



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA DE ALIMENTOS
DEPARTAMENTO DE ALIMENTOS E NUTRIÇÃO

**AVALIAÇÃO DO DESEMPENHO DE QUATRO MÉTODOS DE
ESCALONAMENTO EM TESTES SENSORIAIS DE ACEITAÇÃO
UTILIZANDO MODELOS NORMAIS ADITIVOS DE ANÁLISE DA
VARIÂNCIA E MAPAS INTERNOS DE PREFERÊNCIA**

PARECER

Este exemplar corresponde à redação final da tese defendida por **Nilda Doris Montes Villanueva**, aprovada pela Comissão Julgadora em 31 de julho de 2003.

NILDA DORIS MONTES VILLANUEVA
Bacharel em Estatística
Mestre em Estatística

Campinas, 31 de julho de 2003.


Prof. Dra. Maria Aparecida Azevedo
Pereira da Silva
Presidente da Banca

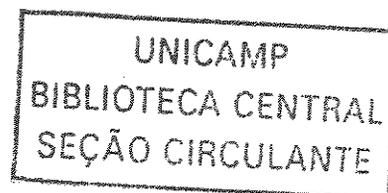
Orientadora: Profa. Dra. MARIA APARECIDA A. PEREIRA DA SILVA

Co-orientador: Prof. Dr. ADEMIR JOSÉ PETENATE

Tese apresentada à Faculdade de Engenharia de Alimentos da Universidade Estadual de Campinas, para obtenção do título de Doutor em Alimentos e Nutrição

Campinas - SP
2003

UNICAMP
BIBLIOTECA CENTRAL



UNIDADE	BC
Nº CHAMADA	Unicamp M765a
V	EX
TOMBO BCI	56164
PROC.	16-12-4103
C	<input type="checkbox"/>
D	<input checked="" type="checkbox"/>
PREÇO	R\$ 11,00
DATA	11/10/03
Nº CPD	

CM00190401-7

Bib id 308186

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA F.E.A. – UNICAMP

M765a Montes Villanueva, Nilda Doris
Avaliação do desempenho de quatro métodos de escalonamento em testes sensoriais de aceitação utilizando modelos normais aditivos de análise da variância e mapas internos de preferência / Nilda Doris Montes Villanueva. – Campinas, SP: [s.n.], 2003.

Orientador: Maria Aparecida Azevedo Pereira da Silva
Co-orientador: Ademir José Petenate
Tese (doutorado) – Universidade Estadual de Campinas.
Faculdade de Engenharia de Alimentos.

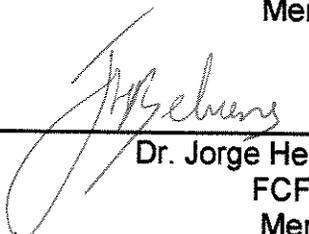
1. Avaliação sensorial – Métodos estatísticos.
2. Consumidores. 3. Análise de variância. 4. Psicometria.
I. Silva, Maria Aparecida Azevedo Pereira da. II. Petenate, Ademir José. III. Universidade Estadual de Campinas. Faculdade de Engenharia de Alimentos. IV. Título.

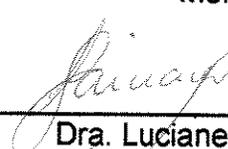
BANCA EXAMINADORA


Dra. Maria Aparecida Azevedo Pereira da Silva
FEA/UNICAMP
Orientadora


Dr. Armando Mário Infante
IMECC/UNICAMP
Membro


Dra. Eliete Vaz de Faria
LAFISE/ITAL
Membro


Dr. Jorge Herman Berhens
FCF/USP
Membro


Dra. Luciane Vieira Garcia
DFMV/UMESP
Membro

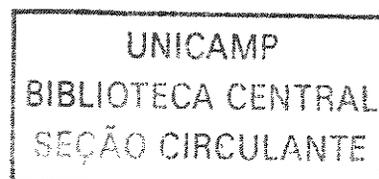
Dra. Helena Maria André Bolini Cardello
FEA/UNICAMP
Suplente

Dra. Rosires Deliza
EMBRAPA
Suplente

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

***Ao Deus da paz, do amor, da vida e da esperança.
Ofereço.***

***Aos meus queridos pais, Ignacio e Gudelia.
Aos meus 9 irmãos e 20 sobrinhos.
Dedico.***



**SEM TI, Ó CRISTO,
SERIA UMA POBRE CRIATURA FINITA.
S. Gregório Nazianzeno, Séc. IV**

O mundo seria um repetir-se sem fim e sem sentido.
Viver para que? Trabalhar para que? Conquistar para que?
Descobrir para que? Inventar para que? Pesquisar para que?
Construir para que? Ajuntar para que?
Fazer música, arte e poesia para que?
Procriar filhos para que?
Eu seria condenado ao nada.

