

VALIDITY OF AN INTUITIVE PERSONALITY SCALE: PERSONAL RESPONSIBILITY AS A PREDICTOR OF ACADEMIC ACHIEVEMENT

JOSÉE MARTEL, STUART J. MCKELVIE, AND LIONEL STANDING
Bishop's University

Fifty undergraduate students (27 female, 23 male) completed measures of verbal intelligence and locus of control, together with the Responsibility scale of the California Psychological Inventory and a student-written scale of Personal Responsibility. On the basis of correlational and multiple-regression analyses, it was found that Intelligence and Personal Responsibility were jointly the most efficient predictors of a subject's mean course grade, and that the latter was the best single predictor. These results add further support to Jackson's (1971) contention that naive item writers can create personality scales of equivalent or greater validity than more formal empirically-derived scales.

In 1971, Jackson offered the provocative suggestion that non-professionals might be able intuitively to create personality test items that display superior empirical validity to those selected by standard psychometric procedures. Two studies (Ashton and Goldberg, 1973; Jackson, 1975) have provided evidence that this challenge has been met, in that items written by psychology students to represent a total of six scales of the California Psychological Inventory (CPI) were found to possess superior or equivalent concurrent validity to those from the official scales. In seeking to generalize Ashton and Goldberg's (1973) results, Jackson (1975)

Responsibility for this paper is shared equally among the three authors, whose names are listed alphabetically. The research was reported in an honours dissertation submitted by the first author to the Department of Psychology, Bishop's University. Reprints may be obtained from L. Standing, Department of Psychology, Bishop's University, Lennoxville, Quebec, J1M 1Z7, Canada.

Copyright © 1987 Educational and Psychological Measurement

utilized a somewhat different population of item writers (undergraduate vs. graduate students), different CPI scales and different methods for obtaining criterion data. However, both studies employed ratings (either of self or of peers), which although often useful as criteria for personality tests where objective data are difficult to obtain, are also subject to a number of errors (Anastasi, 1982, pp. 142, 622–612). One purpose of the present study was to extend Jackson's challenge to a CPI scale (Responsibility) which has promise as a predictor of a more objective criterion on a dimension other than the construct in question: academic achievement as measured by course grades.

The majority of attempts to predict academic performance as measured by instructor grades have utilized tests of general ability, often interchangeably labelled intelligence, academic potential or scholastic aptitude (Anastasi, 1982, pp.305–306, 313). For primary and high school students, validity coefficients have been moderate, ranging from .40 to .60 (Anastasi, 1982, pp.307–321; Butcher, 1968, p.282; Kaplan and Saccuzzo, 1982, pp. 301–305). For university students, they have often been weaker, commonly with values between .30 and .50 for undergraduates (Edward and Waters, 1981; Kaplan and Saccuzzo, 1982, p. 303; Slack and Porter, 1980), and even lower values for graduate students (Boudreau, Killip, MacInnis, Milloy, and Rogers, 1983; Willingham, 1974). It is likely that the latter values reflect a restriction in range (Willingham, 1974) and that non-intellectual variables such as motivation (Cattell, Eber, and Tatsuoka, 1970; Edwards and Waters, 1981; Willingham, 1974) might improve predictive accuracy.

Indeed, Cattell and his associates (Cattell and Butcher, 1968; Cattell, et al., 1970) claim that the addition of personality to ability variables can significantly increase the percentage of achievement variance accounted for with both high school and college students, and their regression equations suggest that the factors of Conscientiousness and Individual Resourcefulness are particularly important. This argument is consistent with the finding by Holland (1959) that validity coefficients for the Scholastic Aptitude Test (SAT) as a predictor of university grades were significantly improved by the addition of scores from the CPI, one of which was Responsibility. As Stroup (1970) points out, however, Holland's sample consisted of exceptionally talented freshmen, thereby restricting the range of the intelligence test scores even further. When Stroup replicated the study with a broader sample of undergraduates, the CPI was not as powerful a predictor as the SAT, although it did provide a slight increase in R (from about .50 to .57). In a multiple regression

analysis of the CPI variables alone (which produced an R of about .35), the Responsibility scale again correlated significantly with grades, although it dropped out when the SAT scores were added to the equation. More recently, Gough and Lanning (1986) derived an equation to predict college grades from six personality scales of the CPI. Commensurate with the value obtained by Stroup (1970), these authors cite correlations of .30 to .40 between scores computed from their equation and grade point average, and they draw attention to the positive weighting of Responsibility and Achievement via Independence, suggesting that "constructive internalization of societal norms" and "ability to work well in ambiguous situations" are important factors in academic performance.

