Big Five Predictors of Academic Achievement

Sampo V. Paunonen

University of Western Ontario, London, Ontario, Canada

and

Michael C. Ashton

Brock University, St Catharines, Ontario, Canada

A large-sample study (N = 717) is described in which two broad Big Five factor measures were compared with two narrow personality trait measures in the prediction of final grades in an undergraduate psychology course. The Big Five factors evaluated as predictors were Conscientiousness and Openness to Experience (or Intellect). The traits evaluated, which were constituents of the respective factors, were need for Achievement and need for Understanding. In each comparison, the lower level trait measure did better than its higher level factor measure in the prediction of course grades. We conclude that the aggregation of narrow trait measures into broad factor measures can be counterproductive from the points of view of both behavior prediction and behavior explanation.

Some researchers believe that the variation in human behavior can be adequately summarized in terms of five broad dimensions of personality. Those dimensions, known as the Big Five personality factors, are typically labeled as Extraversion, Agreeableness, Conscientiousness, Neuroticism (vs Emotional Stability), and Openness to Experience (or Intellect). The Big Five have been recovered in many studies and in many cultures, with the analysis of both comprehensive lists of personality-descriptive adjectives (e.g., De Raad, Perugini, Hrebicková, & Szarota, 1998) and numerous questionnaires comprising diverse personality trait scales (see McCrae & John, 1992).
The Big Five factors capture a large portion of the variance within the domain of measured lower level personality variables. Many of those lower level variables, however, retain a substantial amount of specific, nonrandom variance not related to the Big Five factors, variance that might predict important criterion variables (Paunonen & Jackson, 2000). In fact, it has been shown empirically that the narrow traits that define the Big Five factors can surpass those broad dimensions themselves in the prediction of some real-life criteria.

Some Past Prediction Results

Big Five personality factor measures have sometimes been compared to their narrower facet measures in the prediction of relevant criterion variables. Ashton, Jackson, Paunonen, Helmes, and Rothstein (1995, Table 5) showed that a simple trait measure of orderliness, as an example, predicted a criterion of tidiness of students’ dormitory rooms with a correlation of .44. However, when the orderliness scale was combined with related scales into a methodicalness factor, the accuracy of predicting the tidiness criterion dropped to .26. In another study, Ashton (1998, Table 3) found that trait measures of risk taking and responsibility predicted anonymous self-reports of students’ delinquent behaviors, with correlations of .30 and −.40, respectively. But the best that any Big Five (adjective-based) factor scale could do was −.22 for Conscientiousness. Similarly, Rothstein, Paunonen, Rush, and King (1994) reported that narrow trait measures of achievement, dominance, and exhibition individually predicted business students’ classroom participation grades with correlations reaching .33. The highest correlation for any Big Five factor measure, in contrast, was only −.20 for Agreeableness.

Paunonen (1998) presented many comparisons similar to those described above, with his results generally supporting the conclusion that aggregating lower level trait measures into a higher level factor measure can lead to substantial losses in predictive accuracy (see also Mershon & Gorsuch, 1988). Moreover, using partial-regression techniques, he found that statistically removing the common factor component from the variance of a set of cohort facet scales left residual variances that were, for some of the traits, still predictive of important criterion variables. Such residual variation for any particular trait measure could include, of course, both a random component due to error and a nonrandom component due to trait specificity. The point made by Paunonen (1998; see also Paunonen & Nicol, in press) was that one might want to preserve that trait-specific component of variance because it could be criterion-predictive rather than throw it away in the process of combining facet measures into factor measures.

Despite evidence to the contrary, there seems to be a particularly strident call for the use of broad Big Five factor measures over narrower personality trait measures in Industrial–Organizational psychology. Ones and Viswesv-
aran (1996), for example, have recently reported a meta-analysis of personality variables as predictors of job performance. Their results led them to conclude that ‘general personality traits (large bandwidth) will always have higher predictive validity than specific traits in real world applications’ (p. 616), particularly when the work behavior criterion is also broad. In a detailed reply to the Ones and Viswesvaran article, however, Paunonen, Rothstein, and Jackson (1999) maintained that behavior prediction can easily suffer when either broad personality predictors or broad criteria are favored over their narrower constituent counterparts. They even demonstrated the effect in Ones and Viswesvaran’s (1996) own data. For example, Big Five measures predicted a broad composite criterion of job performance, with a mean correlation of only .17 in their meta-analysis and with no value greater than .22. In contrast, those same Big Five measures predicted specific, narrower job performance criteria, with values averaging .27 and with all values exceeding .23.

