

## EVALUATIVE AND DESCRIPTIVE COMPONENTS OF TRAIT INFERENCES<sup>1</sup>

LENORE A. DE LA LLANA  
*University of the Philippines*

The question of whether attribution of traits to people or objects involve evaluative and descriptive aspects was explored using trait inference tasks. The present investigation replicated Felipe's study with the following modifications: (1) instead of adjectives, nouns pertaining to persons were used as stimulus concepts and (2) additional scales were used as reference vectors to clarify the earlier findings. Results indicated the presence of evaluative and descriptive components in trait inferences and supported the *concept-scale relevance hypothesis*.

When we attribute traits to people or objects, do our inferences reflect evaluation, description or both? A number of investigators have dealt with this question and their findings can be sharply differentiated into two categories, namely: those supporting the evaluative consistency hypotheses and those supporting the descriptive consistency hypotheses.

The semantic differential has been the most widely used instrument in the investigation of this problem. Studies in this area typically involve trait inference tasks requiring subjects to infer the likelihood of a person's (stimulus concept) possessing one or the other of the trait terms defining the bipolar ends of semantic differential scales (Osgood, Suci, and Tannenbaum, 1957; Peabody, 1967; Rosenberg and Sedlak, 1969; Felipe, 1970). In such tasks the evaluative consistency hypothesis predicts that the sign of the inferred trait should be the same as the sign of the stimulus concept. Judging a "stingy" person as "lax" on the *lax-firm* scale follows this hypothesis. On the other hand, the descriptive consistency hypothesis places primary importance on the similarity between, or cor-

respondence of, the descriptive aspects of the stimulus concept and the attributed trait. Judging the same "stingy" person as "thrifty" on the *thrifty-extravagant* scale would be descriptively consistent. "Stingy" and "thrifty" share the common property of being tight with money.

The descriptive consistency hypothesis emerged from Peabody's objection (1967) to the apparent overemphasis on the evaluative components of judgment. He raised the issue of the usual confounding of evaluative and descriptive aspects and devised a method by which this confounding could be eliminated. From his factor analyses of trait inference data, he reported the absence of evaluation. This contradicts the more common contention that trait inferences involve a strong and stable evaluative component. Following Peabody, therefore, judgments we make are primarily, if not only, descriptive.

Felipe (1970) presented evidence contradicting Peabody's findings. He maintained that overwhelmingly descriptive factors were generated in the factor analyses because of the high degree of overlapping of the descriptive attributes of scale and concept. Using Peabody's sets of trait terms as starting points, he compared predictions of the evaluative and descriptive consistency hypotheses. He suggested a dual consistency-producing mechanism appearing in a definite sequence. "When one tests for descriptive consistency and fails to achieve it, he

---

<sup>1</sup>This study was supported by a research grant awarded to Dr. Abraham I. Felipe, Associate Professor of Psychology at the University of the Philippines by the National Research Council of the Philippines.

Dean Bonifacio Salamanca of the University of the Philippines at Manila and Mrs. Carmen Enrile-Santiago of Maryknoll College helped get the subjects for the present study.

generally proceeds to apply the second mechanism, attempting to achieve affective balance, at least" (p.635). This implies the secondary role of evaluative mechanisms but not in Peabody's sense of being derived only from descriptive similarity.

Following Felipe's hypothesis the main focus in trait inferences would be the relation between the stimulus concept and the specific scale. Inferences made on scales which are descriptively relevant to the concept reflect description; inferences made on scales not descriptively relevant to the concept reflect evaluation. This condition is referred to as the *concept-scale relevance hypothesis*. An individual applies first a descriptive criterion and tests whether the scale is relevant to the concept. If it does not work, he then applies the second criterion which is evaluation.

The present investigation replicates Felipe's study with the following modifications: (1) instead of adjectives, nouns pertaining to persons are used as antecedent terms and (2) additional scales are used as reference vectors to clarify further the interpretation of the findings of the earlier study (Felipe, 1970). This paper, however, focuses mainly on the analyses of stimulus concepts.

