Value and Uncertainty as Weighting Factors in Impression Formation

DAVID L. RONIS

University of Michigan

AND

EDMUND R. LIPINSKI

Oakland University

An experiment was conducted to compare mathematical models describing how people combine information to form an impression of another person. Subjects rated how much they would like stimulus persons described by one or two adjectives. Subjects also reported their level of uncertainty about each evaluation. Models using the uncertainty measures to predict integration weights were no more successful than the equal weight averaging model. There was no evidence that extreme adjectives were given high weight. Information integration was best described by a model in which the more negative adjective in a pair is given more weight than the other adjective. © 1985 Academic Press, Inc.

Psychologists have long been concerned with understanding how people combine information when making judgments about others. The traditional task in this research has been to rate the likableness of a stimulus person described by a list of adjectives. Such research has focused on algebraic models predicting overall evaluations from values of each separate adjective. The simplest reasonable model of information integration assumes that the rating of a person described by two adjectives is the average of the ratings of two people each described by one of the adjectives. Equations

for this *equal weight model* and six other models with which it was compared are listed in Table 1.

One focus of research on information integration has been on factors that determine the weight given to items of information. Two general theoretical approaches have been applied to this problem (a) a *value-based* approach in which weight depends on the adjective's location on the evaluative dimension, and (b) an *uncertainty-based* approach in which weight depends on the adjective's power to reduce uncertainty. The current study compared the equal weight model with three value-based models and three uncertainty-based models.

One of the earliest value-based models is part of congruity theory (Osgood & Tannenbaum, 1955). According to the *congruity model,* the integrated impression is a weighted average of the component evaluations where each component is weighted by its extremity. Two other value-based models hypothesize that more negative adjectives are given greater weight. The *range model* (Birnbaum, 1974b) predicts that the most negative item of information in a description is given more weight than the other item(s). The *negativity model* predicts that weight is a negative linear function of value, independent of the other items in the stimulus set.

The three uncertainty-based models examined in this study are based on the assumption that beliefs about likableness are probabilistic (Birnbaum, 1973b; Wyer, 1973, 1974). They predict that some items of information are more informative and are given greater weight. According to two of

TABLE 1
THE SEVEN MODELS[a]

Name	Rating equation	Uncertainty equation
Equal weight	$R_n = \Sigma_i r_i / n$	
Congruity	$R_n = \Sigma_i (|r'_i| r_i) / \Sigma_i |r'_i|$	
Range	$R_n = (w_L r_L + \Sigma_i r_i) / (w_L + n - 1)$	
Negativity	$R_n = \Sigma_i (k - r_i) r_i / \Sigma_i k - r_i$	
Concept identification	$R_n = m_d$	$u_i = -\Sigma_j p_{ij} \log_2 p_{ij}$
Variance	$R_n = \Sigma_i (m_i / v_i) / \Sigma_i (1/v_i)$	$v_n = 1/\Sigma_i(1/v_i)$
Confidence	$R_n = \Sigma_i (c_i r_i) / \Sigma_i c_i$	

[a] R_n is the predicted rating of a person described by n adjectives; r_i is the rating of adjective i; r'_i is the rating expressed on a scale where zero represents a neutral reaction; r_L and w_L are the rating and weight given to the most negative adjective, $w_L > 1$; k is an empirically determined constant; m_d is the mean of a belief distribution derived from the belief distributions of n adjectives; u_i is the predicted uncertainty associated with the rating of adjective i or adjective set i; p_{ij} is the subjective probability that a person described by adjective (set) i falls in likableness category j; m_i and v_i, are, respectively, the mean and variance of the belief distribution associated with adjective i; v_n is the predicted uncertainty in rating a person described by n adjectives; and c_i is the confidence associated with the likableness rating of adjective i. For the range model, the summation is over all but the most negative adjective. Other summations are over all adjectives.

these models, a rating scale response describes the central tendency of a belief distribution. Figure 1, for example, shows a subject's distribution of belief about the likableness of one person described as "nosey" and one person described as "perfectionistic." The subject thinks there is a 60% chance of disliking the nosey person very much and a 20% chance of disliking him or her either a little more or a little less. This distribution represents the subject's implicit theory about what proportion of nosey persons have various degrees of likableness. The subject's rating of the nosey person would be predicted to be a 2. The dispersion of the belief distribution affects the confidence with which a stimulus person is rated. Thus the subject should be more confident in rating the "nosey" person than in rating the "perfectionistic" person.