Taken together, these findings imply that a dimension of "personal self reliance" may contribute to the prediction of scholastic achievement. This conclusion is also consistent with the evidence (Findley and Cooper, 1983; Lefcourt, 1982) linking locus of control with academic achievement. Although some of these studies have produced mixed results, the majority have demonstrated that individuals who have more internal beliefs and "feel personally responsible for things that happen to them" tend to have greater academic success than those who have external beliefs and "feel that outcomes in life are totally beyond their control" (Findley and Cooper, 1983).

The present study therefore investigated both intelligence and personal self reliance as predictors of academic achievement in university students. The former was measured by a standardized test of verbal intelligence (Lorge-Thorndike Intelligence Test; College Edition) and, in view of the preceding research, the latter was measured by Rotter's Social Reaction Inventory, giving locus of control scores, and by the Responsibility subscale of the CPI. However, in light of the two studies meeting Jackson's challenge with CPI scales, a new measure, termed the Personal Responsibility Scale, was added. This was composed, without aid, by a final-year undergraduate psychology student (the first author). In writing the items, an attempt was made to focus on concrete behaviors (e.g., "I pay my bills immediately"; "I drink and drive") representing personally responsible (or irresponsible) actions leading directly to long-term benefits (or harm) to the self and society. Although the notion behind this scale overlaps with those underlying the Social Reaction Inventory and the CPI Responsibility subscale, it differs from the former in consistently emphasizing behavior rather than cognitions, and from the latter in focusing on the respondent's own behavior rather than that of people in general (e.g., one item on the

CPI scale is "We ought to pay our elected officials better than we do"). The new scale provided testees with five response alternatives, ranging from "never" to "always," rather than the dichotomous choice of the previous two scales.

Two hypotheses were proposed: (a) that the addition of the three personality scales to the intelligence scale will improve the predictability of the academic achievement criterion and (b) the new Personal Responsibility subscale will account for at least as much criterion variance as the corresponding CPI scale.

Method

Subjects

Fifty undergraduate students (27 female, 23 male) were randomly selected from the inhabitants of a university residence and were then individually asked to participate; volunteer bias and coercion were thus avoided (Argyris, 1968).

Materials

The following tests were given to each subject.

The Lorge-Thorndike Intelligence Test, College Edition, Level H (I) (Lorge, Thorndike, and Hagen, 1964). Only the verbal battery was used. It consists of five subtests and is timed at 35 minutes. According to the manual, reliability of the verbal scale is high, with a split-half estimate of .93 and alternate form estimate of .90. Although it does not provide data relating test scores to college grades, the general technical manual (Lorge and Thorndike, 1957) cites a correlation of .67 for ninth graders and correlations of .77 or better between scores on non-college levels of the test and other standardized tests of intelligence. In addition, a study by Mendels (1973) found correlations from .42 to .62 between test scores and academic achievement for kindergarten students.

The Social Reaction Inventory (Locus of Control, LOC) (Rotter, 1966). This scale comprises 23 forced-choice items, plus six fillers, a higher score indicating a more external orientation. The reliability of the test is moderate to high, with a split-half estimate of .65, KR20 coefficients of .69 to .76, and test-retest reliabilities after one month of .60, .78, and .83.

Responsibility Scale (R) (Gough, 1975). This subscale of the California Psychological Inventory consists of 42 true-false items in which a higher score indicates a higher level of responsibility. Test-

retest reliability estimates range from .85 to .65 for intervals of a few weeks to one year respectively.

