Philippine Journal of Psychology
1968, Vol. 1, No. 1, 62—67

SOCIAL DESIRABILITY AND ENDORSEMENT ON THE EPPS: PRELIMINARY FINDINGS[1]

ABRAHAM I. FELIPE
University of the Philippines

This paper is a first report from a project evaluating the suitability of the Edwards Personal Preference Schedule (EPPS) for English-speaking Filipinos. Interest in this test is partly based on the fact that it is being used locally for industrial as well as research purposes.

The EPPS is a 225-item personality test with 15 subscales for measuring 15 personality characteristics — achievement, deference, order, exhibitionism, autonomy, affiliation, intraception, succorance, dominance, abasement, nurturance, change, endurance, heterosexuality, and aggression. Each item in a subscale is made up of a pair of statements one of which is expressive of the characteristic being tapped by the subscale and the other member of the pair is irrelevant to this characteristic). The task of the subject in each item is to choose between the two statements which describes him better. The degree to which he is supposed to have a given characteristic is indicated by the number of times he chooses to apply to himself those statements expressive of that characteristic instead of the statements with which they are paired.

One who would like to question the applicability of the EPPS among Filipinos can do so in a number of ways. This paper inquires about this test's possible susceptibility to the common tendency of test-takers to simply give socially desirable answers (the SD variable) instead of answers that are "true' about them.

It is claimed that for Americans the EPPS is relatively free from the SD variable. This "gain" was achieved by using a forced-choice format: statements that have been empirically equated in social desirability are paired. Edwards sought to control the SD variable by this method because he believed that when an individual decides which of two equally desirable characteristics describes him better, his choice will no longer be systematically affected by the tendency to give a socially desirable picture of himself. He and other workers have shown that, whereas around 70% of the variance of test responses in ordinary personality inventories can be predicted from information about the degree of social desirability of the inventory items, only 12% to 26% of the variance of responses on the EPPS can be so predicted (Edwards, 1953, 1957). These results have been interpreted to mean that on American samples, the test has succeeded in minimizing the effects of the SD variable on test responses. It is implicitly assumed that the personality inventory is made more valid when the SD factor is minimized.

When Filipinos take the EPPS, it is reasonable to doubt that for them the SD variable is as well-controlled as it is among Americans. It is doubtful that when Edwards empirically equated his items in social desirability for Americans, he unintentionally succeeded in equating

[1] The financial support of the Philippine Mental Health Association (PMHA) for this project is gratefully acknowledged.

them for Filipinos too. For that to happen, the two groups' judgments of the EPPS items on the social desirability-undesirability dimension will have to be very similar. Few will be willing to grant this.

Three decisions were therefore made: (*a*) to explore how predictable are the EPPS endorsement of Filipino samples from information about the uncontrolled components of the SD variable in the test; (*b*) since there is a good reason to believe that males and females often differ in their judgments of what is socially desirable, to explore whether there are sex differences in the predictability of EPPS endorsements; and (*c*) because of the likelihood that Filipinos and Americans might differ in their social desirability judgments of some traits but not in others, to investigate the predictability of endorsement in each of the 15 EPPS subscales separately.

Method

Scaling of the EPPS Statements

The nine statements relevant to each of the 15 traits in the EPPS, or a total of 135 statements, were first isolated for purposes of getting an estimate of how socially desirable they were (i.e., their SD scale values). All statements were arranged in a questionnaire with the title "Trait Rating Schedule." In this schedule, all references to the first person singular in EPPS statements were removed and the results were described to the Ss as "traits."

The Ss were instructed to indicate on a 9-point scale how desirable or undesirable "in other people" was each of the 135 "traits." Each scale was clearly marked "desirable" at one end and "undesirable" in the other, with the labels "extreme", "strong", "mild", and "neutral" typed on different points of the scale to suggest varying degrees of desirability and undesirability.

Male, female, and combined male-female judgments of the 135 statements which formed the Trait Rating Schedule were subsequently scaled using the method of successive intervals (Edwards, 1957). The purpose of scaling was to generate a "scale value" for each statement which may be interpreted as the location of the statement on the social desir- making the judgment. Thus, three sets of scale values for the 135 statements were generated — the first for the male group, the second for the female group, and the third for all male and female subjects.