According to Wyer's (1973) *concept identification model,* the belief distributions associated with single adjectives can be assessed by self-report. The belief distribution for the combination of two adjectives is the conjunction of the belief distributions for the two adjectives. This is approximated by the area of overlap between the two belief distributions (Wyer, 1973). The overlap between the belief distributions for "nosey" and "perfectionistic" is illustrated in Fig. 1. The mean of this area (2.5) is the predicted rating of a person described by both adjectives. A measure of uncertainty based on information theory predicts the rater's confidence. If the two belief distributions do not overlap, the rater "discounts" the information associated with the higher level of uncertainty. The rating is based entirely on the belief distribution with the lower uncertainty score. Failure to overlap also increases the uncertainty associated with
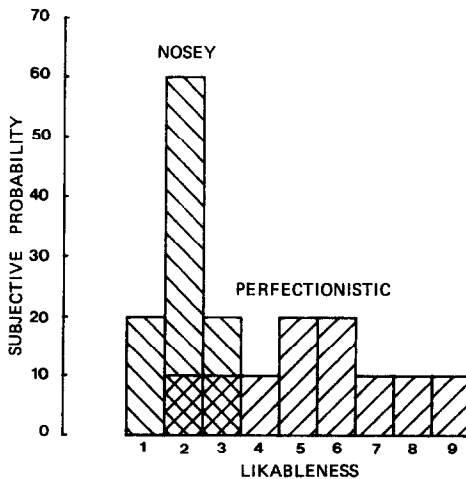

FIG. 1. Hypothetical distributions of belief about the likableness of stimulus persons described as "nosey" (shaded from upper left to lower right) and as "perfectionistic" (shaded from lower left to upper right). The overlap is shaded in both directions.

the judgment. This model is described by the equations in Table 1. See also Wyer (1973, 1974).

The *variance model* hypothesizes that the rating of an adjective set is a weighted average of the means of the separate belief distributions, with each mean inversely weighted by the variance of its belief distribution (Birnbaum, 1973b). This model also makes predictions about the uncertainty associated with a rating based on several items of information (Meyer, 1975). According to the *confidence model,* the rating of an adjective set is the weighted average of the ratings based on each adjective alone, with each rating weighted by the rater's subjective confidence in making it.

Note that unlike the value-based models, these uncertainty-based models are idiographic (Ostrom & Davis, 1979). They allow the possibility that different people will assign very different weights to the same adjective even if they rate the adjective identically. On the other hand, if uncertainty is closely related to value, these models could predict and explain findings predicted by value-based models.

Several studies have identified deviations from the equal weight model, including higher weighting of negative information (Birnbaum, 1974b; Fiske, 1980; Kanouse & Hanson, 1972) and higher weighting of extreme information (Fiske, 1980; Manis, Gleason, & Dawes, 1966; Warr & Jackson, 1975). Deviations from congruity theory have also been noted (Manis et al., 1966; Warr & Jackson, 1975). Consistent with idiographic models, Ostrom and Davis (1979) found that the weight of an adjective varies substantially across people. There have been no published tests of the variance model or the confidence model.

Wyer (1973) tested the concept identification model by having subjects rate stimulus persons described by one or two adjectives, judge their confidence in each rating, and describe their belief distributions based on each adjective separately. The model predicted ratings of stimulus persons described by two adjectives more accurately than either congruity theory or the equal weight model.

The purpose of the current experiment was to replicate Wyer's (1973) study and to extend it to comparisons of the predictive accuracy of the seven models described above, i.e., the equal weight, range, negativity, congruity, concept identification, variance, and confidence models. In addition, attempts were made to overcome possible methodological problems in Wyer (1973). One of these potential problems is the method of analysis: predicted and observed values were averaged over subjects before calculating measures of fit in Wyer (1973). Since averaging can change the form of relationships, this analysis technique may distort the results (Bakan, 1954; Ostrom & Davis, 1979). So measures of fit were calculated separately for each subject in the current study.

The other potential problems are scale-end artifacts. Rating scales

restrict the range of responses that can be given. A common consequence is nonlinearity of the response scale, particularly for ratings of extreme stimuli. In order to minimize scale-end artifacts in the current experiment, the range of the critical stimuli was restricted to exclude the most polarized adjectives, and highly polarized adjectives were used as filler stimuli as recommended by Anderson (1982, p. 8). No such techniques were used in Wyer (1973).

