Modeling hierarchical conjoint processes with integrated choice experiments

Citation for published version (APA):
Oppewal, H., Louviere, J. J., & Timmermans, H. J. P. (1994). Modeling hierarchical conjoint processes with integrated choice experiments. *Journal of Marketing Research*, *31*(1), 92-105. https://doi.org/10.2307/3151949

DOI:
10.2307/3151949

Document status and date:
Published: 01/01/1994

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

HARMEN OPPEWAL, JORDAN J. LOUVIERE, and HARRY J.P. TIMMERMANS

The authors propose and illustrate an extension of the method of Hierarchical Information Integration (HII). HII allows one to handle large numbers of attributes in conjoint tasks by designing subexperiments that include subsets of attributes. It assumes that individuals can use general attributes or decision constructs to summarize their impressions of these subsets, which could be clusters of detailed, managerially relevant attributes. The proposed extension involves the design of subexperiments that include attributes plus summary evaluations of remaining constructs. Advantages are that subexperiments can be analyzed separately but also jointly to estimate one overall preference or choice model; a more flexible and easy task is obtained; and one can test the assumed hierarchical decision structure. The authors illustrate the approach with an application that models consumer choice of shopping center. In this application, results partially support the hierarchical structure and predictive validity. Finally, the authors discuss implications for further research.

Modeling Hierarchical Conjoint Processes With Integrated Choice Experiments

Task sizes in conventional conjoint-like experiments rapidly increase with increasing numbers of attributes and/or levels. Though design techniques such as minimal fractional factorial designs, blocking techniques or hybrid approaches have been suggested to limit the numbers of profiles or choice sets in conjoint tasks, traditional (full-profile) approaches proliferate profile sizes as the number of attributes increases (e.g., Green and Srinivasan 1990; Louviere 1988).

Several approaches have been proposed to handle the problem of large numbers of attributes (Green and Srinivasan 1990), but each has drawbacks: Self-explication of attribute importances requires one to assume that individuals can provide valid and accurate evaluations of attribute weights independent of a specific context. In pairwise trade-off analysis (Johnson 1974), respondents could forget where they are in the trade-off table or adopt patterned responses (Green and Srinivasan 1990) and be forced to make assumptions or inferences about omitted attribute levels (Johnson 1987). This latter problem also could apply to BRIDGER (Albaum 1989; Bretton-Clark 1988), which uses subsets of attributes to construct separate experiments. Separate designs are "bridged" by at least one common attribute, and data from subexperiments are rescaled to a common scale by using OLS regression to calibrate the scale of the bridging attribute(s). This method is rather ad hoc, and its statistical properties are not well understood. These criticisms also apply to Adaptive Conjoint Analysis (Carmone 1987), which asks respondents whether any attribute levels would lead to rejection of an alternative (Green, Krieger, and Bansal 1988) and selects attributes for further presentation through a self-explication task (Green and Srinavasan 1990).

An alternative approach to handling large conjoint problems is Hierarchical Information Integration (HII), proposed by Louviere (1984) and illustrated by Louviere and Gaeth (1987). Louviere's (1984) HII approach (1) categorizes attributes into several non-overlapping sets

Harmen Oppewal is Assistant Professor of Urban Planning, Eindhoven University of Technology, Eindhoven, The Netherlands. Jordan J. Louviere is Professor of Marketing and Park Fellow, University of Utah, Salt Lake City, Utah. Harry J.P. Timmermans is Professor of Urban Planning, Eindhoven University of Technology, Eindhoven, The Netherlands, and Carthy Foundation Professor of Marketing at the University of Alberta, Edmonton, Canada. This research was supported by a grant from the Social and Environmental Research Foundation, which is a part of the Netherlands Organization for Scientific Research (N.W.O.-S.R.O.). The authors thank the editor of JMR and three anonymous reviewers for their helpful comments on earlier versions of this article.

Journal of Marketing Research
Vol. XXXI (February 1994), 92–105

based on theory, logic, empirical evidence or applications demands, such that the sets represent particular constructs such as "quality," "atmosphere," "value for money," or the like; (2) designs and administers separate subexperiments to define each construct in terms of the attributes that categorize it; and (3) develops an "overall" or bridging design based on the constructs, which permits one to concatenate the results of the separate designs and the overall design into one fully specified utility model.

Louviere's (1984) HII approach avoids the need to use the self-explicated weights and scale values and should be less affected by missing information than pairwise trade-offs. Moreover, bridging of designs is based on behavioral and statistical theory rather than an ad hoc calibration. Despite these advantages, Louviere's HII approach also has problems and limitations, which we review in the next section.

Our purpose is to propose, describe, and illustrate an extension of Louviere's (1984) HII approach that overcomes these problems and limitations and also avoids many of the problems associated with the preceding approaches. We first review HII theory and several problems and limitations of the original approach. Next we present the basic idea of our proposed extension and describe procedures for designing integrated experiments consistent with the new theory. Then we discuss the analysis of these experiments and describe ways to test for the hierarchical structure that underlies the estimation of a single choice model from these separate experiments. We then illustrate the approach by applying it to model consumer choice of shopping center. Finally, we draw some general conclusions and suggest some useful topics for further research.

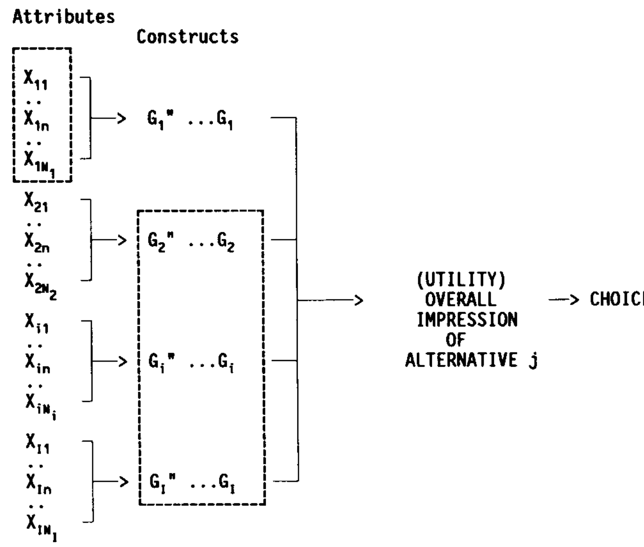
HIERARCHICAL INFORMATION INTEGRATION

Conceptual Background

HII is an extension of Information Integration Theory (Anderson 1981, 1982) to complex decision problems in which one can assume that individuals are likely to group or categorize decision attributes into separate (high-order) decision constructs. HII assumes that individuals categorize attributes with respect to particular constructs and integrate information about attributes to form impressions of alternatives with respect to those constructs and then integrate their separate construct impressions to evaluate alternatives holistically. (Arrows in Figure 1 illustrate this sequence.) Therefore, HII explicitly assumes that individuals simplify the evaluation of alternatives into separate partial and holistic evaluations. Recent research supports the idea that people can use high-order decision constructs when evaluating multiattribute alternatives and choosing among brands or products (Corfman 1991). The idea behind the HII Method is to structure decision tasks to study and analyze each integration process separately and jointly.

