ERIC J. JOHNSON, ROBERT J. MEYER, and SANJOY GHOSE*

Linear compensatory models, which involve tradeoffs between product attributes, have been argued to provide reasonably good predictions of choices made by noncompensatory heuristics, which do not involve tradeoffs. This robustness to misspecification of functional form may fail, however, when there are negative correlations among attributes in a choice set. A Monté Carlo simulation demonstrates that certain noncompensatory rules are poorly fit by linear models, even in orthogonal environments, and that this fit diminishes further in nonorthogonal environments. Two laboratory experiments assess the extent to which such model failure might arise in natural contexts. The first, a process-tracing analysis, examines the decision strategies consumers use in nonorthogonal choice environments. The second explores the ability of a compensatory choice model calibrated on actual choices to predict decisions made in orthogonal and nonorthogonal contexts. The authors conclude with a discussion of the work's implications for current research in applied choice modeling.

When Choice Models Fail: Compensatory Models in Negatively Correlated Environments

Multiattribute preference models have a central role in describing and predicting consumer choice. Their wide appeal stems from a rather simple inherent property: their apparently robust ability to mimic consumer decision processes. It implies they can potentially be used to predict consumer reactions to product attribute changes prior to actual marketplace tests, affording sizable savings in time and money. Because of this possibility, multiattribute models have had numerous applications across a wide variety of marketing contexts, such as pre-test-marketing systems (e.g., Cattin and Wittink 1982; Green, Carroll, and Goldberg 1981; Silk and Urban 1978) and sales-response analyses of mature products (e.g., Corstjens and Gautschi 1983; Guadagni and Little 1983).

In most of these technologies for representing consumers' choice strategies, one assumes that consumers make tradeoffs between all the relevant attributes of a product and form overall evaluations of each alternative. Such decision rules are termed *compensatory* because good features of an alternative can compensate for the bad. If we listen to people describing how they make choices, however, a very different picture emerges. People talk about eliminating alternatives because of objectionable attributes and picking an alternative not because it has the best overall evaluation, but because it is the best on the most important attribute. Even the most casual examination of consumers' verbal reports indicates that decisions are based, at least in part, on *noncompensatory* strategies that do not involve tradeoffs. Supporting evidence for the use of such heuristics pervades the literature, including studies using information display boards (e.g., Bettman and Jacoby 1976; Lussier and Olshavsky 1979; Payne 1976), eye movements (Russo and Dosher 1983), and even in-store verbal protocols (Payne and Ragsdale 1978).

Do these findings suggest that our technology for rep-

*Eric J. Johnson is Associate Professor, The Wharton School, University of Pennsylvania. Robert J. Meyer is Associate Professor, Anderson Graduate School of Management, University of California, Los Angeles. Sanjoy Ghose is Assistant Professor, School of Business Administration, University of Wisconsin, Milwaukee.

The authors thank Kevin Keller and Dick Wittink for comments on a previous draft. The research was supported in part by funds provided by the Office of Naval Research and the Center for Research in Economics and Public Policy, UCLA. Both the Pascal source code used in the simulation and the Mouselab software used in study 2 are available from the first author.

Journal of Marketing Research
Vol. XXVI (August 1989), 255–70

resenting consumer choice strategies is in error? Considerable evidence suggets that there is no cause for alarm. Compensatory models are widely believed to be capable of mimicking a wide range of functional forms, including noncompensatory heuristics, when two conditions are met: (1) attributes are related monotonically to consumers' preferences and (2) there is error or uncertainty about these preferences (Dawes and Corrigan 1974). Because these conditions are thought likely to be satisfied in most applied contexts, compensatory models are widely used even when presumed not to reflect the actual process by which decisions are made (e.g., Cattin and Wittink 1982). Confidence in the model is widespread, as illustrated in Green and Srinivasan's (1978) conclusion about the potential dangers of misspecifying consumers' decision rules:

> . . . the compensatory model of conjoint analysis can approximate the outcomes of other kinds of decision rules quite well . . . even if the respondent's information processing strategy and decision model are complex, the compensatory models can usually produce good predictions.

Despite this optimistic note, there is still reason for concern that compensatory models many not always mimic noncompensatory processes. Several researchers (Curry and Faulds 1986; Curry, Louviere, and Augustine 1981; Einhorn, Kleinmuntz, and Kleinmuntz 1979; Green, Helsen, and Shandler 1988; Newman 1977) have argued that studies demonstrating robustness in linear models have underemphasized a major influence on predictive accuracy: the correlation among attributes in a choice or judgment set.

Interattribute Correlations and the Robustness of Linear Models

Newman (1977) was the first to examine explicitly the effect of interattribute correlations on the fit of linear models, following brief discussions offered by Einhorn (1970) and Goldberg (1971). He noted that, though Wainer's (1976) conjecture that small departures from optimal coefficients in a linear model "don't make no nevermind" will often hold in practice, it *will not* hold when the attributes describing alternatives are negatively correlated.

The intuition behind this conclusion is straightforward. Negative correlations among attributes in a choice set change the performance of models in two ways.

1. They lower the asymptotic descriptive validity (fit) of *any* model applied to the data (holding measurement error constant).
2. They increase the differences in predictive validity *across* alternative models.

The first result has a simple statistical rationale: negative correlations imply that if an option scores highly on one attribute it will tend to score poorly on another. When there are negative interattribute correlations in a choice set, the task of predicting which option is best by a compensatory criterion is more difficult because options will tend to be similar in overall value. An extreme example is when a model predicts that all the alternatives in a choice set are identical in value; here, one could predict choice with only chance levels of accuracy. Hence, negative correlations reduce the fit of models simply by constricting the variance in predicted overall values among options.

Negative correlations also can diminish model performance because they make predictions more sensitive to the particular choice strategy used. Einhorn, Kleinmuntz, and Kleinmuntz (1979, Figure 3) illustrate this effect through the example of a negatively correlated environment containing two options, described on two attributes. The first is good on attribute A and bad on B whereas the second is bad on A and good on B. The option that will be picked in this case depends on how a choice strategy weights the two attributes. A lexicographic strategy that selects the option best on attribute A, for example, would choose the first option. A similar lexicographic strategy looking first at attribute B would pick the second. In contrast, a compensatory rule that weights the two attributes equally would be indifferent. If the two attributes were correlated positively, if the first option were good on both attributes and the second bad on both, one would pick the first option regardless of how the attributes are weighted. Hence, in general, the less the redundancy among attributes in a choice environment (the more negative the average interattribute correlation), the greater the difference in predictions made by different models.

A more precise treatment of this issue was offered by Curry and Faulds (1986), who described the theoretical relationship that should be present between two different linear weighting schemes given different correlational structures. They did not look at the particular problem of the consequences of approximating one *functional form* with another, but they did establish some theoretical outer bounds for the levels of model failure that can occur with misspecification. For example, they noted that when two models are compared in a maximally negative intercorrelational environment, even small deviations in their weights can induce major discrepancies in prediction, approaching a complete rank-order reversal in overall preferences for options.

Conditions Mitigating Model Failure in Practice

Our discussion suggests that simple linear models may *not* be as robust to specification error as popular wisdom often maintains. Past findings of robustness may reflect the forgiving nature of the *choice contexts* used and not the inherent robustness of the models to misspecification. However, the discussion also suggests that the conditions required for model failure are actually rather specialized: the attributes in the choice sets under study must be correlated negatively and consumers must continue to use noncompensatory strategies when faced with these choices. The extent to which model failure occurs in the real world therefore depends on how often these two conditions occur.

First, will negative correlations among attributes arise frequently in real-world settings? There are several reasons why this might occur. Negative correlations are sometimes imposed by production constraints. It is difficult, for example, to make cars that both get good gas mileage *and* have drag-strip acceleration or to produce desserts that are both sinfully rich and low in calories. Negative correlations also naturally arise whenever dominant alternatives (options that are at least as good as others on all attributes and superior on at least one) are removed from consideration in a choice set (see, e.g., Curry, Louviere, and Augustine 1981). Hence, if one believes that either consumers or market forces efficiently screen dominated options, negative correlations should arise with reasonable regularity.[1]

Second, do consumers persist in using noncompensatory rules when faced with negatively correlated sets? Consumers may be implicitly aware that heuristics do not do as good a job of mimicking compensatory rules in negatively correlated environments. When faced with these decision contexts, consumers may increase their use of compensatory rules to make more accurate choices (see Einhorn, Kleinmuntz, and Kleinmuntz 1979; Payne, Bettman, and Johnson 1988). If so, compensatory models might not decline in predictive accuracy because loss in fit would be offset by an increased use of compensatory choice processes by adaptive decision makers.

Anecdotal evidence suggests negative correlations may have the opposite effect. Slovic (1975) and more recently Tversky, Sattah, and Slovic (1987) report that when individuals are faced with choices between equally valued options (which imply negative correlations among attributes), they use a lexicographic strategy to pick the option that is best on the most important attribute. If this tendency is generalizable, it suggests an even greater potential for the failure of compensatory models in negatively correlated environments.

The Research Problem

The purpose of our article is to examine further the robustness of linear models in mimicking noncompensatory processes in correlated environments. Our objective is to address three central questions left unresolved in previous research.

1. To what extent can specific noncompensatory rules be approximated by compensatory models and how are these approximations affected by different levels of negative interattribute correlation?
2. To what extent do consumers use noncompensatory rules when faced with negative interattribute correlations? Do they make greater use of compensatory strategies because of the reduced accuracy of heuristics or, as suggested by Slovic (1975), actually make *greater* use of heuristics?
3. How important is model failure in applied settings when a compensatory model is used to predict choices made in negatively correlated environments?