Personal Responsibility Scale (PR). This scale, written by the first author, consists of 30 positively or negatively worded items, dealing with behavior in a wide range of concrete situations chosen to avoid examples inappropriate to a student population. Five response categories are used whereby the subject indicates how frequently he or she engages in the behavior referred to in each question. Each response is scored from 0 to 4, a high score (maximum 120) representing high personal responsibility. The test requires about 10 minutes, although no time limit is set. With a separate group of 30 student subjects, the split-half reliability (corrected by the Spearman-Brown formula) was .83 and, over a period of one week, the test-retest reliability was .79. The PR scale and scoring key are given in the Appendix.

General Information Questionnaire. This covered the subjects's program and courses taken at the time of testing.

Procedure

The sample was selected early in the fall semester (October). Each subject was given all five questionnaires in a random order, in a single session; testing was performed in small groups (8–12). The subject also gave written permission for the experimenter (first author) to view his/her official transcript of course grades at the end of the semester (December).

Results

Each subject's official course marks (for at least three and usually five semester courses) were first converted to a mean percentage grade. Overall mean grades, together with the mean scores in each of the four predictor scales, are given in Table 1. The zero-order

TABLE 1
Correlation Matrix and Mean Scores for Grades and Four Predictor Variables

	G	I	LOC	R	PR	M	SD
G: Grades	—					62.90	15.32
I: Intelligence	.372**	—				53.48	11.16
LOC: Locus of Control	-.141	-.071	—			11.51	3.81
R: Responsibility	.374**	.199	-.188	—		25.60	3.67
PR: Personal Responsibility	.377**	.172	-.244	.383**	—	84.30	9.46

Note. Maximum scores for each variable were 100 (G), 85 (I), 23 (LOC), 41 (R), and 120 (PR).

** $p < .01$.

TABLE 2
Regression Equation for Best Prediction of Grades, R^2 , Significance of R^2 , and Significance of Decrease in R^2 , as Each Predictor is Removed. ($N = 50$)

Step	Regression Equation	R^2	F and incremental F
1	$G = .40I + .09LOC + .94R + .40PR - 14.7$.28	$F(4, 45) = 4.42^{**}$
2	$G = .40I + .94R + .40PR - 16.5$.28	$F(3, 46) = 6.00^{**}$ $F(1, 45) = .03$
3	$G = .43I + .54PR - 5.9$.24	$F(2, 47) = 7.39^{**}$ $F(1, 46) = 2.69$
4	$G = .64PR + 9.4$.14	$F(1, 48) = 7.95^{**}$ $F(1, 47) = 6.00^*$

Note. In the regression equations, b -weights are given for each variable.

* $p < .05$. ** $p < .01$.

correlations among the five research variables are also given in Table 1. It can be seen that four of them were statistically significant: grades and intelligence, grades and responsibility, grades and personal responsibility, and responsibility and personal responsibility (all $p < .01$).

A stepwise multiple regression analysis of the data was then performed, using the BACKSTEP computer program of Gorman and Primavera (1981). At Step 1, grades were significantly predicted by all four variables, the multiple correlation being .53 ($R^2 = .28$, $p < .01$). The weakest predictor was then dropped at each subsequent step, as shown in Table 2. When the LOC, R, and I variables were successively dropped as predictors, the decrease in proportion of variance accounted for reached significance only at the final step when intelligence was dropped. PR emerged at Step 4 as the best single predictor of grades ($R^2 = .14$, $p < .01$).

Discussion

Both the zero-order correlation coefficients and the regression analyses support the first research hypothesis, that personality variables add significantly to the prediction of grades. Intelligence (I), Responsibility (R) and Personal Responsibility (PR) were all significantly associated with course grades, and the most efficient equation, as defined by the minimum number of predictors with no decrease in R^2 , was at Step 3, where both I and PR were retained. This result is consistent with the previous work (Cattell and Butcher, 1968; Cattell, et al., 1970; Holland, 1959; Stroup, 1970) showing that both personality and intelligence test scores are useful for predicting academic achievement. It does suggest, however, that locus of control is not an important predictor. Although this

conclusion appears at odds with the significant relationship derived by Findley and Cooper (1983), their overall effect size across 275 tests was $r = .18$, which is only "small to medium" (p. 419), and the authors indicate that some studies did show a zero-order relationship. It is also notable that Edwards and Waters (1981) found that locus of control did not contribute to the prediction of grade point average when it was included with a standardized intelligence test.