Method

Subjects

The subjects were 100 students enrolled in introductory psychology classes at Oakland University. Fifty subjects were randomly assigned to each of two replications, but data from 14 of the subjects could not be used because they had failed to answer one or more questions, leaving 45 and 41 subjects in the two replications.

Procedure

The experiment was divided into two parts. In the first part of the experiment, subjects were presented with descriptions of 41 stimulus persons. Each description consisted of one or two adjectives. Subjects were asked to judge how much they would like each person on a scale from 1 to 9: 1 = dislike very very much, 2 = dislike very much, 3 = dislike, 4 = dislike slightly, 5 = neutral—neither like nor dislike, 6 = like slightly, 7 = like, 8 = like very much, 9 = like very very much. In addition, subjects were asked to indicate how confident they were about each rating on a scale from 1 to 9. The ends of this scale were labeled not at all confident (1) and extremely confident (9).

In the second part of the experiment, subjects were presented with 12 adjectives and asked to describe their belief distributions associated with each adjective. They did this by estimating how many out of 100 persons described by the adjective they would like very very much, like very much, etc. The categories of likableness were the same nine used in part 1 of this experiment.

Each participant's instructions, stimulus materials, and response scales were included in a computer-generated booklet. It took each subject approximately 40 min to complete the experiment.

Design and Stimuli

There were two replications of the experiment using different subjects and partially overlapping stimulus sets. Twenty-five of the two-adjective descriptions were constructed from a 5 × 5 factorial design. One set of five adjectives used in this design was the same in both replications: *overcritical, absentminded, solemn, forward,* and *witty.* The likableness of the adjectives on Edwards' (1967) scale from 1 to 7 were 2.13, 3.02, 3.91, 4.78, and 5.66. The five orthogonal adjectives in Replication 1 were *nosey* (2.13), *domineering* (3.02), *lonely* (3.88), *persistent* (4.80), and *levelheaded* (5.66). The five orthogonal adjectives in Replication 2 were *quarrelsome* (2.12), *nonconfident* (3.02), *perfectionistic* (3.92), *idealistic* (4.82), and *individualistic* (5.66). The 10 adjectives presented to each subject in the factorial design were also presented singly.

The remaining six descriptions in part 1 were extreme filler adjectives *friendly* (6.55), *loyal* (6.54), *obnoxious* (1.32), and *dishonest* (1.22), and extreme filler pairs, *friendly and loyal* and *obnoxious and dishonest.* The adjectives for the distribution task (part 2) were the 10 adjectives included in the factorial design and the filler adjectives *friendly* and *dishonest.*

The first two stimuli in the first part of the experiment were the extreme filler pairs. The first two stimuli in the second part of the experiment were the filler adjectives. The order of the other stimuli in each part of the experiment and the order of the two adjectives in each adjective pair were randomly determined for each subject.

RESULTS

Analysis of Variance—Likableness Ratings

Mean likableness ratings of stimulus persons from both replications of the 5 × 5 design are shown in Fig. 2. These ratings were subjected to a 5 (adjectives common to both replications) × 5 (adjectives unique to each replication) × 2 (replications) repeated-measures ANOVA. If information integration is accurately described by an equal weight averaging model, all the lines in Fig. 2 should be parallel and the ANOVA should reveal no interactions between the two adjective factors. Instead, Fig. 2 reveals lines that diverge to the right as value increases. The adjective main effects and the adjective × adjective interaction were statistically
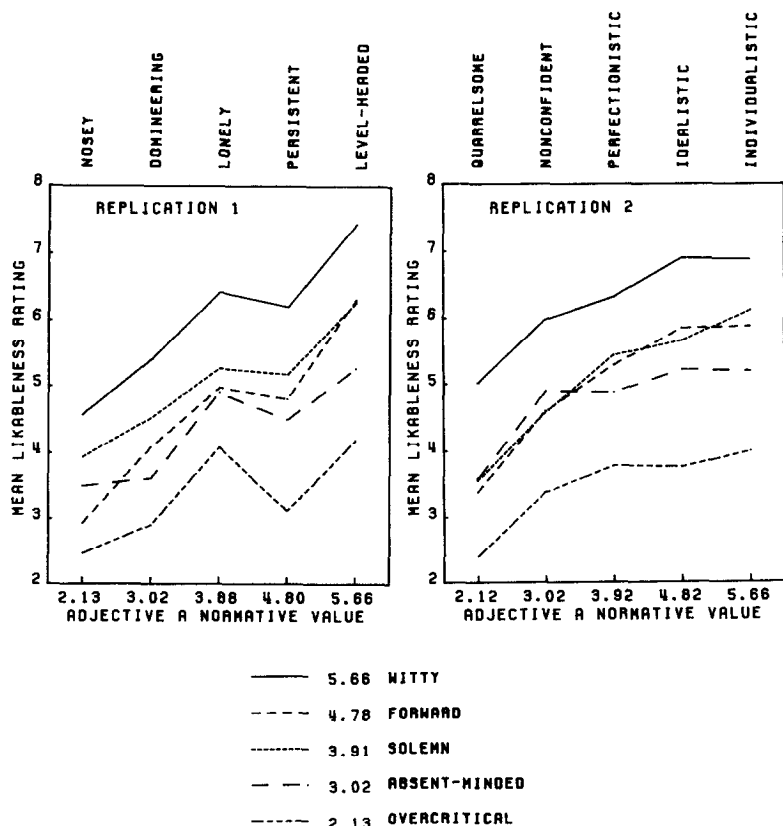