The basic idea of HII was illustrated by Louviere and Gaeth (1987), who applied it to model supermarket shopping behavior. They first identified four high-order supermarket decision constructs using previous research (prices, selection, quality, and shopping convenience). These are indicated in Figure 1 as G_1 to G_I. Nonoverlapping sets of attributes described each construct G_i, and different experimental designs were used to create combinations of attribute levels to define each construct. For example, attributes "prices of meats," "prices of produce," "prices of dairy products," and "prices of packaged goods" were varied in a subexperiment to create profiles of price-related attributes. Subjects evaluated the profiles on an 11-category rating scale (0–10) with respect to the construct "prices." In the other three subexperiments, subjects evaluated attribute profiles that described the other constructs. In Figure 1, construct evaluations in subexperiment i are denoted as G_i''. In a final experiment, each of the four decision constructs was treated as a factor, the levels of which were category values 2, 5, and 8 on the scale that subjects had used to rate the profiles in each separate construct evaluation task. The profiles were described by combinations of hypothetical ratings of decision constructs G_i. Subjects responded to this final task by rating their preference for the construct rating profiles on a 150-mm line-mark response scale. The five experiments were concatenated to

Figure 1
OVERVIEW OF THE PROCESS OF HIERARCHICAL INFORMATION INTEGRATION



Areas within dashed lines indicate attributes and constructs that are included in subexperiment #1 in the new approach; total number of subexperiments is I. The conventional approach to modeling HII would require $I + 1$ experiments, of which subexperiment #1 would include only attributes in the upper shaded area.

obtain one overall utility function by substitution and replacement of terms.

HII also has been applied to other substantive problems, such as telecommunications (Louviere 1984), outdoor recreation (Louviere and Timmermans 1990a), joint decision making (Timmermans et al. 1992), and residential choice (Louviere and Timmermans 1990b). In the latter two studies, the ratings-based bridging task was replaced by a choice experiment, which allowed estimation of multinomial logit (MNL) choice models. Despite encouraging empirical results, these studies also suggested a variety of problems and limitations with the original method.

Problems and Limitations of Previous Approaches to HII Modeling

Among the problems identified previously, the following are germane here:

1. Louviere's (1984) original HII approach does not test the assumed hierarchical decision structure; hence, one must assume the hierarchical structure is correct to concatenate the separate experiments logically.
2. The original HII approach produces several models rather than a single one for which overall measures of goodness- or badness-of-fit and tests of validity can be derived. A concatenated overall model cannot be estimated directly; rather, model parameters are calculated by substitution and replacement of terms in different models that are estimated separately. Hence, the attributes varied in each experiment are not related directly to the final response of interest, whether preference or choice.
3. The values of remaining constructs are not specified in each separate subexperiment because one assumes that they have no systematic effects on evaluations of a particular construct. Consequently, subjects may have to assume or infer values for other constructs (e.g., Johnson 1987). Therefore, attribute effects are tested only over a limited range of values of other decision constructs, and there is no control over respondents' inferences about the values of other decision constructs.
4. The validity of the bridging experiment poses problems because respondents have to evaluate or choose among profiles described in terms of their (hypothetical) profile ratings in the subexperiments. The difficulties of this task are not clear, nor is it clear whether resulting attribute evaluations reflect respondents' real decisions. For example, profiles in bridging tasks typically contain only numerical scores, which could encourage respondents to average construct scores.
5. Though the bridging experiment can be designed as a choice experiment, Louviere's original approach does not allow subexperiments to be framed as choice experiments. Because there could be situations in which one wants to model choice behavior, one might wish to design subexperiments as choice experiments.
6. Finally, interactions between variables that define different constructs cannot be estimated, nor can interactions between attributes and decision constructs.

In the next section, we propose and describe an approach to the design and analysis of HII tasks that avoids most of these limitations and incorporates a more realistic and rich task structure.

INTEGRATED HII CHOICE EXPERIMENTS

Design Considerations

Let us first consider the logic behind the design of integrated HII choice experiments: Suppose a particular choice is influenced by a set X consisting of N total attributes. HII requires one to categorize the N attributes into subsets that define or map into various decision constructs. Categorization of attributes into constructs can be based on logic, theory, or empirical evidence from literature or pre-experimental research. The number of constructs defined should be considered in the light of (1) how respondents categorize product attributes and (2) design requirements imposed by various categorization schemes. Suppose there are I constructs, denoted by G_i ($i = 1,..I$), and assume that each construct is associated with a subset X_i that contains N_i attributes X_{in} ($n = 1,..N_i$). HII applications typically assume that each attribute maps into only one construct; hence, $\Sigma_i N_i = N$. In the conventional HII approach one would design $I + 1$ separate experiments: one subexperiment for each attribute set X_i that defines a decision construct G_i and one bridging experiment. In subexperiment i, subjects would evaluate profiles of attributes from set X_i on a scale that defines the particular decision construct G_i and ignore all remaining decision constructs G_j ($j \neq i$) and corresponding attribute sets X_j ($j \neq i$). In the bridging design, they would evaluate profiles of hypothetical evaluations of all I constructs.

We extend the conventional method by explicitly including summary measures of the other decision constructs G_j ($j \neq i$) as additional design variables in each subexperiment i, as indicated in Figure 1. Hence, experimental profiles describe alternatives as combinations of attribute levels and construct levels. Because each profile potentially describes all I aspects of an alternative that one assumes to be relevant to the respondents, their overall evaluations (whether preference or choice) should provide information about their utilities and preferences regardless of which construct experiment they experience. This extension increases the size of each subexperiment but eliminates the need for a bridging experiment because theoretically equivalent preference or choice models can be estimated from responses to each subexperiment. Moreover, the separate tasks can be concatenated to estimate all attribute parameters simultaneously. If the task is designed as a choice experiment, one can preserve the orthogonality of the design and fix the origin of the utility scale across the different construct experiments by using a common, constant base alternative in all choice sets as recommended by Louviere and Woodworth (1983) and Louviere (1988). One obvious constant alternative might be "other" or "none."

Analysis of Integrated HII Preference and Choice Experiments

Each subexperiment can be analysed independently with standard estimation techniques, such as OLS regression for preference ratings, or MNL regression analysis for discrete choice responses. The data for each subexperiment enable one to estimate a model that contains terms for the utilities of included attributes and scaled decision constructs.

More formally, if one assumes additive linear utility functions (e.g., within a MNL type of discrete choice model), the utility function for alternatives r in subexperiment i can be described as follows:

$$\mathbf{V}_i = \mathbf{X}_{(i)}\,\boldsymbol{\beta}_i + \mathbf{G}_{(i)}\,\boldsymbol{\gamma}_i,$$

where $\mathbf{V}_i$ is the vector of systematic utilities of alternatives $r(r = 1,..R_i)$ in subexperiment i and $\mathbf{X}_{(i)}$ is a matrix whose R_i rows consist of the N_i appropriately coded attributes X_{in} of alternatives r in subexperiment i. $\mathbf{G}_{(i)}$ is a matrix whose R_i rows consist of the $(I - 1)$ coded constructs $G_j(j \neq i)$ that describe the remaining dimensions of alternatives r in subexperiment i; and $\boldsymbol{\beta}_i$ and $\boldsymbol{\gamma}_i$ are unknown parameter vectors for subexperiment i.

Each separate subexperiment supplies estimates of particular attributes and constructs. However, one can estimate a single overall preference or choice model by concatenating all subexperiments and estimating a common vector of parameters across designs. The utility function for such an overall model assumes that the same decision process operates in the separate experiments. Moreover, it assumes that any biases induced by separate experiments cancel out across all experiments and that error variances are equivalent across tasks. The full design matrix $\mathbf{M}$ is a block diagonal matrix that is constructed from the $\mathbf{X}_{(i)}$ and $\mathbf{G}_{(i)}$ matrices, as follows:

$$\mathbf{M} = \begin{bmatrix} \mathbf{X}_{(1)} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{G}_{(1)} & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{(2)} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{G}_{(2)} & \ldots & \mathbf{0} \\ \ldots & & & & & & & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{X}_{(I)} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{G}_{(I)} \end{bmatrix}$$

and the full model can be stated as

$$\mathbf{V} = \mathbf{M}[\boldsymbol{\beta}, \boldsymbol{\gamma}]',$$

where $\mathbf{V}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ are concatenated vectors: $\mathbf{V} = (\mathbf{V}_1, .. \mathbf{V}_i, .. \mathbf{V}_I)'$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, .., \boldsymbol{\beta}_i, .., \boldsymbol{\beta}_I)$, and $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, .., \boldsymbol{\gamma}_i, .., \boldsymbol{\gamma}_I)$.

