

# Empirical test of the Lancaster characteristics model

Berend WIERENGA \*

## 1. Introduction

The 'New Economic Theory of Consumer Behaviour' (Lancaster 1966, 1971) is an important development in the way economists look at the consumer. This approach, also called the Characteristics Model, was especially welcomed (Nicosia 1974; Ratchford 1975) because of its multidimensional orientation. A product is conceived of as a bundle of characteristics that have want-satisfying properties to the consumer. This multidimensional view can directly be linked to the multi-attribute models which have been developed by the behavioral science-oriented consumer researchers.

Although the important contribution of the characteristics model is the improvement of fundamental insight in consumer behavior, the model can also be used as a basis for normative models for the development of profit maximizing new products (Hauser and Simmie 1981) and for the formulation of normative recommendations to defend existing products against competitive entrants

(Hauser and Shugan 1983; Hauser and Gaskin 1983). Various suggestions have been made with respect to empirical applications of the Lancaster characteristics model (Ratchford 1975: 70-71). Up to now a few examples of empirical studies in the Lancaster framework have appeared in the literature: Ryans (1974), Ratchford (1980), Hauser and Gaskin (1983). In these examples the characteristics model has been applied to 'one-shot purchases', i.e. situations where the consumer chooses one product from a set of alternatives. Ratchford (1979) has shown that the Lancaster model can be accommodated to this situation. The real spirit of the model however, is in the explanation of purchases of combinations of different products. McAlister (1979) has reported about an empirical study in this area, without reference to the Lancaster framework though. Up to now no complete operationalization of the Lancaster model in this empirical setting seems to have been carried out.

This paper presents such an operationalization for the case of choice of combinations of vegetables by consumers in The Netherlands. Judging from the several examples about food products given in this book (Lancaster 1971) vegetables should constitute a natural product for application of the Lancaster model. The various elements of the Lancaster model will be measured and we will examine how well this model can explain and predict combinations of vegetables, purchased by consumers for a one-week period.

However, from the outset it was clear that two limitations of the original Lancaster model would probably be too restrictive to provide a successful explanation.

As has been argued by several authors (Nicosia (1974: 166), Ratchford (1975: 67),

\* Author's address: Erasmus Universiteit Rotterdam, (Inter) fakulteit Bedrijfskunde, Postbus 1738, 3000 DR Rotterdam, The Netherlands.

The author wishes to thank Jaap Bijkerk and Gerard Verwey (Agricultural University Wageningen) for their computer assistance and SOCMAR B.V. (Research International) for their cooperation in the data collection process. Useful comments and suggestions were obtained from Leigh McAlister (MIT) and from an anonymous reviewer.

Accepted by Philippe Naert.

Hauser and Simmie (1981: 36)), Lancaster ignores the perception process, i.e. the notion of 'perceived characteristics'. From information processing theory we know that perceived characteristics may well differ from their physical counterparts and that some perceptual dimensions are more the result of socio-psychological cues produced by advertising, consumer education and word-of-mouth communication, than of the physical properties of products. This does not invalidate the Lancaster model but requires an extension with a 'perception module'. The Lancaster model has especially been welcomed because – through its multidimensional orientation – it has offered a framework with which to link the economic theory of consumer behavior with the multi-attribute models developed by the behavioral science-oriented consumer researchers becomes possible. The use of (psychologically based) multidimensional scaling models of product perception within the framework of the Lancaster model, as to be applied in this paper, exemplifies these integration possibilities.

A second element, resulting from recent insights into consumer behavior, not represented in the Lancaster model, is variety seeking behavior which plays a role in many consumer purchasing decisions. Although a variety drive in consumer behavior has been referred to in the past (Berlyne (1960), Howard and Sheth (1969), Wierenga (1974, ch. 6)), the topic has recently received more attention: Faison (1977), Jeuland (1978), Howard (1980), McAlister and Pessemier (1982). It is possible that a consumer acquires utility from the consumption of many different items as such, in addition to the consumption of the characteristics levels implied by the combinations of items. Such an independent value of variety is not accommodated for in the standard characteristics model. This paper presents an extension which makes it possible to determine the importance of this variety drive and integrates it with the characteristics

model. The Lancaster model with these two extensions is operationalized and tested for its explanatory power in this paper.

## 2. The Lancaster model and extensions

### 2.1. The standard model

The Lancaster model assumes a linear relationship between (physical) characteristics and products according to the expression:

$$\begin{aligned}
 y &= Bx, \\
 y &= (r \times 1) \text{ vector of characteristics,} \\
 B &= (r \times n) \text{ 'matrix of consumption technology',} \\
 x &= (n \times 1) \text{ vector of products,} \\
 y_i &= (i = 1, \dots, r) \text{ is the amount of characteristic } i, \\
 x_j &= (j = 1, \dots, n) \text{ is the quantity of product } j, \\
 b_{ij} &= \text{the amount of characteristics } i \text{ per standard unit of product } j. \quad (1)
 \end{aligned}$$

If it is assumed that there is a budget constraint  $K$ , and the prices of the products are  $p_j$  ( $j = 1, \dots, n$ ) the feasible region of characteristics combinations determined by  $K$  and  $p_1, \dots, p_n$  can be indicated in the characteristics space. For the case of two characteristics and four hypothetical products this is illustrated in figure 1. For example, the arrow with the number 1 represents the characteristics combination that can be bought when the whole budget is allocated to product 1. The coordinates of this point are:  $Kb_{11}/p_1$  and  $Kb_{21}/p_1$ . In figure 1 the products, 1, 2 and 4 are efficient. Product 3 is not efficient, because spending the budget on combinations of products 2 and 4 gives higher levels of both characteristics 1 and 2 than buying product 3. The curve, connecting the arrows 1, 2 and 4 is called the 'efficiency frontier'.

According to the theory, only efficient combinations of products are purchased. Which of these combinations are ultimately

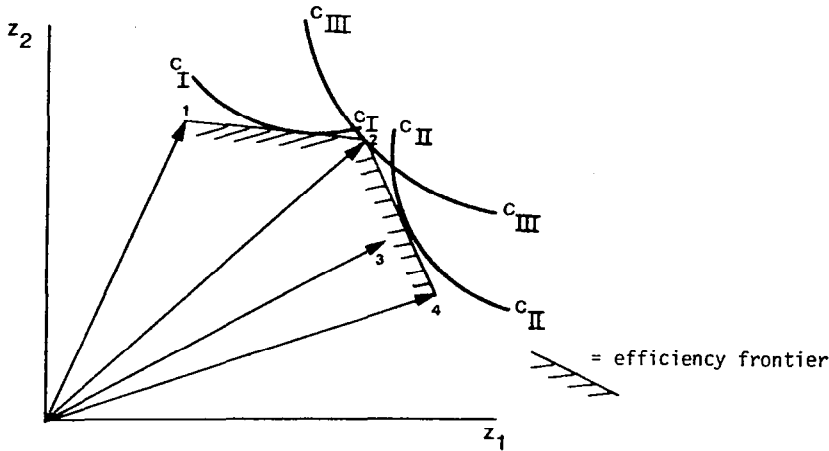


Figure 1. Optimization in the characteristics model with 2 characteristics and 4 products.

chosen depends on the utility functions and the indifference curves of the consumers. For example, in figure 1 a consumer with indifference curve  $C_I-C_I$  would purchase a combination of the products 1 and 2, a consumer with indifference curve  $C_{II}-C_{II}$  buys a combination of products 2 and 4 and a consumer with indifference curve  $C_{III}-C_{III}$  will only buy product 2.

Algebraically, in the Lancaster model a consumer purchases the quantity for which the utility function  $U(y)$  is maximum, given the budget constraint. The optimization problem is solved thus:

$$\max U(y), \text{ such that} \quad (2)$$

$$y = Bx, \text{ and} \quad (3)$$

$$\sum_j^n p_j x_j \leq K, \quad (4)$$

$$x_j, y_i \geq 0. \quad (5)$$

Figure 2 gives a schematic representation of the standard Lancaster model. It shows that the physical features of the products, given by the consumption technology relationship  $y = Bx$ , together with the prices determine the efficient combinations. The optimal combination to be chosen from these efficient combinations is then determined by the utility func-

tion and the budget constraint. With respect to the characteristics, Lancaster assumes 'universality and objectivity' (1971: 18); 'every consumer in the economy is assumed to see the same consumption technology' and '... there is no difference between them as to what collection of characteristics is associated with any specified collection of goods'. Of course, a major problem is to determine the characteristics. Lancaster (1971: 115): 'Our fundamental operational problem is determining the relevant characteristics for choice'. The advantages of the characteristics model are much greater when the number of relevant characteristics is relatively low: Lancaster (1971: 140). Lancaster assumed much more heterogeneity with respect to preferences than with respect to the assessment of products (ibid.: 7): 'Individuals differ in their reactions to different characteristics rather than on their assessment of the characteristics content of various goods collections'. It is important to examine these assumptions in practical applications. It is also important to check two implications of the model: (i) dominated products, i.e. products not on the efficiency frontier will not be bought, and (ii) the number of different product bought by one consumer is smaller than or equal to the number of relevant characteristics.

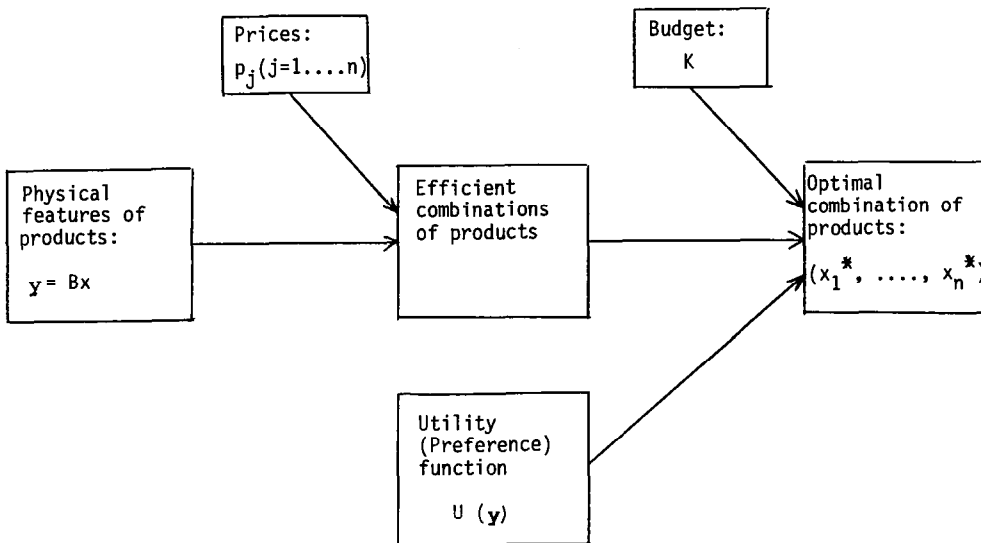


Figure 2. Schematic representation of the Standard Lancaster-model.