Subjects and Procedure

The Trait Rating Schedule was given to 58 male and 65 female students of elementary psychology at the University of the Philippines. One week later, the same Ss were requested to fill out the EPPS. Of the original 65 girls, 63 answered the EPPS, but of the boys only 40 did. Twelve boys who never took the Trait Rating Schedule were then recruited to fill out the EPPS and included in the present sample.

Treatment of the Data

In the EPPS, each of its 225 items occurs as a pair of two statements (between which the S would choose the statement which describes him better). As a result of the successive interval scaling, each EPPS item can be defined for a given sample of Ss by two scale values, one for each of the two statements paired. As a measure of the degree to which the two paired statements have been equated in terms of their social desirability, the difference in their scale values ("scale separation") was taken. This gives a measure of how much more, or less, socially desirable was a statement in comparison with the other statement with which it was paired. This measure of relative social desirability of an EPPS statement was taken from each sample and correlated with the proportion of the sample who endorsed it (i.e., those who applied that statement to themselves). It was expected that the greater the scale separation between two paired statements, the greater the proportion of Ss endorsing the more desirable statement.

Results and Discussion

Equation of the SD Variable

Edwards (1953) found among his American sample that only about 25% of his pairs of statements differed by more than .50 scale value. In the Filipino samples, the corresponding proportions are 46% among males, 44% among females, and 48% in the combined group. In each of the present samples, there are considerably more items than Edwards had in which the scale separations differed by more than .50.

Relation of the SD Variable and Endorsement in the Whole Test

Edwards (1953) found a correlation of .40 between the proportion of the sample who endorsed each statement and the degree of social desirability of the statement relative to the other member of the pair. This means that only 16%

of the variance of EPPS test responses can be predicted from the components of the SD variable that the test construction failed to control. Correlations of comparable magnitude have been reported by others.

For the present sample of Filipinos who probably resemble Americans in values more than the average English-speaking Filipino does, the corresponding correlations are .33 for males, .57 for females, and .52 for both sexes. In terms of variance, these mean that the percentages of the total variance in EPPS test responses that can be predicted from differences in social desirability of paired statements are 11% for males, 32% for females, and 25% for both sexes. As a whole, the correlation among Filipinos is only moderate. Also, the combined male-female Filipino samples and the Americans do not significantly differ in the degree to which endorsement could be predicted from uncontrolled SD factors in the whole test ($z = 1.22$). However, endosement in the whole test of the Filipino female sample is more predictable than endorsement of Americans ($z = 2.34$, $p. < .01$) and of the Filipino male sample ($z = 3.27$, $p < .01$). The Filipino male group and the American sample of both sexes do not differ in degree to which their EPPS endorsement could be predicted ($z < 1.00$).

Relation of the SD Variable and Endorsement in the Subscales

The correlation between scale separations and proportion of endorsements for each of the 15 personality variables in the EPPS are shown in Table 1. The correlations which are significantly different from zero are marked.

The following can be noted from Table 1.

TABLE 1

PEARSON-PRODUCT CORRELATIONS OF DIFFERENCES IN SCALE VALUES AND ENDORSEMENT OF STATEMENTS

Personality characteristics	Sample		
	Males	Females	Combined
1. Ach	.22	.42*	.30
2. Def	.12	.46**	.39*
3. Ord	.23	.40*	.28
4. Exh	.67***	.54**	.63***
5. Aut ± ±	.69***	—.05	.51**
6. Aff ± ±	.22	.81***	.76***
7. Int	.54**	.40*	.51**
8. Suc	.65***	.72***	.72***
9. Dom ±	.80***	.46**	.42†
10. Aba ±	—.10	.39*	.05
11. Nur	.41*	.27	.49**
12. Cha ± ±	.32*	.79***	.64***
13. End	.49**	.69***	.67***
14. Het	.52**	.63***	.58***
15. Agg	.72***	.57***	.72***
Overall	.33***	.57***	.52***

* $p < .05$
** $p < .01$
*** $p < .001$
± significant male vs. female difference $p < .05$
± ± significant male vs. female difference $p < .01$

1. The over-all correlations across all EPPS items do not reflect, but even obscure, the degree of relationships between SD and endorsement in the different subscales. In the male sample, although the over-all correlation was only .33, the range of *r*s is from —.10 to .80; among the females, —.05 to .81; and in the combined sample, .05 to .76. In other words, there is good predictability in some subscales but poor predictability in others.