We report three studies addressing different aspects of these three research questions. The first is a numerical simulation investigating the ability of a linear model to predict the choices made by several noncompensatory rules in environments with varying interattribute correlations. The second is a process-tracing study examining how *consumers,* rather than *models,* react to changes in correlational structure. Of interest is the extent to which consumers alter their choice strategies as interattribute correlation changes. The third study combines the interests of the first two by illustrating the levels of model failure that may arise in applied settings. It examines the changes in predictive validity when a choice model calibrated on a natural sample is used to predict actual choices made within sets differing in size and correlational structure.

STUDY 1. A NUMERICAL SIMULATION

Overview

Prior research suggests that linear models do a surprisingly good job of mimicking noncompensatory processes, both in judgment (Dawes and Corrigan 1974) and choice (Johnson and Meyer 1984). The purpose of the simulation is to examine this ability in choice environments with different interattribute correlations. The simulation differs from previous research in two ways. First, though previous studies have examined the ability of linear models to represent noncompensatory choice rules in orthogonal settings, the evidence for noncompensatory processing has been indirect; for example, researchers have pointed to the presence of interactions in response data or evidence from protocols (e.g., Johnson and Meyer 1984; Olshavsky and Acito 1980). In contrast, we examine the performance of linear models in contexts where data are *known* to be generated by specific *a priori* heuristics. Second, previous investigations of the effect of interattribute correlations have focused on cases in which the locus of error is improperly specified attribute weights in a linear model. Our goal is to examine the more general case in which error is present in both attribute weights and the *functional form* of the decision rule that underlies choice.

We stress that the goal of the simulation is not to identify the level of model failure that might arise in actual applications. Rather, our objective is to explore the kinds

[1]The extent to which real-world product markets are efficient is currently a subject of debate. Curry and Menasco (1979), Curry and Faulds (1986), and Hjorth-Anderson (1985) examined market efficiency using *Consumer Reports*-type measures of product quality. They reported dominated options persisting in several markets, with interattribute correlation matrices thus tending to be *positive* rather than negative. Limiting these studies, however, was their focus on objective rather than subjective measures of product quality. In addition, Tellis and Wernerfelt (1987) report a meta-analysis of nine studies of the relationship between product quality and price. They found consistent support for the predicted positive correlation (implying a negative correlation in utilities) in most cases for these two variables.

of errors that *could* occur when a particular noncompensatory rule is employed. The extent to which failure might arise in practice is the focus of studies 2 and 3.

Method

Overview. We examine the ability of a linear compensatory model to represent four riskless choice heuristics: elimination by aspects (EBA), lexicographic, conjunctive (satisficing), and a phased EBA/compensatory rule, in which EBA initially screens options and a compensatory rule makes the final choice. These rules are selected because they represent heuristics commonly discussed in the literature (e.g., Engel and Blackwell 1982) and represent different dimensions used to categorize heuristics. They use different amounts of information and represent both brand and attribute processing.[2] Hence, our hope is that they provide a set of boundaries for the types of errors likely to arise in practice.

General procedure. The simulation involved three stages: designing a set of choice environments, generating choices within those environments, and assessing the descriptive and predictive validity of compensatory models calibrated on those choices. We used a hypothetical choice problem in which each decision heuristic selected one alternative from a set of four options, each described by four attributes.[3] Our central manipulation was the correlation among attributes within each such set. Four different correlation patterns were examined.

1. Orthogonal sets, or choice sets with no correlation among attributes.
2. Efficient sets from which dominated alternatives were removed, causing a modest negative (−.19) correlation among attributes.
3. Sets with a maximum possible average negative rank-order correlation between all attributes (−.33; see Green and Krieger 1986).
4. Sets with maximum possible negative rank-order correlation (−1) between a single pair of attributes.

Within each correlation condition a total of 100 choice sets were created. The choice sets were generated by a Pascal implementation of the IMSL subroutine GGNML, in which each attribute was transformed to a uniform variate with values ranging from 0 to 1000. For each set, a 4 (number of alternatives) by 4 (number of attributes) matrix was generated having the required pattern of intercorrelation.

The choices that would be made in each decision environment by each heuristic decision rule then were computed. Following Johnson and Payne (1985), we implemented the decision rules as production systems or systems of elementary operations using "if-then" productions. The operationalizations follow.

1. *Elimination by aspects (EBA).* Our representation was a variant of Tversky's (1972) original description, similar to that used by Thorngate (1980). We screened attributes in descending order of importance, eliminating all alternatives below a certain cutoff level. To reflect attribute importance, cutoff levels were proportional to the weights of a compensatory rule (described subsequently). Choices were made either by selecting the last alternative not yet eliminated or by a random selection among those that had passed all cutoffs.
2. *Lexicographic.* This rule picked the alternative that had the highest value on the most important attribute.
3. *Conjunctive (satisficing).* This rule used the same cutoffs as EBA. Here, however, search was by alternative and the first alternative that passed all the cutoffs was chosen.
4. *Phased EBA/compensatory.* The EBA heuristic was applied until the choice set was reduced to two candidates. The chosen option was the one that was best by a weighted linear combination of attributes.

To provide a baseline, we also generated a set of choices made by a random and an additive compensatory rule. The random choice rule selected an option from each set with a constant $1/N$ probability. The additive rule selected the option with the highest weighted linear combination of attribute values.

Each decision rule employed weight parameters of .5, .25, .125, and .125 for each of the four attributes, respectively. The meaning of the parameters depended on the rule. In the *phased* EBA and conjunctive rules, these values defined the cutoff levels. In the compensatory rule, they defined the weights used in the linear combination. Finally, in the EBA and lexicographic rules, they defined the order in which attributes were considered.

Each of the five rules made 100 choices within each of the 100 generated choice sets. To introduce variance in the observed choices, we added random noise to the weight parameters on each trial, sampled from a normal distribution with a zero mean and a variance of .2. This

[2]A larger set of rules might have been considered (such as other types of phased rules), but we felt these were sufficiently representative of the heuristics proposed in the literature to serve as a starting point in a study of error patterns. In our simulations we also explored an additional phased rule, EBA followed by a majority of confirming dimensions (e.g., Russo and Dosher 1983). Because the findings for that policy closely mirror those for the better-known EBA/compensatory policy, we restrict our discussion to the latter rule. Simulation results for the EBA/MCD policy as well as the MCD policy alone are available from the authors upon request.

[3]One potential concern with our decision to focus on the four-attribute case is that, because the largest average negative interattribute correlation that is attainable diminishes with increases in the number of attributes, "real-world" applications employing larger numbers of attributes would involve negative correlations much smaller than that we examine. To explore this issue, we conducted a numerical experiment using Curry and Fauld's (1986) formula for the expected correlation between two linear weighting schemes given a prespecified interattribute correlation matrix of different sizes. This analysis (available from the authors upon request) indicated that the size of the interattribute correlation *per se* is not a critical factor in determining the predictive accuracy of a misspecified linear model. Rather, it is the size of the correlation in relation to the largest attainable for the given number of attributes. Hence, we have reason to expect that the analyses reported here would generalize to experiments with larger numbers of attributes.

level of noise was chosen for two reasons. First, it is sufficient to prevent choice frequencies of 0 or 1 across the 100 trials, which would cause difficulties in least squares estimation of the choice models. Second, it resembles the levels of reliability commonly observed in judgment studies, approximating a reliability, *r*, of .8. Our simulations therefore can be thought of as representing the choices made by a homogeneous segment of decision makers whose decision criteria contain a reasonable and realistic amount of random error.[4]

Compensatory model. Our central interest was assessing the predictive validity of a compensatory choice representation in modeling the choices generated by six decision heuristics: EBA, lexicographic, conjunctive, compensatory, phased EBA, and random. Because our criterion was a discrete choice from a set, we estimated the compensatory model for each combination of heuristic and correlation condition by using the multinomial logit:

$$P(i|j = 1, \ldots, 4)_{kl} = \frac{e^{\beta'_{kl}X_i}}{\sum_{j=1}^{4} e^{\beta'_{kl}X_j}} \tag{1}$$

where $P(i|j = 1, \ldots, 4)_{kl}$ is the proportion of times (in 100) option i was chosen by choice heuristic k in correlation condition l and $\beta'_{kl}X_i$ is a best-fitting linear combination of i's scores on each of four attributes. The parameter vector $\boldsymbol{\beta}_{kl}$ was estimated for each rule and correlation condition via OLS in a manner suggested by Theil (1971).[5] Specifically, the proportion of choices allocated to each option in each set by each rule was taken as a sample estimate of $P(i|j = 1, \ldots, 4)_{kl}$. The log of the ratio of these proportions for each independent pair in a set defined a system of three binary logit equations whose common parameter vector $\boldsymbol{\beta}_{kl}$ could then be estimated.

Results

How well does a compensatory model represent the choices made by heuristics with different patterns of correlation among attributes? We undertook two different approaches to assessing model performance.

1. We noted each model's *estimation fit* or the ability of a compensatory model to represent choices made by each heuristic within each correlational environment.
2. We computed a *cross-validation fit* or the ability of a compensatory model estimated in an orthogonal environment to predict choices in nonorthogonal environments.