## 2.2. The extended Lancaster model

*Perceived characteristics.* The utility a consumer obtains from consuming a product or a combination of products is determined by the characteristics he or she attributes to the products in question. The variables that should enter the utility function are the characteristics as the consumer sees them. As referred to earlier, the perceived characteristics may well be at variance with objective, physical attributes of the products. The transformation of physical product attributes into perceived characteristics can be described by a matrix  $H$  (see also Hauser and Simmie 1981: 37):

$$Y^p = ZH, \quad (6)$$

where  $Z$  is a  $(n \times r)$  matrix of physical attributes,  $H$  is a  $(r \times r_p)$  transformation matrix and  $Y^p$  is the matrix of perceived characteristics, derived from physical attributes. The number of relevant physical attributes is  $r$ , the number of perceived characteristics, derived from physical attributes is  $r_p$ ,  $r_p$  may or may not be equal to  $r$ . Furthermore, through socio-psychological cues (publicity, advertising, consumer education) properties can be

attributed to a product that have no direct relationship with physical attributes: e.g. status-appeal of a product. We indicate such socio-psychological dimensions with the matrix  $Y^s$ . The ultimate perceived characteristics matrix of a product is then: <sup>1</sup>

$$Y = Y^p + Y^s. \quad (7)$$

The matrix  $Y$  indicates how much a product has of a specific characteristic (in the eyes of the consumer). So  $Y$  can be considered as the 'perceived matrix of consumption technology'.  $Y$  can be found by means of multidimensional scaling methods.

*Variety seeking behavior.* Two types of variety seeking consumer behavior have been distinguished (McAlister and Pessemier 1982).

*Derived variation* is variation in consumption patterns not for the purpose of variation as such but as a 'by-product' of other phenomena, for example multiple needs, multiple users, multiple situations.

<sup>1</sup> When there are  $r$  relevant perceived characteristics, of which  $r_p$  are derived from physical attributes and  $r_s$  from socio-psychological cues ( $r = r_p + r_s$ ),  $Y$  can be assumed to be a  $(n \times r)$  matrix of which the last  $r_s$  columns are zero and  $Y^p$  a  $(n \times r)$  matrix of which the first  $r_p$  columns are zero.

*Direct variation* is caused by the need for variety as such. Novelty and change may be pursued as goals in itself. Consumers may vary the products independent of the (perceived) characteristics content of the products. This can be linked to the standard Lancaster model in the following way. In the 'standard' situation the consumer buys a combination only for its characteristics content. The resulting combination of items ( $x_1, \dots, x_r$ ) – the solution of the optimization problem (2)–(5) – is called here the standard model solution. When, apart from characteristics utility, also variety over products plays a role, this can be modeled by adding a term to the objective function (eq. (6)) which expresses the variety level implied by a specific combination ( $x_1, \dots, x_n$ ) of items. Let the variety level be indicated by  $V(x)$ . The optimization problem then becomes:

$$\max wU(y) + (1 - w)V(x), \text{ such that} \quad (8)$$

$$y = Bx, \text{ and} \quad (9)$$

$$\sum p_j x_j \leq K, \quad (10)$$

$$x_j, y_i \geq 0. \quad (11)$$

Here  $w$  ( $0 \leq w \leq 1$ ) expresses the relative weight a consumer attaches to characteristics utility and variety respectively. We call this the Variety Seeking Model. Eq. (8) implies that variety seeking is a matter of degree: for  $w = 1$  there is no variety drive at all (the Standard Model is a special case of the Variety Seeking Model), for  $w = 0$  choice behavior is completely determined by the desire for variety, whatever the implied characteristics of the product. For  $0 < w < 1$  consumer behavior can be seen as a compromise of obtaining utility from characteristics and striving for variety. An important issue is then how to define the variety level  $V(x)$ . A natural choice (Pessemier 1981) is to use an information-theoretical measure, e.g.

$$V(x) = \sum_{j=1}^n -q_j x_j \ln q_j x_j, \text{ where} \quad (12)$$

$$q_j = p_j / K. \quad (13)$$

So  $q_j x_j$  is the proportion of the budget going to product  $j$ . Another possibility is to look directly at the vector  $x$  and to consider the variance over the components  $x_1$  to  $x_n$ . In a situation where the entire purchases are concentrated on one vegetable, e.g.  $x_j = 0$  and  $x_i = 0$  for  $i \neq j$ , the variance in the vector  $x$  is maximum. However, when all vegetables are bought in the same quantities, i.e.

$$x_1 = x_2 = \dots = x_n, \quad (14)$$

the variance in  $x$  equals zero. So a drive for direct variety can be translated into a striving for a low variance in the vector  $x$ . This implies that in eq. (8) the variety measure should be defined as:

$$V(x) = -\text{Variance}(x), \text{ where} \quad (15)$$

$$\text{Variance}(x) = \sum_{j=1}^n (x_j - \bar{x})^2 / n. \quad (16)$$

This specification is similar to Farquar and Rao's Balance Model (1976). Both specifications of  $V(x)$ : eq. (12) and eq. (15) were used in the research reported here.

This completes the description of the Extended Lancaster Model as used in this study. The model is schematically depicted in figure 3 (which can be compared with the standard model of figure 2). Some aspects of figure 3 will become clear after the description of the data collection in the next section.

### 3. Vegetables as a product class; collection of the data

As already stated, the specific situation in which the model is operationalized here is vegetables-buying by Dutch consumers. Vegetables are an important component of the Dutch meal. The typical main course of a hot meal in the Netherlands consists of vegetable(s), potatoes and meat. Much publicity is made about the advantages of eating vegeta-

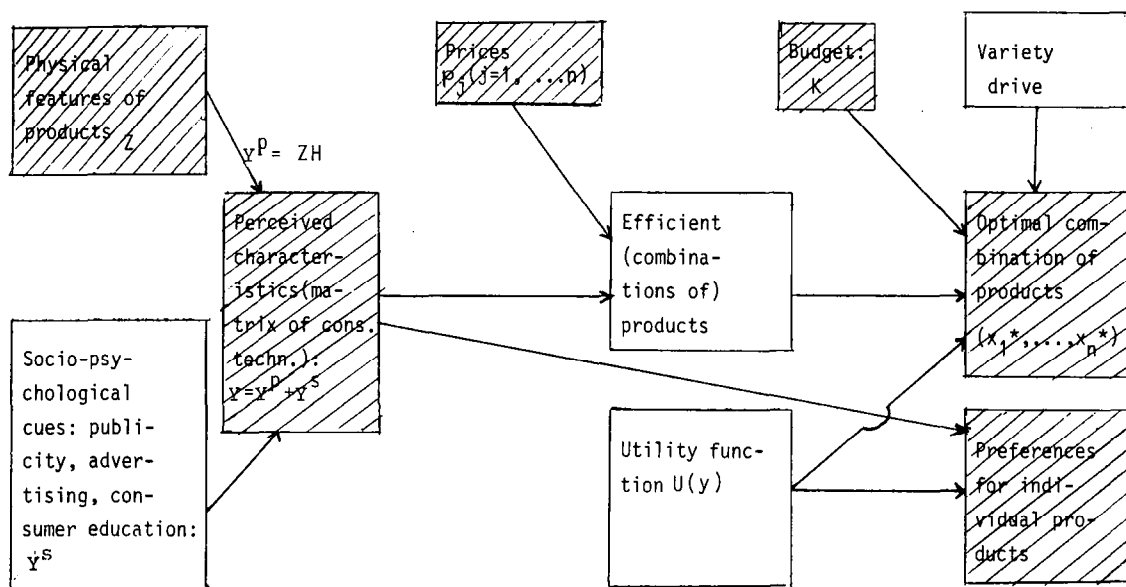


Figure 3. Extended version of the Extended Lancaster Model, used in this study.

bles. This is not only done by interested commercial parties such as the Dutch Central Auctions Organization, but also by independent institutions like the Dutch Consumer Food Education Bureau and the Consumer Union. In these communications much emphasis is put on the role of vegetables in the provision of vitamins and iron (*Fe*). In publications, especially the vitamins *A* and *C* are mentioned. It is estimated (Den Hartog 1972: 282) that in the total Dutch consumption pattern vegetables deliver 85% of total carotene intake (carotene = provitamin *A*), 37% of the total vitamin *C* intake and 9% of the total intake of iron. In our study we used these three components: iron, vitamin *A* and vitamin *C* as physical attributes. These are called here: micro-components, because an individual needs only very small quantities (milligrams) in his daily menu. To the physical characteristics we added two macro-components: carbohydrates and proteins. Although vegetables deliver only a minor part (1,3% and 3,5%, respectively) of these components in the total nutrition, they are important as prominent producers of calories.

The third important macro-component: fat is practically non-existent in vegetables.

When using the Lancaster model for the explanation of combinations of vegetables bought by consumers, the question should be asked if vegetables constitute an 'intrinsic good', i.e. if the decisions of consumers with respect to the characteristics combinations obtained through the consumption of vegetables can be isolated from the decisions with respect to the other products in the daily menu. Since several nutritional components provided by vegetables are also supplied by other food products, a complete separation is not possible. However, for practical purposes, isolating the choice with respect to vegetables does not seem completely unjustified.

The ideal situation would be that a studied product group (here: vegetables), gives rise to a common set of characteristics that would be provided only by this group. See Lancaster's discussion about conditions for isolated analysis of a group: Lancaster (1971, ch. 8). There are characteristics (vitamins) of which vegetables provide a very substantial part and other characteristics for which the contribution of

vegetables is very minor (carbohydrates and proteins). Secondly, the choice of vegetables in the Dutch meal is not a secondary one, i.e. after the other elements of the meal have been determined. A recent market research study reports that 55% of Dutch housewives start the composition of the main meal by choosing a vegetable (25% start with the choice of meat). Thirdly, although there may be a tendency to create an equilibrium of characteristics in the particular meal for one day, when a longer period is considered, e.g. one week, as is done in this study, differences because of the composition of the rest of the meal (meat, soup, dessert) will tend to cancel out.

The data were collected in May 1978 through a Market Research Agency. Respondents were 300 housewives from a panel who come to the agency a few times a year to perform various tests. The interviews were carried out at the agency because of the various aids that had to be used during the interview: coloured photographs of the vegetables, cards with the names of vegetables, lists of prices, etc. Fifteen vegetables were used in the study. These were selected to be representative of the vegetables actually available on the market in the particular period of the year. The vegetables used are listed in table 2. The interview started with a subjective grouping of vegetables to provide data for a perceptual map. The preferences for individual vegetables were asked and subsequently the respondents had to rate the vegetables on a number of scales, including scales referring to the contents of nutritional components: iron, vitamin C, carbohydrates and proteins. The next part of the interview was a buying simulation. Respondents were told they had Dfl. 15,- available and were asked to purchase vegetables for a family of two adults and two children for a one-week period. This was done on a day-to-day basis: they were asked what they would buy for the Monday, then for the Tuesday, Wednesday, etc. To make the exercise more realistic, for each day they were

also asked what they would use with the vegetables: the type of meat, potatoes, rice or an equivalent product, and whether they would have soup and dessert. The prices of the vegetables were chosen as close as possible to the actual retailing prices (which fluctuate greatly in The Netherlands).