O homem nunca cessa de buscar, seja quando é marcado pelo drama da violência, da solidão e da insignificância, seja quando vive na alegria e serenidade, essa busca não cessa jamais. A única resposta que o pode apaziguar nessa procura é aquilo que lhe vem do encontro com Aquele que está na fonte do seu ser e fazer. O caminho é Cristo. Ele é o Caminho, a Verdade e a Vida, que chega a cada pessoa no cotidiano de sua existência. A descoberta desse caminho normalmente nos vem graças à mediação de outros seres humanos. O Cristianismo, antes de ser um conjunto de doutrinas, ou uma regra de salvação, é portanto o “acontecimento” de um encontro.

João Paulo II

Sem o reconhecimento do Mistério presente, a noite avança, a confusão avança e - como tal, no nível da liberdade - a revolta avança, ou a decepção preenche de tal forma a medida que é como se não se esperasse mais nada e se vive sem desejar mais nada. Desde o mistério da Ressurreição de Cristo, uma nova luz invade o mundo e disputa palmo a palmo o terreno com a noite.

Luigi Giussani

AGRADECIMENTOS

Ao Deus vivo e misericordioso doador da vida e fonte de alegria, beleza e paz, fonte de justiça, liberdade e verdade, fonte de graça, esperança e consolo, fonte de amor, perdão e reconciliação, fonte de sabedoria e luz, fonte de comunhão, partilha e amizade fraterna, fonte de plenitude, felicidade e vida eterna.

À Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP), pela concessão da bolsa de doutorado e pelo apoio financeiro sem o qual não teria sido possível a realização deste projeto de pesquisa.

Ao Departamento de Nutrição e Alimentos da FEA – UNICAMP, pela oportunidade de realização deste trabalho.

À minha orientadora Dra. Maria Aparecida Azevedo Pereira da Silva, pelos conhecimentos transmitidos, pelo contínuo incentivo, pelo apoio e dedicação oferecidos e pela competente orientação na construção deste trabalho.

Ao meu co-orientador Dr. Ademir José Petenate, pela disponibilidade e atenção dispensadas, pelas dúvidas esclarecidas e valiosas sugestões que ajudaram a dar maior consistência a este trabalho.

Ao Dr. Armando Mario Infante, pela presença cordial e amigável, pela ajuda sempre oferecida, pelo incentivo, aprendizado, sugestões e contribuições realizadas e, principalmente, por ter ajudado a descobrir o meu caminho profissional, me encorajando e animando a trabalhar numa área aplicada.

Aos integrantes da banca examinadora: Dr. Armando Mario Infante, Dra. Eliete Vaz de Faria, Dra. Helena Maria Cardello, Dr. Jorge Herman Behrens, Dra. Luciane Vieira Garcia e Dra. Rosires Deliza, pelas correções, observações, sugestões e valiosas contribuições na elaboração final deste trabalho.

Ao Dr. Alberto Miele e à equipe de avaliadores especialistas do Centro Nacional de Pesquisa em Uva e Vinho da EMBRAPA, Bento Gonçalves –RS, pela disponibilidade e colaboração na realização dos testes de degustação e pelos importantes comentários e observações.

À minha numerosa e abençoada família: meus queridos pais Ignacio e Gudelia, aos meus irmãos(as), sobrinhos(as), primos(as), tios(as) e cunhados(as). Embora tão longe, se mantiveram sempre perto, me incentivando e apoiando, me encorajando e oferecendo o seu costumeiro carinho.