FIG. 2.   Mean likableness ratings of persons described by adjective pairs.

significant, p's $<$ .001. Consistent with the fanlike appearance of Fig. 2, 60% of the interaction variance was concentrated in the bilinear component. This interaction means that negative information was given more weight than positive information.

Fit of the Models—Likableness Ratings

The free parameters in the range and negativity models were estimated by a least-squares procedure. The optimal equation for the range model was

$$R_2 = .30R_H + .70R_L, \tag{1}$$

where R_2, R_H, and R_L are the ratings of a stimulus person (a) described by two adjectives, (b) described by the adjective given the higher rating, and (c) described by the adjective given the lower rating, respectively. The equation for the negativity model was

$$R_n = \sum_i (12.15 - r_i)r_i \Big/ \sum_i (12.15 - r_i) \tag{2}$$

using the notation of Table 1. The analysis used to estimate the free parameters also indicated that the range model, $F(1, 2149) = 307.4$, $p < .001$, and the negativity model, $F(1, 2149) = 180.3$, $p < .001$, fit better than the equal weight model.

In order to test the fit of the seven models, predicted likableness ratings were computed for each adjective pair separately for each subject using each model. For the equal weight, congruity, range, and negativity models, these predicted ratings were based on the ratings of the adjectives separately. For the confidence model, predictions were based on the single-adjective likableness ratings and the associated confidence ratings. The predictions for the concept identification model and the variance model were based on the belief distributions the subject gave for the two adjectives separately.

