Manipulating Method Variance

Michael D. Biderman

Raven Worthy

University of Tennessee at Chattanooga

Nhung T. Nguyen

Towson University

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 submitted for presentation at the 27th Annual Conference of The Society for Industrial and Organizational Psychology, San Diego, CA. 2012.

Poster
TITLE
Manipulating method variance

ABSTRACT
Method variance in Big Five data was manipulated using faking instructions. A method factor correlated with measures of affect in an honest response condition and with cognitive ability in an instructed faking condition. Validity of conscientiousness factor scores from a method factor model was larger than that of scale scores.

PRESS PARAGRAPH
The topic of common method variance has not been addressed in experimental research. In this study, we investigated method variance and its relationship to external variables and to validity across different response instructions. In the “honest response” manipulation; method variance correlated most highly with affective measures whereas in an “instructed faking” manipulation; method variance correlated most highly with cognitive ability. Adjusting for method variance increased criterion-related validity.

In spite of attempts to create an item pool in which the items uniquely represent five orthogonal dimensions, a common finding in analyses of Big Five personality data is that the scale scores are correlated. One explanation that has been put forth for the correlations is that while each item may indicate only one dimension, the dimensions themselves are correlated. This line of thought has led to a body of literature on higher order factors of the Big Five, from proposals of two higher order factors (e.g., Digman, 1997; DeYoung, Peterson, & Higgins, 2002) to the proposal of a single higher order factor (e.g., Musek, 2007; Rushton & Irwing, 2011).

Another explanation for the correlations is that individual items may not always represent a single concept. The concept of one construct per item is a major tenet in the writing of items for personality questionnaires. For example, the first of five rules for writing items presented by Spector (1992) is, “Each item should express one and only one idea.”. Netemeyer, Bearden, and Sharma (2003) caution to “ . . . avoid double-barrel statements that in effect address two issues . . .”. Although the goal of unidimensional items is a laudable one, it is difficult to achieve it in practice. There are two possible ways that the correlations could be a result of communality of meaning across items. The first is that pairs of items could share meaning simply due to the vagaries of language. If this were the case, it would likely be manifest in correlated residuals across certain pairs of items within a questionnaire when the items were factor analyzed. This possibility was investigated in some detail by Marsh, Lüdtke, Müthén, Asparouhov, Morin, Trautwein, & Nagengast (2010), who estimated correlations between residuals of 57 pairs of items in the NEO-FFM questionnaire and found substantial increases in goodness-of-fit when the residual correlations were estimated.

A second possibility is that the correlations arise, not due to specifics of the language used for pairs of items, but due to respondent characteristics that influence responses to a shared aspect, such as the valence, of *all* the items relatively independently of specific item wording. This second possibility was recently investigated by Biderman, Nguyen, Cunningham, and Ghorbani (2011). Biderman et al. labeled this a common method factor because it was a factor whose values were estimable only when a common method was used for the gathering of data – in their case, self-report. They found substantial increases in goodness-of-fit when a common method factor was estimated. Biderman et al. (2011) were careful to point out, however, that the factor was a personal characteristic whose existence became manifest through the use of a common method and that it was not a characteristics of the method itself.

No one who has performed confirmatory factor analyses of individual responses to personality questionnaires such as Big Five questionnaires will deny that there are idiosyncratic correlations between responses to pairs of items such as those modeled by Marsh et al. (2010). However the presence of such idiosyncratic correlations doesn't preclude the second possibility mentioned above. Evidence for the existence of bias factors affecting all items comes from the literature on applicant faking. Several investigators have provided evidence for the existence of a sixth factor affecting all items of a Big Five personality questionnaire when instructions or incentives to fake are present (e.g., Biderman & Nguyen, 2004; Biderman & Nguyen, 2009; Klehe, Kleinmann, Hartstein, Melchers, König, Heslin, & Lievens, in press; Schmit & Ryan, 1993). Models of such data incorporating the addition of a single common factor seem to be the most parsimonious way of account for the increase in between-dimension correlations that occur when faking instructions or incentives are present. In addition to results from the faking literature, certain results found even when there are no instructions or incentives to fake have

been accounted for by assuming the existence of a single common factor. For example, Biderman, Nguyen, and Cunningham (2011) found that factor scores of a common factor were significantly positively related to self-esteem and significantly negatively related to depression, a result that cannot be explained by an assumption of correlated residuals between item pairs.