## METHOD

### *Selection of Stimulus Concepts*

Stimulus concepts were selected on the basis of the criteria used in the earlier studies (Peabody, 1967; Felipe, 1970). All concepts were (1) significantly positive or negative, using Felipe's method and (2) comprehensible as revealed by a vocabulary test. Each set satisfactorily met the criterion of Peabody's sets, that is, they fit the 2 x 2 schema. The procedure followed for the latter consisted of making pretest subjects select which of the four given alternatives arranged in a multiple choice format satisfied the schema's requirement. The list of stimulus concepts and their evaluative ratings is presented in Table 1. Only two sets were assumed to have the X and Y properties corresponding to Peabody's "tight-loose" description. The rest were not further differentiated and simply labeled as "All Others". The final list consisted of 56 noun concepts, including two incomplete sets of three terms and eleven sets of two terms.

TABLE 1  
SETS OF NOUNS USED AND THEIR  
EVALUATIVE RATINGS

Set No.	X- and Y-bearing			
1	.93	Economizer	1.42	Philanthropist
	-2.01	Miser	-2.39	Squanderer
2	.80	Devotee	1.69	Freethinker
	-1.81	Fanatic	-.70	Agnostic
All Others				
3	1.65	Male	1.94	Female
	-2.03	"Bakla"	-1.79	Lesbian
4	1.47	Admirer	.99	Critic
	-1.05	Flatterer	-1.92	Slanderer
5	.79	Sovereign	1.46	Citizen
	-1.87	Tyrant	-2.00	Slave
6	.99	Ally	.60	Contestant
	-2.21	Collaborator	-.47	Rival
7	.89	Martyr	1.70	Victor
	-1.44	Victim	-2.12	Oppressor
8	1.01	Sympathizer	1.44	Reformer
	-1.15	Bigot		
9	-.91	Spinster	-1.78	Prostitute
	1.95	Virgin		
10	1.42	Saint	-1.44	Sinner
11	2.31	Patriot	-2.14	Traitor
12	.41	Socialist	-.68	Capitalist
13	-1.84	Drunk	-1.14	Prohibitionist
14	1.83	Wit	-1.88	Bore
15	1.72	Scholar	-1.97	Ignoramus
16	2.16	Lady	-2.10	Bitch
17	2.37	Gentleman	-1.96	Rascal
18	.86	Believer	-.80	Pagan
19	1.96	Hero	-1.70	Villain
20	.96	Genteel	-1.90	Barbarian

### *Types of Scales*

Aside from the 24 scales used by Felipe, four types of reference scales were included: four of Rosenberg and Olshan's scales (1970) and Osgood-type scales measuring evaluation, potency and activity. The three major types were derived from two sets of X- and Y-bearing and two sets of M- and N-bearing terms. Type A scales were composed of bipolar terms with the same descriptive attributes but different evaluative signs. These scales eliminate confounding by controlling for descriptive aspects. Type B scales were composed of descriptively bipolar adjectives that had the same sign but differed in descriptive attributes. These scales hold evaluative signs constant and allow only descriptive contents to vary. Type C scales were composed of bipolar terms having different descriptive attributes and different evaluative signs. These scales

were used by Peabody to oppose the predictions of evaluative and descriptive consistency hypotheses to each other. The scales are presented in Table 2.

#### Procedure

A total of 1960 inference items derived from the final list of 56 concepts and 35 scales were divided into eight forms, each administered to 40 freshmen students of introductory psychology at the University of the Philippines and Maryknoll College. For each inference item, subjects judged the likelihood of a stimulus person's possessing one or the other of the trait terms given by the scale. Items were arranged and counter-balanced and concept-scale pairings were systematically rotated. Scale polarities were alternated to minimize the influence of previous ratings.