The first analysis was seen as providing a best-case look at the ability of a compensatory model to mimic a noncompensatory rule in different environments: the coefficients of the compensatory model were those that best represented the observed set of choices within a given correlational structure. The second analysis assessed model performance in a way that mirrored how compensatory models often are estimated and applied in practice. A linear model was estimated in an orthogonal setting and then was used to predict choices made in nonorthogonal environments (e.g., Green, Carroll, and Goldberg 1981; Green, Helsen, and Shandler 1988).[6] Both analyses examined two indices of performance: the correlation between predicted and observed market shares, using log-odds ratios, and the first-choice hit rate or the proportion of decisions in which the model correctly predicted the alternative chosen by the heuristic.

To provide an initial look at the results, we conducted a meta-analysis of the two performance indices across experimental conditions. The objective of the analysis was to test the null hypothesis that the fit of a compensatory model is insensitive to variations in processing rule, environment, and dataset used for assessing fit. Variation in each fit index was expressed as a function of 18 binary variables identifying rule type (EBA, lexicographic, conjunctive, phased, or random), correlational structure (orthogonal or negatively correlated), dataset (estimation or validation), and two-way interactions between correlation, dataset, and rule. The three negative correlation conditions were collapsed into a single level to provide a common directional test of the hypothesis that model performance would be lower in negatively correlated environments. Variations in performance across different levels of negative correlation are examined subsequently.

The analysis of hit rate as a fit criterion posed special problems because it is a censored measure. Though correlation is bounded by the theoretical limits of association between models (either perfect association or perfect disassociation), hit rate is not; observations of association are truncated (by 100% and chance hit rates) before these limits are reached, implying that the relationship between hit rate and model performance is inherently nonlinear. Changes in hit rates near 1 and chance (in this case, .25) hold larger implications for changes in true model performance than changes in hit rates near .5 or .6. Recognizing this problem, we subjected observed hit rates to a logit transformation prior to analysis (see, e.g., Maddala 1983).

[4]In pilot work we also explored higher levels of error. The results mirrored those reported here with the exception of the predictable lower average predictive validity.

[5]Weighted least squares estimation also was pursued, with almost identical results.

[6]Though we might also discuss the cross-validation of nonorthogonal models, such an analysis was of somewhat lesser interest for two reasons. First, because the parameters of nonorthogonal models would tend to have inflated standard errors, the interpretability of predictive accuracies of nonorthogonal models would be clouded by the confounding of error due to true model misspecification with error due to inefficient parameter estimation. Second, this mode of cross validation is less common in real-world settings; compensatory models frequently are estimated in orthogonal environments and used to predict in nonorthogonal ones, but the converse is less common.

The results of a WLS regression analysis for each fit index are summarized in Table 1. The significance levels we report are derived from partial F-tests computed by using the average within-cell theoretical sampling variance of each observed correlation and hit rate. In this table the intercept measures the fit of a compensatory rule in an orthogonal environment in estimation and all other parameters are measures of contrasts with this base level of performance. The main effect of correlation listed in the table, for example, is a measure of the change in the fit of a compensatory model of a compensatory rule when it was applied to a negatively correlated setting.

Though the effects of the experimental treatments vary somewhat by fit index, the table suggests four general conclusions.

1. The fit of all approximations, including the compensatory heuristic, decreased when they were applied in negatively correlated environments.
2. The compensatory model fit noncompensatory models less well than true compensatory processes.
3. Effects 1 and 2 interact for most rules; the negative effect of correlation on the model's performance is worse when approximating a noncompensatory heuristic.
4. There is little effect of estimation versus cross-validation setting on the findings.

The finding of a negative effect of correlation is not surprising, as it largely reinforces the observations of Curry and Faulds (1986) and others that imperfectly specified linear models will decrease in fit when applied in negatively correlated settings. What *is* new is the evidence of an interaction between rule type and correlation; seven of the eight rule-by-correlation interactions (excluding the random rule) are negative in sign, with four being significantly negative, implying that the effect of negative correlation is magnified by modeling noncompensatory choice processes. This accentuated decrease is most pronounced for the correlation of the fit of the lexicographic and conjunctive rules and the hit rate of the conjunctive rule.

To provide a more detailed look at the findings, in Figures 1 and 2 we plot the observed squared correlations and hit rates, respectively, for a linear model fit to each noncompensatory heuristic in each of the four correlational environments, pooling across validation and estimation contexts. To simplify the figures, we omit the random rule, which yields chance levels of prediction. The figures suggest two additional insights about the effects of rule and environment.

1. The primary effect of negative correlation is a contrast between performance in an orthogonal environment and the three negative environments; across rules there are comparatively small differences among the different levels of negative correlation.[7]
2. The compensatory model fails to approximate some noncompensatory rules even in orthogonal environments, with one rule, the conjunctive, failing noticeably by both fit criteria.

One possible reason for the consistently bad fit of the conjunctive rule by the compensatory model is that it was the only policy considered that did not utilize information about *relative* attribute values to at least some degree. In this rule, the decision maker sequentially scanned the set of options and stopped as soon as one exceeded a set of critical threshold values. In contrast, the EBA and the lexicographic rules made at least some effort to establish an option's relative value on an attribute.

Table 1
REGRESSION ANALYSIS OF CORRELATIONS AND HIT RATES, COMPENSATORY APPROXIMATIONS OF HEURISTICS

Source	*Correlation estimate*	*Hit rate[a] estimate*
Intercept	.851[b]	.070
Correlation	−.030	−.594[c]
Dataset	.009	.100
EBA	−.020	−.772[c]
Phased EBA	−.009	−.369[c]
Lexicographic	−.198[c]	−.260
Conjunctive	−.156[c]	−2.020[b]
Random	−.559[b]	−4.050[b]
Data × corr.	−.040	−.070
Data × EBA	−.006	−.078
Data × phased	−.001	−.072
Data × lex.	.006	−.251
Data × conj.	−.034	.278
Data × random	−.137[b]	.125
Corr × EBA	−.075	−.100
Corr × phased	−.026	−.289
Corr × lex.	−.206[b]	.247
Corr × conj.	−.298[b]	−1.070[b]
Corr × random	.086	.125
	Model R^2 .954	Model R^2 .965

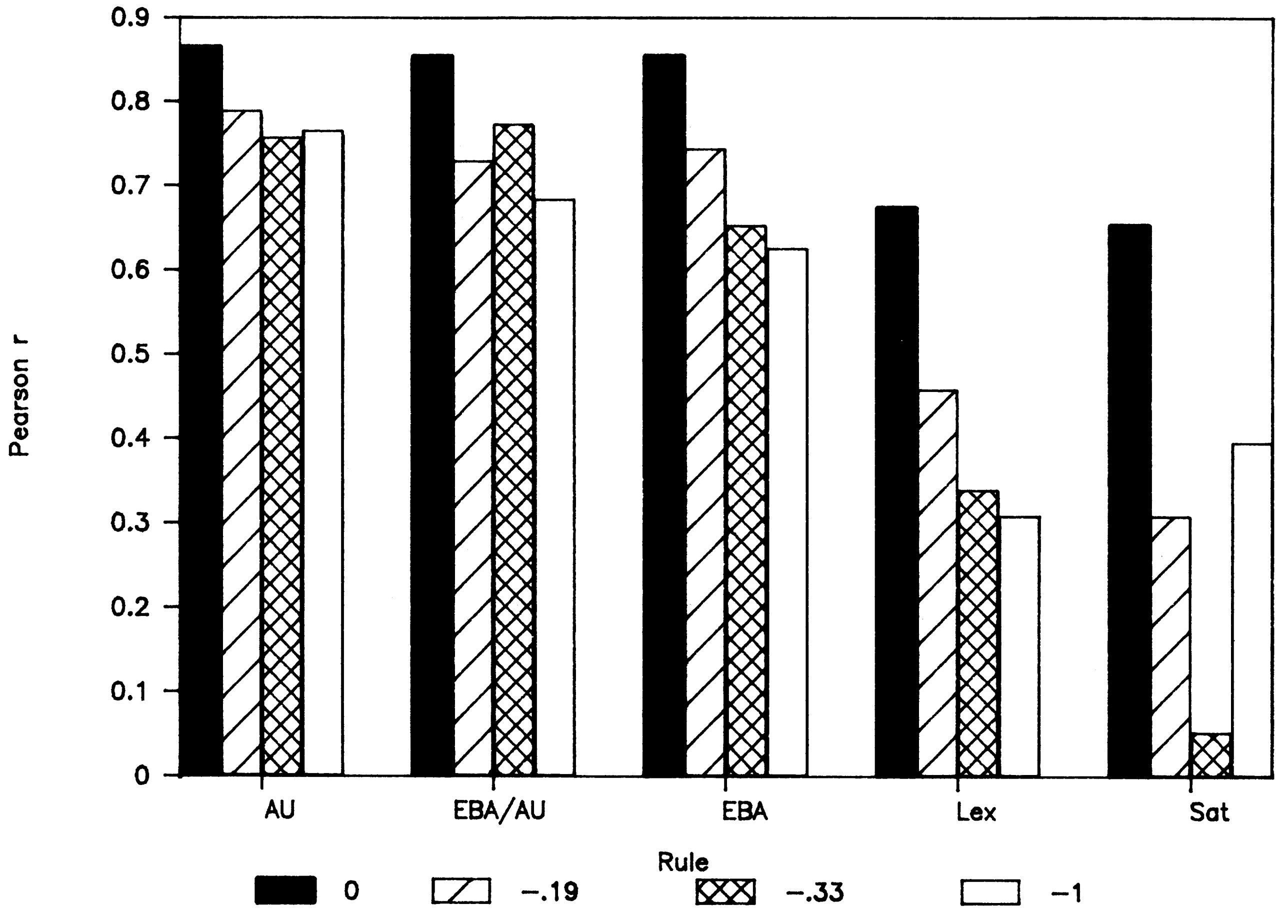
[a]Hit rates were subjected to a logit transformation prior to analysis.
[b]Probability that the true value of the stated coefficient is zero is less than .01.
[c]Probability that the true value of the stated coefficient is zero is less than .05.