Overview of the Present Study

The purpose of this study was to provide another comparison of the predictive validities of broad Big Five personality factors versus the narrower personality variables that make up those factors. The subjects were undergraduate university students enrolled in several sections of a personality psychology course. The criterion to be predicted with the personality variables was an objective measure of academic performance, computed as the students’ final course grades. What makes this study different from most others in this area is the size of the subject sample involved, exceeding 700 students.

METHOD

Subjects and Procedure

This was a large-sample study of 717 undergraduate students (190 men and 527 women) in 18 sections of a 2nd-year psychology course in personality theory and research taught by the first author from 1987 to 1999. All students in that course wrote three examinations containing items of the multiple-choice and short-answer variety. Each examination was graded anonymously by a teaching assistant. Students’ final course grades, forming the criterion variable of this study, were calculated as the arithmetic mean of the three examinations.

As part of a class exercise, all students in each section of the course were requested to complete Form E of Jackson’s (1984) Personality Research Form (PRF). That questionnaire contains 20 16-item trait scales and a Desirability scale, measures which formed the basis of the predictor variables of this study.

Selection of Predictor Variables

We compared the performance of the individual PRF traits with PRF-based Big Five factors in predicting the present criterion of final course grades. But we decided against simply computing all the lower level trait correlations and contrasting those with all the higher level factor correlations. Such a comparison would tend to favor the lower level traits’ correlations because
21 coefficients (i.e., for the 20 PRF content scales plus the Desirability scale) would be compared with only five coefficients (i.e., for the five PRF factors). This could possibly result in a larger number of spuriously significant predictor–criterion correlations occurring in the former case.

Instead of comparing all of the present trait scales against all of the Big Five factor scales in the prediction of course grades, we selected two Big Five factor predictors and compared each of them against only one lower level trait. The two Big Five personality factors that were chosen as potential predictors of academic performance were Conscientiousness and Openness to Experience. Costa and McCrae (1992, p. 16) have noted that high Conscientiousness is associated with academic and occupational achievement, and several researchers have found that Openness to Experience is significantly correlated with measured intelligence (Holland, Dollinger, Holland, & MacDonald, 1995; McCrae, 1993–1994).

For reasons already mentioned, we wanted to select from each of the relevant Big Five domains under consideration (i.e., Conscientiousness and Openness to Experience) only one PRF trait as a comparative predictor. Instead of choosing randomly, however, we decided to select the one trait that would be expected, on rational grounds, to be the best predictor of academic performance. To do this, we gave definitions of the 21 PRF scales (see Jackson, 1984) to five psychology graduate students, who were then asked to rate, on 5-point scales, the extent to which each trait would be expected to predict “a high grade-point average.” The mean intercorrelation of the five judges’ ratings of the expected relations between personality and grade-point average was .63. This corresponds to a very high reliability of .90 for the aggregated judgments of the five-member panel.

The PRF trait rated by the average judge as the best predictor within the Conscientiousness factor was need for Achievement, and the best trait from the Openness to Experience factor was need for Understanding. These selections, which we used as our two lower level trait predictors in this study, are not surprising—Achievement involves a motivation to achieve high levels of performance in many domains including academics, whereas understanding involves an intellectual curiosity about many areas of knowledge.

**Derivation of Factor Measures**

In deriving measures of Conscientiousness and Openness to Experience with the present personality questionnaire data, we wanted to parallel a situation in which Big Five facet scales are simply added together to obtain a Big Five factor scale. This is the procedure used in the scoring of, for example, the NEO-PI-R scales, scores on six facet scales being summed to derive a composite score representing one of the Big Five domains (Costa & McCrae, 1992). To perform such a procedure with our data, we first consulted a study of the PRF in which its scales were claimed to define the Big Five factors.