#### Design

The basic design consisted of two groups of scales (X- and Y-bearing and M- and N-bearing) and two groups of stimulus concepts (X- and Y-bearing and All Others). This basic design was replicated three times according to the type of scales (A, B, and C).

## RESULTS AND DISCUSSION

Each response was made on a seven-point scale scored from -3 to +3 between the favorable and unfavorable trait terms defining the scales. The data obtained were analyzed in two

TABLE 2  
TYPES OF SCALES USED

Type A (same descriptive content, different evaluative sign)	
Stingy(-X)-Thrifty(+X)	Unassured(-M)-Modest(+M)
Extravagant(-Y)-Generous(+Y)	Conceited(-N)-Confident(+N)
Harsh(-X)-Firm(+X)	Impractical(-M)-Idealistic(+M)
Lax(-Y)-Lenient(+Y)	Opportunistic(-N)-Practical(+N)
Type B (different descriptive content, same evaluative sign)	
Stingy(-X)-Extravagant(-Y)	Conceited(-N)-Unassured(-M)
Thrifty(+X)-Generous(+Y)	Confident(+N)-Modest(+M)
Harsh(-X)-Lax(-Y)	Opportunistic(-N)-Impractical(-M)
Firm(+X)-Lenient(+Y)	Practical(+N)-Idealistic(+M)
Type C (different descriptive content, different evaluative sign)	
Extravagant(-Y)-Thrifty(+X)	Unassured(-M)-Confident(+N)
Stingy(-X)-Generous(+Y)	Conceited(-N)-Modest(+M)
Lax(-Y)-Firm(+X)	Impractical(-M)-Practical(+N)
Harsh(-X)-Lenient(+Y)	Opportunistic(-N)-Idealistic(+M)
Reference Scales	
Osgood's Scales	Rosenberg and Olshan's Scales
Passive(-)-Active(+)	Submissive(-)-Dominant(+)
Slow(-)-Fast(+)	Undecided(-)-Decided(+)
Undesirable(-)-Desirable(+)	Introverted(-)-Extroverted(+)
Bad(-)-Good(+)	Inhibited(-)-Impulsive(+)
Weak(-)-Strong(+)	
Soft(-)-Hard(+)	
Cold(-)-Warm(+)	

TABLE 3

## SUMMARY OF INFERENCES ON TYPE A SCALES: RESULTS CONFIRMING AND DISCONFIRMING EVALUATIVE CONSISTENCY HYPOTHESIS

Concepts	No. of Items	Confirming				Disconfirming			
		$\bar{x}$ in Predicted Direction		$\bar{x} > 0$		$\bar{x}$ Opposite Predicted Direction		$\bar{x} > 0$	
		<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
X-, Y-bearing	64	57	89	45	70	7	11	1	02
All Others	384	338	88	261	68	46	12	19	05
<b>TOTAL</b>	<b>448</b>	<b>395</b>	<b>88</b>	<b>306</b>	<b>68</b>	<b>53</b>	<b>12</b>	<b>20</b>	<b>05</b>

Note: Percentages were derived by dividing frequency (*f*) with the corresponding number of items.

ways: (1) for each inference item *t* tests were performed on mean scores of 40 subjects to see whether they differed significantly from zero and (2) the item means were intercorrelated across scales and these intercorrelations were factored by the principal axis method and rotated by the varimax procedure.<sup>2</sup> This second analysis was performed for comparability with Peabody's findings.

#### Analyses of Means

*Inferences on type A scales.* Felipe states that the evaluative consistency hypothesis has unequivocal predictions about inferences on Type A scales: the sign of the stimulus concept that is significantly positive or negative determines the direction of the response. For example, *saint* (a positive concept) is rated *firm* (a positive attribute) on the Type A scale *harsh-firm*. This inference would be evaluatively consistent. Table 3 shows the evidence in favor of the evaluative consistency hypothesis. Out of the 448 inference items, 395 (88%) were in the direction predicted with 306 items (68%) significantly different from zero.