Correcting Misspecification Through Multilinear Model Forms

The failure of compensatory models to mimic noncompensatory processes naturally raises a question: Could fit be improved significantly and the effects of negative correlation overcome by a more complex form of the linear model? Several authors have suggested that some noncompensatory processes, particularly the conjunctive, might be represented better by linear models that include interactions among attributes (e.g., Einhorn 1970; Green and Devita 1975; Louviere 1988; Lynch 1985).

[7]We conducted contrasts of the mean fit rates among the three negative correlation levels, averaging over rules, and were unable to reject a null hypothesis of equality in performance levels among the conditions.

Figure 1
MEAN CORRELATION (PEARSON r) BETWEEN PREDICTED AND OBSERVED MARKET SHARES
(by processing rule and correlation condition)[a]



[a]Pearson r's are averaged across estimation and validation datasets, which give comparable results. AU denotes an additive utility (compensatory) rule, EBA the elimination-by-aspects rule, EBA/AU a phased EBA rule, Lex a lexicographic rule, and Sat a satisficing rule.

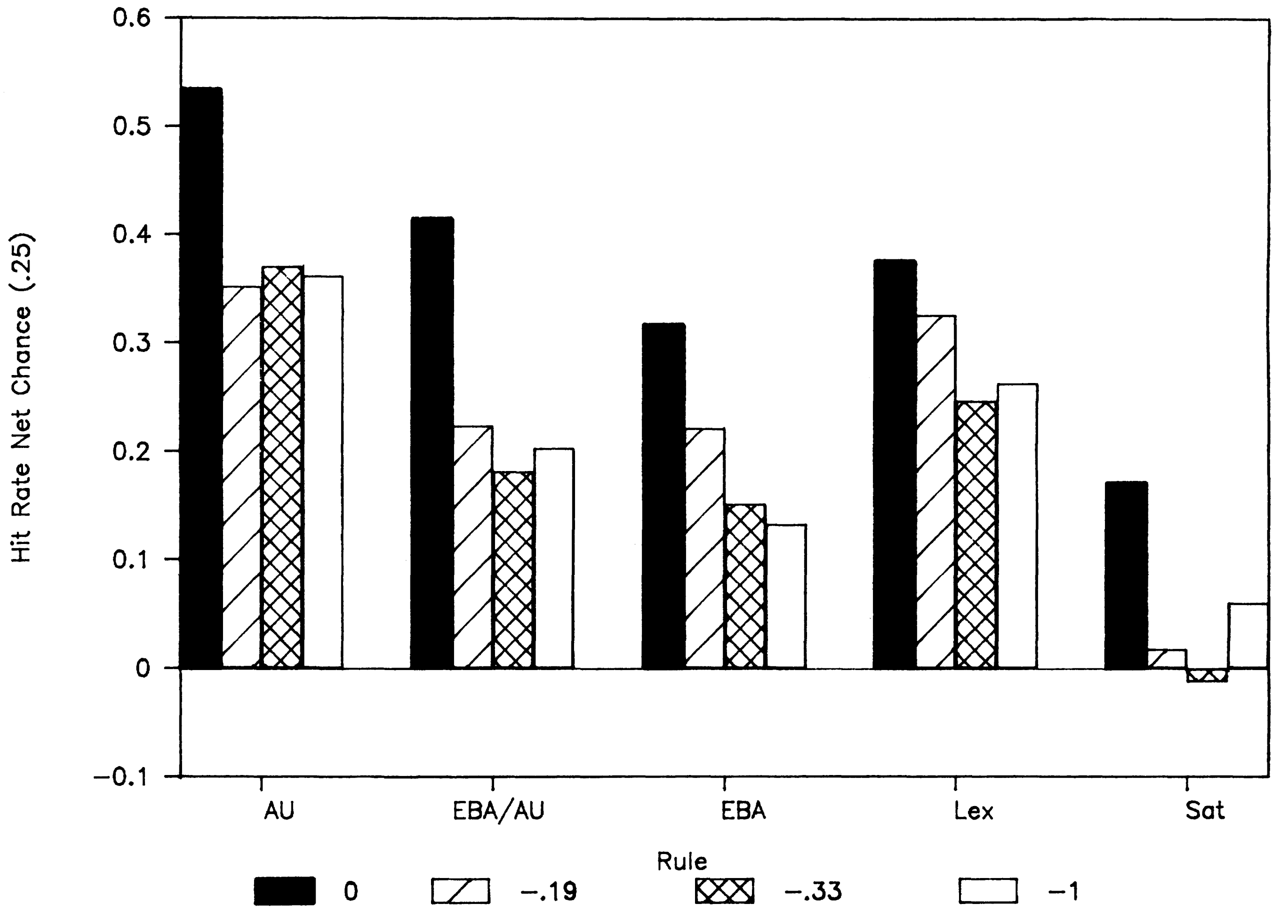
To examine this possibility, we reestimated the compensatory models for each heuristic and correlation condition as before, but expanded each model to include all six two-way interactions created from the four predictors. The results of this analysis are summarized in Figures 3 and 4, in which we plot the mean improvement in correlations and hit rates, respectively, in the validation data when interactions are included.

The results are striking: in the negatively correlated environments, adding interactions has a consistently positive effect across rules, often providing substantial improvements. The average increase in the validation r for the EBA, lexicographic, and conjunctive policies is .12, whereas the average improvement in hit rate for the phased EBA, EBA, and lexicographic policies is .04 (both greater than the improvements expected by chance). Perhaps more importantly, similar increases in fit are not observed in orthogonal environments. Adding interactions in the orthogonal setting significantly helped the approximation of only one rule, the conjunctive (and that by a modest degree). Indeed, for the other processes, adding interactions in the orthogonal settings actually *decreased* validation correlations and hit rates; the average correlation decreased by .01 and the average hit rate decreased by .04. Thus, whereas adding interactions in orthogonal environments appeared simply to result in "overmodeling," interactions substantially helped predictive accuracy in negatively correlated environments.

Discussion

The numerical simulation confirms the basic hypotheses that negative correlation between attributes diminishes the performance of a compensatory model and that the effect is amplified in modeling noncompensatory rules. It also unveils two unexpected results: the low average fit of some compensatory approximations even in

Figure 2
MEAN DISAGGREGATE HIT RATE EXPRESSED AS DEVIATION FROM CHANCE (.25)
(by processing rule and correlation condition)[a]



[a]Hit rates are averaged across estimation and validation datasets, which give comparable results. For abbreviations, see footnote to Figure 1.

orthogonal environments (the lexicographic and conjunctive policies as assessed by correlation and the EBA, phased EBA, and conjunctive policies as assessed by hit rate) and a markedly increased benefit of including interactions in compensatory approximations applied in negatively correlated environments.

The finding of model failure in orthogonal settings is somewhat surprising because it runs counter to the conventional wisdom that compensatory models are robust to process misspecification when applied in orthogonal contexts (e.g., Olshavsky and Acito 1980). The reason for the conflict presumably lies in the fact that previous studies have examined model performance with only *indirect* evidence for noncompensatory processing, such as through verbal protocols. Such procedures undoubtedly establish the existence of heuristic processing, but they do not establish the existence of any specific heuristic used in a uniform way. In contrast, the simulation examines the case in which data are *known* to be generated by a specific heuristic applied with relative homogeneity. Hence, if consumers in a population are strictly following a common noncompensatory heuristic when making decisions—particularly a conjunctive rule—one can expect a poor fit of the linear model even in orthogonal environments.

Our finding that the inclusion of interactions in a compensatory model has an increased beneficial effect in negatively correlated settings is perhaps even more intriguing. In the simulation, interactions that explained little additional variance—and even *hurt* explanation—in an orthogonal environment significantly improved model performance when they were applied in negatively correlated settings. The value of adding interactions to a compensatory model therefore depends not only on the nature of the underlying choice process, but also on the structure of the choice sets to which the model is being applied. When a compensatory model is applied to an orthogonal or positively correlated environment, interactions may do little to help predictive accuracy and may even hurt prediction. When the same model is applied to a negatively correlated environment, however, interactions may yield significant improvements in accuracy. The issue of when to include interactions in a linear model, though not pursued further here, clearly warrants further research.