However, one also can constrain the matrices $\mathbf{G}_{(i)}$ to be "generic" across subexperiments. This requires that zero-valued column vectors of length R_i are inserted in the matrices $\mathbf{G}_{(i)}$ to represent the missing construct i; call these extended matrices $\mathbf{G}_{(i)}^{+0}$. One then can construct a design matrix $\mathbf{M}_g$ as follows:

$$\mathbf{M}_g = \begin{bmatrix} \mathbf{X}_{(1)} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{G}_{(1)}^{+0} \\ \mathbf{0} & \mathbf{X}_{(2)} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{G}_{(2)}^{+0} \\ \ldots & & & & & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{X}_{(I)} & \mathbf{G}_{(I)}^{+0} \end{bmatrix}$$

and the full model is stated as

$$\mathbf{V} = \mathbf{M}_g[\boldsymbol{\beta}, \boldsymbol{\gamma}_g]',$$

where $\mathbf{V}$ and $\boldsymbol{\beta}$ are vectors as defined previously and $\boldsymbol{\gamma}_g$ is a single vector of parameters for the I constructs. Hence, in each subexperiment, attributes or constructs that do not appear in that experiment are coded zero in the design matrix. If all attribute and construct effects are centered around a zero mean and attributes in each subexperiment are orthogonal, the structural zeroes will not affect the parameter estimates. The estimated utility function, therefore, will contain separate utility parameters for each attribute and construct. Attribute parameter vectors $\boldsymbol{\beta}_i$ and "specific" construct parameter vectors $\boldsymbol{\gamma}_i$ are estimated only from subexperiment i. If construct parameters are constrained to be generic ($\boldsymbol{\gamma}_g$), then construct estimates are derived from I minus 1 subexperiments. Only the constant, if included, is based on all I subexperiments.

The overall model estimated from the concatenated experiments can be used to predict the utility of new or existing alternatives. This requires that values of either a particular construct or its corresponding attributes are set to zero in the design matrix. In this way, utilities for new alternatives can be predicted for cases in which attribute values are known but decision construct evaluations are not or, conceivably, construct evaluations are known but attribute values are not.

Assessing the Validity of the HII Model Structure

An important advantage of our extended HII approach over Louviere's (1984) approach is that we can test the validity of the assumed hierarchical model structure. We outline four tests and illustrate them in our example application. The first three involve the use of OLS regression and require that respondents evaluate (judge) the constructs described by the attribute profiles. To ensure measurement compatibility, respondents must evaluate constructs on the same rating scales used by the experimenter to define the levels of these constructs in the other subexperiments. The first test of the assumed hierarchical model structure is to check whether all attributes are statistically significant in a regression of these construct evaluations $\mathbf{G}_i''$ on attributes from set X_i and other constructs $\mathbf{G}_j(j \neq i)$. If a construct is well defined by its attributes, all its attributes should have significant effects. Moreover, if the ratings of a particular construct are independent of the levels of other constructs, none of the effects of the other construct measures (main effects or interactions with attributes) should be significant. More formally, this involves the estimation of the following regression equation for each subexperiment i:

$$\mathbf{G}_i'' = \mathbf{X}_{(i)}\,\boldsymbol{\beta}_i'' + \mathbf{G}_{(i)}\,\boldsymbol{\gamma}_i'',$$

where matrices $\mathbf{X}_{(i)}$ and $\mathbf{G}_{(i)}$ are as defined previously, and $\mathbf{G}_i''$, $\boldsymbol{\beta}_i''$ and $\boldsymbol{\gamma}_i''$ have similar dimensions as the previously defined vectors $\mathbf{V}_i$, $\boldsymbol{\beta}_i$ and $\boldsymbol{\gamma}_i$, respectively. Hypotheses are that (1) $\boldsymbol{\beta}_i'' <> \mathbf{0}$ and (2) $\boldsymbol{\gamma}_i'' = \mathbf{0}$.

Second, one can compare the OLS construct regression parameters of attributes with the overall choice (or preference) model results. All attributes that have significant effects on the utilities $\mathbf{V}_i$ in the overall preference or choice models also should have significant effects on the construct evaluations $\mathbf{G}_i''$. If an attribute is significant in the overall model, but not significant in construct-regression models, it is not well represented in the specific construct. Hence, the regressions indicate whether hierarchical processes within subexperiments are independent and include all relevant attributes.

A third application of OLS construct regression results involves a test of process equality across subexperiments. The separate construct regression equations based on observed values $\mathbf{G}_i''$ can be used to predict construct values $\mathbf{G}_i$ associated with attribute profiles (or sets) $\mathbf{X}_i$ in subexperiment i. One can substitute vectors of these predicted values for the zero-valued column vectors in matrices $\mathbf{G}_{(i)}^{+0}$ to obtain matrices that we denote by $\mathbf{G}_{(i)}^{+p}$, and estimate a generic "constructs-only" model across subexperiments. A modified Chow test (e.g., Kmenta 1986, p. 421) can be used to test whether the effect of a construct defined as a profile of attributes in a particular subexperiment is equal to the effect of the corresponding hypothetically defined construct evaluation in the other subexperiments. Dummies are used to create sets of parameters that represent potential construct parameter deviations across particular subexperiments as follows:

$$\mathbf{M}_c = \begin{bmatrix} \mathbf{G}_{(1)}^{+p} & +\mathbf{G}_{(1)}^{+p} & \mathbf{0} & .. & \mathbf{0} \\ \mathbf{G}_{(2)}^{+p} & \mathbf{0} & \mathbf{G}_{(2)}^{+p} & .. & \mathbf{0} \\ \cdots & & & & \\ \mathbf{G}_{i(I-i)}^{+p} & \mathbf{0} & \mathbf{0} & .. & +\mathbf{G}_{(I-1)}^{+p} \\ \mathbf{G}_{(I)}^{+p} & -\mathbf{G}_{(I)}^{+p} & -\mathbf{G}_{(I)}^{+p} & .. & \mathbf{G}_{(I)}^{+p} \end{bmatrix}$$

and the full model is stated as

$$\mathbf{V} = \mathbf{M}_c[\boldsymbol{\gamma}_g, \boldsymbol{\gamma}_g^1, \boldsymbol{\gamma}_g^2, .., \boldsymbol{\gamma}_g^{I-1}]',$$

with the $\boldsymbol{\gamma}_g^i$ being vectors of Chow test parameters, and other terms as defined previously. Some of these Chow test parameters represent tests of whether effects of profiles of attributes are equal to effects of corresponding hypothetical construct evaluations. None of these should be significant. The remaining Chow test parameters represent another test of the equivalence of construct effects across subexperiments.

This fourth test can be performed even if no separate construct evaluations (ratings) were obtained. It involves testing whether parameter estimates for the same hypothetically defined constructs differ between different subexperiments. In theory, if each subexperiment represents the same choice process, the parameters of the decision constructs should be the same except for sampling error and differences in error variability across the different experiments. Therefore, one expects no significant construct effect differences between subexperiments that both included the construct as a hypothetical rating. Significant Chow test parameters represent context effects that induce construct parameter differences between subexperiments. In the case of choice tasks, these tests assume that the "Independence of Irrelevant Alternatives" assumption holds and utility functions estimated from different subexperiments have equivalent scale units or can be rescaled to common units (cf. Swait and Louviere 1993).