*Inferences on type B scales.* Terms defining the Type B scales vary in descriptive content

but have the same evaluative sign. Inferences on these scales, therefore, are predicted by the descriptive consistency hypothesis and not by the evaluative consistency hypothesis. The stimulus term *philanthropist* rated *extravagant* along the *stingy-extravagant* scale is descriptively, but not evaluatively consistent. In this study, however, description is not expected to be strongly supported since the condition is such that a majority of the items have low concept-scale relevance. Heterogeneity of concepts (unlike in Felipe's study where the characteristics of antecedent terms are clearly defined) could account for the weakness of the descriptive dimension. Table 4 summarizes the evidence for the descriptive consistency hypothesis. Out of the 309 items (69%) in the predicted direction, only 34% reached significance. These figures are comparatively lower than what Felipe obtained (82% were in the direction predicted with 51% reaching significance). The table shows that the descriptive consistency hypothesis was most successful in predicting inferences from the X- and Y-bearing concepts onto the X- and Y-bearing scales. Concepts (All Others) whose descriptive attributes were not known seem to be also involved descriptively in inferences. There were more errors for the descriptive consistency hypothesis on Type B scales (31%) than for the evaluative consistency hypothesis on Type A scales (12%).

<sup>2</sup>Analyses were done at the U.P. Computer Center. Grace de Vera's invaluable help is gratefully acknowledged.

TABLE 4

## SUMMARY OF INFERENCES ON TYPE B SCALES: RESULTS CONFIRMING AND DISCONFIRMING DESCRIPTIVE CONSISTENCY HYPOTHESIS

Inference Items			Confirming				Disconfirming			
			$\bar{x}$ in Predicted Direction		$\bar{x} > 0$		$\bar{x}$ Opposite Pre-Direction		$\bar{x} > 0$	
Concepts	Scale Type	No. of Items	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
X-, Y-bearing	X-, Y-bearing	32	27	84	18	56	5	16	00	00
	M-, N-bearing	32	18	56	12	37	14	44	10	31
All Others	X-, Y-bearing	192	120	62	52	27	72	38	28	15
	M-, N-bearing	192	144	75	72	38	48	25	38	20
TOTAL		448	309	69	154	34	139	31	76	17

Note: Percentages were derived by dividing frequency (*f*) by the corresponding number of inference items.

TABLE 5

## PARTIAL SUMMARY OF INFERENCES ON TYPE C SCALES: ERRORS ON CRITICAL ITEMS OF EVALUATIVE AND DESCRIPTIVE CONSISTENCY HYPOTHESES

Critical Inference Items			Errors of Evaluative Hypothesis				Errors of Descriptive Hypothesis			
			Errors in Direction		$\bar{x} > 0$		Errors in Direction		$\bar{x} > 0$	
Concepts	Scale Type	No. of Items	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%	<i>f</i>	%
X-, Y-bearing	X-, Y-bearing	16	7	44	5	31	9	56	5	31
	M-, N-bearing	16	4	25	5	12	12	75	11	70
All Others	X-, Y-bearing	96	39	41	17	18	57	59	34	35
	M-, N-bearing	96	24	25	10	10	72	75	51	53
TOTAL		224	74	33	34	15	150	67	102	46

Note: Percentages were derived by dividing frequency (*f*) with the corresponding number of items.

*Inferences on type C scales.* These were used by Peabody to oppose both hypotheses. For half of the inference items, both hypotheses have joint predictions. Following Felipe's criterion of the "least number of sign changes" (the number of sign changes required to make an obtained subset of means balanced both evaluatively and descriptively), 78% of the means jointly predicted differed significantly from zero.

The remaining 224 items constitute the crucial items to test both hypotheses. Again, following the criterion mentioned earlier, the evaluative consistency hypothesis had fewer errors (33%) as compared to the descriptive consistency hypothesis (67%). In sets where scales were assumed to be descriptively relevant there seems to be an almost equal chance for inferences to reflect evaluation and description.