2. In autonomy and dominance, responses are significantly more predictable among males than among females. On the other hand, in affiliation, abasement, and change, they are significantly more predictable among females than among males.

3. Using a correlation of .50 (which implies than 1/4 of the variance in test responses can be predicted from the scale separations of items) as an arbitrary cut-off point between "high" and "low" correlations, a number of patterns of correlations can be observed. (*a*) In some subscales, the correlations for all samples remain low, i.e., achievement, deference, order, abasement, and nurturance; (*b*) In some, they are all high for all the samples i.e. exhibitionsim, succorance, heterosexuality and aggression. (*c*) In others, the correlation in one sample may be low but high in another sample, i.e., autonomy, affiliation, intraception, dominance, change, and endurance.

Relation of Test-Item Characteristics and the Correlations

Edwards' (1953) SD hypothesis suggests that where the correlations in Table 1 are "low", the SD variable had been well-controlled, and where they are "high", the SD variable had not been well-controlled. To test the SD hypothesis as the explanation for the correlations in Table 1, it was therefore necessary to get some measure of the degree to which the SD variable was controlled for the sample. The scale separation for each item was the best available measure for this purpose.

As the first step in the search for an explanation, it was decided to look for the magnitude of scale separation that is decidely reliably associated with endorsement of the more desirable item. Although the SD hypothesis predicts that "popularity" of statements increases directly with social desirability, it is possible that the scale separation might have to read a certain minimum value for one statement to be considered reliably more popular than the other. Therefore, all scale separations were categorized into five levels: those that differed by .00 to .25 in scale value, by .26 to .50, by .51 to .75, by .76 to 1.00, and by more than 1.00 scale value. As the measure of the "popularity" of the more desirable statements at any given level of separation (say, .51 to .75), a ratio is formed between the number of more desirable statements on that level of separation chosen significantly more often[2] and the number of more desirable statements on the same level of separation chosen significantly less often. The results are given in Figure 1. Of the 10 points plotted, only one was erratic, otherwise both curves (drawn freely) would be smooth positively accelerating curves. It will be noted that when the scale separation is of the order of .51 to .75, the ratio of number of more desirable statements chosen significantly more often to number of more desirable statements chosen significantly less often is 2:1 for the two sexes. This ratio accelerates, and for statements separated by at least 1.00 scale value, it is 15:1 for girls and 18:1 for boys.

[2] On a binominal test, with a 50-50 chance of choosing either statement as the null hypothesis.

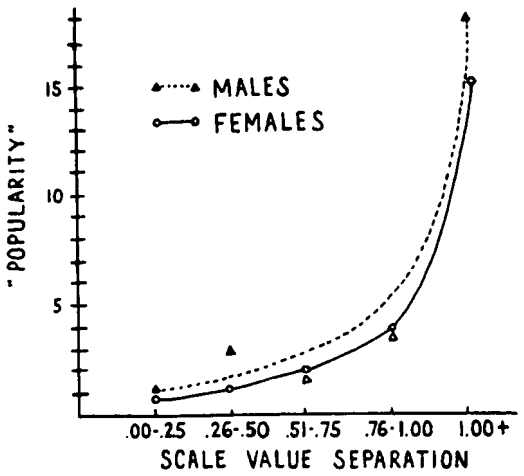

FIG. 1. Ratio between the number of more desirable statements chosen significantly more often and the number of more desirable statements chosen significantly less often on the same level of separation.

Figure 1 suggests that the magnitude of scale separation indexes the predictability of EPPS endorsement from information about uncontrolled SD. Three levels of scale separation were, therefore, explored as indices of the degree to which endorsement in a subscale could be predicted: (*a*) the number of statements in each subscale that are separated by at least .51 of scale value; (*b*) the number separated by at least .76 scale value; and (*c*) the number separated by more than 1.00 scale value.