If the data fail these two Chow tests, the models are context dependent, that is, choice processes are not equal across subexperiments. Though this requires caution when interpreting the results, indices of relative importances of attributes and constructs within separate subexperiments are as valid as any other (non-HII) single choice experiment. Moreover, the concatenated model still could be useful for predictive purposes because one can set construct effects to zero and use only attribute effects for prediction. There is insufficient experience to know whether context-dependent models transfer well to other (e.g., real-world) choice situations; however, context dependence could average out across different subexperiments, posing little problem for prediction. Which outcome is more likely in a real decision context is an empirical question.

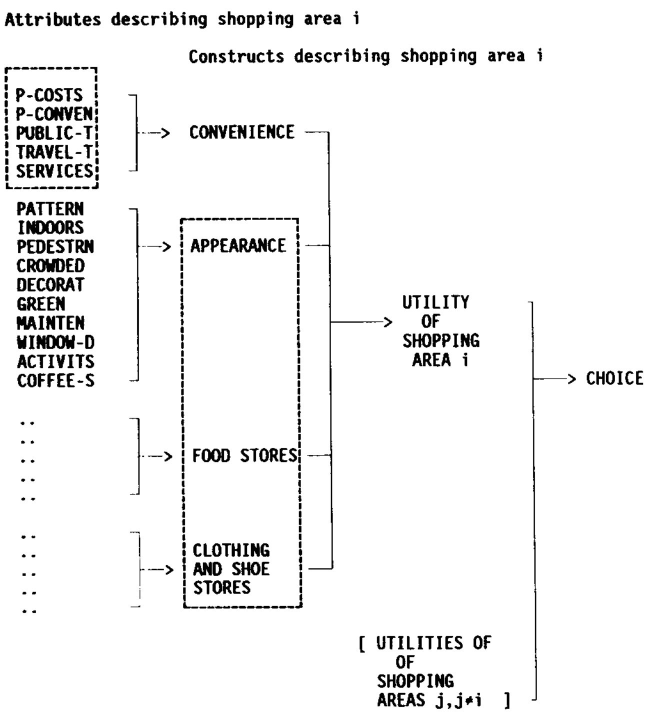
AN APPLICATION TO CONSUMER CHOICE OF SHOPPING CENTER

Introduction to the Problem

Though consumer choice of shopping destination continues to be a research problem of considerable academic and applied interest, only a limited number of conjoint studies have focused on retail destination choice. They entail either consumer store choices (Louviere and Gaeth 1987; Louviere and Johnson 1991; Meyer and Eagle 1982; Moore 1990; Recker and Schuler 1981; Verhallen and De Nooij 1982) or shopping center choices (Timmermans 1982; Timmermans, Van der Heijden, and Westerveld 1984). Predictor variables typically were limited to small sets of broadly defined constructs, such as "selection," "prices," "atmosphere," and/or "distance." Hence, models quantify the relative influence of such constructs on choice and allow one to predict changes in consumer choices, if one knows how to change the subjective values of these constructs by managerial actions.

Though useful for positioning and broad strategy inferences, the applicability of subjective construct-based choice models probably is limited because the set of physically manipulable (actionable) variables that influence the choice of a store or shopping center could be much larger than the small set of constructs typically in-

Figure 2
OVERVIEW OF THE HIERARCHICAL STRUCTURE ASSUMED IN OUR ILLUSTRATIVE APPLICATION



Areas within dashed lines indicate attributes and constructs that were included in the "convenience" subexperiment; attribute names correspond to labels in Tables 1a and 1b.

cluded in previously published models (cf. Timmermans, Van der Heijden, and Westerveld 1982; Zimmer and Golden 1988). In addition, in many practical situations the values of constructs are neither known, nor forecastable; rather, values are known only for physically measurable attributes that (at the present time) are related to subjective constructs in ill-understood ways. Indeed, managers are usually interested in this larger set of potentially manipulable physical attributes, but the latter set is often too large or ill-defined to study in conventional conjoint tasks.

Our proposed extension of HII was used to accommodate a large number of physically manipulable attributes in a conjoint study of consumer patronage of shopping centers. On the basis of a literature review, interviews, and considerations about managerial relevance, a large number of attributes of shopping centers were generated. Attributes were categorized into four groups using logical and practical considerations, and were assumed to correspond to particular decision constructs as shown in Figure 2. Separate choice experiments were designed to represent attribute variation within these groups. Implementation of our integrated design approach involved that in addition to its "own" set of attributes, each experiment included three summary (scaled) measures that represent variation in levels of the three remaining decision constructs.

In our application, decision constructs and subexperiments were defined as follows:

1. *Location convenience and accessibility*. In the corresponding subexperiment ("convenience" subexperiment), this construct was defined by five attributes, as listed in Table 1a. The supply of (nonretail) services was included as an attribute because it could affect the shopping convenience of destinations. In the remaining subexperiments, location convenience and accessibility was one factor with four levels ("− − − very inconvenient," "− moderately inconvenient," "+ moderately convenient," and "+ + + very convenient").
2. *Appearance, layout, and furnishings*. In the corresponding subexperiment ("appearance") this construct was defined by ten attributes, as listed in Table 1b. The construct was disassociated deliberately with the label "atmosphere" to avoid confoundment with characteristics of stores. In the remaining subexperiments the construct was defined as "− − − very unpleasant," "− moderately unpleasant," "+ moderately pleasant," or "+ + + very pleasant."
3. *Selection of stores for food and packaged goods*. In the subexperiment ("food stores") this construct was defined by nine attributes that described the range, quality, and variety of stores that offer these goods. In the remaining subexperiments it was defined as "− − − very bad." "− moderately bad," "+ moderately good," or "+ + + very good."
4. *Selection of stores for clothing and shoes*. Attributes and construct were similar to the "food stores" subexperiment, but explicitly referred to stores for clothing and shoes ("clothing and shoe stores").

Choice data were collected for two different contexts: shopping for foods and packaged goods and shopping for clothing and shoes. However, we use only the food and packaged goods choice data to limit the length of the discussion. Because our present purpose is illustrative rather than substantive, we also confine the discussion to results from the "convenience" and "appearance" subexperiments only. Results for the selection of stores subexperiments led to similar conclusions regarding HII.[1]

Design, Sample, and Procedure

For the "convenience" subexperiment 128 attribute and construct profiles were generated from an orthogonal fraction of a $4^5 \times 2^3$ factorial design. In the "appearance" subdesign, 128 profiles were generated from an orthogonal fraction of a $4^8 \times 2^5$ factorial design. Profiles in each subexperiment were combined into sets of either two or three profiles according to a separate choice set generating design, which generated 48 sets of either two or three profiles for each subexperiment. A constant "not

[1]These conclusions are available from the authors on request.

Table 1a

ATTRIBUTES AND LEVELS THAT DESCRIBE LOCATION CONVENIENCE AND ACCESSIBILITY OF SHOPPING AREAS

p-costs	PARKING COSTS	1 = free 2 = 1 NGL/hr 3 = 2 NGL/hr 4 = 3 NGL/hr
travel-t	TRAVEL TIME in minutes for a single trip (from your home, with the most appropriate means of transport for you, waiting times included)	1 = 5 min 2 = 20 min 3 = 35 min 4 = 50 min
p-conven	PARKING CONVENIENCE	1 = easy 2 = difficult
public-t	PUBLIC TRANSPORT ACCESSIBILITY	1 = good 2 = bad
services	NUMBER OF OTHER SERVICES, FACILITIES OR OFFICES (bank, post office, library, travel agent, etc.)	1 = many 2 = few

First column displays variable labels that are used in Tables 2 to 4.

in stores" base alternative was added to accommodate catalog and other nonstore shopping options. Each respondent evaluated three choice sets from either the "convenience" or "appearance" subexperiment.[2]

Data were collected in Maastricht, The Netherlands, whose 1991 population was approximately 117,000. Streets were selected randomly in each of 18 postal zones. Locally hired and trained interviewers randomly selected households residing on these streets, and interviewed the person in the household who did most of the shopping. Of 428 persons who agreed to be interviewed, 396 completed all parts of the survey including the HII experiment.