TABLE 6  
VARIMAX FACTOR MATRIX OF 56 NOUNS

Set No.	CONCEPTS	Evaluative Content	Rotated Factors				h <sup>2</sup> *
			i	ii	iii	iv	
<i>X-, Y-bearing</i>							
1	Economizer	+	-63	-63	01	03	84
	Miser	-	-30	23	-80	00	91
	Philanthropist	+	-28	-04	65	26	83
	Squanderer	-	17	72	05	-03	76
2	Devotee	+	-41	-43	33	02	88
	Fanatic	-	-58	37	-32	-11	88
	Freethinker	+	-78	-03	17	-39	90
	Agnostic	-	09	46	-59	01	85
<i>All Others</i>							
3	Male	+	-92	17	06	-18	92
	"Bakla"	-	64	54	24	03	92
	Female	+	27	-36	73	-49	88
	Lesbian	-	01	32	-22	07	90
4	Admirer	+	-52	26	64	08	91
	Flatterer	-	-18	89	-20	01	93
	Critic	+	-95	-10	-02	-07	95
	Slanderer	-	-24	72	-25	03	76
5	Sovereign	+	-94	-01	-01	28	95
	Tyrant	-	-60	46	-41	-04	84
	Citizen	+	-42	-50	49	-01	86
	Slave	-	51	-44	00	-03	83
6	Ally	+	-54	-06	72	-12	91
	Collaborator	-	-44	29	-23	-06	87
	Contestant	+	-87	00	-01	05	80
	Rival	-	-75	36	-40	-04	91
7	Martyr	+	-41	-38	26	-03	90
	Victim	-	30	-12	01	08	84
	Victor	+	-88	18	-05	16	89
	Oppressor	-	-69	24	-26	-15	85
8	Sympathizer	+	-28	-08	71	00	80
	Bigot	-	-32	31	-74	-18	80
	Reformer	+	-88	05	01	-42	93
9	Spinster	-	-06	-20	-75	-42	84
	Virgin	+	-08	-08	24	-01	88
	Prostitute	-	10	73	-14	-10	84
10	Saint	+	-34	-48	38	09	94
	Sinner	-	08	-58	-04	-17	89
11	Patriot	+	-68	12	-02	-05	90
	Traitor	-	-30	34	-47	04	74
12	Socialist	+	-77	-19	-31	-20	91
	Capitalist	-	-87	16	-21	15	89

Out of the 16 X- and Y-bearing inference items 44% constitute errors in the direction of the evaluative consistency hypothesis and 56% constitute the errors in the direction of the descriptive consistency hypothesis. The results are shown in Table 5. Considering all the items the percentage of errors in the direction of the descriptive consistency hypothesis is greater.

#### Factor Analysis of Stimulus Concepts

The basic data for this analysis consisted of item means of the 56 concepts correlated across the 35 scales. These intercorrelations were factored by the principal axis method and rotated by the varimax procedure. The rotation resulted in a nine-factor matrix; however, only the first four are shown in Table 6.

Factor I, accounting for 31% of the total variance, might be called a "dominance" factor defined as "the exercise of control, authority, or influence on others". Nouns with loadings greater than .60 on this factor include *sovereign*,

*critic*, *prohibitionist*, *victor*, *wit*, *capitalist*, *scholar*, *male*, *oppressor*, *hero*, *rival*, and *socialist*.

Factor II, explaining 18% of the total variance, might be interpreted as an evaluative factor. In sets of four, loadings of concepts with the same evaluative sign should follow the same direction. For example, the X- and Y-bearing concepts *economizer* and *philanthropist* loaded negatively and *miser* and *squanderer* loaded positively, indicating consistency in the evaluative direction.

Factor III, explaining 15% of the total variance, might be interpreted as "cynicism or misanthropy vs. love for mankind". Illustrative concepts include *miser*, *villain*, *barbarian*, *agnostic*, *spinster*, *philanthropist*, *sympathizer*, *ally*, and *admirer*.