The present research was designed to explore further the manifestations of individual differences in the characteristics of common method factors. We did so by estimating a common method factor across three manipulations – honest responding, instructed faking, and incentives to fake. We computed factor scores for the common method factor and examined correlations of those scores with external variables. Our expectation was that there would be systematic differences in the nature of the correlations between items across three conditions, correlations that would be most usefully accounted for by the assumption of a common factor affecting all the items. We expected the correlations of that factor with external variables to depend on the condition in which it was estimated. Specifically, we expected that when no instructions or incentives to fake were present, the common factor would correlate with measures of self-concept, as reported in Biderman et al. (2011). In conditions in which instructions to fake were present, we expected the common factor to represent respondents' attempts to follow those instructions. Since instructions to fake create a problem solving situation and persons with higher cognitive ability general do better in problem solving tasks, we expected that those better at faking in this condition would be higher in cognitive ability. Thus we expected the common method factor to correlate positively with cognitive ability in the instructed faking condition. Finally, in the condition with incentives to fake, it was our expectation that those who distorted the most would also be those who scored highest on a common measure of social desirability. Thus we entertained the following three hypotheses.

H1: A common method factor in an honest response conditions will correlate positively with self-esteem and negatively with depression.

H2: A common method factor in an instructed faking condition will correlate positively with cognitive ability.

H3: A common method factor in a condition with incentives to fake will correlate positively with social desirability.

Finally, because response distortion affecting all items adds variance that is irrelevant to the dimensions of the Big Five, we investigated this potential problem by including a criterion – academic performance operationalized as cumulative undergraduate Grade Point Average (UGPA) – and examined the correlation of UGPA with conscientiousness scores from the Big Five. We compared the validity of conscientiousness measured in two ways - as conscientiousness scale scores and as conscientiousness factor scores computed from a model in which a common method factor was estimated. Our expectation was since the irrelevant variance due to the common method factor was removed from the factor scores, the validity of the factor scores would be higher than the validity of conscientiousness scale scores. Specifically,

H4: Conscientiousness factor scores from a common method bias model will correlate more highly with UGPA than conscientiousness scale scores.

Since cognitive ability is well known to be a predictor of UGPA, we included a measure of cognitive ability and controlled for it in the analyses conducted here.

METHOD

Participants

Participants were 328 undergraduates at a medium sized public university in the southeast. Respondents participated for extra credit in psychology classes. Respondent characteristics were: 29.6% male, 68.0% White, 24.1% Black, with the remaining 7.9% indicating they were Asian, Hispanic, or of mixed ethnic background. Mean age was 20.3 with standard deviation 5.3.

Design

The experiment was a three-condition randomized groups design. About one-third of the participants served in either an Honest response condition, a condition in which instructions to fake were given, or a condition in which incentives to fake were present.

Measures

Big Five. Items from the 100-item Big Five questionnaire from the IPIP web site (www.ipip.org) were used. The items chosen were those from the 100-item questionnaire that were not included in the 50-item sample questionnaire. The response scale was seven points, with response alternatives labeled from "Completely Inaccurate" to "Completely Accurate". Individual Big Five scale scores were computed as the mean of the 10 items from each dimension after reverse-scoring the negatively worded items.

Wonderlic Personnel Test. The paper and pencil version of WPT Form II was administered (Wonderlic, 1999).

Academic performance (UGPA). At the end of the academic year in which respondents participated, their end-of-year GPAs were recorded from the student information system and used as criterion scores.

Depression. The Costello and Comrey (1967) depression scale was used as the measure of depression. Sample items include: "When I wake up in the morning I expect to have a

miserable day." and "I wish I were never born." Responses were on the same seven-point scale and higher scores reflected higher levels of chronic depression.

Self-Esteem. The Rosenberg self-esteem scale (Rosenberg, 1965) was taken from Carmines and Zeller (1979). Sample items include: "I feel I have a number of good qualities." and, "I wish I could have more respect for myself." Responses were on the same seven-point scale and higher scores reflected higher levels of self-esteem.