The first part of the survey asked respondents questions about shopping habits and perceptions of frequented shopping centers. These data were used to assess the predictive validity of the model as discussed subsequently, but are also relevant to the context of the experiments: (1) For both "food and packaged goods" and "clothing and shoes," we asked how frequently respondents patronized each shopping area and how often they purchased these types of products from a number of nonstore alternatives (e.g., door-to-door, at a market, by mail or telephone, or at wholesalers). The nonstore alternative is the base choice in the tasks; hence, we wanted to ensure that respondents noted these alternatives and the shopping context. In addition, respondents allocated 20 points among all shopping areas and nonstore alternatives to represent the total household budget they spent on each of the two product categories monthly. To compare task results and make the context of the experiments as relevant as possible, we used this same allocation task in the choice experiment. (2) Shopping centers frequented by respondents were evaluated on all experimental constructs and attributes to familiarize them with the complete attribute space and provide a specific context for the experimental task.

Respondents then reviewed a summary list of all attributes and example choice sets. A first example contained two hypothetical shopping areas defined only in terms of decision constructs. The interviewer explained the hypothetical nature of the choice situation, which involved "a residential environment described in terms of its shopping possibilities"; he or she explained that each construct level could serve as a summary measure of an associated profile of underlying attributes. Following this introduction and task warm-up phase, respondents evaluated the experimental choice situations, six from the selection of stores experiments and three from either the "convenience" or "appearance" subexperiment. Figure 3 shows an example choice set from the "convenience" subexperiment and illustrates the general structure of the choice task in this application. In each choice situation respondents first evaluated each construct profile on a 9-category ratings scale, which ranged from "− − − −" (extremely inconvenient, unpleasant, or bad) to "+ + + +" (extremely convenient, pleasant, or good). Next, respondents twice allocated 20 budget-points among the profile alternatives and the "not in stores" alternative. The first allocation was for food and packaged goods, the second for clothing and shoes. However, in the "food stores" subexperiment, only the first allocation was collected, and in the "clothing and shoe stores" subexperiment we only obtained the second allocation. We therefore cannot include this latter subexperiment in our analyses of food and packaged goods allocations.

[2]Further information on the design can be obtained from the authors on request.

Table 1b

ATTRIBUTES AND LEVELS TO DESCRIBE THE APPEARANCE, LAYOUT AND FURNISHINGS OF SHOPPING AREAS

pattern	PATTERN OF STORE LOCATIONS (walking routes, interruption of store fronts, surveyability)	1 = dispersed 2 = compact
indoors	Proportion of shopping area INDOORS	1 = 0% 2 = 30% 3 = 60% 4 = 90%
pedestrn	Proportion of shopping area that is reserved for PEDESTRIANS	1 = 10% 2 = 40% 3 = 70% 4 = 100%
crowded	CROWDING of shopping area	1 = very uncrowded 2 = moderately uncrowded 3 = moderately crowded 4 = very crowded
decorat	FURNISHINGS AND DECORATIONS in the shopping area (signs and displays, stalls, benches, flags, etc.)	1 = few 2 = many
green	GREENERY	1 = little 2 = much
mainten	MAINTENANCE of streets, hallways and buildings	1 = very bad 2 = moderately bad 3 = moderately well 4 = very well
window-d	Proportion of store fronts with ATTRACTIVE WINDOW-DISPLAYS	1 = 0% 2 = 30% 3 = 60% 4 = 90%
activits	Number of STREET ACTIVITIES (markets, musicians, parades, etc.)	1 = few 2 = many
coffee-s	Number of COFFEESHOPS, CAFÉS AND RESTAURANTS	1 = few 2 = many

First column displays variable labels that are used in Tables 2 to 4.

Analysis of HII Data

To illustrate the analysis of HII data, we first estimate separate MNL choice models for each subexperiment from monthly budget allocations, assuming that these allocations are summed counts and reports of the long run frequency distribution of choices. Treating allocations as proportions is consistent with MNL and has been used as such by many others (e.g., Batsell 1980; Hauser and Shugan 1980; Louviere and Woodworth 1983). Next we regressed the construct ratings on their attribute-plus-construct design matrices and compared them with the choice model results. Finally, we performed the Chow tests described previously. We limit the scope of the discussion by presenting main effects only models; however, our design strategy permits one to estimate particular interactions among attributes and constructs. Similarly, one can estimate particular cross-effects to test for IIA, but this also is not illustrated here.

All attributes and constructs were coded for analysis using orthogonal polynomials. That is, attributes with two levels (1,2) were coded $(-1,+1)$ and attributes with four levels (1,2,3,4) were coded $(-3,-1,+1,+3; +1,-1,-1,+1$; and $-1,+3,-3,+1)$ to produce independent linear, quadratic, and cubic components (e.g., Louviere 1988; pp. 44, 62).

Choice models for separate subexperiments. The logit estimated utility components for the "convenience" subexperiment are displayed in the left-hand columns of Table 2. The MNL choice model fits the (aggregated) choice data well (ρ^2 = .82) and all parameter estimates have the expected signs. Partworth utilities derived from these parameters reveal that, for the current range of factor levels, selection of stores for food and packaged goods and

Figure 3
EXAMPLE CHOICE SET FROM "CONVENIENCE" SUBEXPERIMENT

SHOPPING ENVIRONMENT IN RESIDENTIAL SITUATION	SH. AREA #1	SH. AREA #2
COSTS OF PARKING	1 NGL/hr	free
CONVENIENCE OF PARKING	difficult to park	easy to park
PUBLIC TRANSPORT ACCESSIBILITY	p.t.access. is bad	p.t.access. is good
SINGLE TRIP TRAVEL TIME (from your home, waiting times included)	5 minutes	20 minutes
OTHER SERVICES, FACILITIES OR OFFICES (bank, post-office, library, travel agent)	few other facilities	many other facilities
(extremely unfavourable) ---- = *a*, --- = *b*, -- = *c*, - = *d*, 0 = *e*, + = *f*, ++ = *g*, +++ = *h*, ++++ = *i* *(extremely favourable)* LOCATION CONVENIENCE AND ACCESSIBILITY (FROM YOUR HOME) IS *INDICATE YOUR IMPRESSION:* =>	....	....
APPEARANCE, LAYOUT AND FURNISHINGS OF AREA IS	very pleasant +++	very pleasant +++
SELECTION OF STORES FOR FOOD AND PACKAGED GOODS IS	very bad ---	moderately bad -
SELECTION OF STORES FOR CLOTHING AND SHOES IS	moderately bad -	moderately good +

ALLOCATE 20 BUDGET-POINTS:	NOT IN STORES	AREA #1	AREA #2
for foods and packaged goods	... *points*	... *points*	... *points*
for clothing and shoes	... *points*	... *points*	... *points*

Columns within dashed lines each describe one total shopping center.

travel time had the most influence on monthly budget allocations for food and packaged goods, whereas other attributes related to location convenience and accessibility had somewhat less effect. The constructs appearance, layout and furnishings, and selection of stores for clothing and shoes were less important, though significant.

Logit estimates for the "appearance" subexperiment are displayed in the right-hand columns of Table 2. This model also fit its aggregate choice data well ($\rho^2 = .80$). The partworth utilities indicate that the summary constructs "selection of stores that offer food and packaged goods" and "location convenience and accessibility had large effects on the allocation of purchases of these goods," whereas "selection of stores for clothing and shoes" had a much smaller effect. The attributes in this subexperiment were less influential than the attributes in the "convenience" subexperiment, but all except "street decorations" contribute significantly to the explanatory power of the model. Relative attribute effects in order of size, respectively, appear to be "extent to which the shopping area is indoors"; "maintenance of streets, hallways and buildings"; "attractive window-displays"; and "street-activities."