LABORATORY EXPERIMENTS

Overview

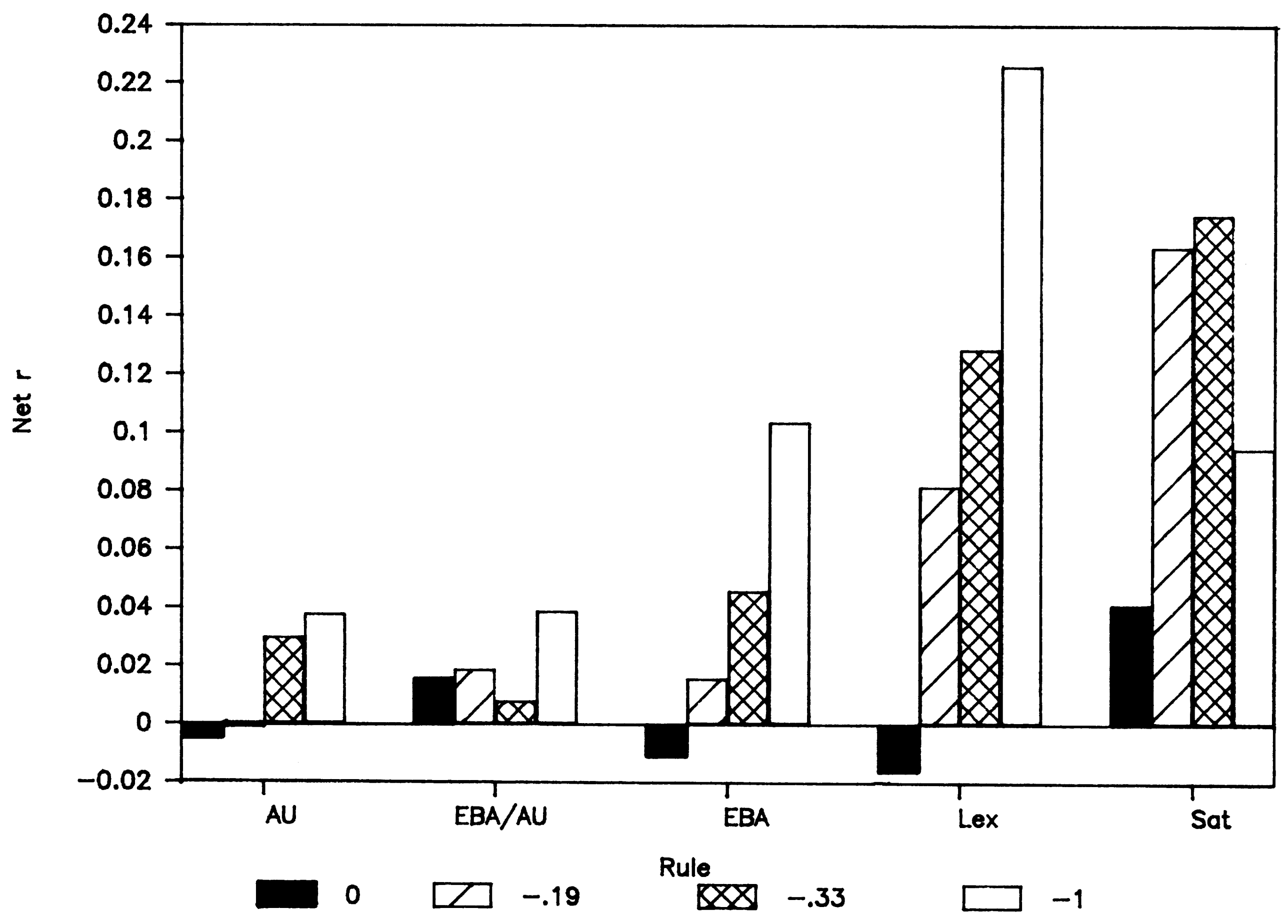
The simulation gives a clear message: the predictive validity of compensatory models of consumer choice processes *can* be affected by the structure of underlying choice rules and the pattern of correlation among attributes in a choice set. However, study 1 merely provides an existence proof; the level of model failure that is likely to occur in natural contexts is unclear. The degree of model failure will depend on the amount and nature of noncompensatory processing, the empirical pattern of correlation among attributes, and other influences not considered in the simulation.

We conducted two investigations to study actual, not simulated, decisions in correlated environments. Study 2 was a process-tracing analysis of choice strategies used by consumers when faced with different patterns of interattribute correlation. The study examined whether consumers indeed use noncompensatory heuristics when placed in negatively correlated environments or adapt their choice strategies in response to changes in correlation. Study 3 explored the ability of an aggregate compensatory model to predict choices made in orthogonal and nonorthogonal contexts. Our interest was in illustrating the level of model failure that may arise in studies with natural heterogeneous samples.

Figure 3

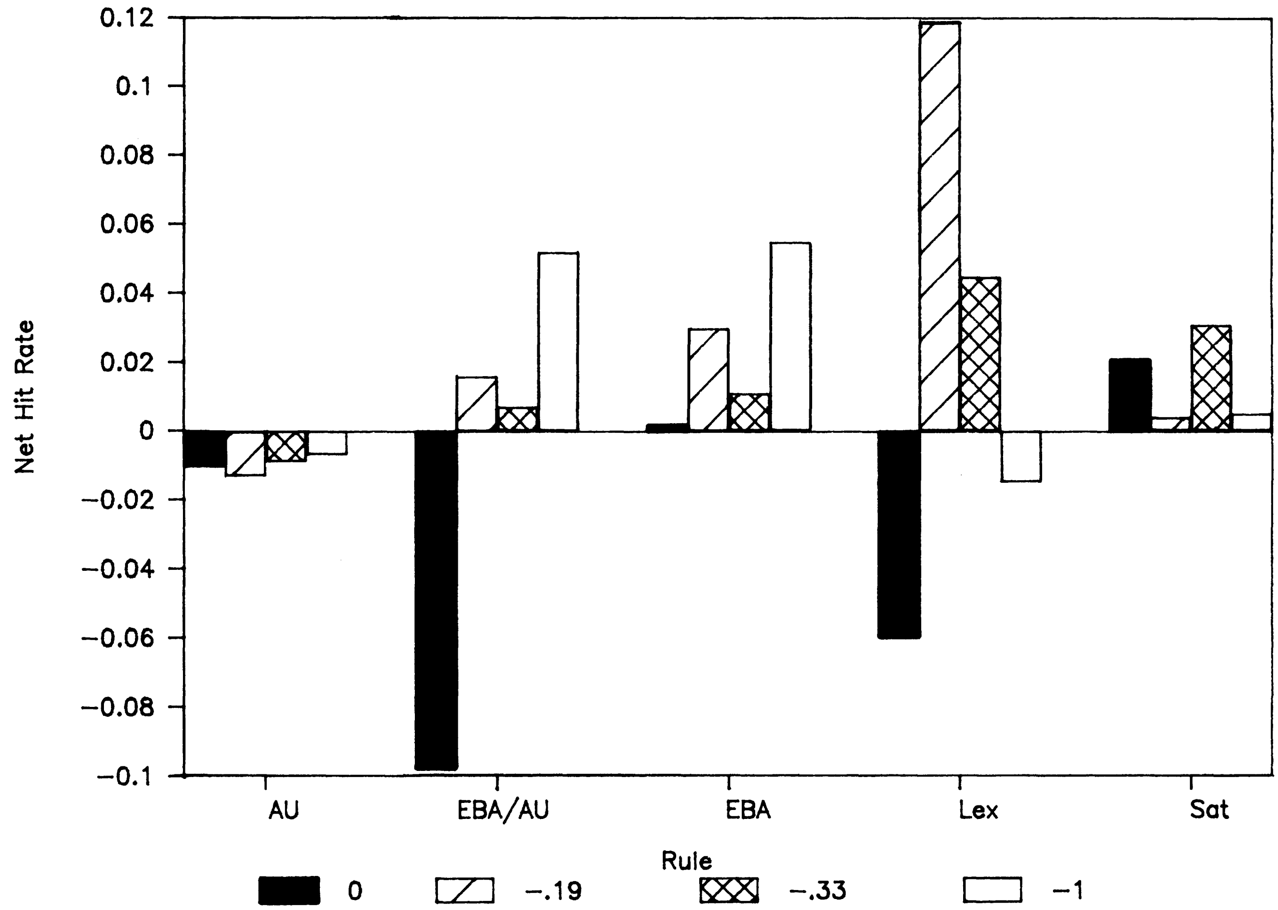
IMPROVEMENT IN CORRELATION BETWEEN PREDICTED AND OBSERVED VALIDATION MARKET SHARES OBTAINED BY ADDING TWO-WAY INTERACTIONS TO A LINEAR MODEL

(by processing rule and correlation condition)[a]



[a]For abbreviations, see footnote to Figure 1.

Figure 4
IMPROVEMENT IN DISAGGREGATE VALIDATION HIT RATE OBTAINED BY ADDING TWO-WAY INTERACTIONS TO A LINEAR MODEL
(by processing rule and correlation condition)[a]



[a]For abbreviations, see footnote to Figure 1.

Study 2. A Process-Tracing Study of Choice with Correlated Attributes

Overview. The experiment centered on subjects' preferences for hypothetical apartment alternatives. Mirroring the stimulus profiles used in a previous study (Johnson and Meyer 1984), the apartment descriptions gave values on four attributes: rent (dollars per month), distance to campus (in minutes walking time), level of maintenance (a verbal description), and appearance of the apartment and its neighborhood (a verbal description) (see Johnson and Meyer 1984 for examples of the stimuli). Our objective was to monitor the process by which subjects make choices given sets differing in terms of two criteria, the number of options (2 or 8) and the pattern of interattribute correlation.

To monitor choice processes we used the Mouselab system (Johnson et al. 1988), which presented the choices on a computer display in an alternative-by-attribute matrix. The values for the alternatives and attributes were concealed behind a blank box. To examine an attribute value, the subject moved a cursor to the box by using a mouse. The display then instantaneously displayed a value, which remained visible until the cursor left the box. The software unobtrusively recorded the order, duration, and frequency of search along with the choice made by the subject. After a brief training period with the choice task and mouse, subjects reported that they found the process and pointing device very natural. The mouse was a particularly attractive device in these applications, because it is fast, quickly learned, and provides data approaching the density of eye-movement recording.