The present findings also support the second hypothesis that PR would be a better predictor than the R scale. Although they were correlated with each other, the latter dropped out between Steps 2 and 3, and R was the best single predictor of all four variables, having the highest simple correlation with grades and being the last variable to remain. Although this result is consistent with Holland's (1959) finding that personality is a more powerful predictor of grades than is intelligence, it is somewhat surprising in view of the previous research (Stroup, 1970) relating intelligence to grades in a broadly-based student sample, as was used here. However, it provides further strong support for Jackson's (1971) contention that items produced by a naive writer can be more successful predictors than empirically-derived ones. It also corroborates the specific findings of Ashton and Goldberg (1973) and Jackson (1975) that CPI scales can easily be improved upon.

It is notable that the equations at Step 1, with all four predictors, and Step 3, with only intelligence and personal responsibility, respectively accounted for 28% and 24% of the variance in course grades ($R = .53$ and $.49$). Despite being somewhat attenuated by the absence of a nonverbal test, which is usually included in measures of general intelligence (Anastasi, 1982, pp. 315, 348, 375), these are similar to the best values reported by Stroup (1970) using both verbal and mathematical sections of the SAT and three CPI variables, and to the results typically obtained from attempts to predict grade point average from the SAT plus previous grades (Goldman & Slaughter, 1976). Moreover, although the present criterion is more objective than the ratings utilized by previous investigators of Jackson's challenge, it represents a mean of grades in diverse courses that involve different grading practices, rendering it inherently difficult to predict (Goldman and Slaughter, 1976). Goldman and Slaughter (see also Willingham, 1974) argue that, for prediction in higher education, it may be necessary to use and develop different criteria of success such as single course grades or grade point average based on a combination of courses weighted according to dimensions such as difficulty. In the present context, one useful criterion might be a standardized achievement test such as the

Graduate Record Advanced score in the student's major area of specialization. It would be interesting to observe whether the PR scale would demonstrate improved predictive power for this criterion, particularly if combined with verbal and nonverbal intelligence, and previous grades.

In summary, the present study shows that a measure of responsibility created by a naive item writer can contribute significantly to the prediction of university grades. More specifically, it suggests that the successful student possesses not only intellectual skills but also habits of personal self reliance.

REFERENCES

- Anastasi, A. (1982). *Psychological testing*, (5th ed.). New York: Macmillan.
- Argyris, C. (1968). Some unintended effects of rigorous research. *Psychological Bulletin*, 70, 185-197.
- Ashton, S. G. and Goldberg, L. R. (1973). In response to Jackson's challenge: The comparative validity of personality scales constructed by the external (empirical) strategy and scales developed intuitively by experts, novices and laymen. *Journal of Research in Personality*, 7, 1-20.
- Boudreau, R. A., Killip, S. M., MacInnis, S. H., Milloy, D. G., and Rogers, T. B. (1983). An evaluation of Graduate Record Examinations as predictors of academic success in a Canadian context. *Canadian Psychology*, 24, 191-199.
- Butcher, H. J. (1968). *Human intelligence: Its nature and assessment*. London: Methuen.
- Cattell, R. B. and Butcher, H. J. (1968). *The prediction of achievement and creativity*. New York: Bobbs-Merrill.
- Cattell, R. B., Eber, H. W., and Tatsuoka, M. M. (1970). *Handbook for the Sixteen Personality Factor Questionnaire (16 PF)*. Champaign, IL: Institute for Personality and Ability Testing.
- Edwards, J. E. and Waters, L. K. (1981). Moderating effect of achievement motivation and locus of control on the relationship between academic ability and academic performance. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 41, 585-587.
- Findley, M. J. and Cooper, H. M. (1983). Locus of control and academic achievement: A literature review. *Journal of Personality and Social Psychology*, 44, 419-427.
- Goldman, R. D. and Slaughter, R. E. (1976). Why college grade point average is difficult to predict. *Journal of Educational Psychology*, 68, 9-14.
- Gorman, B. S. and Primavera, L. H. (1981). BASIC BACKSTEP: A simple backward-selection multiple regression program for mini-computers and microcomputers. *Behavior Research Methods and Instrumentation*, 13, 703.
- Gough, H. G. (1975). *Manual for the California Psychological Inventory*. Palo Alto: Consulting Psychologists Press.