Figure 3 shows observed ratings as a function of predicted ratings for the 50 pairs of stimulus adjectives (averaged over subjects) for each of the models. The equal weight and confidence models were approximately correct in the predicting ratings of combinations of positive adjectives. Ratings of other stimulus combinations were more negative than predicted by these models. The data showed similar deviations from the congruity predictions—but the deviations were somewhat larger for the more neutral adjective pairs and somewhat smaller for the extreme pairs. The concept identification, variance, range, and negativity models were approximately correct in predicting ratings of the more neutral adjective pairs. However, ratings of the other adjective pairs were more extreme than predicted. These deviations were substantially smaller for the range and negativity models than for the concept identification and variance models.
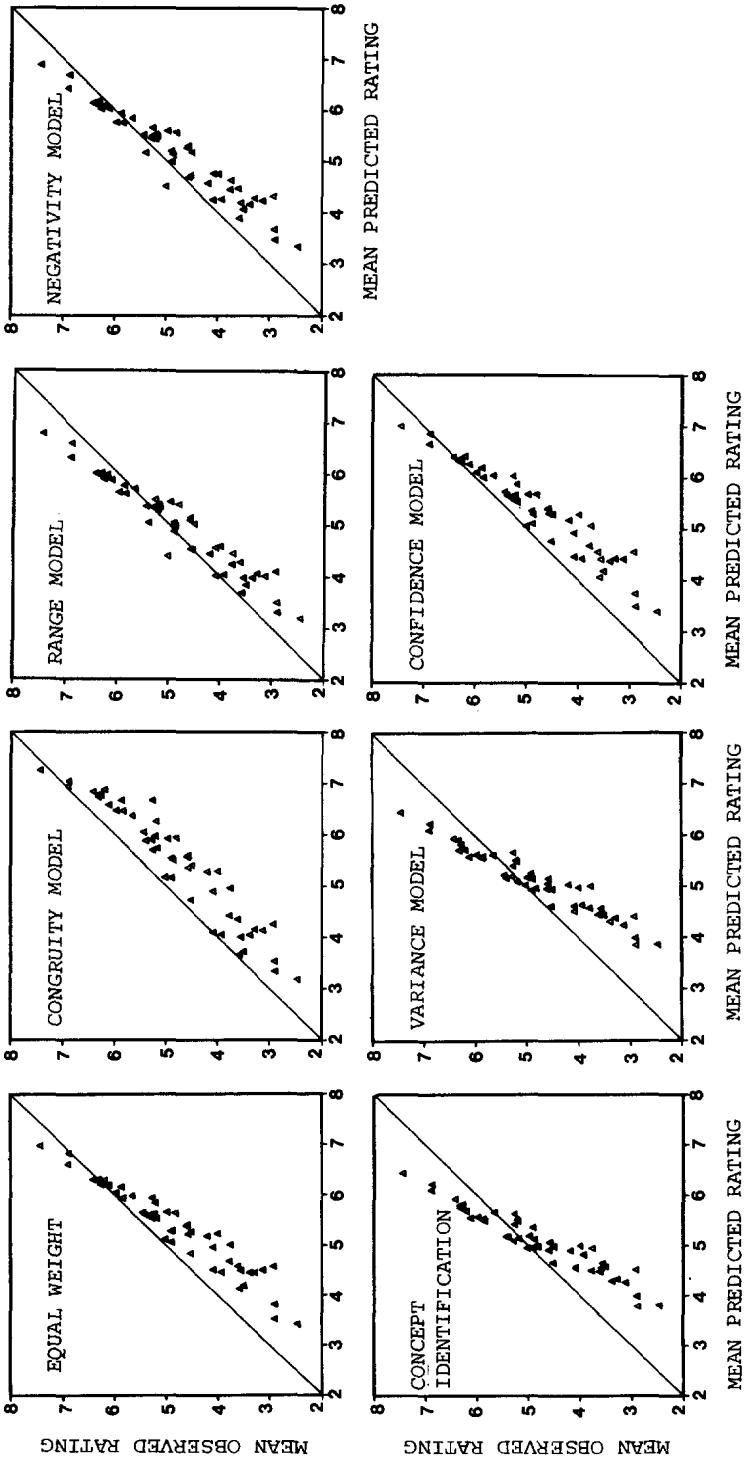
FIG. 3. Mean likableness ratings of persons described by adjective pairs as a function of predicted values based on seven models.

Two indices of fit between the predicted and observed ratings were computed for each subject: the standard error of prediction and the Pearson correlation. These indices had both been used by Wyer (1973). Dependent t tests were used to compare the fit of the seven models. The results are shown in Table 2. Fisher's r to z transformation was performed on the correlations before conducting the t tests. Mean z scores were transformed back to correlations for presentation in the table.

Three findings are notable in Table 2. First, the best performance was shown by models that give greater weight to negative information. Second, neither the congruity model nor the uncertainty-based models outperformed the equal weight model. Third, the standard error of prediction was more sensitive than the correlation to differences among the models. This last finding is not surprising since the correlation coefficient is not affected by either the slope or intercept of the relationship.

Analysis of Variance—Confidence

Mean ratings of confidence associated with the likableness ratings from both replications of the 5 × 5 design are shown in Fig. 4. Analysis of variance identified significant quadratic main effects ($p < .05$) and bilinear interaction ($p < .001$). These effects reflect the impact of four adjective pairs which were rated with particularly high confidence: the combination of two very positive adjectives or two very negative adjectives. These adjective pairs stand out clearly in Fig. 4.

Fit of the Models—Confidence Ratings

The variance model and the concept identification model predict the uncertainty associated with the likableness ratings. Uncertainty should be negatively correlated with confidence ratings. Pearson correlations among the confidence ratings and uncertainty predictions were computed for each subject, separately for the single- and double-adjective descriptions.