The thinking process, so generated within the respondents, is a very natural one for a Dutch housewife who usually goes to the supermarket or the greengrocer to buy vegetables for her family for a couple of days. The housewives did not have any great problem in making their decisions, so it may be expected that their purchasing decisions in the buying simulation are representative of their actual daily behavior. Although a budget for vegetables of Dfl. 15,- should be adequate, respondents were asked whether they found this amount of money sufficient. More than two thirds (68%) found the budget sufficient. The shaded boxes in figure 3 indicate the variables in the model on which data are available. Apart from the data on perceptions, preferences and vegetable-combinations bought, as obtained in the interview, data are also available with respect to the physical characteristics of the vegetables, while the prices and the budget were set by the researcher. The various parts of the model will be analysed in the following sections. The analysis was performed on data from 150 (randomly selected) of the 300 respondents. The data from the other 150 respondents were used for validation purposes.

#### **4. Determining the perceived characteristics of vegetables**

##### *4.1. Different methods for perception analysis*

Since the detection of the right dimensions in the Lancaster framework is of critical importance, five different methods of perception analysis were used.

Table 1  
Frequencies of vegetable scores on attribute statements and average scale values for nutritional components ( $n = 150$ ).

	endive	asparagus	cauliflower	mushrooms	cucumber	sweet pepper	leeks	rhubarb	red cabbage	lettuce	French beans	spinach	onion	white cabbage	carrots
real vegetable for those doing heavy work	147	52	140	5	32	6	100	14	125	120	142	141	44	115	26
well-known vegetable	120	7	82	2	12	3	72	3	129	48	108	115	74	110	72
high nutritional value	150	65	150	54	111	47	132	82	148	144	148	148	123	140	148
expensive	142	28	79	32	31	51	90	27	91	96	104	131	63	77	106
festive meal	7	140	42	102	6	68	10	8	0	3	46	12	1	0	10
a lot of preparation	5	136	63	129	23	76	4	20	1	45	52	18	3	0	31
for high income	81	60	15	16	1	10	16	18	45	23	18	111	16	24	35
many ways to serve	3	109	21	62	8	38	7	7	3	6	16	8	4	1	10
for modern people	68	74	54	119	69	117	37	15	22	94	15	10	90	26	25
for a side dish	33	46	34	81	58	95	36	27	15	61	35	30	42	21	30
iron	2	55	4	95	133	89	10	114	1	122	5	3	44	4	19
vitamin A	3.00	1.45	1.47	1.27	1.33	1.73	1.94	1.76	1.67	2.31	1.90	3.83	1.57	1.57	1.63
vitamin C	2.52	1.83	2.00	1.76	1.83	2.21	2.09	1.87	1.94	2.47	2.27	3.63	1.86	1.81	2.71
carbohydrates	2.96	1.92	2.09	1.87	2.18	2.85	2.31	2.25	2.01	3.11	2.33	3.25	2.13	1.91	2.72
proteins	2.01	1.75	2.05	1.77	1.45	1.67	1.97	1.72	2.22	1.72	2.18	2.03	1.83	2.31	2.01
	2.06	1.93	1.94	2.04	1.34	1.59	1.75	1.42	1.51	1.79	1.96	2.11	1.57	1.66	1.83



One of these methods is of the decompositional type (for a discussion of this typology see Hauser and Koppelman 1979; Huber and Holbrook 1979): multidimensional scaling of similarity data (MDSCAL). The other methods were of the compositional type: factor analysis on individual data (FACTIN), discriminant analysis on individual data (DIS-

CIN), factor analysis on aggregate data (FACTAG) and MDPREF. Similarity data on the 15 vegetables were obtained using the subjective grouping method.

For the compositional methods the respondents were asked to rate the vegetables on sixteen attribute scales. These referred to the five nutritional components and eleven psy-

Table 2a

Perceptual matrix  $Y$  of 15 vegetables obtained by factor analysis of individual attribute scores (FACTIN) 2 (rescaled average factor scores).

	Dim. 1	Dim. 2	Dim. 3
endive	0.132	1.816	0.912
asparagus	2.283	0.848	0.140
cauliflower	0.676	1.653	0.337
mushrooms	1.699	0.206	0.064
cucumber	0.082	0.116	0.000
sweet pepper	1.050	0.000	0.421
leeks	0.111	1.274	0.474
rhubarb	0.093	0.112	0.044
red cabbage	0.000	1.758	0.289
lettuce	0.179	0.802	0.777
French beans	0.597	1.784	0.587
spinach	0.233	1.868	0.885
onions	0.040	0.825	0.066
white cabbage	0.003	1.589	0.189
carrots	0.217	1.428	0.606

Table 2b

Correlations of dimensions with attributes ( $|r| \geq 0.6$ ).

	expensive:	0.995	real vegetable:	0.940	vitamin A:	0.925
	festivity:	0.964	heavy work:	0.936	vitamin C:	0.913
	high income:	0.980	for a side dish:	-0.912	iron:	0.898
	well-known vegetable:	-0.698	carbohydrates:	0.832	high nutritional value:	0.894
			high nutritional value:	0.830	real vegetable:	0.721
			well-known vegetable:	0.830		
			for modern people:	-0.730		
			many ways to serve:	-0.618		
Name:	'Distinction'		'Energy'		'Micro-components'	



notations as well as perceived nutritional characteristics.

#### 4.2. The perceptual dimensions found

In FACTIN there were four eigenvalues greater than one, which together explain 52% of the variance. This percentage seems rather low but this is caused by the variance in individual data. FACTAG produced a very similar factor structure, with three eigenvalues greater than one, explaining together 91% of the variance. In DISCIN there were two eigenvalues greater than one, explaining together 79% of the variance. With a third discriminant factor added, the percentage explained was 86%. In MDPREF where the attributes of table 1 were the vectors and the vegetables were represented by points, three factors explained 83% of the variance. The stress of the MDSCAL solutions in 1, 2, 3 and 4 dimensions is 0.350, 0.169, 0.080 and 0.042, respectively. This shows an 'elbow' at a dimensionality of three. So a dimensionality of three turned out to be appropriate given these results.

Using C-match, see Cliff (1966), the mutual correspondence between the different percep-

tual configurations was explored. Table 3 shows that essentially all perceptual matrices obtained have the same three dimensions. From the five perceptual matrices a choice had to be made for use in the analyses to follow. This choice is a bit arbitrary because of the similarities between them. We decided to use the perceptual matrix of table 2 which was obtained by factor analysis of individual attribute scores (FACTIN). Hauser and Koppelman (1979) report the best result for this method in predicting preferences (as compared to discriminant analysis and MDS-CAL). Moreover, through factor analysis the relationship between the original attribute ratings and the dimensions obtained is retained the best, which is especially important for the nutritional components. The matrix of table 2 is the perceived 'matrix of consumption technology' used further in the study.

Table 2 gives the perceptual matrix obtained by FACTIN (scaled so that per dimension the lowest value is zero). In the lower part of table 2 the correlation coefficients with the attribute scales are given (if they are larger – absolutely – than 0.6). Dimension 1 is related to expensive, festive, high income, not very well-known vegetable (exclusive). Therefore, this dimension is called: *distinc-*

Table 3  
Correlation coefficients between dimensions of perceptual matrices obtained by different methods (after C-match).

		FACTIN	DISCIN	FACTAG	MDPREF	MDSCAL
FACTIN	{ distinction	–				
	{ energy	–				
	{ micro components	–				
DISCIN	{ distinction	0.996	–			
	{ energy	0.951	–			
	{ micro components	0.922	–			
FACTAG	{ distinction	0.911	0.924	–		
	{ energy	0.833	0.901	–		
	{ micro components	0.765	0.702	–		
MDPREF	{ distinction	0.873	0.901	0.826	–	
	{ energy	0.856	0.932	0.863	–	
	{ micro components	0.956	0.871	0.755	–	
MDSCAL	{ distinction	0.617	0.619	0.632	0.662	–
	{ energy	0.847	0.908	0.925	0.895	–
	{ micro components	0.768	0.839	0.698	0.799	–

Table 4

Kendall's coefficient of concordance  $W$  with related statistics for perceived nutritional characteristics.

	$W$	$\chi^2_{14}$	$p$ -level
iron	0.418	877.48	0.00
vitamin A	0.159	332.93	0.00
vitamin C	0.234	491.04	0.00
carbohydrates	0.126	264.38	0.00
proteins	0.134	281.39	0.00

*tion*. The notion of distinction/festive of dimension 1 is supported by the fact that in the simulation study vegetables with a high value on dimension 1 tend to be eaten at the weekend and vegetables with a low value on weekdays. Dimension 2 is related to aspects such as real vegetable, heavy work, not for a side dish, carbohydrates, etc. We use the label *energy* for this dimension. Dimension 3 is closely related to vitamins and iron. We call this the *micro-components* dimension.

#### 4.3. Universal characteristics?

To examine how far respondents have similar perceptions with respect to the vegetables first the nutritional components iron, vitamin A, vitamin C, carbohydrates and proteins are

considered. Table 4 gives Kendall's coefficient of concordance  $W$  and related statistics indicating how much agreement there is with respect to the scores of the vegetables on nutritional components. There are individual differences with respect to the amounts of nutritional components attributed to the vegetables ( $W$  is considerably less than 1): However from the significance of  $W$  it can be concluded that there is a basic common pattern in the perceptions of the characteristics. This basic common pattern is represented by the average scores in the last five lines of table 1. From table 4 it can be concluded that agreement among respondents is highest with respect to iron and lowest with respect to carbohydrates. As a measure of agreement for the socio-psychological variables we defined the coefficient of agreement:

$A_{ij}$  = the number of vegetables on which respondent  $i$  and respondent  $j$  agree with respect to the score on a specific attribute.

Here agreement means: either both respondents think that the attribute statement (see table 1) is applicable to the vegetable or both respondents think that the statement is not applicable. The average coefficients of agree-

Table 5

Coefficients of agreement among respondents with respect to attribute scores (see text).