- Gough, H. G. and Lanning, K. (1986). Predicting grades in college from the California Psychological Inventory. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 46, 205–213.
- Holland, J. L. (1959). The prediction of college grades from the California Psychological Inventory and the Scholastic Aptitude Test. *Journal of Educational Psychology*, 50, 135–142.
- Jackson, D. N. (1971). The dynamics of structured personality tests: 1971. *Psychological Review*, 78, 229–248.
- Jackson, D. N. (1975). The relative validity of scales prepared by naive item writers and those based on empirical methods of personality scale construction. *EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT*, 35, 361–370.
- Kaplan, R. M. and Saccuzzo, D. P. (1982). *Psychological testing. Principles, Applications, and Issues*. Belmont, CA: Wadsworth.
- Lefcourt, H. M. (1982). *Locus of control* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Lorge, I. and Thorndike, R. L. (1957). *Technical Manual for the Lorge-Thorndike Intelligence Tests*. Boston: Houghton Mifflin.
- Lorge, I., Thorndike, R. L., and Hagen, E. (1964). *Manual for Administration of the Lorge-Thorndike Intelligence Test*. Boston: Houghton Mifflin.
- Mendels, S. B. (1973). The predictive validity of the Lorge-Thorndike Intelligence Tests at the kindergarten level. *Journal of Educational Research*, 66, 320–322.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external locus of control of reinforcement. *Psychological Monographs*, 80, 1–26.
- Slack, W. V. and Porter, D. (1980). The Scholastic Aptitude Test: A critical appraisal. *Harvard Educational Review*, 50, 154–175.
- Stroup, A. L. (1970). The prediction of academic performance from personality and aptitude variables. *Journal of Experimental Education*, 38, 83–86.
- Willingham, W. W. (1974). Predicting success in graduate education. *Science*, 183, 273–278.

APPENDIX

The Personal Responsibility Scale

Instructions

Please read carefully and answer these questions as honestly as you can. Thank you.

Indicate the proper answer by circling the appropriate letter:

- A) Never
- B) Seldom
- C) Sometimes
- D) Often
- E) Always

Items

1. If a car broke down that I was driving in, I would leave the location and get someone else to take care of the situation. (N)
2. If I saw someone steal something, I would report it to the authorities. (P)
3. I drink and drive. (N)
4. I give things back that I borrow. (P)
5. When I say things to people that I do not mean, I take it back. (P)
6. If I were to engage in sexual relations, I would use some form of contraceptives. (P)
7. I do not care what I say to people as long as I am honest with them. (N)
8. I miss classes. (N)
9. I run out of money. (N)
10. I miss buses, trains, rides, planes etc. . . . (N)
11. I have overdrawn my bank account. (N)
12. I miss meetings which I am supposed to attend. (N)
13. When I break something not belonging to me, I replace it. (P)
14. I drive without car insurance. (N)
15. I pay my bills immediately. (P)
16. I hand in papers late. (N)
17. I break appointments. (N)
18. I initiate romantic relationships, I do not let them happen. (P)
19. My family is responsible for me attending university. (N)
20. When someone does something wrong to me, I do not do anything about it. (N)
21. I go to church because my parents want me to. (N)
22. I always read a paper over before I hand it in. (P)
23. I let other people pick up after me. (N)
24. I take credit for things other people have done. (N)
25. I put a seat belt on when I enter a car. (P)
26. I jay-walk. (N)
27. I have cheated on an exam. (N)
28. If I thought I would not get caught stealing something, I would steal it. (N)
29. I have regular check-ups with the doctor. (P)
30. I return books my library books on time. (P)

Scoring

The letters P and N refer to items indicating personally responsible and not responsible behavior, respectively. In actual administration, they are replaced by A B C D E for each question. For P items, 0, 1, 2, 3, or 4 points respectively are awarded for responses A through E. For N items, 0–4 points are given for responses E through A. These points are then summed to yield a total, with a possible range of 0–120.