Table 2
LOGIT MODELS FOR "CONVENIENCE" AND "APPEARANCE" SUBEXPERIMENTS

		Convenience		*Appearance*	
		Parameter Estimates	*Asymptotic t-Stat*	*Parameter Estimates*	*Asymptotic t-Stat*
CONSTANT		1.0442	36.02	1.1778	39.14
(Constructs:)					
CONVEN	L			.1266	21.76
	Q			.0645	4.15
	C			−.0090	−1.43
APPEAR	L	.0437	6.68		
	Q	−.0019	−.14		
	C	−.0192	−3.18		
CLOTH&S	L	.0408	7.56	.0183	3.73
	Q	.0301	2.34	.0084	.61
	C	−.0046	−.72	.0076	1.25
FOOD&P	L	.2685	44.42	.2789	40.41
	Q	.0393	2.75	.0006	.05
	C	−.0003	−.04	−.0007	−.11
(Attributes:)					
P-COSTS	L	−.0854	−13.41		
	Q	−.0020	−.15		
P-CONVEN	L	−.1785	−14.16		
PUBLIC-T	L	−.1406	−10.15		
TRAVEL-T	L	−.1881	−31.40		
	Q	.0104	.81		
	C	−.0268	−4.69		
SERVICES	L	−.1020	−8.26		
PATTERN	L			.0378	2.81
INDOORS	L			.0611	10.54
	Q			−.0317	−2.40
PEDESTRN	L			.0211	3.65
	Q			−.0004	−.03
CROWDED	L			−.0150	−2.60
	Q			−.0335	−2.38
DECORAT	L			−.0040	−.31
GREEN	L			.0324	2.50
MAINTEN	L			.0596	8.54
	Q			−.0316	−2.23
WINDOW-D	L			.0550	9.40
	Q			−.0485	−3.64
ACTIVITS	L			.1011	7.69
COFFEE-S	L			.0677	5.30
Model Statistics:					
# Choice sets			48		48
# Cases			176		176
LL(0)			−3745.92		−3287.74
LL(B)			−687.54		−667.87
RhoSq			.816		.797
−2[LL(0) − LL(B)]			6116.76		5239.75
Chi-Sq d.f.			18		25

Attribute names correspond to labels in Tables 1a and 1b. "L," "Q," and "C" are appropriately coded linear, quadratic, and cubic components of attributes.

Table 3
MULTIPLE REGRESSIONS OF CONSTRUCT EVALUATIONS IN "CONVENIENCE" AND "APPEARANCE" SUBEXPERIMENTS

		Convenience		*Appearance*	
		Parameter Estimates	*t-Stat*	*Parameters Estimates*	*t-Stat*
CONSTANT		5.3908	123.29	5.9815	160.11
(*Constructs:*)					
CONVEN	L			−.0090	−.53
	Q			.0855	2.03
	C			.0013	.08
APPEAR	L	−.0203	−1.03		
	Q	.0030	.07		
	C	−.0281	−1.45		
CLOTH&S	L	.0169	.86	.0101	.60
	Q	.0391	.90	.0273	.73
	C	−.0321	−1.64	−.0055	−.33
FOOD&P	L	−.0227	−1.16	.0039	.24
	Q	.1297	2.97	.0218	.58
	C	.0096	.49	.0134	.80
(*Attributes:*)					
P-COSTS	L	−.1441	−7.37		
	Q	.0845	1.93		
P-CONVEN	L	−.2919	−6.68		
PUBLIC-T	L	−.3860	−8.83		
TRAVEL-T	L	−.4304	−22.09		
	Q	−.0032	−.07		
	C	−.0376	−1.91		
SERVICES	L	−.3441	−7.87		
PATTERN	L			.0777	2.08
INDOORS	L			.0761	4.57
	Q			.0084	.23
PEDESTRN	L			.1148	6.87
	Q			−.0302	−.81
CROWDED	L			−.0475	−2.84
	Q			−.0316	−.84
DECORAT	L			.0880	2.36
GREEN	L			.1622	4.34
MAINTEN	L			.1786	9.57
	Q			−.0395	−1.06
WINDOW-D	L			.1759	10.57
	Q			−.0686	−1.84
ACTIVITS	L			.1940	5.19
COFFEE-S	L			.1373	3.68
Model Statistics:					
Multiple R			.557		.421
R Square			.310		.177
Adjusted R Square			.303		.165
Standard Error			1.802		1.535

Attribute names correspond to labels in Tables 1a and 1b. "L," "Q," and "C" are appropriately coded linear, quadratic, and cubic components of attributes.

Analysis of construct ratings and comparison with choice models. The left-hand columns in Table 3 display the results of a regression of the "location convenience and accessibility" construct evaluations on the attributes and constructs that were used in the "convenience" subexperiment. The model fits the data fairly well (R^2 = .31), considering the disaggregate (individual) level of this analysis. All attributes in this model are significant, and, except one quadratic component, no constructs are significant. All attributes also were significant in the logit analysis. As discussed previously, this result supports the assumption that in this subexperiment the attributes represent or define the construct and are independent from other constructs in the hierarchical structure.

The right-hand columns of Table 3 contain the regression results for the "appearance, layout and furnishings" construct evaluations. The estimated model fits the data less well than the model for the "convenience" subexperiment (R^2 = .18), but is still significant. All attribute effects, but no construct effects, are significant (except one quadratic construct component); all attributes significant in the choice models are also significant in this regression. This latter result again supports the assumed hierarchical structure.

In summary, therefore, both regression and choice model results support the conclusion that within each subexperiment each construct represents well its particular attributes independent of the other constructs. Similar results were obtained for the "selection of food stores" subexperiment. Therefore, within subexperiments the choice data seem to satisfy the hierarchical structure assumption.

Equivalence of construct parameters across construct experiments. In the two Chow tests previously outlined we also include the "food stores" subexperiment to increase their relevance. The construct regression equations reported here were used to predict evaluations of the one missing construct in each subexperiment. Next, data from the three construct experiments were concatenated, and dummy codes were used to create parameters that represent variations between construct experiments. Positive dummies were used for the "appearance" and "food stores" subexperiments; dummies were coded negative for the "convenience" subexperiment. Hence, the "convenience" subexperiment was used as a contrast. The original design matrix was multiplied with these dummies to estimate two sets of parameters that represent differences between subexperiments, in addition to generic or pooled construct effects.

An MNL model was estimated from the concatenated data. The results, presented in Table 4, indicate that the effects of constructs defined as profiles of attributes were generally larger than effects of corresponding constructs defined as hypothetical evaluations. For example, the effect of predicted "location convenience and accessibility" evaluations in the "convenience" subexperiment was significantly larger than the effect of the corresponding hypothetical construct evaluation in both other subexperiments. Similarly, the effect of predicted "appearance, furnishings and layout" evaluations in the "appearance" subexperiment was larger than the effect of the corresponding (hypothetical) construct evaluation in both other subexperiments; and the effect of predicted "selection of stores for foods and packaged goods" evaluations was larger than the effect of hypothetical evaluations of the same construct.