TABLE 2
FIT OF THE SEVEN MODELS

Model	Mean standard error of prediction	Mean correlation
Range	1.25[a]	.72[a]
Negativity	1.29[b]	.72[a]
Equal weight	1.33[c]	.73[a]
Confidence	1.35[c]	.72[a]
Variance	1.50[d]	.68[a,b]
Congruity	1.51[d,e]	.69[b,c]
Concept identification	1.54[e]	.64[c]

Note. Within a column, means not sharing a superscript are significantly different at $p < .001$. This α level was selected to compensate for the large number of pairwise comparisons.
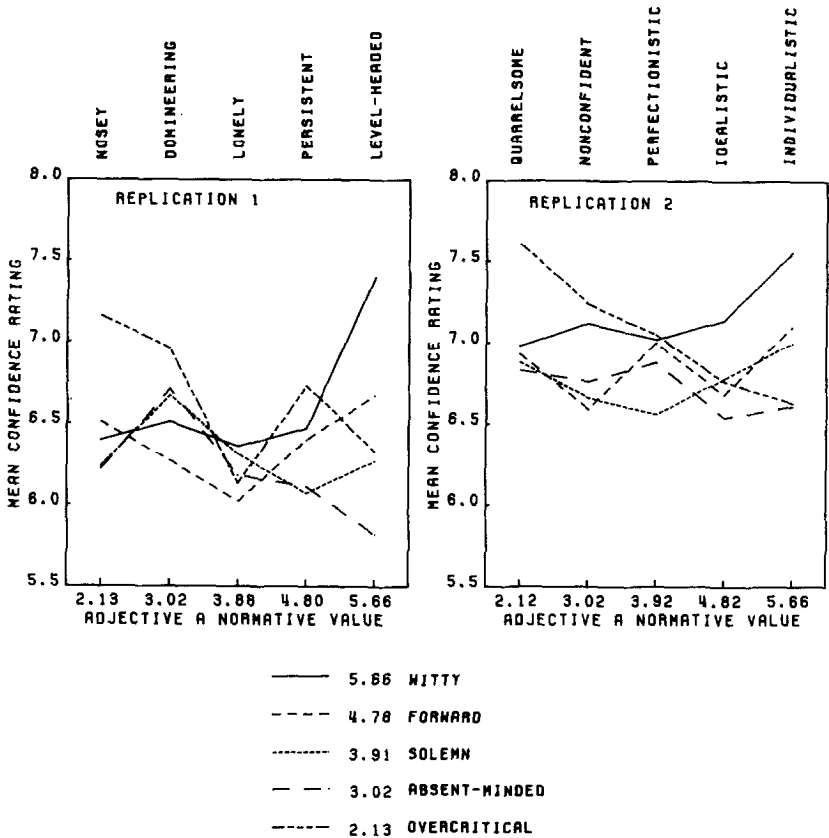
FIG. 4. Mean confidence in judging the likableness of persons described by adjective pairs

Fisher's r to z transformation was used before averaging and mean z scores were transformed back to correlations for presentation in Table 3. The predictions of the two models were reliably correlated for both single- and double-adjective descriptions. The correlation between confidence and predictions of the concept identification model for the single-adjective descriptions was also significant. None of the other correlations approached significance. At their very best, the uncertainty predictions were weakly correlated with subjective ratings of confidence.

DISCUSSION

Subjects judged the likableness of stimulus persons described by single adjectives and by adjective pairs, and rated their confidence in each judgment. They also created belief distributions to describe their beliefs about persons described by the single adjectives. Various mathematical models were used to predict ratings in the two-adjective conditions from measures in the single-adjective conditions.

TABLE 3

CORRELATIONS AMONG CONFIDENCE RATINGS AND UNCERTAINTY PREDICTIONS

	Concept identification model	Variance model	Confidence subjective rating
Uncertainty prediction — concept identification model	— —	$\bar{r} = .88^{**}$ $t(85) = 22.65^{**}$	$\bar{r} = -.10^*$ $t(82) = 2.41^*$
Uncertainty prediction — variance model	$\bar{r} = .57^{**}$ $t(85) = 12.97^{**}$	— —	$\bar{r} = -.05$ $t(82) = 1.04$
Confidence— subjective rating	$\bar{r} = .00$ $t(83) = .04$	$\bar{r} = -.03$ $t(83) = .96$	— —

Note. Correlations above the diagonal are for the single-adjective descriptions. Correlations below the diagonal are for the double-adjective descriptions.

