

1995

Students' sophistication level and confidence in testing procedures predicts susceptibility to the Barnum effect

Thomas R. Rutledge
University of Alaska at Anchorage

Follow this and additional works at: <https://scholar.utc.edu/mps>



Part of the [Psychology Commons](#)

Recommended Citation

Rutledge, Thomas R. (1995) "Students' sophistication level and confidence in testing procedures predicts susceptibility to the Barnum effect," *Modern Psychological Studies*: Vol. 3 : No. 2 , Article 7.
Available at: <https://scholar.utc.edu/mps/vol3/iss2/7>

This articles is brought to you for free and open access by the Journals, Magazines, and Newsletters at UTC Scholar. It has been accepted for inclusion in Modern Psychological Studies by an authorized editor of UTC Scholar. For more information, please contact scholar@utc.edu.

Students' Sophistication Level and Confidence in Testing Procedures Predicts Susceptibility to the Barnum Effect

Thomas R. Rutledge

University of Alaska at Anchorage

Abstract

The present study examined the relationship between students' level of sophistication and confidence in personality assessment procedures with their susceptibility to the Barnum effect—the tendency for individuals to accept highly generalized, ambiguous profiles as accurate descriptions of their personality. Thirty-five university students (22 females, 13 males, mean age 26) completed a brief personality questionnaire under the impression they would be receiving an interpretation from a) a masters level clinician, b) a clinical psychologist (PhD), or c) a form of computer assessment. A pretest asking for the subject's age, sophistication, and their perceptions of the three assessment sources accompanied the questionnaire. Subjects received one of two profiles categorized by moderate or high favorability and were asked to rate both the accuracy and degree to which the profile described their unique personality. While neither the main effects for feedback source or favorability proved significant, multiple regression analyses found subjects' sophistication and initial confidence in the personality assessment procedures to be effective predictors of their accuracy and uniqueness ratings. The results suggest that cognitive variables can mediate susceptibility to the Barnum effect.

Since Forer's (1949) original article, researchers investigating what has been labeled the "Barnum effect"—after P.T. Barnum's famous phrase, "There's a sucker born every minute"—have repeatedly demonstrated that people will accept personality interpretations comprised of vague, highly generalized statements as accurate descriptions of

their personalities (Carrier, 1963; Dana & Fouke, 1979; Fichter & Sunerton, 1983; Halperin & Snyder, 1979; Lattal & Lattal, 1967; Snyder & Newburg, 1981). Furthermore, the effect appears robust across different assessment procedures (e.g., interviews, projective tests, and objective tests), test interpreters (e.g., clinicians, computers, and undergraduate students) and occupational backgrounds (Dana & Graham, 1976; Dmitruk, Collins, & Clinger, 1973; Forer, 1949; Snyder, 1974; Snyder & Larsen, 1972; Snyder, Larsen, & Bloom, 1976).

The experimental procedures for the overwhelming majority of studies in this area have been the same. Subjects, typically undergraduate students, complete some form of assessment measure under the impression their responses will be analyzed by a trained interpreter and are at a later date given an ambiguous profile and asked to evaluate the accuracy of the interpretation. Although these methods remain characteristic of Barnum research, the past two decades have brought a subtle shift in the nature of the experimental questions posed. Research efforts in recent years have explored a variety of situational and intrapersonal factors not in an attempt to determine if the effect occurs, but why the effect occurs. (Furnham & Schofield, 1987; Snyder, Shenkel, & Lowery, 1977). As a result of these efforts, the Barnum effect appears primarily to be a product of both the relevance (meaning the degree to which the profile is perceived to be intended for the specific subject) and favorability of the bogus profiles (Dickson & Kelly, 1985).

Profile favorability (i.e., how complimentary the profile is for the subject) in particular has been a frequent topic of investigation (Furnham & Schofield, 1987). Studies in which the favorability of the feedback has been directly manipulated indicate not only that subjects accept favorably worded feedback more than negatively phrased profiles but that subjects' attitudes towards the assessment source are positively influenced as well (Collins, Dmitruk, & Ranney, 1977; Snyder & Shenkel, 1976).