Table 4

CONSTRUCTS-ONLY LOGIT MODEL FOR THREE CONCATENATED SUBEXPERIMENTS, INCLUDING CHOW TEST TERMS FOR EQUIVALENCE ACROSS SUBEXPERIMENTS

		Parameter Estimates	*Asymptotic t-Stat*	*Construct-definition in subexperiment*	
CONSTANT		1.0100	21.38		
CONVEN	L	.2406	46.29		
	Q	.0435	3.31		
APPEAR	L	.1079	7.42		
	Q	.1074	3.64		
CLOTH&S	L	.0312	11.37		
	Q	.0000	.01		
FOOD&P	L	.3305	40.64		
	Q	.0030	.21		
"Appearance" versus "Convenience" subexperiment attributes:				*Appearance*	*Convenience*
CONSTANT		.2525	2.98		
CONVEN	L	−.1035	−16.91	Hypo	Pred
	Q	−.0029	−.19		
APPEAR	L	.0982	3.42	Pred	Hypo
	Q	.2003	3.44		
CLOTH&S	L	−.0172	−4.42	Hypo	Hypo
	Q	.0056	.54		
FOOD&P	L	−.0449	−4.97	Hypo	Hypo
	Q	−.0191	−1.21		
"Food stores" versus "Convenience" subexperiments attributes:				*Food stores*	*Convenience*
CONSTANT		−.1888	−3.30		
CONVEN	L	−.0785	−13.90	Hypo	Pred
	Q	−.0278	−1.97		
APPEAR	L	−.0483	−3.29	Hypo	Hypo
	Q	−.0918	−3.06		
CLOTH&S	L	.0134	3.72	Hypo	Hypo
	Q	−.0099	−1.19		
FOOD&P	L	.0987	6.41	Pred	Hypo
	Q	.0442	1.69		
Model statistics:					
Number of choice sets			192		
Number of cases			704		
LL(0)			−14162.37		
LL(B)			−2783.20		
RhoSq			.803		
−2[LL(0) − LL(B)]			22758.34		
Chi-Sq d.f.			27		

The two columns on the right indicate whether a construct score was a hypothetical evaluation or a predicted attribute profile evaluation.

Other terms in Table 4 represent tests of the equivalence of construct effects across both experiments in which a construct was presented as a hypothetical construct evaluation. These results suggest that "selection of stores" constructs had larger effects in the "convenience" subexperiment than the "appearance" subexperiment. Furthermore, the "appearance" construct had a larger effect in the "convenience" subexperiment than the "food stores" subexperiment. In contrast, the "selection of stores for clothing and shoes" construct had a larger effect in the "food stores" subexperiment than in the "convenience" subexperiment.

These results suggest that construct effects were not equal across subexperiments; however, we did not account for potential scale differences among the variou subexperiments, so caution should be used in interpret ing these results. Though the scale parameter in the util ity function of the MNL model is arbitrarily set to on for one sample of data (e.g., Ben-Akiva and Lerma 1985), if different samples are compared or concate nated, scale parameter ratios should reflect difference in variances among the samples (see Swait and Louvier 1993). To investigate this alternative explanation (cor struct effect differences among subexperiments resultin from variance differences), we analyzed pairs of sut experiments to obtain optimal rescaling factors (i.e., th ratio of the scale parameters from both experiments Design matrices for these constructs-only MNL mode were rescaled to equate variances and the models were reestimated. It appeared that even when variation in variances of choices is accounted for, construct effects vary significantly between construct experiments.

Hence, the Chow tests suggest that the hierarchical processes induced by the task structure were not equal across subexperiments, even though regression results supported the independence assumptions within subexperiments. This suggests caution in using the concatenated choice model for predictive purposes because the model may be context sensitive. Having said that, it could be that context effects "average out" across subexperiments, but we cannot directly test this possibility with our data. Rather, we can test the predictive validity of the model as a less satisfactory test. We therefore present the results of the predictive validity tests in the next section.

Predictive Validity of the Concatenated Choice Model

The three subexperiments were concatenated to estimate a choice model that includes all linear and quadratic attribute main effects components. Recall that in the survey, before respondents received any information on the experiment, real market budget allocations were collected and shopping destinations were evaluated on all constructs and attributes used in the experiment to describe hypothetical shopping destinations. Hence, data are available with which to compare budget allocations predicted with the experimental model. To handle missing data problems, except for travel time, attribute ratings of destinations were aggregated to obtain average ratings for the 30 shopping destinations within the urban area. Scores were recoded in accordance with the linear and quadratic components of the experimental attributes to obtain predictor values. Allocations were aggregated across respondents in each postal zone to avoid idiosyncracies in choice set composition. Each zonal choice set consisted of the alternatives that had been mentioned by one or more respondents residing in that zone. The estimated choice model was used to predict distributions within zonal choice sets. The Pearson product moment correlation between predicted and observed market shares within zones was .424. This correlation is substantially higher than those obtained when various restricted models were used to obtain predictions: The correlation with predictions from a null model that included the constant representing purchases "in stores" versus "not in stores" was only .270; a model including this constant and the parameters for size of center (from the "selection of food stores" subexperiment) and travel time (from the "convenience" subexperiment) resulted in a correlation of only .386; and a model including all attributes and remaining constructs from the "convenience" and "appearance" subexperiments resulted in a correlation of .406. Use of the mean absolute error as a measure of predictive validity showed a similar pattern of results (with values 38.13, 41.00, 39.56, and 38.16, respectively).

These results indicate that inclusion of more attributes increases the predictive ability of experimentally based models. The predictive accuracy of all experimental models may seem low, but this is a disaggregate prediction level. Various factors could have reduced our ability to better fit predicted and observed allocations in our study. First, there could have been substantial measurement errors in our validation data. For example, the use of budget allocations could have introduced error because this is a rather difficult task. Second, we did not account for heterogeneity among respondents, though our sample was representative of the total urban area. Third, we used main-effects models only, whereas particular attributes could interact in real markets. For example, the disutility of alternatives with high parking costs and low parking convenience could have been underestimated. We also did not take into account the spatial structure of our validation area, though it is well known that size differences among shopping centers can lead to particular violations of IIA. Finally, the variability in choices in the experimental and validation data sets could have differed, which would affect the unit of the utility scale in the MNL model.

DISCUSSION

We extended the method of HII and illustrated this extension with an empirical application. Our approach involves the design of separate conjoint preference or choice experiments that include both attributes pertaining to a particular decision construct and summary evaluative measures of all other constructs not defined by attributes particular to the experiment. We describe several potential advantages of the extended approach to modeling HII relative to other approaches used to handle choice situations involving many attributes, including previous approaches to modeling HII.

The proposed method was applied to a study of consumer choice of shopping center. Analysis with logit choice models showed that all defined attributes (except one) had significant effects on allocations among shopping centers of monthly household budgets for foods and packaged goods, with signs as expected. Data from the subexperiments were concatenated easily to permit the estimation of one overall choice model permitted that direct estimation of all attribute partworths.

Two types of tests were used to assess the validity of the assumed hierarchical decision process. The first type involves the relation between attributes and constructs within subexperiments and involves tests of (1) whether and how construct ratings are functionally related to all the assumed attributes, independently from the context defined by other remaining constructs and (2) whether this functional relationship of attributes with constructs differs from the direct relationship of attributes with utility or choice. In our application, the choice data passed both these tests.

The second type of test involves whether the influence of constructs on choice is equivalent across (1) subexperiments and (2) conditions in which the construct is *a* *priori* defined as a single measure as well as conditions in which the construct is defined by a profile of attributes. Our data did not pass these latter tests. First, for all constructs, hypothetical single evaluation scores had smaller effects on choice than profiles of attributes. Second, the sizes of effects of hypothetical construct evaluations differed across construct experiments.

A possible explanation for the nonequivalence of construct effects and joint effects of attributes is that each additional attribute in a profile captures some share of the attention that subjects pay to profiles. Attribute sets and constructs that are theoretically equivalent, therefore, are not equivalent in their effects on utility. This resembles the effects that have been observed in conjoint ranking tasks for increasing numbers of levels of attributes (Wittink, Krishnamurti, and Reibstein 1989).