* $p < .05$.

** $p < .001$.

Before discussing the major findings it is necessary to address one methodological limitation of the study: the dependence upon scatter plots, standard errors of estimates, and correlations as indicators of fit. Several authors (Anderson & Shanteau, 1977; Birnbaum, 1973a, 1974a) have pointed out that these methods (especially the correlation) are rather crude techniques for determining the underlying functional form. The strongest criticism is that it is possible for a model of the wrong form (e.g., additive) to fit better than a model of the correct form (e.g., multiplicative) if incorrect assumptions are made about the scaling of the measures. Thus our confidence in interpreting the results depends on our faith in the interval scaling (linearity) of likableness and the ratio scaling of confidence.

In this study, two techniques were used to avoid nonlinearities in the likableness measure: exclusion of extreme values as critical stimuli and inclusion of extreme filler items. In addition, the distribution of stimuli across the scale was fairly symmetric. Range–frequency theory (cf. Mellers & Birnbaum, 1983) suggests that a symmetric distribution facilitates linearity of the response scale. Because of these design features, linearity of the likableness scale is a very reasonable assumption. So we can be confident about the relative success of the range, negativity, equal weight, variance, congruity, and concept identification models. The ratio scaling of confidence is, however, quite doubtful. So we cannot be quite as confident about the fit of the confidence model.

Failure of the Uncertainty-Based Models

The confidence model, variance model, and concept identification model predict that items of information that individually lead to evaluations with low uncertainty will be given greater weight. None of these models

predicted ratings more accurately than the equal weight model. Two of the uncertainty-based models also made predictions about subjects' confidence in rating persons described by adjective pairs. The accuracy of these predictions was no better than chance. These findings are contrary to Wyer's (1973) findings that the concept identification model (a) predicted ratings more accurately than the unweighted average and (b) made uncertainty predictions that were reliably correlated with confidence ratings.

Three methodological differences may explain the discrepancies. First, Wyer (1973) used more extreme stimulus adjectives than were used in the current study. Their scale values ranged from 1.38 to 6.39 on Edwards' (1967) scale from 1 to 7, while the scale values in the current study ranged from 2.12 to 5.66. Second, the current study used extreme filler stimuli to avoid scale-end artifacts. Because some of the stimuli in Wyer (1973) were themselves extreme, fillers that were more extreme could not be used. Thus Wyer (1973) may have found more support for the concept identification model because it is more accurate when a wider range of scale values is included or because of ceiling and floor effects and artifactual constrictions of highly polarized belief distributions.

The third methodological difference between Wyer (1973) and the current study is that the present study calculated fit for each subject separately and then averaged over subjects. In Wyer (1973) the predicted and observed values for each stimulus were averaged over subjects before calculating measures of fit. When the current data were reanalyzed by that method, the fit of all models improved tremendously, and the variance model performed better than the equal weight model. Thus, averaging responses and predictions across subjects produced some misleading results, but did not lead to a replication of Wyer's findings.

In light of the above considerations it seems appropriate to conclude that the concept identification model and the other uncertainty-based models are not useful for predicting how people will integrate most impression formation stimuli (i.e., the broad middle range of scale values). While this does not mean that people ignore the relative informativeness of information, it does indicate that any gain in accuracy of prediction from considering belief distributions or confidence was more than offset by the measurement error associated with these variables. Studies reviewed by Slovic, Fischhoff, and Lichtenstein (1977) suggest that estimates of probability and reports of probability density functions are unreliable and strongly dependent on the method of elicitation. The weak correlations between the distribution-based uncertainty measures and confidence ratings suggest that these problems apply to the measures used in the current study.

If measurement error is responsible for the failure of uncertainty-based models in the current study, these models should be more successful if differences in the informativeness of information are increased. Evidence

for this possibility comes from studies in which the reliability of information was experimentally manipulated (Birnbaum, 1976; Birnbaum, Wong, & Wong, 1976; Surber, 1981). In these studies, judgments were well described by an averaging model in which items of information were weighted by their reliabilities. These findings support the general category of uncertainty-based models. They are consistent with the hypothesis that differences in the perceived informativeness of adjectives in the current study were insufficient to overcome the effects of measurement error.