Costa and McCrae (1988) factored the 21 PRF scales together with the five NEO-PI-R domain scales. They found five orthogonal personality factors that they interpreted as strongly resembling the traditional Big Five (see their Table 4), two of those factors being particularly pertinent to this study. Their Conscientiousness factor (on which the NEO-PI-R Conscientiousness scale loaded .77) was well defined by PRF Achievement, Cognitive Structure, Desirability, Endurance, Order, and (low) Impulsivity. Their Openness to Experience factor (on which the NEO-PI-R Openness to Experience scale loaded .78) was defined by PRF Autonomy, Change, Sentience, Understanding, and (low) Harmavoidance.

We decided to use the results by Costa and McCrae (1988) in this study as a basis for calculating composite scores on the Big Five factors of Conscientiousness and Openness to Experience. That is, we simply added together unit-weighted trait scores on the relevant PRF scales (keyed in the appropriate direction). First, however, we thought it necessary to verify that the present PRF data conformed at some level to the PRF data collected independently by Costa and McCrae. Accordingly, we factored our 717 students’ 21 PRF scale scores using
the method of principal components. We then rotated five factors to the pattern matrix reported by Costa and McCrae (1988, p. 263) using an orthogonal Procrustes procedure (Schönemann, 1966). The two factor structures were extremely alike. Only the Desirability scale did not have its highest loading on the same factor in both sets of data, and even that scale split across the same two factors in both analyses. Coefficients of congruence (Harman, 1976, p. 344) computed for the corresponding pairs of factors were consistently high, averaging .96 across the Big Five (Extraversion = .97, Agreeableness = .95, Conscientiousness = .95, Neuroticism = .94, and Openness to Experience = .97).

Our factor replications indicate an empirical stability to the PRF scales in terms of their five-factor structure (see also Jackson, Paunonen, Fraboni, & Goffin, 1996; Paunonen & Ashton, 1998; Rothstein et al., 1994). Accordingly, we proceeded with our plans and computed scores for the subjects in this study on Conscientiousness and Openness to Experience by simply adding together their scores on those factors’ salient PRF traits. Thus, Conscientiousness scores were derived as the sum of PRF Achievement, Cognitive Structure, Desirability, Endurance, Order, and (negative) Impulsivity. Openness to Experience scores were derived as the sum of PRF Autonomy, Change, Sentience, Understanding, and (negative) Harm-avoidance.

RESULTS

Final grades correlated significantly with students’ sex in our large sample of 717, with a value of $r = .11$ ($p < .01$), such that women had slightly higher grades than did men. Although that correlation was not substantial, student gender was nonetheless used as a covariate in all the correlations reported below. This ensured that the estimated validities of our personality measures were not inflated by sex differences common to predictor and criterion (see also Paunonen, 1989, p. 541).1

As far as the two selected Big Five factors predicting grades is concerned, the Conscientiousness composite (i.e., the sum of PRF scales of Achievement, Cognitive Structure, Desirability, Endurance, Order, and negative Impulsivity) was moderately correlated with final grades at $r = .21$ ($p < .001$). However, the single scale of Achievement, selected from the Conscientiousness factor on the basis of our judges’ ratings, did slightly better, correlating .26 ($p < .001$) with the same criterion. Although the difference between the two correlations was not significant ($z = 1.0, ns$), it is clear that adding the other five lower level traits to Achievement in deriving the Conscientiousness composite did nothing to improve on the efficiency of the latter in predicting final grades in this study.

1 The PRF item responses of the students in this study were scored on the 21 trait scales using a hand-scoring template. Consequently, the individual item data were not available for reliability analysis of those measures. The 16-item PRF scales, however, have a reported reliability averaging .72 (Jackson, 1984, p. 41). Note that the personality factor scale scores we calculated for the students are almost certain to have much higher reliabilities due to their greater lengths. In fact, Paunonen and Nicol (in press) derived Big Five factor scales similar to those of the present study by summing several PRF trait scale scores, finding reliabilities that averaged .90 for their composites.
More telling in the present instance were the data on the Openness to Experience factor. The Openness to Experience composite (the sum of PRF Autonomy, Change, Sentience, Understanding, and negative Harmavoidance), was not significantly correlated with grades in the present sample. In fact, that factor actually correlated slightly negatively ($r = -0.04$, $ns$) with that criterion. In contrast, with regard to the relatively narrow, constituent trait scale of Understanding, which was perceived by our judges as particularly important in this context, we found a correlation of $0.23$ ($p < .001$) with final grades. Thus, as in the case of the Conscientiousness factor described above, combining several lower level trait scales into a broader Openness to Experience factor scale resulted in a decrease in criterion predictability. In this case, however, the size of the decrease was highly significant ($Z = 5.18$, $p < .001$).