Other Measures. The Hartman Value Profile (Hartman, 1967), the Myers-Briggs Type Inventory (MBTI, Myers & McCaulley, 1985), and the 50-item IPIP sample Questionnaire were administered as part of another research project. The data of those questionnaires were not analyzed here.

Method Factor Model. A three-group confirmatory factor analysis (CFA) model with an additional common method factor influencing all indicators was fit to the data of the Big Five questionnaire using Mplus V6.11 (Müthén, & Müthén, B.O.,1998-2011). For this model, covariances between factors were assumed to be equal across the three conditions. The common method factor was assumed to be orthogonal to the Big Five factors. Factor loadings were estimated separately for each condition. The result was that the method factor for each condition was essentially a different factor, indicated by the loadings of the items on it only within that condition.

Procedure. Respondents participated in groups of one to 15 in a computer lab. After signing informed consent forms, a paper and pencil form of the WPT was administered. Participants were then directed to the web site at which the HVP, Big Five, and MBTI were administered in the described order. After completion of the MBTI, participants were directed to the SurveyMonkey web site where all scales analyzed in this study were presented. The

questionnaire items were presented in groups of 10 or more per screen, in a fashion similar to the appearance of questionnaire items on paper and pencil questionnaires. Respondents responded to all items on the screen, then clicked a button to move to the next screen.

Participants were given all questionnaires except the last IPIP questionnaire with instructions to respond honestly. Prior to the last questionnaire, in the honest condition participants were told simply to respond "as honestly as you can." In the instructed faking condition, participants were told to "respond as if you were someone who needed the job so bad that if stretching or embellishing yourself a bit would get you the job, you would go ahead and do it." In the incentive condition, participants were told that "based on the responses to this section, the twenty participants who would make the best candidates for employment will be entered into a drawing for a $50 gift certificate."

RESULTS

Table 1 presents means, standard deviations, and correlations between the study variables.

Table 2 presents summary statistics from model application. Three models were applied. Model 1 was a three-group CFA with no common method factor. Separate loadings on the Big Five factors were estimated for each condition. This model served as a baseline against which to compare models with a method factors. Model 2 was a CFA with a common method factor estimated assuming loadings were equal across three conditions. That is, for Model 2, the data were treated as a single group with respect to estimation of the method factor, although trait loadings were estimated separately for each condition as in Model 1. Finally, Model 3 was a CFA with separate method factor loadings estimated for each condition.

Table 2 shows that Model 2 fit the data significantly better than Model 1, indicating, as has been reported by others, that fit of CFAs to Big Five data can be improved by adding a common method factor. More importantly for the present research, Model 3 fit significantly better than Model 2, indicating that the nature of the response distortion represented by the common method factor differed across the three conditions.

The top panel of Table 3 presents correlations of the common method factor scores with depression and self-esteem from each condition. As the table shows, there was a consistent pattern of correlations across three conditions, with negative correlations with depression and positive correlations with self-esteem in all three conditions. The correlations were the weakest in the instructed faking condition. The pattern of correlations gives tentative support to Hypothesis 1, although the strength of the correlations in the incentive conditions was not expected.

The middle panel of Table 3 presents correlations of the common method factor scores with cognitive ability. As expected in Hypothesis 2, the correlation was positive in the instructed faking condition and negligible in both the honest and incentive conditions.

The bottom panel of Table 3 presents correlations of the common method factor scores with BIDR self-deception and impression management scores. The self-deception scores were positively correlated with the method factor across conditions. The correlations with impression management scores were not statistically significant in any condition. Although, as expected, the correlations involving both self-deception and impression management were numerically largest in the incentive condition the differences between those correlations and those from the other conditions were not statistically significant.

To test the hypothesis that removing the variance associated with the common method factor from conscientiousness scores would result in incremental validity, regressions of UGPA onto conscientiousness were performed. In order to control for possible differences between conditions and differences in cognitive ability, the regressions controlled for Wonderlic scores and also controlled for condition using group coding techniques. To compare conditions, two contrast codes were formed. The first compared the two faking conditions to the honest condition. The second compared the incentive faking condition to the instructed condition. To insure that differences between the conditions did not moderate the validities, product variables representing the product of condition-coding variables and conscientiousness were included in all the regressions. Two sets of regression analyses were conducted. In the first, conscientiousness scale scores were treated as the predictor. Results of these analyses are presented on the left side of Table 4. In the second, conscientiousness factor scores from the CFA were treated as the predictor as shown on the right side of Table 4.