Favorability has also been found to interact with the status of the interpretation source. Halperin, Snyder, Shenkel, and Houston (1976) found that acceptance effects differ between positively and negatively worded profiles only when low status assessors administer the feedback. Under high status conditions alone, both negative and positive feedback are rated as accurate. These findings place considerable responsibility upon the clinician, who is likely to be perceived as a high status source of feedback by the client (Weinberger & Bradley, 1980).

However, more recent interpretations of this literature suggest that the impact of favorability may be better understood as a covariate of the bona fide accuracy of the profiles and not as a direct contributor to the acceptance ratings of the feedback itself (Dickson & Kelly, 1985). In short, this position contends that more favorable profiles are more truthful. In one of a handful of studies that have addressed this notion (Snyder & Shenkel, 1976), no significant differences were found in the acceptance of profiles weighted by level of favorability after controlling for base rate accuracy, indicating that the specific role favorability plays in the Barnum effect has yet to be explained.

An important area that has received little attention in the literature is the conjectured moderating effects of "sophistication" variables in the clients' acceptance of Barnum profiles. Theoretical literature suggests that more knowledgeable or experienced clients would be less susceptible to the Barnum effect (Lattal & Lattal, 1967; Stagner, 1958). Two independent studies largely account for the attempts to explore these parameters. Forer (1949) and Stagner (1958) reported no differences in acceptance effects across subjects' ages or occupational backgrounds, supporting the hypothesis that the Barnum effect would generalize to populations outside of academic settings. Taking a separate approach, other researchers have compared acceptance ratings across educational levels and between graduate and introductory level psychology

students. These studies have found reliable group differences (Greene, 1977; Greene, Harris, & Macon, 1979; Schroeder & Lesyk, 1976).

However, it is arguable whether either group of variables optimally defines the sophistication construct and not an underlying covariable. Increased experience in educational settings, particularly with clinically related training, is likely to increase students exposure to and awareness of the many imperfections in psychological testing, making them somewhat more skeptical of personality assessment in general. If this supposition is valid, then attitude-related variables may influence an individual's susceptibility to the Barnum effect. The results of several studies offer support for this position (Glick, Gottesman, & Jolton, 1989; Snyder & Shenkel, 1976). In a comparison of different assessment procedures, Snyder and Shenkel (1976) found a significant relationship between subjects' confidence in the diagnostician's skills and their subsequent acceptance of the interpretation. The conclusion drawn from this research is that subjects' faith and confidence in the assessment procedures can affect their ensuing evaluation of the feedback they receive.

In a related area, Greene (1977) found that when questioned directly, subjects could identify the generality of the Barnum profile even while rating it as accurate. Although the methodology and conclusions of this study have been criticized by some (Baillargeon & Danis, 1984; Snyder, Handelsman, & Endelman, 1978), the general finding that subjects are able to discriminate between accuracy and uniqueness under particular circumstances has been replicated (Harris & Greene, 1984).

The purpose of the present study was to further explore the relationship between profile favorability, the subject's level of sophistication and confidence in the assessment procedures, and their acceptance of Barnum profiles on dimensions of accuracy and uniqueness. Specifically, it was hypothesized that (a) sophistication variables, operationalized in this study to include grade point average

(GPA), number of years in college, and GPA x years in college interaction, would predict subjects' subsequent accuracy and uniqueness scores, and (b) subjects indicating higher levels of confidence in the testing procedures would rate their profiles to be both more accurate and more uniquely descriptive of their personalities.

Method

Subjects and Design

Thirty-five students (22 females and 13 males, age range from 18-51, $M = 26$), enrolled in three laboratory sections of the Experimental Psychology (PSY 261) course at the University of Alaska at Anchorage, participated in the study in return for credit on a laboratory exercise. Individuals were informed that, through random assignment, their personality questionnaires were to be analyzed and scored by either a masters level clinician, a clinical psychologist (PhD), or a form of computer interpretation. Standard and highly favorable profiles were randomly assigned in equal numbers to students in each of the laboratory sections, creating a 2 x 3 (levels of favorability x computer, masters level, and doctoral level assessment source) design.