We also correlated the simple sum of the Achievement plus Understanding trait scales with final grades, resulting in a value of $0.31$ ($p < .001$). In contrast, the sum of the Conscientiousness plus Openness to Experience factors correlated only $0.15$ ($p < .001$) with grades, a significantly lower level of prediction ($Z = 3.20$, $p < .01$). Even the best predicting linear combination of Conscientiousness and Openness to Experience only managed a multiple correlation of $0.22$ ($p < .001$) with the criterion (or $0.23$ with student sex in the equation). These differences in predictive efficacy are substantial and noteworthy, and they strongly favor the narrower trait measures over the broader factor measures.

The relatively poor predictive performance of the broad Big Five factor variables relative to the narrow trait variables can be better understood by considering the individual correlations of relevant PRF trait scales with grades. As shown in Table 1 with regard to the Conscientiousness-related PRF traits, Achievement had the highest correlation with final grades. With the exception of Order, the other Conscientiousness-related traits mostly had significant, although somewhat smaller, correlations with that criterion. Moreover, those correlations were all in the expected direction. (Note that Impulsivity correlated in the expected negative direction with grade-point average).

In contrast to the Conscientiousness-related traits’ results shown in Table 1, things were quite different for the Openness to Experience-related traits. Specifically, only the PRF trait of Understanding was positively correlated with final grades (see Table 1). Sentience was essentially uncorrelated with that criterion, and the remaining Openness to Experience traits were actually correlated in the ‘‘wrong’’ direction with final grades. (Note that the positive correlation for Harmavoidance was expected to be negative.) Such differences in the direction of these five variables’ criterion correlations clearly contraindicate their combination into a broader composite measure as far as predicting the present criterion variable is concerned.
TABLE 1
Correlations of Grades with PRF Scales Related to Conscientiousness and Openness to Experience

<table>
<thead>
<tr>
<th>Conscientiousness-related scales</th>
<th>Conscientiousness composite</th>
<th>Openness to Experience-related scales</th>
<th>Openness to Experience composite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement (A)</td>
<td>.26***</td>
<td>Understanding (U)</td>
<td>.23***</td>
</tr>
<tr>
<td>Endurance (E)</td>
<td>.19***</td>
<td>Harmavoidance (H)*</td>
<td>.14***</td>
</tr>
<tr>
<td>Impulsivity (I)*</td>
<td>-.17***</td>
<td>Change (C)</td>
<td>-.10*</td>
</tr>
<tr>
<td>Desirability (D)</td>
<td>.16***</td>
<td>Autonomy (A)</td>
<td>-.08*</td>
</tr>
<tr>
<td>Cognitive Structure (C)</td>
<td>.11**</td>
<td>Sentence (S)</td>
<td>-.01</td>
</tr>
<tr>
<td>Order (O)</td>
<td>-.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conscientiousness composite</td>
<td>(A + E - I + D + C + O)</td>
<td>.21**</td>
<td></td>
</tr>
</tbody>
</table>

Note. Subject sex partialled from each correlation; N = 717. PRF = Personality Research Form.

* Personality scale expected to correlate negatively with grades.

** p < .01.

*** p < .001.

DISCUSSION

The results of this study suggest that academic performance can be better predicted by narrow personality traits such as achievement motivation and intellectual curiosity than by broad Big Five personality factors such as Conscientiousness and Openness to Experience, even though the factors subsume the traits. This finding is consistent with other results that have shown narrow traits to surpass broad factors in the prediction of many, although certainly not all, criteria (e.g., Ashton, 1998; Ashton et al., 1995; Mershon & Gorsuch, 1998; Paunonen, 1998; Paunonen & Nicol, in press; Paunonen et al., 1999; Rothstein et al., 1994).