Because none of the product variables was significant in the analyses, only the results controlling for WPT and condition are presented in Table 4. As can be seen in the table, the incremental validity of conscientiousness factor scores was consistently and numerically greater than that for scale scores. Formal statistical comparison of scale and factor scores was complicated by the fact that since both WPT and the condition coding variables contributed significantly to prediction of the criterion the appropriate comparison is of the partial regression coefficients – the partial regression coefficient of C scale scores on the left vs. the coefficient of C factor scores on the right of Table 4. It was decided that the most feasible comparison of the relative validity of scale and factor scores as predictors of the criterion was to include them both in a multiple regression analysis, controlling for WPT and condition codes.

The bottom panel of Table 4 presents the results of the analysis pitting scale scores vs. factor scores in the same regression. As was the case with the previous regression analyses, the interactions of the conscientiousness scales with the condition codes were assessed, and, as in the previous analyses, were not significant and are not presented in Table 4. Thus, the table presents only the results controlling for WPT and those controlling for WPT and conditions. As shown in the bottom panel of Table 4, when both conscientiousness scale scores and factor scores are in the regression equation, the factor scores exhibit significantly incremental validity over both faking condition and WPT while the scale scores are not significant in any of the analyses.

DISCUSSION

This study explored the relationship of common method variance to both personality variables and performance variables in conditions in which respondents were expected to respond honestly, in conditions in which they were instructed to fake, and in conditions in which they were given a small incentive to fake. The results suggest that the nature of the dissimulation represented by the method bias factor estimated here depends on the situation. In situations in which there are no instructions to fake, the common factor is related to the self-concept of respondents, with persons high on the common factor showing greater self-esteem and less depression than persons low on the factor. This surprising result replicates those of Biderman et al. (2011) and Biderman, Nguyen & Cunningham (2011). It suggests that there are subtle aspects of respondent behavior when responding to personality questionnaires that may be useful in assessing the person beyond the personality dimensions ostensibly represented.

On the other hand, when respondents were given explicit instructions to fake, relative standing on the common factor was determined to some extent by cognitive ability of the respondents, with those high in cognitive ability being those who were highest on the factor.

Moreover the correlations of method bias with self-concept in the instructed faking condition were minimal. Thus, simply adding a couple of sentences to the instructions radically changed the nature of respondent dissimulation, from a veiled report of self-concept to dissimulation that represented the cognitive capabilities of the respondents.

Finally, we found generally positive relationships of the method factor to the self-deception scale of the BIDR. This finding suggests that in part the relationships to self-esteem and depression may have involved elements of self-deception on the part of the respondents. The absence of correlations with the BIDR IM scale suggests that the self-report represented by the common factor was not a result of respondent intent as represented by the impression management scale

It is important to note that dissimulation was found in all conditions of the present research. This result dispels the notion that in the absence of instructions or incentives to fake, respondent dissimulation is minimal or absent. On the contrary, dissimulation is present and it is related to personal characteristics of the respondents. In the absence of instructions or incentives to fake, persons who would be characterized as "fakers" are those with high self-esteem and low depression. But when instructions to fake are given, those would be characterized as "fakers" are those who are best able to solve the problem given by the situation – those with the highest cognitive ability.

As expected, the analyses shown in Table 4 suggest that removing unwanted variance from the measure of conscientiousness made it a better predictor of academic performance than the raw conscientiousness scale scores. There is a widely held belief that social desirability correlations for dissimulation do not lead to increases in validity. For example, Griffith & Peterson (2008) said, " . . . when measures of SD are used as a proxy variable for applicant

faking behavior, faking does not appear to substantially affect criterion-related validity nor do score corrections improve the accuracy of selection decisions." The results found here suggest that the problem in the assessment of faking may have been in the use of social desirability measures. When direct measurement of dissimulation is made on the selection instrument itself, as with the common method factor estimated here, that dissimulation does moderate the validity of a personality predictor of performance. The results of this comparison were consistent with previous research (e.g., Biderman, Nguyen, and Sebren, 2008) and clearly warrant further research on the measurement of Big Five factors with the effects of removing unwanted variance.

