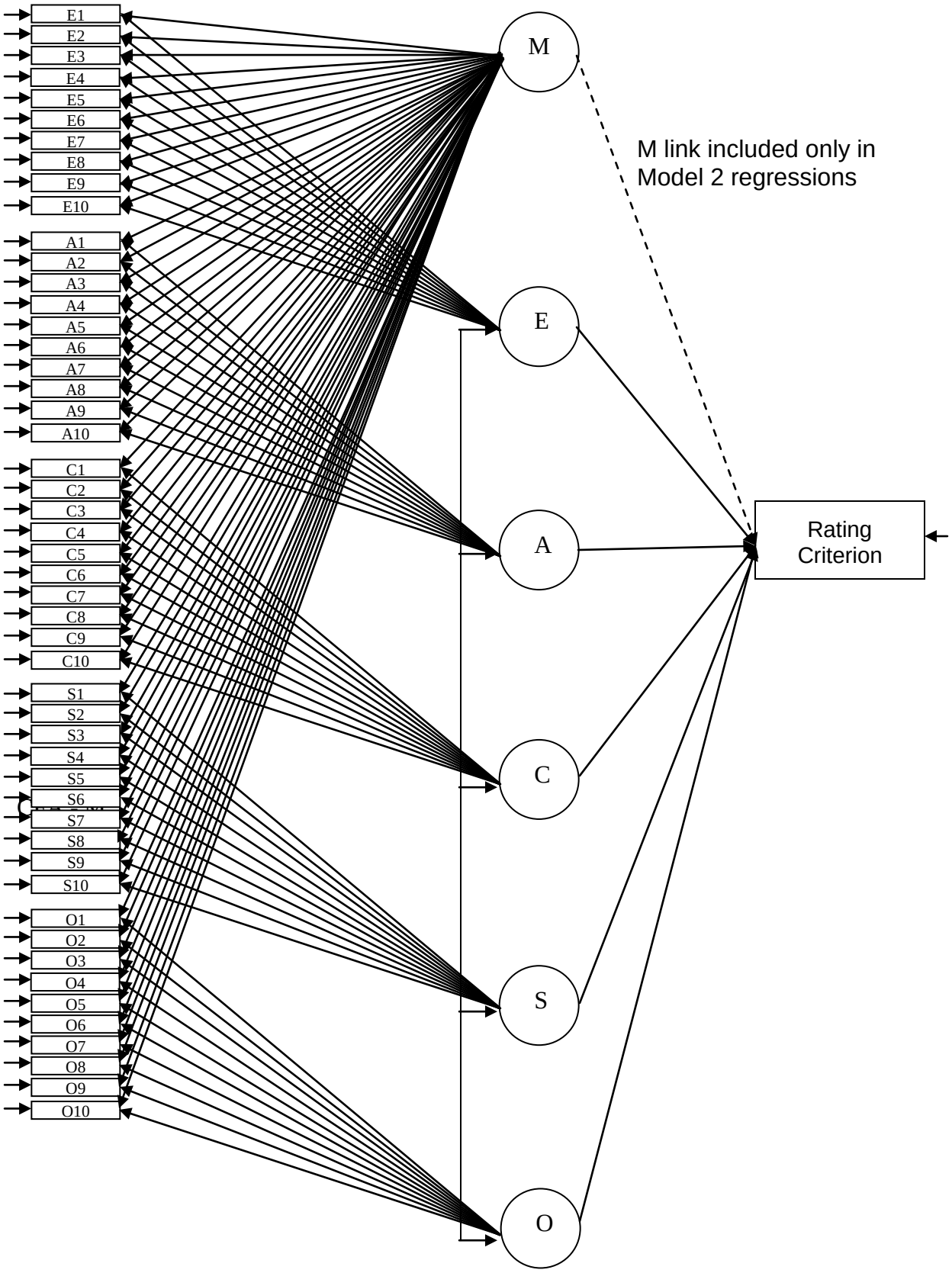
Figure 1. Model 1: Confirmatory factor analysis of 50-item Big Five scale. Big Five dimensions are labeled E, A, C, S, and O for extraversion, agreeableness, conscientiousness, stability, and openness/intellect respectively.

Figure 2. Model 2: Confirmatory factor analysis with method factor indicated by all 50 individual items. Method factor is labeled M.

Figure 3. Form of the structural model. Specific criteria were customer service rating, sales rating, collections rating, and overall rating. M was included as a predictor in regressions involving Model 2.


E1
E2
E3
E4
E5
E6
E7
E8
E9
E10
E
A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
A
C1
C2
C3
C4
C5
C6
C7
C8
C9
C10
C
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S
O1
O2
O3
O4
O5
O6
O7
O8
O9
O10
O


E1
E2
E3
E4
E5
E6
E7
E8
E9
E10
A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
C1
C2
C3
C4
C5
C6
C7
C8
C9
C10
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
O1
O2
O3
O4
O5
O6
O7
O8
O9
O10
M
E
A
C
S
O


E1
E2
E3
E4
E5
E6
E7
E8
E9
E10
A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
C1
C2
C3
C4
C5
C6
C7
C8
C9
C10
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
O1
O2
O3
O4
O5
O6
O7
O8
O9
O10
M
E
A
C
S
O
M link included only in
Model 2 regressions
Rating
Criterion