Reasons for Predictive Differences

The reason for the superiority of the present narrow traits over broad traits in predicting the grades criterion, we believe, has to do with the nature of the content underlying those measures. A broad aggregate of several narrow
variables might not be predictive of a criterion even though some of the constituent variables are. This is because the individual variables each have their own component of trait-specific variance. And, the specific components of variance in the variables that do not predict the criterion could dilute, or even cancel out, by the process of aggregation, those components in the variables that do. Thus, such a broad measure would not be expected to predict the criterion as well as would some of its constituent narrow measures.

There are situations, of course, when an aggregate of several discrete measures would be expected to predict a criterion better than would any of the measures themselves. One situation would be when the variables aggregated into a composite are psychometrically parallel (see Paunonen & Gardner, 1991). In such a case, the variables would be composed of only common variance and error variance, with no specific variance. The more such parallel variables, each with an equal (nonzero) criterion correlation, are aggregated, the better the prediction by the composite. The improved prediction, however, would not be due to an increase in predictor breadth, because the underlying content domain would be unchanged (see Paunonen et al., 1999). Instead, the increase in prediction would be due to a decrease in error variance following aggregation.

There are also circumstances under which an aggregated personality measure that is very broad in terms of its content would be expected to predict better than would any of its constituent narrower measures. An obvious scenario is when the broad measure represents some sort of regression-weighted composite of the narrow measures, the weights chosen according to the variables’ independent contributions to criterion prediction. But even without using regression weights, a broad measure might generally be expected to predict more broadly. That is, because it covers a wider range of content, there is a greater likelihood that a broad measure of personality would better predict multiple criteria across different behavior domains than would any one narrow measure, which might do better at predicting only a few of those criteria. Some examples of these prediction differences favoring broad measures can be found in Paunonen (1998), Paunonen and Nicol (in press), and Paunonen et al. (1999).

Our point in this section is that we do not claim broad measures are always a poor choice for criterion prediction. Our thesis, instead, is that much important information can be lost when one’s focus on personality is exclusively at the Big Five factor level (e.g., see Ones & Viswesvaran, 1996). In contrast, with facet-level assessments nothing is lost because one can profit from the personality information available both at that lower level and, through the proper aggregation of those measures, at a higher level of assessment as well.
We do recognize that the prediction results of this study are specific to the present psychology course and measure of academic achievement. It is conceivable that different outcomes might ensue if, say, personality were used to predict grades in a statistics course or in a modern art course. It is also conceivable that the prediction results could have been different if our criterion of academic achievement were related to, say, grades based on classroom participation instead of grades based on written work. Relevant to these observations, Rothstein et al. (1994) found that none of the PRF personality scales predicted grades based on written work in a large sample of Master of Business Administration students (although verbal and quantitative ability scales did). In contrast, three PRF trait scales strongly predicted grades based on the students’ levels of class participation. Moreover, the broader Big Five factors overlying those narrow trait scales did not do nearly as well at predicting classroom participation, a result that is consistent with the main thesis of this article.

A comment or two is warranted at this point about the absolute sizes of the correlations between personality and grades reported in this study, which might be considered small by some. First, we note that those coefficients are consistent with similar results reported in other studies of personality and academic achievement (e.g., Goff & Ackerman, 1992; Rothstein et al., 1994). Second, the present results are based on grade-point averages in only one course—aggregating grades over several courses should lead to a more reliable criterion of academic achievement and, consequently, higher correlations with the personality predictors. And third, personality is only one plausible component of academic achievement. With regard to this latter point, the most salient other factor is cognitive ability, which can operate independently of personality in the prediction of grades. For instance, Rothstein et al. (1994) found that scores on the Verbal and Quantitative subscales of the Graduate Management Admission Test (Shrader, 1979) significantly predicted classroom participation grades in a graduate business course ($R = .22$), but that two PRF trait scales added independently and substantially ($p < .001$) to that prediction ($R = .41$).

Another independent contributing factor to course grades might be class attendance. In fact, attendance was noted during the first 4 days of class for some of the sections of the present psychology course. We were thus able to derive an absenteeism variable, ranging from 0 days to 4 days, for some of the students ($N = 232$). In a multiple-regression analysis, Absenteeism added substantially ($p < .001$) to the sum of the trait scales of Achievement and Understanding in predicting grades, resulting in a strong incremented multiple correlation climbing from .31 to .42 (or .43 with subject sex in-
cluded as a predictor). We would expect that adding a measure of cognitive ability to the equation would further raise this level of prediction.