	$A$ (actual)	$\bar{A}$ (random)	maximum $A$ (random)	$A$ (actual)– $\bar{A}$ (random)
real vegetable	11.42	7.69	7.74	3.73
for those doing				
heavy work	10.49	7.66	7.74	2.83
well-known vegetable	11.90	10.12	10.15	1.78
high nutritional value	9.18	7.50	7.55	1.68
expensive	12.38	10.16	10.20	2.22
festive meal	11.51	9.09	9.13	2.42
a lot of				
preparation	10.99	9.89	9.93	1.10
for those with a				
high income	12.57	11.50	11.52	1.07
many ways to serve	9.72	7.99	8.03	1.73
for modern people	9.44	8.85	8.87	0.59
for a side dish	11.69	8.57	8.63	3.12

ment (averaged over all pairs of respondents) indicated as  $A$  (actual) are presented in the first column of table 5. Considering that the maximum value is fifteen (the number of vegetables) there seems to be a fair amount of agreement among respondents.  $A$ -random was obtained by simulation where the locations of ones and zeros were determined by chance: (100 runs per attribute), keeping the number of ones given by each individual fixed at its actual number. This constitutes a non-parametric test. From a comparison of  $A$  (actual) with  $A$  (random) it can be concluded that for  $\alpha = 0.01$  the hypothesis of no agreement is rejected.

Using the Tucker and Messick (1963) approach and carrying out INDCSAL-analyses according to pre-determined categories such as degree of urbanization, social class, age, working/not working and education, it turned out that no subgroups with perceptual dimensions, different from those mentioned earlier, could be found.

In conclusion, the results of this section show that there are differences among respondents with respect to the perceived attributes of vegetables but that there is also a significant common perceptual structure with three basic dimensions: distinction, energy and mi-

cro-components. So, as far as the basic perceptual configuration is concerned, it can be said that vegetables have universal characteristics.

#### 4.4. Relationship between perceived characteristics and physical attributes

Are the dimensions (characteristics) found also 'objective' as Lancaster assumes? Table 6 gives the matrix  $H$  that maps the physical features  $Z$  into perceived characteristics  $Y$ . The numbers are regression coefficients (betaweights) obtained by regressing the perceived characteristics on the objective characteristics of the vegetables, i.e. the true contents with respect to iron, vitamin  $A$ , vitamin  $C$ , carbohydrates and proteins. The figures for the true contents are obtained from the official tables of the Dutch Consumer Food Education Bureau, which are reprinted in many Dutch cookery books.

In the upper part of table 6 individual perceived characteristics (averages over respondents) are regressed on their physical counterparts ( $n = 15$ ). The lower part of table 6 gives the regression coefficients for the basic perceptual dimensions: distinction, energy and micro-components as found in the analysis.

Table 6

Estimated matrix  $H$  that maps the physical features  $Z$  into perceived characteristics (betaweights of regression coefficients).

	Z = physical features (true contents)					
	iron	vitamin <i>A</i>	vitamin <i>C</i>	carbo- hydrates	proteins	<i>R</i> <sup>2</sup>
<i>perceived contents of:</i>						
iron	0.805 **	0.112	0.037	− 0.082	− 0.196	0.73
vitamin <i>A</i>	0.021	0.848 **	0.111	− 0.059	− 0.078	0.71
vitamin <i>C</i>	0.311	0.516 **	0.208	− 0.175	− 0.218	0.60
carbohydrates	− 0.004	0.257	0.072	0.353	0.531 *	0.32
proteins	0.223	0.373	0.083	0.000	0.702 **	0.76
<i>dimensions of perceptual matrix Y:</i>						
distinction	− 0.135	0.029	0.333	− 0.108	0.632 **	0.44
energy	0.214	0.285	0.014	0.264	0.352	0.30
micro components	0.219	0.575 **	0.177	− 0.076	0.200	0.52

\* Significant at  $\alpha = 0.10$ .

\*\* Significant at  $\alpha = 0.05$ .

The betacoefficients in table 6 show that the perceived contents of iron, vitamin *A* and protein are well predicted by their physical counterparts (proteins were seen not to be a basic element in the perception of vegetables though). For vitamin *C* and carbohydrates there is not such a relationship: perceived vitamin *C* content is better predicted by the true vitamin *A* content than by the true vitamin *C* content and perceived carbohydrates are better predicted by true proteins than by true carbohydrates. The socio-psychological attribute distinction is to some extent predicted by true protein content (partly due to the high protein content of mushrooms). Perceived energy content is not related to any physical characteristic at all (although it has a correlation of 0.83 with *perceived* carbohydrate content, table 2). Perceived scores on micro-components are most closely related to true vitamin *A* content.

4.4.1. *Chunking of information*

Table 7 gives the correlation coefficients between perceived and corresponding true characteristics as well as within the group of

perceived and true characteristics respectively. The rectangular matrix (upper-right) confirms the results of table 6. A comparison of corresponding numbers from both triangular matrices shows an interesting phenomenon. For example, the correlation between the perceived content of vitamin *C* and the perceived content of iron is +0.80 whereas actually these two nutritional components are negatively related with  $r = -0.29$ . The correlation in the perceived matrix of vitamin *C* and vitamin *A* is as high as +0.89, the corresponding correlation coefficient in the true matrix is  $-0.20$ . So it seems that a phenomenon of generalization occurs: when a vegetable scores high on one micro-component the consumer thinks that this vegetable is also favourable as far as the other micro-components are concerned. The phenomenon can be considered as chunking of information: the consumer is not capable of memorizing all the different micro-component scores. Therefore, this information is stored together as a chunk: a vegetable is either good or bad on micro-components. In this situation vitamin *A* happens to be the anchor point for this informa-

Table 7  
Correlations among perceived and true characteristics.

	perceived					true				
	iron	vit. A	vit. C	carbo.	prot.	iron	vit. A	vit. C	carbo.	prot.
<i>perceived:</i>										
iron	1.00					0.82	0.55	-0.20	-0.38	0.06
vit. A	0.71	1.00				0.47	0.83	-0.09	-0.11	-0.04
vit. C	0.80	0.89	1.00			0.52	0.67	0.02	-0.25	-0.18
carbo.	0.21	0.20	-0.09	1.00		0.13	0.15	-0.02	0.22	0.39
prot.	0.49	0.48	0.30	0.36	1.00	0.61	0.37	-0.18	-0.28	0.71
<i>true:</i>										
iron						1.00				
vit. A						0.61	1.00			
vit. C						-0.29	-0.20	1.00		
carbo.						-0.42	-0.05	0.13	1.00	
prot.						0.32	-0.14	0.18	-0.25	1.00

tion. To some extent this also occurs with macro-components. According to table 7 perceived proteins are positively related to perceived carbohydrates ( $r = 0.36$ ), whereas in reality there is a negative correlation of  $r = -0.25$ .

Tables 6 and 7 demonstrate that the consumer has not the slightest idea about the true vitamin C content of vegetables. It is interesting that in the importance ratings given by the respondents with respect to the attributes of vegetables, vitamin C obtained the highest importance score: an average of 4.43 on a 5-point scale.

The conclusion is that although there are relationships between a number of perceived and physical characteristics, the mapping from the physical features into the basic perceived dimensions is far from complete, which is due to the limited information-processing capabilities of consumers. To apply the Lancaster model two ways are open now.

(i) Use the perceptual dimensions found as the 'characteristics' in the sense of the model, not concerned about the question whether these dimensions correspond with true physical properties or not: 'If everyone believes that snake oil has special medical properties, we would analyze behavior as though this were indeed true' (Lancaster 1971: 18).

This is the first path we will follow in the following sections. One important problem then is that we have not measured these dimensions on a ratio-scale, as is required in the Lancaster model. Therefore we will also follow a second route.

(ii) Take those physical characteristics, which correlate well with the perceptual dimensions and use these as the basic characteristics in the Lancaster model. These characteristics are: iron, vitamin A and proteins (table 6). Although we do not have a complete representation of the perceptual dimensions in this way, the measurements on these characteristics do satisfy the ratio scale requirement.

## 5. Efficient products

According to the Lancaster model the perceived matrix of consumption technology and the prices of the products determine whether or not a specific product is efficient. Product  $j$  is efficient if it is not dominated by one of the other products. To examine dominance, the maximum amounts of the characteristics that can be obtained when the budget is completely spent on product  $j$  is determined. If, by spending the same budget on a different product or combination of products, a combination of characteristics can be obtained which is not worse in any of the characteristics and better in at least one characteristics, it is said that such a product or combination of products dominates product  $j$ . One would expect that dominated products are not purchased. This can be verified for the vegetables case. In order to do this, we first have to take a closer look at the operationalization of the matrix of consumption technology.

### 5.1. Perceived matrix of consumption technology

The numbers in the matrix of consumption technology should indicate the amount of a characteristic per unit product on a ratio scale: distinction per kilogram endive, energy per kilogram endive, etc. The numbers on the perceived matrix of consumption technology  $Y$  (table 2a) were obtained through scores on attribute scales. Although measurement devices for obtaining ratio-scale perception may be developed in the future (Hauser and Simmie 1981: 42), as stated earlier in this paper we used a cruder attribute scale for which no interval-properties are assured. The scale used for the nutritional components – 'does not contain the nutritional components' (1) ... 'contains a lot of the nutritional component' (5) – may come close to a ratio-scale though. McAlister (1979) used a similar scale. To have only positive quantities of characteristics, the dimensions of the perceptual matrix in table

Table 8  
Standard units per vegetable and prices per standard unit.

Vegetable	Standard unit	Price per standard unit (Dfl.)
endive	kilogram	1.50
asparagus	500 gram	5.00
cauliflower	one ( = 750 gr.)	2.00
mushrooms	1 tray ( = 200 gr.)	2.00
cucumber	one ( = 500 gr.)	0.75
sweet pepper	one ( = 100 gr.)	1.00
leeks	kilogram	1.25
rhubarb	kilogram	1.75
red cabbage	kilogram	1.00
lettuce	1 head ( = 350 gr.)	0.50
French beans	kilogram	4.00
spinach	kilogram	1.50
onions	kilogram	0.75
white cabbage	kilogram	1.00
carrots	kilogram	1.00

2a have been scaled in such a way that per dimension the lowest value is zero. Especially for the dimension distinction this zero point is more or less arbitrary. A very similar procedure to determine a zero point as an attribute scale was followed by Hauser and Gaskin (1983). Apart from the assumption of a ratio scale in table 2a, we also have to decide to which quantity units of vegetables the numbers refer. The most natural choice is the unit in which a vegetable is usually bought. These 'standard units' differ per vegetable: for endive it is 1 kilogram, for mushrooms a tray of 200 gram, for lettuce one head, etc. These standard units are indicated in table 8. These quantities correspond to the numbers of units most often bought per purchasing occasion in the buying simulation (modus of the distribution). Although it was not explicitly stated in the interview we assumed that these are the quantities respondents had in mind when giving their attributes scores for the vegetables. So from now on we assume that the numbers of the perceived matrix of consumption technology (table 2a) are the quantities of characteristics per standard unit of a vegetable.

### 5.2. Dominated products

The prices stated in the purchase simula-

tion were expressed in the same standard units as referred to above.<sup>2</sup> These prices are given in the last column of table 8. To determine efficiency, in principle it is possible to all efficient goods combination (see figure 1). However, since this is a rather cumbersome operation, we restricted ourselves to the detection of domination by other products only and not by combinations of other products. This is a somewhat weaker analysis: a product may not be dominated by any of the other products and nevertheless not be efficient (product 3 in figure 1).