Measures

The questionnaire packet completed by each subject consisted of three sections. Section I contained items concerning the subjects' gender, age, major, number of psychology credits, estimated GPA, and the number of years they had spent in college. The subjects were informed that this information would be used by the interpretation source in supplement to the personality test.

Section II asked each subject to rate the three assessment modalities (computers, masters level clinicians, and doctoral level clinicians) on a scale from 1 to 10 (*higher scores indicating greater confidence*) in terms of the confidence the subject would place in an interpretation from a specific source. These separate scores were summed and averaged to

produce an overall indicator of confidence.

The final section was comprised of the NEO-Five Factor Inventory (Costa & McCrae, 1992) a 60-item test measuring five fundamental dimensions of personality, including neuroticism, extraversion, openness, agreeableness, and conscientiousness. The standard interpretation profile was taken from Forer's (1949) original study. Several statements were modified by the researcher to enhance its apparent favorability for the highly favorable version.

Procedure

Following a 45 min lecture on issues and research in personality by the experimenter (a male laboratory assistant), the students were asked to participate in a research project examining the relative accuracy of different forms of personality assessment. Students were asked to complete a questionnaire packet and were informed which assessment source would be providing them with an individual interpretation in the following lab. It was explained that each had been randomly assigned to the interpretation conditions, and that the sincerity of their responses and profile evaluations were essential to producing meaningful results.

The packets were returned in the next class session with an attached interpretation profile and a second sheet asking the subject to respond to two items concerning the profile (a) "Give an overall rating of the accuracy of your profile in percentage terms. State your answer in numbers from 0-100% according to how well your profile describes you", and (b) "regardless of the source of your interpretation, an effort was made to create a profile that fit you as a unique individual. Specify to what degree do you feel that this personality profile uniquely describes you, from 1 (*not at all*) to 5 (*extremely well*)."

Subjects were given 10 min to complete the exercise and asked to remain quiet during this period. Following their completion, the experimenter collected the

questionnaires and provided a complete debriefing of the true research design, an explanation of the necessity of the deception involved, and a discussion of the Barnum effect.

Results

Manipulation Check

A one-way analysis of variance (ANOVA) was conducted initially in order to verify that subjects held varying levels of confidence in the three assessment sources. As expected, this result proved significant, $F(2, 32) = 8.41$, $p < .01$. Follow-up analyses were performed to clarify the nature of these differences. Use of Scheffe's post hoc comparison procedure revealed that subjects placed significantly higher confidence in the clinical psychologists ($M = 7.2$ on a 10 point scale) than in either the masters level ($M = 5.5$) or computer assessment ($M = 4.9$) sources ($ps < .05$ & $< .01$ for these differences respectively). The marginal differences in confidence levels subjects expressed for the masters level clinician and computer assessment sources were not significant. A 2 (high favorable and standard profile) x 3 (computer, masters level, and doctoral level assessment source) analysis of variance was performed on the average of the subject's accuracy and uniqueness scores (justification for this conglomeration is given below). Neither the main effects for assessment source, $F(2, 28) = .12$, favorability, $F(1, 28) = .67$, or their interaction, $F(2, 28) = .19$, approached significance.

Multiple Regression Procedure

A hierarchical multiple regression analysis was conducted to examine the hypothesis that the subjects' GPA, number of years in college, and GPA x years in college interaction would predict their accuracy and uniqueness ratings. Forced entry procedures were followed in preference to stepwise methods for strictly theoretical reasons and to ensure that the interaction variable entered the equation last (whereby any incremental validity

contributed by this construct could be evaluated while controlling for the GPA and years in college variables). Contrary to expectation, the correlation between the accuracy and uniqueness variables proved highly significant, $r(35) = .76$, $p < .01$. Due to the redundancy of these constructs, and the subsequent similarity in the results of the regression equations, the separate regression analyses were collapsed into a single table.