**Beyond the Big Five**

One of the most interesting results of the present study was the finding that some PRF traits relevant to Openness to Experience showed significant positive correlations with our measure of academic performance, but that other traits relevant to that factor showed significant negative correlations (see Table 1). Specifically, whereas the trait of Understanding was positively correlated with final grades, three other Openness to Experience-relevant PRF traits (Change, Autonomy, and low Harmavoidance) were negatively correlated with that criterion. This is one of the clearest examples yet of the narrow traits defining the same Big Five factor showing contrary correlations with a relevant criterion.

Should one expect a Big Five factor’s constituent personality traits all to have similar correlations with other variables? Not necessarily. Considering the nature of the openness traits referred to in this study, and the nature of our academic achievement criterion, our results do not seem particularly surprising to us, and they are certainly not inexplicable. Perhaps there is no good reason to expect change, autonomy, and harmavoidance to be positively related to overall performance in the present psychology course because those traits have to do more with an openness to different sensations than with an openness to intellectual ideas. But a point of our study is that the differential operation of these personality traits in the prediction of final grades would never have been discovered had we used only Big Five personality factors as predictors of the criterion.

What the present results underscore is a declaration we have made in the past—that much can be lost in both behavior prediction and behavior explanation when combining lower level traits into their higher level factors. This is not to say that, for example, the Openness to Experience factor has no utility for prediction. As suggested earlier, that factor could conceivably do well at predicting a broader academic achievement criterion than that used here or at predicting a grades criterion for which the sensation aspect of the openness domain plays a greater role. However, the relative predictive merits of a Big Five personality factor versus the facets of that factor must be evaluated empirically in each case and not just assumed.

Our discussion at this juncture raises an interesting question. Are the variables implicated by the Five-Factor Model of personality structure sufficient in their entirety, either at the facet level or factor level, as predictors of any and all personality-relevant criteria? Certainly if the Big Five variables cover the complete domain of personality, and if a criterion variable’s determinants include a personality component, then some Big Five variable should predict
that criterion. We think, however, that such is not always the case. Our rea-
soning is based on some of our own recent work with variables of personality
that do not appear to fit well into the traditional five-factor space.

Ashton, Lee, and Son (2000) have presented cross-replicated lexical data
supporting the view that a dimension of behavior related to honesty might
be considered a (6th) basic personality factor. And, Paunonen and Jackson
(2000) have reported evidence, based on a reanalysis and reinterpretation of
data published by Saucier and Goldberg (1998), that at least 9 or 10 personal-
ity traits have empirical and theoretical relations with the Big Five so low
as to question their allegiance to that factor space. Now, if some personality
variables actually exist beyond the space of the Big Five, it means that they
have substantial amounts of variance independent of those factors and facets.
It is conceivable, therefore, for that variance to predict criteria not predicted
by the traditional Big Five variables.

The search for personality variables that have little overlap with Big Five
dimensions is important for both practical and theoretical reasons. The practi-
cal reason is that such variables can add significantly to our ability to predict
socially important criteria beyond the level achieved by Big Five factors and
facets alone. The theoretical reason is that such variables, should they exist,
raise questions about the presumed adequacy of the Five-Factor Model as a
complete description of personality structure.

CONCLUSIONS

There is a growing tendency for researchers evaluating the determinants
of behavior to rely exclusively on measures of the Big Five personality fac-
tors. In this article, we have argued that such a strategy is short-sighted be-
cause those factors, although they account for much of the variation in human
behavior, do not account for all of that variation. The data of the present
study, together with data published elsewhere, support our claim that sub-
stantial benefits can be obtained in both behavior prediction and behavior
explanation by looking within and beyond the Big Five factors at other vari-
ables. Those other variables include the narrow facets of the Big Five that
are aggregated to form the broad factors themselves. The facets can contain
nontrivial amounts of trait-specific variance, independent of their common
factor variance, that should be exploited for criterion prediction. Also to be
considered are personality variables that have been identified as possibly not
belonging within the factor space of the Big Five. Such variables can add
substantially to the levels of behavior prediction realized solely by the tradi-
tional Big Five factors and facets.

REFERENCES


Paunonen, S. V., Rothstein, M. G., & Jackson, D. N. (1999). Narrow reasoning about the use