Design and procedure. Twelve subjects made 24 choices, four in each cell of a factorial design defined by two levels of choice set size (2 or 8) and three levels of correlation between attributes. We focused on three levels of interattribute correlation: (1) orthogonal ($r = 0$) choice sets, (2) sets with a correlation of -1 between two attributes, and (3) sets with the maximum possible average negative correlation between all pairs of attributes ($-.33$). In condition 2 the two attributes that were correlated inversely were the ones identified as being the most important across subjects in a previous study (Johnson and Meyer 1984) with the same stimulus set: rent and distance.

Though the primary interest in the study was the correlation manipulation, a set size manipulation was included to provide a baseline for assessing the relative strength of the correlation effect. Specifically, set size has long been known to induce observable changes in the way decisions are made, with noncompensatory strategies being more prevalent in larger choice sets (e.g., Johnson and Meyer 1984; Payne 1976). By including this manipulation, we could see whether interattribute correlations had an absolute effect on decision strategies, as well as the size of this effect in relation to a better-known context manipulation.

Subjects were undergraduate and graduate students who were paid for their participation in the experiment. Each was seated at an IBM PC equipped with a mouse and they were instructed that they were to participate in an investigation of apartment preferences. Subjects then were given the 24 choice scenarios. The scenarios were presented as six blocks of four with the order of the contexts (set sizes and correlations) being randomized among subjects.

The specific attribute levels describing each apartment were generated randomly, much as in the simulation. Attribute levels were assigned by using a subroutine that generated attribute scores under a prespecified covariance structure.

Results. Previous research has demonstrated that shifts from compensatory to noncompensatory strategies can be detected by monitoring the pattern and amount of information acquired by decision makers (Johnson and Meyer 1984; Lussier and Olshavsky 1979; Payne 1976; Payne, Bettman, and Johnson 1988). Noncompensatory strategies should be characterized by:

—an increase in the proportion of transitions within an attribute and a corresponding decrease in transitions within an alternative,

—an increased concentration of search on the alternative eventually chosen by the consumer, and

—changes in these effects as the decision progresses. Specifically, consumers seem to shift toward brand-based processing after eliminating several undesirable alternatives. The result is an increase in brand processing and in the concentration of search toward the end of the decision.

To examine possible changes in processing strategy due to intercorrelation and set size, we computed indices based on these search behaviors.

We examined changes in search patterns by calculating the normalized indices for same-brand and same-attribute transitions suggested by Bettman and Jacoby (1976). The means of these indices, reported in Table 2, show that attribute processing increased as the number of alternatives increased and that attribute processing was significantly more prevalent in the first half of the decision, particularly for larger set sizes. There are also corresponding decreases in the brand transition index (SBI). Though these results largely replicate the standard shift to noncompensatory strategies as set sizes increase, interattribute correlation appears to have no effect on these data.

A repeated-measure ANOVA confirmed these observations. It showed, for same-attribute transitions (SAI), significant effects of set size and phase. Neither measure, however, had a significant effect on interattribute correlation ($p > .10$).

A second indicator of processing strategy that we examined was the search concentration. We adopted a measure of the concentration of search used in previous decision-making studies (e.g., Payne 1976). We defined the concentration of search, P, as

$$P = \frac{CA \cdot NA}{TA} - 1 \qquad (2)$$

where CA is the number of acquisitions of the chosen alternative, NA is the number of alternatives, and TA is the total number of acquisitions across all alternatives. If search is distributed equally across alternatives, this index has a value of zero; positive values indicate that search is concentrated on the chosen alternative. Increases in P result from early elimination of undesirable alternatives, a consequence of strategies such as EBA or satisficing.

Large values of P were found, for the most part, only in the second half of decisions where eight alternatives were being considered. Table 2 gives the mean values of P by set size, interattribute correlation, and phase. An ANOVA showed a sizable effect of set size ($F(1,66) = 66.56$, $p < .001$) and smaller effects of phase, interattribute correlation, and their interaction ($F(1,66) = 15.38$,

Table 2

MEAN INDICES FOR SAME-ATTRIBUTE TRANSITIONS (SAI) AND SEARCH CONCENTRATION (*P*) BY SET SIZE, DECISION PHASE, AND CORRELATION, EXPERIMENT 1

		Two alternatives		*Eight alternatives*	
Interattribute r	*Decision phase*	*P*	*SAI*	*P*	*SAI*
0	First	.016	.720	.024	.752
	Second	.057	.695	.229	.504
−.33	First	.009	.752	.058	.784
	Second	.075	.701	.226	.532
−1	First	.025	.804	.074	.759
	Second	.048	.809	.198	.591

$p < .05$; $F(2,66) = 10.89$, $p < .001$; $F(1,66) = 27.24$, $P < .001$), respectively. The latter effects suggest a slight increase in the concentration of search, given negatively correlated attributes, as N increases and in the early phases of a decision. Though the effects are small, they are consistent with Slovic's (1975) suggestion that subjects may shift to more noncompensatory strategies when faced with difficult choices; they are inconsistent with the shift toward compensatory strategies suggested by a cost-benefit perspective (e.g., Johnson and Payne 1985).

Several other measures also were examined, including the amount of information acquired, the total time per decision, and the amount of time spent on each attribute. There were often sizable effects for set size, but none of these measures produced even marginal results for manipulations of intercorrelation. Hence, changes in the pattern of correlation between attributes in choice sets did not appear to produce marked shifts in decision-making strategies or evidence for a shift to compensatory rules with negative interattribute correlations. Though more research is needed before we can generalize this result, it suggests that major changes in strategies do not occur in response to changes in correlation. This result is in contrast to findings of other research showing that changes in strategies *do* occur in response to other manipulations in similar within-subject designs, such as set size and the dispersion of importance weights (Payne, Bettman, and Johnson 1988).

Study 3. A Group-Level Test of Predictive Validity

Overview. The results of the process analysis suggested that some of the necessary conditions for model failure were present: noncompensatory choice processes often were used, even with negative interattribute correlations. They did not, however, suggest how severe this failure would be in a natural setting. In particular, if a compensatory model were used to represent a heterogeneous group of decision-making strategies, how would its predictive validity be affected by changes in correlational structure? Examining the possible severity of this effect was the purpose of study 2.

Basic design. Again, we examined students' preferences for hypothetical apartment alternatives described in terms of rent, distance from campus, maintenance, and appearance. Our focus was on the predictive validity of a compensatory model of apartment choice in predicting choices made across a variety of orthogonal and nonorthogonal contexts. As in the process-tracing study, we included another context manipulation, set size, to provide a baseline for assessing the relative size of the correlation effect. The overall experimental design had nine different groups of choice sets, each containing eight different decisions. One of the cells asked subjects for preference ratings for each of eight alternatives generated by an orthogonal experimental design, much as in standard applications of conjoint analysis. The remaining cells were created by combining two levels of set size (2 and 8 alternatives) with four different levels of interattribute correlation.

The eight options in each set were generated by one of two processes. First, to mirror normal procedures used in conjoint analysis, we drew the eight apartments in the judgment condition from a main-effects (8-cell) fraction of a 2^4 factorial design. Specific attribute values were drawn randomly from ranges of prescaled high and low levels of each attribute. Second, the apartments in the eight groups of choice sets were created by a process similar to that used in the process-tracing study, random assignment of attribute values from a uniform distribution, to create the desired pattern of correlations.

The four patterns of intercorrelation were a mean *positive* (rank) correlation of .33, zero correlation, a mean *negative* correlation of .33, and a maximum single-attribute correlation of -1. Once again, the maximum correlation was between the attributes rent and travel time to campus. A unique set of apartment profiles was generated for each subject. The design yielded 72 decision problems, presented in a booklet, with order of presentation of tasks, blocks, and alternatives being randomized among subjects.

Subjects and procedure. Seventy-seven students enrolled in an MBA program participated and were run in groups of 10 to 15 per session. Each subject was paid eight dollars to participate in the experiment, which on average took 45 minutes to complete.

For choice sets of size two or eight, subjects chose their most preferred apartment. For the eight judgment problems (choice sets of size one), they rated each option's desirability on a 10-point scale with extremes of the best and worst apartments they could imagine (the same methodology used by Johnson and Meyer 1984).

Method of analysis. This experiment examined the ability of compensatory models estimated in orthogonal choice sets to predict choices made by actual decision makers faced with correlated attributes. To increase the comparability of models derived by using the judgment data with those derived by using the choice data, before analysis the ratings in the judgment condition were converted to discrete categories by the transformation $y = 1$ if rating $>$ mean rating, 0 otherwise.[8] Compensatory models for set sizes of two and eight were obtained by estimating the multinomial logit model,

$$Pr(i|j = 1, \ldots, N) = \frac{e^{\beta' X_i}}{\sum_{j=1}^{N} e^{\beta' X_j}}, \tag{3}$$

where $Pr(i)$ is the relative frequency with which apart-

[8]The analysis conducted with the raw ratings produced little difference in results. The analyses based on the discrete transformed ratings are presented to ensure maximum comparability across the different set size analyses.

ment i was selected from a set of $j = 1, \ldots, N$ options and $\beta' X_i$ is a linear combination of that apartment's value on each attribute. A compensatory model for the judgment condition was obtained by estimating the binary logit,

$$Pr(i) = \frac{1}{1 + e^{-\beta' X_i}} \tag{4}$$

where $P(i)$ is the relative frequency with which option i was given a rating higher than the mean apartment rating.[9] Exploring the performance of interactive compensatory models (as in the initial simulation) might also have been of interest, but was precluded by the sparsity of replications per cell for each subject (8). Hence, in our design we elected to forego the opportunity to look at the effect of adding interactions to a linear model in order to examine the performance of a simple model across a larger number of set size and correlation conditions.