Assumptions

The examination of a plot comparing standardized residuals and predicted values failed to reveal distinct patterns or clusters, giving no cause to question the assumptions of linearity and homoscedasticity. Similarly, the tenability of the normality and independence assumptions was supported by the lack of outlying values within a histogram of standardized error scores.

Outliers

Weisberg's (1980) t statistic was used to probe the accuracy and uniqueness variables for the presence of outliers. Type I error inflation was controlled using the Bonferroni procedure. Two values surpassed the minimum values for significance found in tables provided by Stevens (1986). Cook's distance values for each of these outliers exceeded 1, indicating that the impact of neither score was disproportionately influential on the regression equation. Mahalanobis D^2 values for the predictor variables remained within acceptable limits (GPA, years in college, and GPA x years in college variables).

The predictor variables were entered into the equation in the order presented above. As Table 1 shows, the subject's GPA and number of years in college reliably predicted accuracy-uniqueness scores, $F(1, 32) = 10.15$, $p < .001$, accounting for approximately 40% of the total variance. The GPA x years in college interaction failed to significantly predict accuracy-uniqueness scores after controlling for the two factors independently.

Table 1

Results of Multiple Regression Analysis Predicting Accuracy and Uniqueness Ratings of Barnum Profiles.

Step	Factor	R ²	F	Sig. of F	F Change	Sig. of F
1	GPA	.20	8.21	.007	8.21	.007
2	Years School	.39	10.15	<.001	9.83	.004
3	GPA x School	.40	6.71	.001	.300	.588

Cross-Validation Issues

A primary concern when utilizing regression procedures is the stability of the results across samples. Because regression analysis weights the predictor variables in such a way that they are maximally correlated with the dependent measure(s) the results are inevitably to some degree sample specific (Stevens, 1986). Sample size and the number of predictor variables are the major determinants of the equation's generalizability. Herzberg's formula provides the most rigorous and widely endorsed estimator of the amount of shrinkage to be expected under cross validation of the equation (i.e. replication). The R^2 value in the above equation remains highly significant after this statistical adjustment, $R^2 = .23$, $p < .01$. The reliability of these results are further supported by reference to Park and Dudycha's (1974) tables, which project the shrinkage to be less than or equal to .1 with a 95% probability under replication.

To explore the relationship between subjects' expressed confidence in the assessment procedures and their evaluation of the profiles, individuals were categorized into low and high confidence groups according to a median split and compared on their accuracy and uniqueness scores (results were once again collapsed due to the high correlation between these variables). Subjects in the high confidence group attributed significantly higher ratings than those in the low confidence group (M 's of 85.9 vs. 72.2 on the accuracy measure, and 3.9 vs. 3.1 on the uniqueness variable), $t(32) = 2.55$, $p = .016$.

Discussion

In accordance with previous research (Dickson & Kelly, 1985; O'Dell, 1972; Weinberger & Bradley, 1980), this study found no significant differences in subjects' accuracy ratings as a result of being told their tests were interpreted by discrete assessment sources. Similarly, the absence of a difference in perceived accuracy among the subjects receiving highly favorable and standard profiles is in agreement with current literature suggesting that favorability may be confounded with the generality and base rate accuracy of the profile statements (Furnham & Schofield, 1987). An alternative explanation for the failure of these two profiles to produce different accuracy ratings is that some form of ceiling effect is operating upon the profiles, whereby the potency of efforts to discriminate between groups is diminished by testing within a reduced variation range. Manipulation of the base rate truthfulness of the profiles, rather than the perceived favorability, appears to hold more potential in future research. This suggestion is supported by Baucom and Greene's (1979) study demonstrating that the base rate accuracy of many commonly used profiles was considerably less than perfect.