Table 9a gives the attainable combinations of characteristics levels for each of the fifteen vegetables, given a budget restriction of Dfl. 15,-. Domination is indicated by x-es in the right-hand part of table 9a. For example: endive is dominated by lettuce and carrots since these vegetables both permit higher levels of all characteristics than endive. Three groups of vegetables can be distinguished:

(a) vegetables that are not dominated at all: asparagus, sweet pepper, red cabbage and lettuce;

(b) vegetables that are dominated by a few (1 to 4) other vegetables;

(c) vegetables that are dominated by the majority of other vegetables: cucumber and rhubarb.

Both latter vegetables do not contain much of any of the characteristics (the reason why they are dominated). Their function to the meal is complementary; they are used to trim a dish but are practically never a substantial part of the meal. Another result is that lettuce dominates many of the other vegetables. This is in agreement with its wide use.

Of course one would not expect in an empirical study like this, dominated vegetables not to be purchased at all. If it is taken into account that perceptions are, strictly speaking heterogeneous, the efficiency frontier is not in

<sup>2</sup> With the exception of asparagus for which in the buying simulation the standard unit was 100 gram. The modal quantity bought, however, was 500 gram.



Table 9a

Dominance analysis: perceived characteristics.

Nr. Vegetable	Attainable combinations			Dominated by vegetable nr.:															Number of times dominated	Number of respondents buying the vegetable
	Distinction	Energy	Micro components	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
1. endive	1.30	18.20	9.10										X					X	2	115
2. asparagus	6.85	2.54	0.42																0	20
3. cauliflower	5.10	12.38	2.55										X						1	136
4. mushrooms	12.75	1.58	0.45																0	59
5. cucumber	0.16	0.24	0.00	X	X	X	X			X	X	X	X	X	X			X	11	114
6. sweet pepper	15.75	0.00	6.30																0	58
7. leeks	1.32	15.24	5.64										X		X			X	3	71
8. rhubarb	0.77	0.94	0.34	X	X	X	X			X		X	X	X	X				9	42
9. red cabbage	0.00	26.40	4.35																0	76
10. lettuce	5.40	24.00	23.40																0	147
11. French beans	2.25	6.68	2.21				X					X	X					X	4	103
12. spinach	2.30	18.70	8.90									X						X	2	106
13. onions	0.80	16.60	1.40	X								X	X					X	4	72
14. white cabbage	0.00	23.85	2.85									X							1	32
15. carrots	3.30	21.45	9.15									X							1	123

fact a line, but a region, determined by the distribution of the perception parameters in the population (Hauser and Simmie 1981: 42–44). So there will always be consumers for whom a specific vegetable is not dominated. However, when the efficiency frontier is taken as a central tendency, one would expect at least that a vegetable is purchased less the more often it is dominated by other vegetables. Looking at the last column of table 9a this does not seem to be the case: dominated and non-dominated vegetables are consumed in the same amounts.

One can speculate about the reasons for this. One reason might be that additional characteristics play a role in the determination of choice. This is a very real possibility since an important dimension has so far been left out: taste. Taste is a subjective attribute and could not be included in the common perceived matrix of consumption technology  $Y$ ; in the preference analysis taste will be included though. A second reason might be: satiation. The concept of efficiency is based on the assumption that all characteristics are of 'the more the better' type. However, when this is not the case: when there is an ideal

point beyond which further increases of a characteristic are disliked instead of liked, a consumer may not strive for as much of a characteristic as possible for the budget given. A third possible reason is the need for variety. We referred to the variety drive earlier. It can be mentioned here that one of the slogans in the campaigns of the Dutch Consumer Food Education Bureau is 'Make sure that there is enough variety in your meal'. The existence of satiation effects and variety seeking in the purchasing of vegetables will be examined in later sections of this paper.

The results of the dominance analysis depend on the zero points chosen for the attributes. Table 9a is based on the consumption technology matrix of table 3, where for each attribute the zero point is – more or less arbitrarily – set at the level of the vegetable with the lowest score. Sensitivity analyses have been carried out where the zero point was decreased with 0.1 and 0.5 respectively. The results are given in the second and third column of table 9b (the first column is the situation of table 9a.). In an alternative approach, the amounts of perceived characteristics of the consumption technology matrix

Table 9b

Dominance analysis for different specifications of attributes and zero points on attribute scales.

Nr. Vegetable	Number of times dominated: attainable levels computed for				
	amount of perceived char./guilder zero point at lowest vegetable	amount of perceived char./guilder zero point 0.1 lower	amount of perceived char./guilder zero point 0.5 lower	amount of perceived character + price	amount of physical char./ guilder
1. endive	2	2	2	0	1
2. asparagus	0	1	7	0	11
3. cauliflower	1	1	3	0	9
4. mushrooms	0	0	2	0	11
5. cucumber	6	3	1	1	8
6. sweet pepper	0	0	0	0	5
7. leeks	3	3	4	1	3
8. rhubarb	9	10	11	5	9
9. red cabbage	0	0	1	0	1
10. lettuce	0	0	0	0	1
11. French beans	4	5	10	0	5
12. spinach	2	2	2	0	0
13. onions	4	3	1	0	0
14. white cabbage	1	1	1	1	1
15. carrots	1	1	1	1	0

were taken as such (not divided by price) and the price was added as a separate characteristic for the determination of dominance. In this case only interval scale properties are required for the perceived attributes. Finally a dominance analysis was carried out with the amounts of physical characteristics iron, vitamin *A* and proteins as attributes. These characteristics are truly ratio-scaled. Table 9b shows that the results of the dominance analysis clearly depend on the specification of the characteristics and the way the attainable levels are computed.

All 15 vegetables are very common products in The Netherlands with considerable consumption volumes. Therefore the analysis with price as a separate variable has the greatest face validity: in this case the number of dominancies is lowest. This might indicate that consumers tend to look at characteristic levels and prices as such and not at characteristics per guilder. This would have important consequences with respect to the assumptions of the Lancaster model. From the first three columns of table 9b, it can be concluded that

there is a limited sensitivity of the results with respect to the choice of the zero point.

## 6. Utility functions

As indicated in figure 3, we have information about the respondents' preferences with respect to individual products. This information is in the form of pairwise preference statements. First it was examined per respondent how consistent she was in her preference statements. For this purpose for each respondent the coefficient of consistency ( $\gamma$ ) was computed, which is a function of the number of circular triads in the preferences: David (1963). This coefficient ranges from 0 (no consistency at all) to 1 (complete consistency). Our respondents appeared to be quite consistent:  $\gamma = 0.89$ ; st. dev. ( $\gamma$ ) = 0.12. Also the  $X^2$ -test, given by David (ibid.: 38) was applied. Only for one respondent could the null hypothesis of random preference statements not be rejected at the 0.01 level. These pairwise preferences were the input for the LIN-

MAP-program (Srinivasan and Shocker 1973), which was used to estimate the parameters of the preference functions. The configuration used in LINMAP was the perceived matrix of consumption technology (table 2a), with one attribute added: taste. This is a 0–1 variable: 1 = good taste; 0 = not good taste. This attribute of course is idiosyncratic per respondent.

Two specifications of the utility function were used: a linear specification (corresponding with the ORDREG version of LINMAP) and a function with the possibility of an ideal point per dimension. In case this ideal point goes to infinity, it is automatically replaced by a linear (vector) specification for that dimension. The latter specification corresponds to the MIXED-MODE version of LINMAP.

It has been observed (Green and Srinivasan 1978: 188) that sometimes consumer decision processes have a multistage character. For example, it may be that first the product alternatives are divided into two groups on the basis of one attribute and that subsequently other attributes are used to determine a preference order within these groups. If this is the case, the LINMAP-output shows only one attribute with a weight factor different from zero, the other weight factors are zero and the fit is perfect. In our analyses this very often happened with the attribute taste. Evidently consumers tend to form a preference order with all the vegetables of which they like the taste on top and all the vegetables of

which they do not like the taste in the lower part of their preference order.

When analysing the data with the ORDREG-version of LINMAP, 108 of the 150 respondents appeared to have a two-stage process. With LINMAP-MIXED MODE the number of 'two-stage respondents' was 122 (containing 105 of the 108 respondents who were also 'two-stage' in the ORDREG-version). To estimate the utility parameters for attributes other than taste, a second estimation was carried out for the two-stage respondents. Here the attribute taste was left out and the pairwise preference data used were only those referring to vegetables in the same set: the set of vegetables of which the taste is liked and the set of vegetables of which the taste is not liked. The pairwise preferences for vegetables not in the same set were treated as missing values.

The results for the linear utility model and the mixed utility model are given in table 10 and 11 respectively. The  $r_s$  in the right-hand column of both tables is the Spearman Rank correlation coefficient (root-mean-square over respondents), indicating the reproduction of the original preferences by the utility model. The fit level – about 0.80 – is quite satisfactory. For the mixed utility model the fit is about 2 points better than for the linear model. Looking at the linear model, the characteristic taste has the largest weight, followed by micro-components, distinction and energy in that order. Of course the weight for taste could

Table 10  
Estimation results: linear utility model.

		dim. 1 (distinction)	dim. 2 (energy)	dim. 3 (micro-components)	dim. 4 (taste)	$r_s$ (fit) (RMS)
one-stage respondents $N = 42$	mean					
	weight	0.43	0.25	0.47	0.73	0.788
	$\sigma$ weight	0.37	0.30	0.58	0.29	
two-stage respondents $N = 108$	mean					
	weight	0.18	0.15	0.53		0.792
	$\sigma$ weight	0.62	0.51	1.14		
all respondents						0.796

Table 11

Estimation results: mixed utility model.

			dim. 1 (distinction)	dim. 2 (energy)	dim. 3 (micro-components)	dim. 4 (taste)	$r_s$ (fit) (RMS)
one-stage model ( $N = 28$ )	vector	number of cases	15	7	26	18 **	0.783
	speci-	mean weight	-0.58 *	-0.45	-0.46	-0.73	
	fication	$\sigma$ weight	0.30	1.95	0.59	0.42	
	ideal	number of cases	13	21	2		
	point	mean i.p.	1.38	1.44	1.15		
	speci-	$\sigma$ i.p.	0.60	1.02	0.30		
	fication	mean weight	0.50	0.59	0.84		
		$\sigma$ weight	0.30	0.43	0.30		
two-stage model ( $N = 122$ )	vector	number of cases	56	47	103		0.837
	speci-	mean weight	-0.47	-0.04	-0.42		
	fication	$\sigma$ weight	0.55	0.52	0.94		
	ideal	number of cases	66	75	19		
	point	mean i.p.	1.18	1.33	-0.60		
	speci-	$\sigma$ i.p.	0.79	2.45	5.95		
	fication	mean weight	0.88	0.73	1.59		
		$\sigma$ weight	0.58	0.47	1.72		
all res- pondents ( $N = 150$ )	number of cases vector specification		71 (47%)	54 (36%)	129 (86%)	18 ** (100)%	0.827
	number of cases ideal point specification		79 (53%)	96 (64%)	21 (14%)	0 (0%)	

\* A negative weight in the vector specification of the mixed mode model implies: the larger the amount of the characteristics, the smaller the 'distance', i.e. the higher the preference.