The results suggest that responses to individual items are influenced by at least one factor other than those the items have been written to represent. Our data suggest that the factor is not specific to a few items but is one that influences all items in the same way. Although Bäckström, Björklund, & Larrson (2009) have shown that the influence can be ameliorated by writing items so that they have little obvious valence, to the best of our knowledge none of the Big Five questionnaires in use at the present time has been written with the intent to minimize such valance. In fact, the results of the present study suggest that it may not be necessary to remove the effects of the common method factor estimated here. Rather, it may be better to embrace that influence, measure it through the methods used here, and use it to enhance validity and to better understand the variables measured in personality theory.

References

Bäckström, M., Björklund, F. & Larsson, M. (2009). Five-factor inventories have a major general factor related to social desirability which can be reduced by framing items neutrally. *Journal of Research in Personality, 43, 335- 344.*

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

Biderman, M. D. & Nguyen, N. T. (2009, April). Measuring faking propensity. Paper presented at the 24th annual conference of The Society for Industrial and Organizational Psychology, New Orleans, LA.

Biderman, M. D., Nguyen, N. T., & Cunningham, C. J. L. (2011). A method factor measure of self-concept. Paper presented at the 26th Annual Conference of The Society for Industrial and Organizational Psychology, Chicago, IL.

Biderman, M. D., Nguyen, N. T., Cunningham, C. J. L., & Ghorbani, N. (2011). The ubiquity of common method variance: The case of the Big Five. *Journal of Research in Personalit*y, 45, 417-429.

Biderman, M.D., Nguyen, N.T., & Sebren, J. (2008). Time-on-task mediates the Conscientiousness – Performance Relationship. *Personality and Individual Differences, 44:* 887-897.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment.* Newbury Park, CA: Sage.

Costello, C. G., & Comrey, A. L. (1967). Scales for measuring depression and anxiety. The *Journal of Psychology, 66,* 303-313.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2002). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual differences, 33*, 533-552.

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

Griffith, R. L., & Peterson, M. H. (2008). The failure of social desirability measures to capture applicant faking behavior. *Industrial and Organizational Psychology, 1*, 308-311.

Hartman, R. S. (1967). *The Structure of Value: Foundations of scientific Axiology.* Southern Illinois Press.

Klehe, U.-C., Kleinmann, M., Hartstein, T., Melchers, K., König, C., Heslin, P. & Lievens, F. (in press). Responding to Personality Tests in a Selection Context: The Role of the Ability to Identify Criteria and the Ideal-Employee Factor. *Human Performance*.

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the Big Five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22*, 471-491.

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of Research in Personality, 41*, 1213-1233.

Müthén, L.K. and Müthén, B.O. (1998-2011). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Müthén & Müthén

Myers, I. B., & McCaulley, M. H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling Procedures: Issues and Applications.* Thousand Oaks, CA: Sage.

Rosenberg, M. (1965). *Society and the adolescent self image.* Princeton, NJ: Princeton University Press.

Rushton, J. P., & Irwing, P. (2011). The general factor of personality. In Chamorro-Premuzic, von Stumm, & Furnham, A. (Eds). *The Wiley-Blackwell Handbook of Individual Differences.* West Sussex: UK.

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.

Spector, P. E. (1992). Summated Rating Scale Constructon: An Introduction. Newberry Park, CA: Sage.

Wonderlic, Inc. (1999). *Wonderlic's Personnel Test manual and scoring guide*. Chicago: IL: Author.

Figure Captions

Figure 1. Common method variance model.

Figure 1.