Harris and Greene (1984) found that students could discern between real and shammed personality feedback when asked the appropriate questions. More specifically, by having subjects evaluate Barnum profiles in terms of their uniqueness and usefulness in addition to the standard accuracy ratings, they found them to be capable of recognizing the generality of the pseudo interpretations.

As evidenced by the strong correlation between subjects' accuracy and uniqueness ratings in the present study, our results fail to corroborate Harris and Greene's conclusion. As suggested by Snyder et al. (1978) reply to Greene (1977), small but meaningful differences in the testing situations may account for the disparity of findings. The type of assessment used (e.g., personality measures vs. aptitude instruments, the latter of which the subject may be much more aware of his or her individual strengths and deficits), the specific wording, and the order of the acceptance questions can operate as subtle demand characteristics, arousing subjects' suspicions toward the procedure.

The differences found between groups categorized by low and high degrees of confidence in the assessment procedures appear to substantiate an avenue of research originally explored by Snyder et al. (1976) in a study comparing the relative acceptance of psychological, graphological, and astrological interpretations. After controlling for differences in sample size, the strength of the relationship found in the present study coincided closely with that observed by Snyder et al. (1976). In the present study, correlations of .38 and .39 were observed between the measures of confidence and accuracy and confidence and uniqueness ($p < .05$) variables, whereas Snyder et al. (1976) reported a value of .12 ($p < .05$) between confidence and acceptance. Attempts to gain an understanding of the source(s) of individual faith differences are advocated. Confidence in the assessment procedures was not found in this study to be related to the subject's GPA or number of years in college, suggesting its influence is derived from an independent source.

As hypothesized, operationalized sophistication variables were able to statistically account for variance on subjects' accuracy and uniqueness ratings. Previous studies (Harris & Greene, 1984; Stagner, 1958) have examined occupational backgrounds and educational levels (including possibly confounded comparisons of undergraduates with

clinical graduate students and psychiatric residents), omitting concern for more specific subject variables such as their GPA and amount of educational experience. Each of the latter contributed significantly to the regression equation while the interaction of the two failed to account for variance separately. Further investigations are encouraged to clarify the potentially moderating effects of these variables.

References

- Baillargeon, J., & Danis, C. (1984). Barnum meets the computer: A critical test. *Journal of Assessment, 48*, 415-419.
- Baucom, D. H., & Greene, R. L. (1979). The universality of generalized personality statements. *Journal of Personality Assessment, 43*, 497-500.
- Carrier, N. A. (1963). Need correlates of "gullibility". *Journal of Abnormal and Social Psychology, 66*, 84-86.
- Collins, R. W., Dmitruk, V. M., & Ranney, J. T. (1977). Personal validation: Some empirical and ethical considerations. *Journal of Consulting and Clinical Psychology, 45*, 70-77.
- Costa, P. T. Jr., & McCrae, R. R. (1992). *The Revised NEO Personality Inventory Manual*. Odessa, FL: Personality Assessment Resources.
- Dana, R. H., & Graham, E. D. (1976). Feedback of client-relevant information and clinical practice. *Journal of Personality Assessment, 40*, 464-469.
- Dana, R. H., & Fouke, H. (1979). Barnum statements in reports of psychological assessment. *Psychological Reports, 44*, 1215-1221.
- Dickson, D. H., & Kelly, I. W. (1985). The "Barnum effect" in personality assessment: A review of the literature. *Psychological Reports, 57*, 367-382.
- Dmitruk, V. M., Collins, R. W., & Clinger, D. I. (1973). The "Barnum effect" and acceptance of negative personal evaluation. *Journal of Consulting and Clinical Psychology, 41*, 192-194.
- Fichter, C. S., & Sunerton, D.

(1983). Popular horoscopes and the "Barnum effect". *Journal of Psychology*, 114, 123-124.

Forer, B. R. (1949). The fallacy of personal validation: A classroom demonstration of gullibility. *Journal of Abnormal Psychology*, 44, 118-123.