Success of the Range and Negativity Models

Likableness ratings were most accurately predicted by the range and negativity models. In these models, greater weight is given to more negative information. The good fit of these models is consistent with the results of Birnbaum (1974b), and with many other studies showing the high weight that is given to negative information (cf. Anderson, 1981; Fiske, 1980; Kanouse & Hanson, 1972; Ronis, 1980; Wyer, 1973).

Uncertainty-based models can explain the negativity effect if negative information is seen as highly informative. In the current study, ratings of single adjectives should be negatively correlated with confidence and positively correlated with the distribution-based uncertainty measures. In fact, likableness ratings were not reliably correlated with confidence ratings. Higher confidence was associated with extreme ratings. But ratings did have the predicted positive correlations with the uncertainty measures, $\bar{r} = .14$, $t(85) = 2.77$, $p < .01$, for the variance model and $\bar{r} = .16$, $t(85) = 3.38$, $p < .002$, for the concept identification model. Averaging the measures across subjects increased the correlations to .41 $(df = 13, p > .10)$ and .61 $(df = 13, p < .05)$, respectively. These correlations provide modest support for an uncertainty-based explanation of the negativity effect. Future research on this theoretical possibility will require improved measures of uncertainty.

REFERENCES

Anderson, N. H. (1981). *Foundations of information integration theory*. New York: Academic Press.

Anderson, N. H. (1982). *Methods of information integration theory*. New York: Academic Press.

Anderson, N. H., & Shanteau, J. (1977). Weak inference with linear models. *Psychological Bulletin*, **84**, 1155–1170.

Bakan, D. (1954). A generalization of Sidman's results on group and individual functions. *Psychological Bulletin*, **51**, 63–64.

Birnbaum, M. H. (1973a). The devil rides again: Correlation as an index of fit. *Psychological Bulletin*, **79**, 239–242.

Birnbaum, M. H. (1973b). The nonadditivity of impressions (Doctoral dissertation, University of California, Los Angeles). *Dissertations Abstracts International*, **33**, 3333B. (University Microfilms No. 72-33897, 174)

Birnbaum, M. H. (1974a). Reply to the devil's advocates: Don't confound model testing and measurement. *Psychological Bulletin*, **81**, 854–859.

Birnbaum, M. H. (1974b). The nonadditivity of personality impressions. *Journal of Experimental Psychology Monograph*, **102**, 543–561.

Birnbaum, M. H. (1976). Intuitive numerical prediction. *American Journal of Psychology*, **89**, 417–429.

Birnbaum, M. H., Wong, R., & Wong, L. (1976). Combining information from sources that vary in credibility. *Memory & Cognition*, **4**, 330–336.

Edwards, J. D. (1967). *Revised likableness ratings of 554 personality-trait adjectives.* Mimeo available from the Ohio State University Social Psychology Training Program, Columbus, OH.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, **38**, 889–906.

Kanouse, D. E., & Hanson, L. R. (1972). Negativity in evaluations. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior.* Morristown, NJ: General Learning Press.

Manis, M., Gleason, T. C., & Dawes, R. M. (1966). The evaluation of complex social stimuli. *Journal of Personality and Social Psychology*, **32**, 278–282.

Mellers, B. A., & Birnbaum, M. H. (1983). Contextual effects in social judgment. *Journal of Experimental Social Psychology*, **19**, 157–171.

Meyer, S. L. (1975). *Data analysis for scientists and engineers.* New York: Wiley.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review*, **62**, 42–55.

Ostrom, T. M., & Davis, D. (1979). Idiosyncratic weighting of trait information in impression formation. *Journal of Personality and Social Psychology*, **37**, 2025–2043.

Ronis, D. L. (1980). Repetition and agreement with opposing arguments: A delayed action effect. *Journal of Experimental Social Psychology*, **16**, 376–387.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. *Annual Review of Psychology*, **28**, 1–39.

Surber, C. F. (1981). Effects of information reliability in predicting task performance using ability and effort. *Journal of Personality and Social Psychology*, **40**, 977–989.

Warr, P., & Jackson, P. (1975). The importance of extremity. *Journal of Personality and Social Psychology*, **32**, 278–282.

Wyer, R. S., Jr. (1973). Category ratings as "subjective expected values": Implications for attitude formation and change. *Psychological Review*, **80**, 446–467.

Wyer, R. S., Jr. (1974). *Cognitive organization and change: An information-processing approach.* Hillsdale, NJ: Erlbaum.