\*\* The dimension taste was only relevant for 18 of the 28 'one-stage respondents', since 10 respondents had the same score for taste for all vegetables.

only be determined for one-stage respondents.

In the mixed utility model the number of respondents with an ideal point differs per characteristic: 53% for distinction, 64% for energy, 14% for micro-components and 0% for taste. Apparently there is an important difference between both nutritional components: micro-components are predominantly of 'the more the better' type whereas energy (carbohydrates) is seen by most respondents as a characteristic of which one can get too much. Also for many respondents distinction is only good up to a specific amount. As expected, taste is a 'the more the better' characteristic.

It is clear that satiation effects, as referred to in the last section are possible. From tables 10 and 11 it can also be concluded that re-

spondents differ greatly in their preference parameters: often the standard deviations are of the same order of magnitude as the mean parameter values. This is in agreement with the Lancaster theory where different consumers are assumed to have different utility functions.

## 7. Combinations of products bought

After having determined (i) the perceived matrix of consumption technology and (ii) the utility functions of individual respondents, this section ties together the various elements of the research for the ultimate goal of the study: explanation and prediction of combinations of items purchased by respondents (in the buying simulation).

### 7.1. Utility evaluation of combinations actually chosen

According to the Lancaster model, a consumer will choose that combination of vegetables that maximizes his utility. Therefore, a preliminary check of the validity of this model, as operationalized in this study, is to examine whether the utility of a combination actually bought by a consumer (i.e. in the buying simulation) is higher than the utilities of other combinations the respondents might buy with the same budget. For each respondent in principle an infinity of other combinations are possible; many of which would not make sense as a menu of vegetables for a week. We choose to consider as a set of sensible alternative combinations the combinations of vegetables actually bought by other respondents.

To make a utility evaluation of a given vegetable combination for a particular respondent, we proceeded as follows. Using the perceived matrix of consumption technology of table 2a expanded with a fourth idiosyncratic characteristic 'taste', the combination of vegetables, represented as a vector  $(x_1, \dots, x_{15})$  was transformed into levels of the characteristics: distinction, energy, micro-components and taste. For this four-dimensional characteristics vector  $(y_1, \dots, y_4)$  the utility is computed, using the function  $U(y)$  estimated in the last section. Two points have to be made with respect to the use of this utility function. First, for the respondents with a two-stage utility model no estimate for the weight factor of the characteristic taste in the utility model was obtained in the LIN-MAP-procedure. The basic characteristic of two-stage respondents is, that in their preference order all vegetables with a good taste have a higher position than all vegetables of which they do not like the taste. Using this, we devised an indirect estimate of the effect of a good taste in the utility function.

First the utilities of individual vegetables were determined using the utility model ob-

tained for the three characteristics: distinction, energy and microcomponents only. Then the quantity  $\Delta$  was computed:  $\Delta$  is the difference in utility between (i) the highest utility vegetable from the set: vegetables that do not have a good taste and (ii) the lowest utility vegetable from the set: vegetables with a good taste. Subsequently, a quantity was added to  $\Delta$  which is equal to the mean utility difference between consecutive vegetables (in utility order) in the set of 'good taste' as well as the set of 'not good taste' vegetables. Adding this 'good taste effect'  $\Delta$  to the utility value of a vegetable with a good taste ensures that such a vegetable will always be preferred above a 'not good taste' vegetable. This quantity was  $\Delta$  was used as the weight factor for taste ( $y_4$ ) in the utility model.

Secondly, in the mixed mode model ideal-points were estimated with respect to the levels of characteristics. However, these utility parameters were obtained from preferences for individual vegetables and not for menus of vegetables over a one-week period. It is assumed that the reference situation for the respondent was (implicitly) one serving, i.e. one meal. Of course, the ideal levels of the characteristics for a period of one week are higher than for one meal. Therefore, when transferring the utility function from individual products to combinations of products bought over a one-week period, the ideal points were multiplied by a factor of 7: the number of meals per week. In the linear model the absolute level of a characteristic does not matter, therefore we do not have to make similar adaptations there.

For each respondent we took the utility function of that particular respondent and used it to evaluate (i) the combination of vegetables actually chosen by the respondent and (ii) the combinations chosen by the 149 other respondents in the sample. To examine the validity of the model we counted per respondent the number of combinations with a higher utility value than the respondent's

own combination. This is called: number of dominating combinations. In a perfect situation the utility of the own combination would be higher than the utilities of all other combinations (number of dominating combinations = 0). If the model has no validity at all the utility of another combination would be as often above as below one's own combination. In that case the expected number of dominating combinations is 74.5. The result is, however, that over all respondents the average number of dominating combinations for the linear utility model is 51.5 with a standard deviation of 3.42. For this mixed model the average number is 65.4 with a standard deviation of 3.60. So for both models the number of dominating combinations is significantly lower than would be expected under random conditions. The conclusion is that the characteristics model as operationalized here has validity. It should be realized that with this analysis we are in the standard Lancaster model: i.e. with no variety seeking component included. It is surprising that the linear model performs better here than the mixed mode model, notwithstanding the somewhat better fit of the latter model in the last section. In the linear model a smaller number of parameters have to be estimated than in the mixed mode model. The vector model has one parameter per characteristic (the weight factor), the ideal point model two: weight factor and ideal point. For a limited number of alternatives this may affect the accuracy and stability of the estimates and produce better estimates for the linear model. Also an ideal point model may suffer a bias, due to the phenomenon of inventory of an attribute (McAlister 1982).

## 7.2. Optimal combinations of products

### 7.2.1. Standard model

According to the Lancaster model, in the vegetables situation a respondent chooses the combination of vegetables ( $x_1, \dots, x_{15}$ ) which

is the outcome of the mathematical program (2) to (5). With the perceived matrix of consumption technology used before and the individual utility function, per respondent the optimal combination can be found by solving this mathematical program. In the case of a linear model this is almost trivial. Since in such a model  $dU/dx_j$  is constant, in the optimal combination only one vegetable will be bought: the vegetable with the highest utility per guilder spent. In the case of the mixed mode model the utility function is quadratic and the program (2) to (5) can be solved by a quadratic programming code. For this purpose we used Beale's algorithm: Beale (1967). The interesting question is how close the optimal combination so computed is to the combination actually bought by a respondent. Five different statistics were used to compare optimal and actual purchase vectors: correlation coefficient, Theil-U statistic, average deviation between optimal and actual quantities,  $p$ -level of the hypergeometric distribution and a composite coefficient of agreement  $c$ . These statistics are described in the Appendix.

The results for the standard Lancaster model (linear and mixed model utility function, respectively) are given in the first two columns of table 12. The numbers are average statistics over the 150 respondents in the analysis sample. The standard deviations of these averages (not given in table 12) ranged from 0.01 to 0.04. Although there is a significant positive correlation between optimal and actual quantities (without any relationship  $r$  would have been 0), the agreement is far from perfect, especially when the other statistics are considered. Most notably the standard Lancaster model very much underpredicts the variety level in the vegetables combination. With a linear utility function the average number of different vegetables bought is 1.00, with a mixed mode utility function: 1.96, but the actual number is 8.49 (much higher than the number of different characteristics, as is

Table 12

Agreement between optimal and actual combinations of vegetables: averages over respondents (analysis sample).

	Standard Lancaster model		Variety seeking model	
	utility function linear	utility function mixed mode	utility part linear	utility part mixed mode
$r$ (Pearson) <sup>b</sup>	0.35 <sup>a</sup>	0.29	0.42	0.32
Theil – U <sup>c</sup>	0.82	0.74	0.47	0.49
AVDEV (average deviation) <sup>c</sup>	2.26	1.85	0.87	0.90
$p$ (hypergeome- tric) <sup>c</sup>	0.61	0.59	0.32	0.41
$c$ = composite criterion <sup>b</sup>	–1.72	–1.49	–0.66	–0.78
$w$ = weight factor in eq. (8)	1.00	1.00	0.25	0.20
Number of different vegetables <sup>d</sup>	1.00	1.96	8.00	8.51

<sup>a</sup> Numbers in table are averages over 150 respondents.<sup>b</sup> Larger (algebraically) values imply better fit.<sup>c</sup> Smaller values (absolute) imply better fit.<sup>d</sup> Actual number is 8.49.

predicted by the standard Lancaster model). A comparison of column 1 and column 2 of table 12 shows that for the standard Lancaster model the mixed mode utility function produces purchase vectors that are somewhat more in agreement with actual purchases than the linear utility function. This may be due to the phenomenon of satiation, implied by the mixed mode model.

### 7.2.2. Variety seeking model

To arrive at the optimal combination, for this model the mathematical program eq. (8) to eq. (11) has to be solved. For the variety function  $V(x)$  first the variance expression (16) was taken. Variance is a quadratic function and therefore the quadratic programming code, referred to earlier, could be used to find the solution. To determine  $w$ , the optimization process was carried out for different levels of  $w$  (with steps of 0.05), and we chose that value for which the average predicted number of different vegetables was closest to the actual number of 8.49. The same value of  $w$  was used for all respondents.

The results obtained with the variety seeking model are given in the third and fourth columns of table 12. It is clear that the optimal combinations obtained are much closer to the actual combinations than with the Standard Lancaster Model. The higher level of variation not only induces larger numbers of different vegetables, but also vectors of vegetables that are much more in agreement with actual vectors ( $r$ , Theil-U and AVDEV) and more often contain the right vegetables: see  $p$  (hyper geometric). The weight factor for the intrinsic utility part of the model is as low as 0.20 to 0.25, leaving for the variety part a weight of 0.75 to 0.80, which suggests an important role of variety seeking in the purchasing of vegetables. Here the linear model gives a better explanation than the mixed mode model. This raises the question whether satiation with characteristics really occurs: in the variety seeking model with a linear utility part the change-over with respect to vegetables is completely due to direct variation seeking and not to satiation.

A comment with respect to the choice of

the variety function is in order. Besides Variance ( $x$ ), which was used to obtain the results in table 12 and further tables, the use of the information measure of eq. (12) was also examined. For this purpose an iterative optimization program was written (based on the incremental search principle with the possibility of undoing earlier allocations). In a number of exploratory runs the resulting optimal purchase vectors were very similar to those obtained when Variance ( $x$ ) is used as the variety measure ( $r = 0.95$ ). Therefore, we no longer used the information measure in this study. It seems that the resulting purchase vector is not very sensitive to the exact specification of the variety measure.