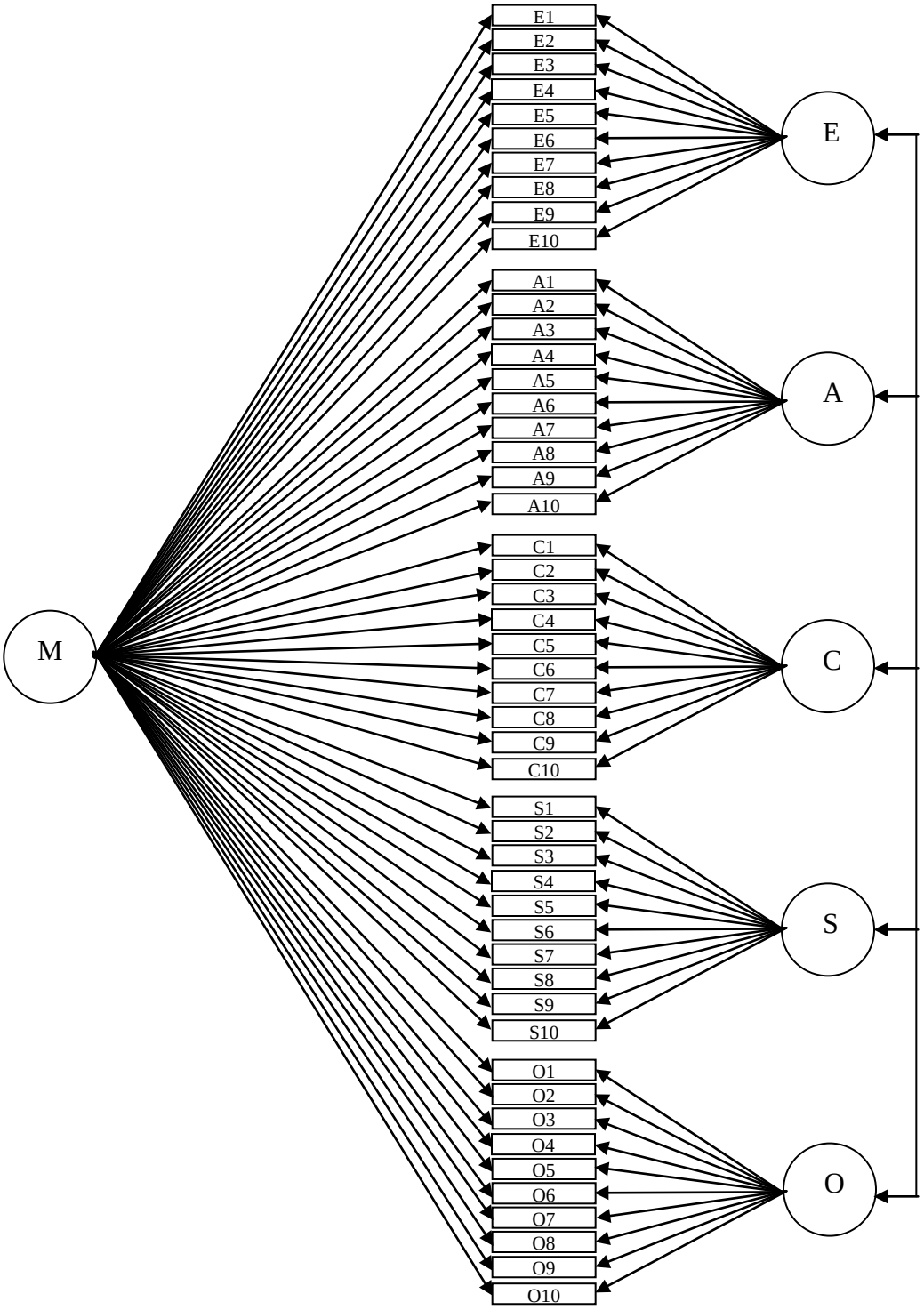
E1
E2
E3
E4
E5
E6
E7
E8
E9
E10
A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
C1
C2
C3
C4
C5
C6
C7
C8
C9
C10
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
O1
O2
O3
O4
O5
O6
O7
O8
O9
O10
M
E
A
C
S
O

Table 1. Means, standard deviations and correlations between scales. N = 328. Entries on the diagonal are coefficient alpha values. Diagonal entries for factor scores are coefficients of determination for the three conditions.