Furnham, A., & Schofield, S. (1987). Accepting personality test feedback: A review of the Barnum effect. *Current Psychological Research & Reviews*, 6, 162-178.

Glick, P., Gottesman, D., & Jolton, J. (1989). The fault is not in the stars: Susceptibility of skeptic and believers in astrology to the Barnum effect. *Personality and Social Psychology Bulletin*, 15 (4), 572-583.

Greene, R. L. (1977). Student acceptance of generalized personality interpretations: A reexamination. *Journal of Consulting and Clinical Psychology*, 45, 965.

Greene, R. L., Harris, M. E., & Macon, R. S. (1979). Another look at personal validation. *Journal of Personality Assessment*, 43, 419-423.

Halperin, K. M., & Snyder, C. R. (1979). Effects of enhanced psychological test feedback on treatment outcomes: therapeutic implications of the Barnum effect. *Journal of Consulting and Clinical Psychology*, 47, 140-146.

Halperin, K., Snyder, C. R., Shenkel, R. J., & Houston, B. K. (1976). Effects of source status and message favorability on acceptance of personality feedback. *Journal of Applied Psychology*, 61, 85-88.

Harris, M. E., & Greene, R. L. (1984). Students' perception of actual, trivial, and inaccurate personality feedback. *Journal of Personality Assessment*, 48, 179-184.

Lattal, K. A., & Lattal, A. D. (1967). Students' "gullibility": A systematic replication. *Journal of Psychology*, 67, 319-322.

O'Dell, J. W. (1972). P. T. Barnum explores the computer. *Journal of Consulting and Clinical Psychology*, 38, 270-273.

Park, C., & Dudycha, A. (1974). A cross validation approach to sample size

determination for regression models. *Journal of American Statistical Association*, 69, 214-218.

Schroeder, H. E., & Lesyk, C. K. (1976). Judging personality assessments: Putting the Barnum effect in perspective. *Journal of Personality Assessment*, 40, 470-474.

Snyder, C. R. (1974). Acceptance of personality interpretations as a function of assessment procedures. *Journal of Consulting and Clinical Psychology*, 42, 150.

Snyder, C. R., Handelsman, M. M., & Endelman, J. R. (1978). Can clients provide valuable feedback to clinicians about their personality interpretations? A reply to Greene. *Journal of Consulting and Clinical Psychology*, 46, 1493-1495.

Snyder, C. R., & Larsen, G. R. (1972). A further look at student acceptance of general personality interpretations. *Journal of Consulting and Clinical Psychology*, 38, 384-388.

Snyder, C. R., Larsen, D. L., & Bloom, L. J. (1976). Acceptance of general personality interpretations prior to and after receipt of diagnostic feedback supposedly based on psychological, graphological, and astrological assessment procedures. *Journal of Clinical Psychology*, 32, 258-265.

Snyder, C. R., & Newburg, C. L. (1981). The Barnum effect in a group setting. *Journal of Personality Assessment*, 45, 622-629.

Snyder, C. R. & Shenkel, R. J. (1976). Effects of "favorability", modality, and relevance upon acceptance of general personality interpretations prior to and after receiving diagnostic feedback. *Journal of Consulting and Clinical Psychology*, 44, 34-41.

Snyder, C. R., Shenkel, R. J., & Lowery, C. R. (1977). Acceptance of personality interpretations: The "Barnum effect" and beyond. *Journal of Consulting and Clinical Psychology*, 45, 104-114.

Stagner, R. (1958). The gullibility of personnel managers. *Personnel Psychology*, 11, 347-352.

Stevens, J. (1986). *Applied multivariate statistics for the social sciences* Hillsdale, New Jersey: Lawrence

BARNUM EFFECT

Erlbaum Associates.

Weinberger, W. S., & Bradley, L. A. (1980). Effects of favorability and type of assessment device upon acceptance of general personality interpretations. *Journal of Personality Assessment*, 44, 44-47.

Weisberg, S. (1980). *Applied linear regression*. New York: Wiley.