### 7.2.3. Individual differences in explanation

The statistics in table 12 are averages over respondents. Of course the agreement between optimal and actual vegetable vectors differed among respondents. An attempt was made to relate different fit levels to different socio-demographic characteristics: age, education, social class and occupational status. No significant relationships were found. However, the results per respondent were consistent in the sense that respondents with a higher composite coefficient of agreement for one model generally have a higher coefficient of agreement for the other model and also have a lower number of dominating combinations as defined in section 7.1. The level of explanation of the vegetable combinations bought, is related to the quality of the preference information that served as input: the higher the coefficient of consistency in the pairwise preference judgments (see section 6), the higher the composite coefficient of agreement. Perhaps there is something like a basic consistency trait in individuals, not related to socio-demographic characteristics, which makes that some consumers are more able to give consistent preference judgments, and act accordingly, than other consumers. Another possibility is that some consumers were more

consistent simply because they were more involved in the experiment.

### 7.2.4. Reproduction of aggregate purchase vector

It is true, as the results of the last subsection demonstrate, that the variety seeking model is a closer approximation of consumer choice behavior than the standard Lancaster model at the level of the individual consumer. It might then be argued that if the purpose is to predict purchases at the aggregate level, i.e. the market, individual differences tend to cancel out and the prediction of the standard Lancaster model are as good as those of the variety seeking model. The upper half of table 13 gives the agreement statistics between the aggregate predicted vector of vegetables (corresponding with the various models) and the actual aggregate purchase vector. For a specific model the aggregate predicted vector is simply the sum of the 150 individual optimal purchase vectors with all elements subsequently divided by 150, i.e. it is an average vector. For the comparison of aggregate purchase vectors only the statistics  $r$  (Pearson), Theil-U and AVDEV are applicable.

The numbers in table 13 clearly show, that also when the purpose is to predict purchases at the aggregate level, the variety seeking model gives much better results than the standard model. Although some cancelling out of differences between respondents takes place with the standard model, the aggregate vector produced by the variety seeking model is much closer to the actual vector. When comparing table 12 with the upper half of table 13 a general phenomenon to be observed is the fact, that – also for the variety seeking model – the reproduction of the average vector is much better than the reproduction of individual purchase vectors. This is important since the marketer is ultimately interested in aggregate sales and not so much in sales to individual customers.

Another relevant question is: How im-



portant is to have utility functions of individual customers? What are the results if a common utility function is used and the optimal purchase factor is determined for the 'representative consumer'? The lower part of table 13 gives the results when a common utility function (derived from the overall preference of all 150 respondents) was used for the optimization. It is clear that the predictions using the representative consumer are considerably less precise than when individual utility functions are used. This applies even more to the variety seeking model than to the standard model. Of course, in practical applications it may not always be possible to carry out the analysis in all the details one would desire from a scientific point of view. It is important to note here the tendency in table 13 that it is more disastrous to eliminate the variety seeking part from the model, than to refrain from the estimation of individual utility functions and work instead with a 'representative consumer'.

### 7.2.5. Results for validation sample

As is often the case in empirical research, in this study the development and implementation of the model used to predict combinations of vegetables purchased by consumers has taken place in an interaction with the

Table 14

Agreement between optimal and actual combination of vegetables: averages over respondents (validation sample).

	Standard Lancaster model		Variety seeking model	
	Utility function linear	utility function mixed mode	utility part linear	utility part mixed mode
$r$ (Pearson)	0.34	0.30	0.40	0.27
Theil-U	0.82	0.75	0.48	0.52
AVDEV	2.23	1.83	0.88	0.95
$p$ (hypergeometric)	0.61	0.60	0.34	0.44
$c$ = composite criterion	-1.71	-1.48	-0.70	-0.86
$w$ = weight factor in Eq. (8)	1.00	1.00	0.25	0.20
number of different vegetables	1.00	1.98	8.04	8.11

data (i.e. the analysis sample). The perceived matrix of consumption technology was obtained from the same respondents for whom afterwards this matrix was used to predict purchases. Also, the weight factors of the variety seeking model ( $w$ ) were determined for a specific data base. Therefore, it is important to see how well the model predicts purchase vectors of consumers of which no information was used in the development in the model. For this purpose, the 150 respon-

Table 13

Reproduction of aggregate purchase vector.

		Standard Lancaster model		Variety seeking model	
		utility function linear	utility function mixed mode	utility part linear	utility part mixed mode
individual utility functions	$r$ (Pearson)	0.76	0.77	0.80	0.62
representative utility function ('representative consumer')	Theil-U	0.80	0.68	0.30	0.33
	AVDEV	1.93	1.51	0.49	0.56
	$r$ (Pearson)	0.77	0.30	0.70	0.37
	Theil - U	0.84	0.70	0.55	0.36
	AVDEV	2.43	1.58	1.23	0.60

dents of the validation sample were used. The only information from these respondents used to predict their combinations of vegetables are their (idiosyncratic) pairwise preferences for the fifteen vegetables in the study and their judgments with respect to the taste of these vegetables (a 0–1 variable). The agreement between optimal (model predictions) and actual combinations are given in table 14. From a comparison with table 12 it is clear that for these 'fresh data' the prediction results are practically as good as for the analysis data. For the best fitting model: the variety seeking model with a linear utility function; the decay in agreement is somewhat less than for the mixed mode specification.

## 8. Alternative approaches

### 8.1. Physical characteristics in the analysis

In the analysis up to now the perceived characteristics were used, where the perceptual matrix of table 2 was taken as the matrix of consumption technology. As announced before, because of the problem that these

perceived characteristics do not have ratio-scale properties, the complete analysis was also carried out with physical characteristics. Those physical characteristics (iron, vitamin A and proteins) were taken as 'proxys', which were clearly correlated with the perceived characteristics. As in the former analysis the idiosyncratic attribute taste was included as fourth attribute.

The fit of the utility function, using iron, vitamin A and proteins, was a fraction lower than with the perceived characteristics distinction, energy and micro-components in the utility function. The (root-mean-square) rank correlation coefficients between observed and reproduced preference over all respondents was 0.718 for the linear utility model and 0.752 for the mixed utility model. These numbers should be compared to 0.790 (table 10) and 0.827 (table 11), which are the corresponding fit coefficients when the perceived characteristics were used in the utility function.

However, the more important question is whether or not physical characteristics are a better means to explain actual combinations of vegetables bought. The answer is given by

Table 15  
Agreement between optimal and actual combinations of vegetables: physical instead of perceived characteristics.

	Standard Lancaster model		Variety seeking model	
	utility function linear	utility function mixed mode	utility part linear	utility part mixed mode
<i>r</i> (Pearson)	0.31	0.33	0.36	0.33
Theil – U	0.79	0.76	0.47	0.50
AVDEV (average deviation)	2.00	1.87	0.83	0.93
<i>p</i> (hypergeometric)	0.61	0.50	0.35	0.40
<i>c</i> = composite criterion	–1.58	–1.45	–0.68	–0.79
<i>w</i> = weight factor in eq. (8)	1.00	1.00	0.20	0.20
number of different vegetables	1.00	1.00	8.81	8.21

table 15, which can directly be compared to table 12. All corresponding numbers of both tables are in the same order of magnitude. So the conclusion is that using the physical instead of the perceived characteristics has no advantages. Apparently, the fact that these attributes have ratio-properties does not enough compensate for the disadvantage that the relationships between perceived and physical characteristics are rather weak (consumers know very little about true physical characteristics of food).

### 8.2. Equal expenditures model

The preceding sections have demonstrated that the capacity of the standard Lancaster model to explain purchased combinations of vegetables is very modest. In the extended model the variety component contributed much more to the explanation than the characteristics component. Given then that there are important difficulties to measure the perceived characteristics on a ratio-scale, a natural question is whether a model which only contains a variety component and does not include characteristics at all, would product a good fit. Such a model, which might be called 'naive' is a model which assumes that a con-

sumer purchases only vegetables rated as tasting 'good' and that he spends equal expenditures on all vegetables within this category.<sup>3</sup>

Purchases can easily be simulated with such a model and the purchase vectors can be compared to the actual ones. This was done here and the same statistics were used as before to compare optimal and actual vectors. The results, which are quite revealing, are presented in table 16. It turns out that this simple equal expenditures model gives a better fit than all the several versions of the Lancaster models dealt with before. This poses the question of the whole applicability and relevance of this Lancaster model and situations like we have here. This will be discussed in the last paragraph.

## 9. Conclusions and discussion

First the most important findings of this study will be summarized. Then we will discuss the implications, especially for the applicability of the Lancaster model.

### *Relevant characteristics.*

Two different types of data: subjective grouping as well as attribute scores and five different analysis techniques (decompositional as well as compositional): multidimensional scaling, factor analysis (two versions), discriminant analysis and MDPREF were used to find the most important characteristics consumers use in their judgment of vegetables. The results were unequivocal and consistent among the different approaches. Three dimensions emerged: distinction, energy and micro-components. Although there are differences in perception among respondents, there is a basic common perceptual structure. In that sense these three dimensions can be called 'universal characteristics'. Considerable differences were found between perceived characteristics and their physical counter-

Table 16  
Agreement between optimal and actual combinations of vegetables using different models.

	Equal Expenditures Model	Best fitting Extended Lancaster Model with perceived character	Best fitting Extended Lancaster Model with physical character
<i>r</i> (Pearson)	0.45	0.42	0.36
Theil – U	0.38	0.47	0.47
AVDEV	0.63	0.87	0.83
<i>p</i> (hypergeometric)	0.29	0.32	0.35
<i>c</i> = composite criterion	–0.46	–0.66	–0.68
number of different vegetables	10.43	8.00	8.81

<sup>3</sup> This very useful specification was suggested by a reviewer.

parts, most notably for vitamin C and carbohydrates. Due to limited capacity to process information, consumers simplify multidimensional product stimuli and 'chunk' different attributes together. Only the true physical levels of the nutritional components iron, vitamin A and proteins were significantly related to the perceptual dimensions found. Although the respondents find vitamin C the most important nutritional component in vegetables; the *true* vitamin C content is not related to one of the perceptual dimensions. (Whereas the *perceived* level of vitamins C has a correlation of 0.91 with the perceptual dimension micro-components.)

#### *Utility functions.*

From the preference data collected in the survey individual utility functions were estimated, using the LINMAP procedure. Two versions of the preference model were applied: linear utility model and mixed mode utility model (where respondents can have ideal points). The reproduction of the preference orders of the vegetables for the individual respondents was quite satisfactory. For the linear model the (root-mean-square) correlation coefficients between original and reproduced rank order is: 0.79, for the mixed mode: 0.83.

#### *Explanation of combinations of vegetables bought.*

When the utility functions and characteristic scores are known, the combination of vegetables that is optimal for a respondent, given his budget restriction, can be determined. If the Lancaster model is valid, this optimal combination should be close to the combination actually bought. With the Standard Lancaster model this fit is very low: especially the model severely underpredicts the number of different vegetables bought. Clearly there is a drive for variety in consumer vegetables buying which is not included in the Lancaster model. When the model is extended with a 'variety seeking module' the prediction of actual combinations

improves considerably. From the weight factor for this variety seeking model it is clear that variety seeking is much more important than the utility obtained from the characteristics. Finally a 'naive' Equal Expenditures Model was explored. The model assumes that a consumer allocates his total expenditure equally over those vegetables of which he likes the taste. So this model completely emphasizes variation in vegetables and does not pay any attention to the characteristics of the vegetable. The performance of this model in predicting purchased combinations of vegetables was superior to all the more sophisticated alternatives.