Factor IV, accounting for only 4% of the total variance, was not clearly interpretable since the loadings were relatively low.

The foregoing analyses show that evaluative and descriptive components are involved in trait

Table 6 (continued)

Set No.	CONCEPTS	Evaluative Content	Rotated Factors				h <sup>2</sup> *
			i	ii	iii	iv	
14	Wit	+	-.86	.23	.35	.13	.95
	Bore	-	.79	.00	-.23	-.37	.92
15	Scholar	+	-.85	-.32	.12	.21	.93
	Ignoramus	-	.61	.00	.02	-.31	.76
16	Lady	+	-.12	-.57	.65	.08	.92
	Bitch	-	-.20	.88	-.16	.01	.94
17	Gentleman	+	-.57	-.49	.48	.16	.92
	Rascal	-	.26	.57	-.58	-.13	.85
18	Believer	+	-.69	-.31	.17	-.08	.91
	Pagan	-	.13	.10	-.16	-.90	.90
19	Hero	+	-.83	-.24	.19	.09	.90
	Villain	-	-.46	.65	-.50	.01	.95
20	Genteel	+	-.35	-.53	.60	.19	.91
	Barbarian	-	-.18	.65	-.51	.06	.80
%Total $\sigma^2$			.31	.18	.15	.04	.86
%Common $\sigma^2$			.36	.22	.18	.05	

\*Decimal points omitted for loadings. \*The communalities represent all nine factors, not only the first four.

inferences. This finding argues against Peabody's earlier contention that descriptive aspects were always decisive over evaluative ones. Osgood's claim that evaluation is a major factor was supported but not in the sense of its being the most important aspect of judgment. No evidence in this study warrants such conclusion.

The *concept-scale relevance hypothesis* was also clearly supported. Concept-scale relevance was relatively low in this study and this accounts for the greater use of the evaluative dimension. This is evident in the first analysis. Where descriptive-type inferences are not possible, judgments are necessarily evaluative. In factor analysis, increasing the number of nonrelevant concepts would bias factor structures in favor of evaluation; increasing the number of relevant concepts would bias factor structures in favor of description. This phenomenon of concept-scale interaction has been studied by Darnell (1966) and Nordenstreng (1969). Scales shift in meaning or relevance as a function of the nature of concepts being rated. If such is the case, then factor structures would necessarily change from concept to concept. These effects, however, could be more clearly demonstrated by an analysis of the scales involved (Felipe *et al.*, 1972).

Relevance here has been defined in a very limited sense. Other aspects of it can be explored and techniques formulated to appropriately test, at the operational level, its role in the primacy of evaluation or description in trait inferences and human judgment in general.

## REFERENCES

- DARNELL, D. K. Concept-scale interaction in the semantic differential. *Journal of Communication*, 1966, 16, 104-115.
- FELIPE, A. I. Evaluative versus descriptive consistency in trait inferences. *Journal of Personality and Social Psychology*, 1970, 16(4), 627-638.
- FELIPE, A. I., DE LA LLANA, L., DE LA CRUZ, A. and DE VERA, G. Evaluative and descriptive aspects of semantic differential ratings. Unpublished manuscript, Department of Psychology, University of the Philippines, 1972.
- NORDENSTRENG, K. Changes in meaning of the semantic differential scale: A measurement of concept-scale and subject-scale interaction effects. The Research Institute of the University of Tampere, Tampere, Finland, 1969.
- OSGOOD, C. E., SUCI, G. J. and TANNENBAUM, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.
- PEABODY, D. Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology Monograph*, 1967, 7 (Whole No. 644).
- ROSENBERG, S. and SEDLAK, A. Structural representation of perceived personality trait relationship. Paper presented in a seminar on scaling and measurement held at the University of California at Irvine, June, 1969.
- ROSENBERG, and OLSHAN, K. Evaluative and descriptive aspects in personality perception. *Journal of Personality and Social Psychology*, 1970, 16(4), 619-626.