Besides these measures, two others were taken which can be defended as reasonably possible indices of the degree to which endorsement could be predicted. The first of these two was the absolute value of the sum of scale separations of the relevant statement in each subscale. This was taken by subtracting from the scale value of each statement relevant to a subscale (R_i where i is the EPPS number containing the statement) the scale value of the paired non-relevant statement (N_j), adding algebraically the resulting scale separations within a subscale and getting the absolute value of the sum —/ Σ $(R_i - N_j)$/. The resulting number is interpreted as an index of the "net social desirability" of all statements relevant to a subscale. The other measure was the sum of the absolute scale separations in each subscale. This was taken by getting the absolute scale separation for each item of a subscale and adding them— Σ /$R_i - N_j$/. The resulting number is interpreted as the "total scale separation within a subscale."

The five measures derived from scale separation were correlated with the magnitude of the obtained correlations between social desirability and endorsement. The purpose is to explain the correlation. The results are given in Table 2. None of the correlations differed significantly from zero.

TABLE 2

RANK CORRELATIONS BETWEEN *z* AND INDICES OF UNCONTROLLED SD IN SUBSCALES

Sample	X_1	X_2	X_3	X_4	X_5
Male	—.03	.01	.08	—.10	.43
Female	.44	.12	.11	.00	.20
Combined	.03	—.11	.19	—.14	—.05

X_1 = Number of items within subscale with at least .51 scale separation.
X_2 = Number with at least .76 scale separation.
X_3 = Number with more than 1.00 scale separation.
X_4 = Σ $(R_i - N_j)$: net scale separation of statements relevant to a subscale.
X_5 = Σ / $(Ri - Nj)$/: total scale separation within a subscale.

CONCLUSION

1. The present Filipino sample and the sample reported by Edwards (1957) are comparable in the degree to which their endorsement of items in the whole EPPS could be predicted from information about the SD variable in the test. However, the overall correlation between endorsement and scale separation in the whole test does not seem to reflect accurately the complex relation of these

variables in the different subscales. (*a*) Filipino females' EPPS responses are significantly more predictable than those of males. (*b*) Responses in some subscales appear to be more predictable than in others. (*c*) The degree to which endorsement in a subscale could be predicted sometimes varies between the two sexes. Considering that the present subjects probably resemble in values Americans more than the average English-speaking Filipino, these findings suggest that *the validity of many EPPS subscales for Filipino subjects is open to question.*

2. The present findings have implications on any plan to adapt the EPPS among Filipinos. The investigator envisioned an EPPS for Filipinos that will be relatively free of SD. For this reason, an explanation for the complex set of correlations between social desirability and endorsement in the present study was sought in the hope that some manipulable, controllable characteristics of test stimuli which affected the correlations might be discovered. This would permit the formulation of a general rule on how to minimize the effects of the SD variable in a forced-choice test like the EPPS.

In this study, however, the obtained correlations could not be explained in terms of manipulable and controllable characteristics of test stimuli. It can be concluded that the vulnerability of an EPPS subscale to the influence of SD is not due to the number of pairs of statements in the subscale which are separated by at least .51 scale value, or those separated by at least.76 scale value, or those separated by more than 1.00 scale value. It is also not due to how much more desirable (or undesirable) the statements revelant to a subscale are in comparison with the statements with which they are paired. It is not even due to the total amount of difference in social desirability of the paired statements. Hence, the results suggest that one will not be able to free the EPPS from SD by setting a minimum number of items with a given size of scale separations, nor by reducing the difference in social desirability of statements relevant to a subscale and statements with which they were paired, nor by minimizing the total amount of differences in social desirability of paired statements.

It does not follow that there is no single variable which can explain the correlations. Whatever is this variable, however, it is certainly not obvious. On the other hand, it appears that all reasonable and defensible explanations from data on scale values have been exhausted. An intuitive possibility exists that the dispersion of social desirability judgments on each of the different statements might have affected the obtained correlations. As of the moment, however, the way to test this possibility is not yet clear.

In the absence of an explanatory variable and knowing that the correlation between endorsement and social desirability is high in some subscales even though there is little scale separation in the items, the investigator is forced to tentatively conclude that some subscales are sensitive even to minute scale separations but others are sensitive only to sizable ones. It appears that to construct an EPPS for Filipinos that is free from SD effects, one will have to tediously free each individual subscale separately.

REFERENCES

EDWARDS, A. L. *The social desirability variable in personality assessment and research.* New York: Holt, Rinehart and Winston, Inc., 1957.

EDWARDS, A. L. The relationship between the judged desirability o fa trait and the probability that a trait will endorsed. *Journal of Applied Psychology,* 1953, **37**, 90-93.