Scale	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1 WPT	21.76	5.23	.999															
2 E Scale	5.14	0.92	.181[b]	.828														
3 A Scale	5.65	0.78	.149[b]	.634[c]	.795													
4 C Scale	5.19	1.95	.130[a]	.520[c]	.579[c]	.876												
5 S Scale	4.79	1.17	.247[c]	.498[c]	.481[c]	.428[c]	.893											
6 O Scale	5.15	0.89	.316[c]	.590[c]	.564[c]	.576[c]	.549[c]	.830										
7 E FS	0.28	0.92	.063	.763[c]	.549[c]	.222[c]	.323[c]	.253[c]	.889									
8 A FS	0.10	0.91	.041	.495[c]	.826[c]	.373[c]	.341[c]	.251[c]	.727[c]	.891								
9 C FS	-0.02	0.88	.032	.220[c]	.386[c]	.781[c]	.251[c]	.244[c]	.237[c]	.487[c]	.877							
10 S FS	0.14	0.93	.163[b]	.291[c]	.315[c]	.188[b]	.861[c]	.298[c]	.355[c]	.385[c]	.268[c]	.921						
11 O FS	0.23	0.94	.232[c]	.196[c]	.262[c]	.272[c]	.360[c]	.607[c]	.231[c]	.282[c]	.415[c]	.448[c]	.925					
12 Method FS	0.26	0.98	.167[b]	.597[c]	.557[c]	.663[c]	.385[c]	.730[c]	.122[a]	.115[a]	.126[a]	.058	.077	.942				
13 Dep	1.98	0.93	.033	-.236[c]	-.151[b]	-.102	-.142[b]	-.093	-.228[c]	-.142[b]	-.054	-.119[a]	.040	-.121[a]	.932			
14 RSE	5.55	1.05	-.059	.225[c]	.115[a]	.134[a]	.148[b]	.151[b]	.195[c]	.082	.067	.135[a]	.027	.176[c]	-.720[c]	.898		
15 BIDR SD	4.40	0.55	-.072	.280[c]	.093	.233[c]	.274[c]	.232[c]	.165[b]	.056	.130[a]	.223[c]	.056	.237[c]	-.291[c]	.433[c]	.615	
16 BIDR IM	3.78	0.66	-.034	.075	.185[b]	.178[b]	.111[a]	.107	.070	.177[b]	.203[c]	.112[a]	.127[a]	.055	-.062	.115[a]	.251[c]	.696
17 GPA	2.96	0.65	.299[c]	.018	.055	.163[b]	.016	.071	-.010	.049	.201[c]	.005	.090	.011	.044	-.027	-.120[a]	.067

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

Table 2. Results from application of the confirmatory factor analysis model. Chi-square differences compare a model with the model in the row above.

Model	Chi-square	df	CFI	RMSEA	SRMR	ΔChi-square	df
1 (No M)	7090.07[c]	3605	.600	.094	.139		
2 (One M)	6582.65[c]	3553	.652	.088	.132	507.42[c]	52
3 (Three Ms)	6411.12[c]	3453	.661	.089	.091	171.53[c]	100

[c] $p < .001$

Table 3. Correlations of common method factor scores with depression and self-esteem within each condition.

	Honest	Instructed	Incentive
Depression	-.108	-.103	-.141
Self-Esteem	.218[a]	.115	.222[a]
Cognitive ability	.009	.301[b]	.107
BIDR SD	.228[a]	.260[b]	.268[b]
BIDR IM	-.022	.041	.172
Sample Size	110	108	110

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

Table 4. Moderated regression analyses predicting GPA.

Predictor	Standardized Coefficient	ΔR^2	Predictor	Standardized Coefficient	ΔR^2
WPT	.283[c]	.078	WPT	.293[c]	.086
C Scale Scores	.126[a]	.016	C Factor Scores	.192[c]	.037
WPT	.293[c]	.084	WPT	.308[c]	.093
C Scale Scores	.190[b]	.032	C Factor Scores	.222[c]	.048
Faking v Honest	-.185[b]	.032	Condition Code 1	-.172[b]	.029
Incent F v Instr F	.093	.008	Condition Code 2	-.072	.005

Predictor	Std Coef	ΔR^2
WPT	.300[c]	.087
C Scale Scores	-.067	.004
C Factor Scores	.244[b]	.060
WPT	.306[c]	.090
C Scale scores	.020	.000
C Factor Scores	.208[a]	.016
Faking v Honest	-.174[b]	.028
Incent F v Instr F	.075	.005

---

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