Being concerned with both the aggregate and individual-level predictions made by the models, we estimated the parameter vector $\boldsymbol{\beta}'$ at both the individual and group levels. Maximum likelihood estimates were obtained by the SAS supplemental procedures MLOGIT and LOGIST for equations 2 and 3, respectively.

Aggregate predictive validity. The design provided three separate orthogonal sets, which we used to estimate separate logit models: the judgment condition and two different choice conditions with set sizes of two and eight. In Table 3 we report the derived choice hit rate, expressed as deviations from chance (percent correct − $(1/N)$), for all three orthogonal models. The models do fairly well in the orthogonal choice sets and in the set with an average positive correlation, but deteriorate badly in attempting to predict choices when the attributes have a negative correlation.

This failure is at least as dramatic as that indicated by the simulation. For example, when we use any one of the three compensatory models to predict choices from orthogonal pairs of apartments, on average we select the correct alternative 60% of the time—a reasonable success rate ($p < .05$). However, for an efficient set defined by two attributes with −1 correlation, the same model is worse than chance, selecting the correct alternative 46.9% of the time. Similar decreases occur for the predictions to choice sets of size eight. The data therefore suggest a marked decrease in the compensatory model's ability to predict choices when faced with negatively correlated attributes.

One other notable aspect of the results is the finding of superior predictive accuracies for compensatory models in choice sets of size eight versus size two (an ANOVA on proportions suggested the probability of the effect arising by chance is less than .01). This finding is somewhat counterintuitive as our process-tracing analyses had suggested that noncompensatory processing is more likely (hence the compensatory model is less appropriate) in larger set sizes. The most likely explanation was offered previously (Johnson and Meyer 1984) when a similar result was found in a study of set size effects on choice models. When faced with larger set sizes, subjects tended uniformly to adopt a simplified choice policy that picked the apartment with the lowest rent. Hence, the decrease in fit due to the increased use of a noncompensatory heuristic may have been offset by increased homogeneity in the policies being used to make choices.

Table 3

HIT RATE EXPRESSED AS A DEVIATION FROM CHANCE, AGGREGATE SPECIFICATION OF CHOICE MODELS

		Size of forecast set		
Interattribute r	Size of estimation set	1	2	8
.33	1		.135[a]	.334[a]
	2		.151[a]	.360[a]
	8		.154[a]	.218[a]
0	1	.239[a]	.135[a]	.334[a]
	2	.189[a]	.151[a]	.360[a]
	8	.189[a]	.154[a]	.218[a]
−.33	1		−.017	.061[a]
	2		.042[a]	.146[a]
	8		.034	.079[a]
−1	1		−.031	.071[a]
	2		.002	.108[a]
	8		.002	.036

[a]Probability that the reported hit rate reflects chance variation is less than .05.

Individual-level predictive validity. In many applications, market-share predictions are based not on a single aggregate model, but on a system of models estimated for each individual (e.g., Green, Carroll, and Goldberg 1981). The hypothesis is that individual-level models eliminate errors due to aggregation, reducing bias and improving forecasts. Evidence suggestive of this effect is provided by Wittink and Montgomery (1979).

We explored the possibility that model failures would be less severe in disaggregate-based forecasts by estimating 77 individual logit models using subjects' ratings in the judgment condition and predicting individual choices in each choice condition. Our analysis was based only on the judgment models because data sparsity in the choice conditions precluded efficient estimation of individual-level models. Additionally, our previous analysis (Table 3 and Johnson and Meyer 1984) suggested that little difference in predictive validity would be found between models estimated with different set sizes.

The results are summarized in Table 4. The table closely mirrors the aggregate analysis. We see a marked degradation in predictive validity as we move from positive to negative correlational environments and, again, choices from sets of size eight were predicted more accurately

[9]This form of the logit can be equated to that given in expression 3 simply by multiplying the numerator and denominator of that form by $e^{-\beta' X_i}/e^{-\beta' X_i}$.

Table 4
HIT RATE EXPRESSED AS A DEVIATION FROM CHANCE, DISAGGREGATE CHOICE MODELS

	Size of forecast set	
Interattribute r	*2*	*8*
.33	.088[a]	.306[a]
0	.046[a]	.188[a]
−.33	−.061[a]	.072[a]
−1	−.040	.071[a]

[a]Probability that the reported hit rate reflects chance variation is less than .05.

(in relation to chance) than those from sets of size two.

Perhaps the most intriguing feature of this table in comparison with the aggregate analysis is that the disaggregate predictions tend to be worse. This finding contrasts sharply to the common wisdom (Wittink and Montgomery 1979) that individual-level estimation would lead to increased predictive validity, but is consistent with Hensher's (1984) finding of a lack of a predictive advantage for disaggregate models.[10]

DISCUSSION

Applied work in choice modeling often involves an important assumption: the cognitive processes underlying choices may be complex, contingent, and noncompensatory, but they can be modeled well by simple compensatory models (e.g., Green and Srinivasan 1978). Though perhaps counterintuitive, the assumption is supported by a relatively large literature arguing that such models are robust, even when the underlying process is misspecified (e.g., Dawes and Corrigan 1974; Johnson and Meyer 1984).

Our studies contribute to a recent stream of work that has questioned the generality of this result. We were motivated by the work of such researchers as Curry and Faulds (1986) and Newman (1977), who found that specification errors in linear models that seem small when studied in orthogonal environments can be amplified when applied in negatively correlated environments. We addressed three sets of research issues that had not been examined in previous work:

1. Can linear models approximate the choices made by different noncompensatory heuristics and is this approximation affected by different patterns of interattribute correlation?
2. Do consumers adjust their choice strategies when placed in negatively correlated environments?
3. Do compensatory models calibrated on natural respondent pools give poorer predictions when applied in negatively correlated environments?

In answer to the first question, a computer simulation replicated the now well-known finding that the performance of even correctly specified linear models is diminished when they are applied in negatively correlated environments. New are the findings that this sensitivity is amplified when the linear model is being used to approximate a noncompensatory choice process and that some noncompensatory rules—most notably the conjunctive—are not well approximated even in orthogonal contexts. Though these results seem discouraging for analysts seeking to apply linear models in judgment analysis, the simulation also offered a result suggesting cause for optimism: interactions that contributed little to model performance in an orthogonal environment yielded substantial improvements in fit when applied in negatively correlated environments.

The simulation results therefore seem to underscore a need to develop decision models that more accurately capture the process underlying consumers' decisions. The performance of a model in an orthogonal environment may provide only a limited guide to its performance in a negatively correlated one; the less redundancy among attributes in a given prediction context, the greater the need for a more precise representation of the consumer's judgment process. However, a caveat must be added to this advice: identifying and modeling functional form should help to *reduce* prediction errors in negatively correlated environments, but may not fully *eliminate* them. A noteworthy finding of the simulation is that negative correlations served to amplify modeling error from *all* sources, not just that due to misspecification of functional form. As there will always be unsystematic error in measurement of preferences in practice, the high levels of prediction commonly observed in orthogonal environments may be inherently more difficult to obtain in negatively correlated ones. Better representations of decision processes would be useful in eliminating systematic errors that arise in these environments. They cannot, however, be expected to restore fully the types of accuracy levels found in orthogonal contexts.

The second interest of the research was the way in which *consumers,* rather than *models,* behave in negatively correlated environments. Previous research had been ambiguous on this issue: it gave a normative rationale for switching to compensatory rules (e.g., Johnson and Payne 1985), but limited evidence that noncompensatory rules might be even more pronounced (e.g., Slovic 1975). Our research yielded a perhaps surprising finding: there was no evidence that any shift in processing toward compensatory rules occurred in response to correlation changes.

Before this finding is generalized to other contexts, however, we must add a cautionary note: one possible

[10]Recall that the traditional argument favoring disaggregate models is that they avoid aggregation error due to individual variation in weights. The drawback of disaggregate analysis, however, is that individual-level model estimates tend to be less efficient (have higher variances) than those obtained in aggregate models because of small sample sizes. In comparatively homogeneous samples, therefore, aggregate models might well outperform disaggregate models simply by virtue of their increased statistical efficiency. This may well account for our findings.

explanation for the finding is that it was encouraged by the within-subject nature of the experimental design. In particular, the changes in correlation may not have been noticed by subjects during the task or, even if they were, the psychological costs of switching rules during the task may have outweighed the perceived benefits. Hence, we might have observed a strategy change if subjects had been required to make a much larger number of judgments within each correlational condition because a strategy shift might then have been cognitively worthwhile.

Supporting the generality of the result, however, is the fact that we *did* observe strategy changes in response to set size, which also was manipulated on a within-subject basis. Hence, a prudent interpretation of the findings is that strategy changes may occur in response to changes in intercorrelational structure, but the magnitude of those changes is likely to be small in relation to that usually associated with changes in numbers of alternatives. Exploring the robustness of this finding will be an important area for future research.