An attention effect, however, would not explain nonequivalence across subexperiments of effects of *a priori* defined hypothetical construct ratings. These differences seem to indicate that choices were context dependent in our application. However, there also could have been other reasons why differences in construct effects were found: The IIA property could have been violated, there could be specification errors due to omitted interaction terms in the utility functions, or scales of utility functions could have been unequal across construct experiments. The latter of these alternative explanations was tested and rejected, but further research is needed to investigate the other possible alternative explanations.

Further research also should focus on the external validity of conjoint choice models and of HII models in particular, because it remains an open empirical issue as to what extent subjects' decision-making processes outside laboratories conform to hierarchical structures and whether preferences and choices are unaffected if subjects are forced to process information according to such a task structure. These are important research questions for not only HII but for conjoint-type models in general. Indeed, our tests suggest that context dependency could be a larger problem than commonly is assumed. The advantage of our approach over other conjoint-type methods is that it provides indicators of these effects.

However, because our present focus is on HII as a way to design experiments that allow one to predict choices, another important research objective was to test the predictive ability of choice models derived from HII experiments. Such tests can be performed independently of tests of the assumed hierarchical structure because all construct indicators in the estimated utility function can be set to zero. We therefore used the concatenated model to predict choice, regardless of whether construct effects differed between construct experiments. Model predictions corresponded to the observed "revealed" budget allocations reasonably well, suggesting that our proposed HII extension deserves further research attention as a possible way to develop utility functions when many attributes are of interest.

Finally, it should be noted that the current application involved only one possible way to implement HII. For example, our approach also can be applied to preference tasks, which might allow one to estimate individual level models. Furthermore, numerous variations in the design of preference and choice experiments can be implemented. For example, one could link attributes to more than one construct; tasks in which two or more constructs are substituted with profiles of attributes could be implemented (and compared to full profile implementations); tasks in which respondents are not instructed about the potential hierarchical structure of attributes and constructs could be used if more critical tests of the mapping of attributes into constructs are required. (Even "filler" attributes that are expected to have no relation with constructs could be used.) Finally, the effects of missing attributes on construct evaluation and choice could be investigated with tasks in which the absence of attributes is varied systematically. This range of potential designs and research questions illustrates the flexibility of the proposed HII approach. This and other advantages demonstrated here suggest that integrated conjoint experiments could provide a valuable tool to estimate partworth utilities and predict choice for situations that involve many attributes.

REFERENCES

Albaum, Gerald (1989), "Bridger, Simgraf" (review), *Journal of Marketing Research*, 26, 486–8.

Anderson, Norman H. (1981), *Foundations of Information Integration Theory*. New York: Academic Press.

——— (1982), *Methods of Information Integration Theory*. New York: Academic Press.

Batsell, Richard R. (1980), "Consumer Resource Allocation Models at the Individual Level," *Journal of Consumer Research* 7, 78–87.

Ben-Akiva, Moshe, and Steve Lerman (1985), *Discrete Choice Analysis*. Cambridge, Mass.: MIT Press.

Bretton-Clark Inc. (1988), *BRIDGER*. New York.

Carmone, Frank (1987), "ACA System for Adaptive Conjoint Analysis," *Journal of Marketing Research* 24, 325–7.

Corfman, Kim P. (1991), "Comparability and Comparison Levels Used in Choices Among Consumer Products," *Journal of Marketing Research*, 28, 368–74.

Green, Paul E., Abba M. Krieger, and Pradeep Bansal (1988), "Completely Unacceptable Levels in Conjoint Analysis: A Cautionary Note," *Journal of Marketing Research* 25, 293–300.

——— and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice," *Journal of Marketing*, 54, 3–19.

Hauser, John R. and Steven M. Shugan (1980), "Intensity Measures of Consumer Preference," *Operations Research*, 28, 278–320.

Johnson, Richard M. (1974), "Trade-off Analysis of Consumer Values," *Journal of Marketing Research*, 11, 121–7.

Johnson, Richard D. (1987), "Making Judgments When Information Is Missing: Inferences, Biases, and Framing Effects," *Acta Psychologica*, 66, 69–82.

Kmenta, Jan (1986), *Elements of Econometrics*. New York: MacMillan.

Louviere, Jordan J. (1984), "Hierarchical Information Integration: A New Method for the Design and Analysis of Complex Multiattribute Judgement Problems," *Advances in Consumer Research*, 11, 148–55.

——— (1988), *Analyzing Decision Making: Metric Conjoint Analysis*. Sage University Paper Series on Quantitative Applications in the Social Sciences, series no.07-067. Beverly Hills, CA: Sage Publications.

——— and Gary J. Gaeth (1987), "Decomposing the Determinants of Retail Facility Choice Using the Method of Hierarchical Information Integration: A Supermarket Illustration," *Journal of Retailing*, 63, 25–48.

——— and Richard D. Johnson (1991), "Using Conjoint Analysis to Measure Retail Image" in *Spatial Analysis in Marketing: Theory, Methods, and Applications* (Research in Marketing, Supplement 5), A. Ghosh and C. Ingene, eds. Greenwich, CT: JAI Press.

——— and Harry J.P. Timmermans (1990a), "Using Hierarchical Information Integration to Model Consumer Responses to Possible Planning Actions: Recreation Destination Choice Illustration," *Environment and Planning A*, 22, 291–308.

——— and Harry J.P. Timmermans (1990b), "Hierarchical Information Integration Applied to Residential Choice Behavior," *Geographical Analysis*, 22, 127–45.

——— and George G. Woodworth (1983), "Design and Analysis of Simulated Choice or Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, 20, 350–67.

Meyer, Robert J. and Thomas C. Eagle (1982), "Context-induced Parameter Instability in a Disaggregate Stochastic Model of Store Choice," *Journal of Marketing Research*, 19, 62–71.

Moore, Laurence (1990), "Segmentation of Store Choice Models using Stated Preferences," *Papers of the Regional Science Assocation*, 69, 121–31.

Recker, Wilfred W. and Harry J. Schuler (1981), "Destination Choice and Processing Spatial Information: Some Empirical Results with Alternative Constructs," *Economic Geography*, 57, 373–83.

Swait, Joffre and Jordan J. Louviere (1993), "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models," *Journal of Marketing Research*, 30 (August), 305–14.

Timmermans, Harry J.P. (1982), "Consumer Choice of Shopping Centre: An Information Integration Approach," *Regional Studies*, 16, 171–82.

———, Aloys Borgers, Joan van Dijk, and Harmen Oppewal (1992), "Residential Choice Behaviour of Dual Earner Households: A Decompositional Joint Choice Model," *Environment and Planning A*, 24, 517–33.

———, Rob E.C.M. van der Heijden, and Hans Westerveld (1982), "Cognition of Urban Retailing Structures: a Dutch Case Study," *Tijdschrift voor Economische en Sociale Geografie*, 73, 2–12.

———, ———, and ——— (1984), "Decision-Making Between Multiattribute Choice Alternatives: A Model of Spatial Shopping Behaviour Using Conjoint Measurements," *Environment and Planning A*, 16, 377–87.

Verhallen, Theo M.N. and Gert Jan de Nooij (1982), "Retail Attribute Sensitivity and Shopping Patronage," *Journal of Economic Psychology*, 2, 39–55.

Wittink, Dick R., Lakshman Krishnamurthi and David J. Reibstein (1989), "The Effect of Differences in the Number of Attribute Levels on Conjoint Results," *Marketing Letters*, 1, 113–23.

Zimmer, Mary R. and Linda L. Golden (1988), "Impressions of Retail Stores: A Content Analysis of Consumer Images," *Journal of Retailing*, 64, 265–93.