#### *Implications.*

The first weakness of the Lancaster model demonstrated by the model was the fact that perceived characteristics cannot directly be related to physical characteristics as Lancaster assumes. However, this is not too surprising, given our present knowledge of consumer information processes. A much more intriguing point is that, although the relevant characteristics of vegetables can be unequivocally determined, these characteristics play at most a very subordinate role in the choice of combinations of vegetables. Variety seeking, not included in the standard Lancaster model, is much more important.

Two possible explanations of this minor role of characteristics in the purchasing process can be speculated about. The first one is that the characteristics are important, but that their impact did not emerge because of measurement limitations in this study. As was discussed before, for the Lancaster model, the characteristics have to be measured at ratio level, i.e. with a fixed zero point. We did not have measurements of distinction, energy and micro-components of this type and for practical reasons have set the zero point of a characteristic at the level of the vegetable with the lowest score on scale for this characteristic. A theoretical justification for this procedure is the 'anchoring effect': consumers evaluate

products relative to the worst product along each dimension (Huber, Payne and Puto 1982). In one specific application Hauser and Gaskin (1983) found little sensitivity of market-share with respect to the chosen zero point. Because of this undefined zero point of the perceptual characteristics we also carried out the analysis with those physical characteristics (iron, vitamin A and proteins) that were correlated with the perceptual dimensions. With these characteristics, for which ratio-scale properties are assured, the predictive results of the (extended) Lancaster model did not improve. Nevertheless, in further studies it should be examined whether the recent progress in measurement methods: Hauser and Simmie (1981), Hauser and Shugan (1980), Moscovitz, Jacobs and Firtle (1980) can help to measure the scores on the perceptual characteristics better. Maybe in that case the role of characteristics in choice of combinations of products will show up more clearly. Also the use of utility functions estimated on individual products for the evaluation of product combinations, as was done in this study, has limitations. A comparison with method applied by McAlister (1979), where respondents give preference statements for combinations of products, seems useful.

A second explanation is that – although with better measurement methods the role of characteristics in the choice processes may become somewhat more pronounced – in a situation like vegetables the role of product characteristics is truly small. When multiple choices are made from a product class, with very short interpurchase times (one day) the consumer might not worry much about characteristic levels of a product bought at a single day. If, for example, the level of vitamin C in the vegetable I buy today would be low, this will be easily compensated by the vegetables I will eat tomorrow and in the other days to come. Believing that this averaging out will take place automatically, the consumer can limit himself to his striving for variety. In fact

this directly corresponds to one of the slogans of the Dutch Consumer Food Education Bureau.<sup>4</sup>

In contrast with this situation of buying multiple items in a product class is the situation of a 'one-shot' purchase. In this case, where the consumer has to 'stake everything in one throw' he will be more concerned with the characteristics level of the particular item chosen. Most research up to now has been devoted to the 'one-shot' situations. It would be interesting to see whether the phenomenon of the dominance of the variety drive also occurs in other product fields. For situations with much variety seeking (such as vegetables) the Lancaster Model might have very limited value in general.

It is hoped that the research reported here will stimulate further empirical work in this area. For the marketing of many products the interest is not so much in single purchases, but in combinations of products bought by consumers over a certain time period.

## Appendix

Statistics used to compare optimal and actual vectors of vegetables (computed per respondent).

- (1)  $r(\text{Pearson})$  = the (Pearson) coefficient of correlation between actual and optimal quantities of the vegetables (all quantities in standard units).
- (2) Theil-U =

$$\frac{\sqrt{\frac{1}{15} \sum_{j=1}^{15} (x_j^* - x_j)^2}}{\sqrt{\frac{1}{15} \sum_{j=1}^{15} x_j^{*2}} + \sqrt{\frac{1}{15} \sum_{j=1}^{15} x_j^2}} \quad (\text{A.1})$$

<sup>4</sup> The Consumer Information Brochure of this Bureau in the paragraph on vegetables says: 'Put variety in your choice: this is the way to get enough of all nutritional components'.

This 'inequality-coefficient' defined by Theil (1961: 32) lies between 0 (complete agreement between predicted and actual quantities) and 1 (no agreement at all between predicted and actual quantities).

- (3)  $AVDEV$  = the average deviation between optimal and actual quantity, computed over all 15 vegetables =

$$\left\{ \sum_{j=1}^{15} |x_j^* - x_j| \right\} / 15.$$

- (4)  $p(\text{hypergeometric})$  = the  $p$ -level of the hypergeometric distribution. Here the interest is in the extent to which the optimum set and the actual set contain the same vegetables. Let the actual set contain  $k$  vegetables and the optimal set  $n$ . Then, under conditions of random choice, the number of vegetables in the optimal set that are also in the actual set ( $= NBOTH$ ) is a hypergeometric variable with the probability distribution:

$$\text{prob}(NBOTH) = \frac{\binom{k}{NBOTH} \binom{15-k}{n-NBOTH}}{\binom{15}{n}}.$$

As a measure for the agreement between optimal and actual combinations we take the probability of observing – under conditions of random choice – a number of  $NBOTH$  equal to the number actually found or higher.

$$NBOTH \leq \min(k, n)$$

$$NBOTH \geq \max[0, n - (15 - k)]$$

- (5)  $c$  = composite criterion  
 $= 0.446 r$  (Pearson) –  $0.589$  Theil-U –  
 $0.492 AVDEV$   
 $= 0.456 p$  (hypergeometric)

$c$  is the principle component of the statistics (1) to (4), explaining 62% of their variance.

The agreement between optimal and actual purchase vectors is better the higher the value of  $r$  and  $c$  (algebraically) and

the lower the values of Theil-U,  $AVDEV$  and  $p(\text{hypergeometric})$ .

## References

- Beale, E.L.M. (1967), Numerical methods. In: J. Abadie (ed.), *Nonlinear programming*. Amsterdam: North Holland Publishing Co.
- Berlyne, D.E. (1960), *Conflict, arousal and curiosity*. New York: McGraw Hill.
- Cliff, N. (1966), Orthogonal rotation to congruence. *Psychometrika* 31 (March), 33–42.
- David, H.A. (1963), *The method of paired comparisons*. London: Griffin.
- Den Hartog, C. (1972), *Nieuwe voedingsleer*. Utrecht: Het Spectrum.
- Faison, E.W.J. (1977), The neglected variety drive. *Journal of Consumer Research* 4(3) (December) 1972–1975.
- Farquhar, P., V.R. Rao (1976), A balance model for evaluating subsets for multiattributed items. *Management Science* 22 (January), 528–539.
- Green, P.E., V. Srinivasan (1978), Conjoint analysis in consumer research: Issues and outlooks. *Journal of Consumer Research* (September), 103–123.
- Green, P.E., D.S. Tull (1978), *Research for marketing decisions* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Green, P.E., M.T. Devita (1975), An interaction model of consumer utility. *Journal of Consumer Research* Vol. 2 (September), 146–153.
- Hauser, J.R., F.S. Koppelman (1979), Alternative perceptual mapping techniques: Relative accuracy and usefulness. *Journal of Marketing Research* Vol. XVI (November), 495–506.
- Hauser, J.R., S.P. Gaskin (1983), *Application of the 'defender' consumer model*. Working Paper 1404–1483. Alfred P. Sloan School of Management, MIT.
- Hauser, J.R., S.M. Shugan (1980), Intensity measures of consumer preference. *Operations Research* 28(2), 278–320.
- Hauser, J.R., S.M. Shugan (1983), Defensive marketing strategies. *Marketing Science* 2(4) (Fall), 319–360.
- Hauser, J.R., P. Simmie (1981), Profit maximizing perceptual positions: An integrated theory for the selection of product features and price. *Management Science* 27 (January), 33–56.
- Howard, J.A. (1980), Promotion in a static market. In: K.B. Monroe (ed.), *Advances in Consumer Research* Vol. VIII, 9–11.
- Howard, J.A. and J.N. Sheth (1969), *The theory of buyer behavior*. New York: Wiley.
- Huber, J., M.B. Holbrook (1979), Using attribute ratings for product positioning: Some distinctions among composite approaches. *Journal of Marketing research* Vol. XVI (November), 507–516.
- Huber, J., J.W. Payne, C. Puto (1982), Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research* 9 (1) (June), 31–44.

- Jeuland, A.B. (1978), Brand preference over time: A partially deterministic operationalization of the notion of variety seeking. *AMA Educator's Proceedings*, Subashi Jain (ed.).
- Lancaster, K. (1966), A new approach to consumer theory. *Journal of Political Economy* 74 (April), 132–157.
- Lancaster, K. (1971), *Consumer demand: A new approach*. New York: Columbia University Press.
- McAlister, L. (1979), Choosing multiple items from a product class. *Journal of Consumer Research* 6 (December), 213–224.
- McAlister, L., E.A. Pessemier (1982), Variety seeking behavior: An interdisciplinary review. *Journal of Consumer Research* 9(2) (December), 311–322.
- Moskowitz, H.R., B. Jacobs, N. Firtle (1980), Discrimination testing and product decisions. *Journal of Marketing Research* Vol. XVII (February), 84–90.
- Nicosia, F.M. (1974), Toward and empirical theory of consumer behaviour based on the economics of goods-characteristics (Book Review). *Journal of Marketing Research* Vol. XI (February), 115–118.
- Pessemier, E.A. (1981), *Varied consumer behavior: Some theory and measurement methods*. Working Paper Purdue University.
- Ratchford, B.T. (1975), The new economic theory of consumer behaviour: An interpretative essay. *Journal of Consumer Research* 2 (September), 65–75.
- Ratchford, B.T. (1979), Operationalizing, economic models of demand for product characteristics. *Journal of Consumer Research* (June), 76–85.
- Ratchford, B.T. (1980), The value of information for selected appliances. *Journal of Marketing Research* 17 (February), 14–25.
- Ryans, A.B. (1974), Estimating consumer preference for a new durable brand in an established product class. *Journal of Marketing Research* Vol. XI (November), 434–443.
- Srinivasan, V., A.D. Shocker (1973), Linear programming techniques for multidimensional analysis of preferences. *Psychometrika* 38 (September), 337–369.
- Theil, H. (1961), *Economic forecasts and policy*. (2nd ed.) Amsterdam: North Holland Publishing Co.
- Tucker, L.R., S. Messick (1963), An individual differences model for multidimensional scaling. *Psychometrika* 28 (December), 333–367.
- Wierenga, B. (1974), *An investigation of brand choice processes*, Rotterdam/London: Rotterdam University Press/Gower Press.