Our final analysis combined the interests of the first two. We examined the extent to which a compensatory model calibrated on actual choices would fail when applied in negatively correlated environments. The rationale for the study was that evidence suggesting compensatory models may fare poorly in compensatory environments was based largely on numerical simulations (in addition to our work, Curry and Faulds 1986). Left unclear was the level of model failure likely to be encountered in studies with natural subject populations. Supporting the simulation, the degrading effect of a change in correlational structure was substantial, with predictive accuracy approaching that expected by chance alone.

Combined, our three studies seem to tell a fairly coherent story about the behavior of compensatory models in negatively correlated environments. When a model is misspecified, whether because of an incorrect inference about process or a simple inability to measure preferences perfectly, its prediction will be sensitive to the correlational structure of the environment under study. The more negative the interattribute correlation among attributes, the more "apparently small" specification errors will be amplified.

Though this conclusion seems pessimistic for modelers of applied choice, the research also offers two notes of optimism. First, *models* may be highly sensitive to correlational structure, but *consumers* appear to be less so. Hence, though an analyst should expect prediction rates to fall when a model is transferred to a negatively correlated context, this degradation should not be exacerbated by additional changes in the consumers' decision structures. Second, our simulation suggests that even this natural decrement in predictive ability might be mitigated though the use of multilinear models that more accurately mimic the process underlying choices.

Finally, our research had an additional motivation apart from central focus on correlational structures: to illustrate the potential benefits of combining behavioral and quantitative approaches to the study of consumer decision making. By combining numerical simulations, process tracing, and revealed preference analyses, we were afforded insights into how decisions and models of decisions are affected by decision context—which would be impossible by any one method alone. Hence, we hope that in addition to motivating further work on the robustness of linear models, our research will serve to encourage broader cooperation between behavioral and quantitative researchers in marketing.

REFERENCES

Bettman, James R. and Jacob Jacoby (1976), "Patterns of Processing in Consumer Information Acquisition," in *Advances in Consumer Research,* Vol. 3, Beverlee B. Anderson, ed. Ann Arbor, MI: Association for Consumer Research, 315–20.

Cattin, Philippe and Dick R. Wittink (1982), "Commercial Use of Conjoint Analysis: A Survey," *Journal of Marketing,* 46 (Summer), 44–53.

Chakravarti, Dipankar and John G. Lynch (1983), "A Framework for Exploring Context Effects in Consumer Judgment and Choice," in *Advances in Consumer Research,* Vol. 10, Richard P. Bagozzi and Alice M. Tybout, eds. Chicago: Association for Consumer Research, 289–97.

Corstjens, Marcel L. and David A. Gautschi (1983), "Formal Choice Models in Marketing," *Marketing Science,* 2 (Winter), 19–56.

Curry, David J. and David J. Faulds (1986), "Indexing Product Quality: Issues, Theory and Results," *Journal of Consumer Research,* 13 (June), 134–45.

———, Jordan J. Louviere, and Michael Augustine (1981), "Comment on the Sensitivity of Brand Choice Simulators to Attribute Importance Weights," *Decision Sciences,* 12 (July), 502–16.

——— and Michael Menasco (1979), "Some Effects of Differing Information Processing Strategies on Husband-Wife Joint Decisions," *Journal of Consumer Research,* 6 (September), 192–203.

Dawes, Robin M. and Bernard Corrigan (1974), "Linear Models in Decision Making," *Psychological Bulletin,* 81 (March), 95–106.

Einhorn, Hillel J. (1970), "The Use of Nonlinear, Noncompensatory Models in Decision Making," *Psychological Bulletin,* 73 (3), 221–30.

———, Donald N. Kleinmuntz, and Benjamin Kleinmuntz (1979), "Linear Regression and Process-Tracing Models of Judgment," *Psychological Review,* 86 (September), 464–85.

Engel, James F. and Roger Blackwell (1982), *Consumer Behavior*. Hillsdale, NJ: The Dryden Press.

Goldberg, Louis R. (1971), "Five Models of Clinical Judgment: An Empirical Comparison Between Linear and Nonlinear Representations of the Human Inference Process," *Organizational Behavior and Human Performance,* 6, 458–79.

Green, Paul E., J. Douglas Carroll, and Steven M. Goldberg (1981), "A General Approach to Product Design Optimization Via Conjoint Analysis," *Journal of Marketing,* 45 (Summer), 17–37.

——— and Michael T. Devita (1975), "An Interaction Model of Consumer Utility," *Journal of Consumer Research,* 2 (March), 146–53.

———, Kristiaan Helsen, and Bruce Shandler (1988), "Conjoint Internal Validity Under Alternative Profile Representations," *Journal of Consumer Research,* 15 (December), 392–7.

——— and Abba M. Krieger (1986), "The Minimal Rank Correlation, Subject to Order Restrictions, with Application to the Weighted Linear Choice Model," *Journal of Classification,* 3 (1), 67–95.

——— and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research,* 5 (September), 103–123.

Guadagni, Peter M. and John D. C. Little (1983), "A Logit Model of Brand Choice Based on Scanner Data," *Marketing Science,* 2 (Summer), 203–38.

Hensher, David A. (1984), "Achieving Representativeness of the Observable Component of the Indirect Utility Function in Logit Choice Models: An Empirical Revelation," *Journal of Business,* 57 (April), 265–80.

Hjorth-Anderson, Chr. (1984), "The Concept of Quality and the Efficiency of Markets for Consumer Products," *Journal of Consumer Research,* 12 (September), 708–18.

Johnson, Eric J. and Robert J. Meyer (1984), "Compensatory Models of Non-Compensatory Choice Processes: The Effect of Varying Context," *Journal of Consumer Research,* 11 (June), 528–41.

——— and John W. Payne (1985), "Effort and Accuracy in Choice," *Management Science,* 31 (April), 395–414.

———, ———, David A. Schkade, and James R. Bettman (1988), "Monitoring Information Processing and Decision: The Mouselab System," working paper, Graduate School of Industrial Administration, Carnegie-Mellon University.

Louviere, Jordan J. (1988), *Analyzing Decision Making: Metric Conjoint Analysis.* Newbury Park, CA: Sage Publications, Inc.

Lussier, Dennis A. and Richard W. Olshavsky (1979), "Task Complexity and Contingent Processing in Brand Choice," *Journal of Consumer Research,* 6 (September), 154–65.

Lynch, John G. (1985), "Uniqueness Issues in Decompositional Modeling of Multiattribute Overall Evaluations: An Information Integration Perspective," *Journal of Marketing Research,* 12 (February), 1–19.

Maddala, G. S. (1983), *Limited Dependent and Qualitative Variables in Econometric Applications.* Cambridge, MA: The MIT Press.

Newman, J. Robert (1977), "Differential Weighting in Multiattribute Utility Measurement: Where It Should and Where It Does Make a Difference," *Organizational Behavior and Human Performance,* 20 (December), 312–25.

Olshavsky, Richard W. and Franklin Acito (1980), "An Information Processing Probe into Conjoint Analysis," *Decision Sciences,* 11 (July), 451–70.

Payne, John W. (1976), "Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis," *Organizational Behavior and Human Performance,* 16 (August), 366–87.

———, James R. Bettman, and Eric J. Johnson (1988), "Adaptive Strategy Selection in Decision Making," *Journal of Experimental Psychology: Human Memory, Learning and Cognition,* 14 (July), 534–52.

——— and E. K. Easton Ragsdale (1978), "Verbal Protocols and Direct Observation of Supermarket Shopping Behavior: Some Findings and a Discussion of Methods," in *Advances in Consumer Research,* Vol. 5, H. Keith Hunt, ed. Chicago: Association for Consumer Research, 571–7.

Russo, Jay E. and Barbara A. Dosher (1983), "Strategies for Multiattribute Binary Choice," *Journal of Experimental Psychology: Learning, Memory and Cognition,* 9 (October), 676–96.

Silk, Alvin J. and Glen L. Urban (1978), "Pre-Test-Market Evaluation of New Packaged Goods: A Model and Measurement Methodology," *Journal of Marketing Research,* 15 (May), 171–91.

Slovic, Paul (1975), "Choice Between Equally Valued Alternatives," *Journal of Experimental Psychology: Human Perception and Peformance,* 1 (July), 280–7.

Tellis, Gerard J. and Birger Wernerfelt (1987), "Competitive Price and Quality Under Asymmetric Information," *Marketing Science,* 6 (Fall), 240–53.

Theil, Henri (1971), *Principles of Econometrics.* New York: John Wiley & Sons, Inc.

Thorngate, W. (1980), "Efficient Decision Heuristics," *Behavioral Science,* 25 (May), 219–25.

Tversky, Amos (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review,* 79 (July), 281–99.

——— and Shumel Sattah (1979), "Preference Trees," *Psychological Review,* 86 (November), 542–73.

———, ———, and Paul Slovic (1987), "Contingent Weighting in Judgment and Choice," *Psychological Review,* 95 (July), 371–84.

Wainer, H. (1976), "Estimating Coefficients in Linear Models: It Don't Make No Nevermind," *Psychological Bulletin,* 83 (March), 213–17.

Wittink, Dick R. and David B. Montgomery (1979), "Predictive Validity of Trade-Off Analysis for Alternative Segmentation Schemes," in *Educators' Conference Proceedings,* Neil Beckwith et al., eds. Chicago: American Marketing Association, 68–73.

Wright, Peter R. (1973), "Use of Consumer Judgment Models in Promotion Planning," *Journal of Marketing,* 37 (October), 27–33.

Reprint No. JMR263100