À minha amiga de todas as horas Stela Maris Rinke, pelo seu apoio constantes, pelas suas mãos sempre estendidas, pela sua bondade, acolhimento e amizade que fazem que ela seja um sinal concreto e visível da presença amorosa e terna de Deus na minha vida.

À minha comunidade de Comunhão e Libertação de Campinas, em especial a Hana, Rinke, Bete e César, Paulo, Suzie e Ester, por me oferecerem seu carinho e sua amizade com tanta simplicidade e gratuidade.

À minha Fraternidade São José, por ser Companhia Vocacional e lugar de amizade com Cristo, em especial às amigas Solange, Lucia e Marcinha, por me acolherem sempre de coração aberto.

Ao Pe. Vando Valentini, pela sua presença paternal e amorosa em todos estes anos da minha vida, pela sua paciência, compreensão e acolhimento, pelas suas indicações e orientações claras nos momentos de indecisão e dúvida, pelo conforto e ânimo nas horas difíceis.

A todos os amigos da república, de modo especial ao Elcio, Rogério, Lidermi e Nilde, pela partilha, pelos momentos agradáveis de convívio, pela ajuda mutua e pela sincera amizade.

Aos meus amigos de sempre, com imensa gratidão e carinho: Doris, Emma, Marina, José e Elizabeth, Andrés e Flavia, Dener e Bete, Maria Elena e Célio,

Serginho e Silvia, Germano, Fabitcho, Jorge Mendez, Ana Paula, Patrícia, Eliane, Lucy, Luciane e Maria Fachim.

À Norita, “o meu braço direito”, pela sua valiosa ajuda na realização dos experimentos, pela sua disponibilidade e amizade, pelos momentos compartilhados: de entusiasmo ou desânimo, de fadiga ou descanso, de serenidade ou atribulação.

Aos amigos do Laboratório de Análise Sensorial: em especial a Selma, Regina, Jorge, Alessandra, Rosa e Kátia, pelo carinho e companheirismo, pela solidariedade e amizade que fizeram do laboratório um ambiente familiar e alegre.

À Eliete e Nice (Lab. de Análise Sensorial), à Cidinha e Eduardo (Secretaria do DEPAN), à Fátima (Sala de Micro), ao Estefani, Marçal e Paulo (Centro de Informática), ao Cosme (Secretaria da Pós-graduação) e a todos os funcionários da Biblioteca, pela gentil disponibilidade, habitual apoio e muitos serviços oferecidos.

Aos provadores que participaram dos testes sensoriais, pela boa vontade, entusiasmo, colaboração, paciência e disponibilidade na realização dos testes, que fizeram possível a concretização deste trabalho de pesquisa.

Finalmente, desejo manifestar a minha eterna gratidão por tudo e por todos aqueles que direta ou indiretamente contribuíram para a realização deste trabalho.

ÍNDICE GERAL

RESUMO GERAL	xiii
GENERAL ABSTRACT	xvii
INTRODUÇÃO GERAL	1
CAPÍTULO 1: REVISÃO BIBLIOGRÁFICA	7
1. ESCALAS AFETIVAS E SUAS PROBLEMÁTICAS.....	7
2. ESTUDOS COMPARATIVOS SOBRE DESEMPENHO DE ESCALAS.....	15
3. A PROBLEMÁTICA DA ANÁLISE ESTATÍSTICA DE DADOS COLETADOS ATRAVÉS DE ESCALAS.....	23
4. REFERÊNCIAS BIBLIOGRÁFICAS.....	26
CAPÍTULO 2: AVALIAÇÃO DO DESEMPENHO DE TRÊS MÉTODOS DE ESCALA E DIAGNÓSTICO DO MODELO ANOVA USUAL ADITIVO EM ESTUDOS COM CONSUMIDORES	29
Resumo.....	29
Abstract.....	30
1. INTRODUÇÃO.....	30
2. MATERIAIS E MÉTODOS.....	33
2.1 Produtos.....	33
2.2 Provedores.....	33
2.3.Delineamento Experimental e Modelo Estatístico	33
2.4 Métodos Sensoriais Afetivos	34
2.5 Análise Estatística.....	37
3. RESULTADOS.....	38
3.1 Escala Autoajustável x Escala hedônica.....	38
3.1.1 Análise exploratória dos dados.....	38
3.1.2 Efeito de ordem de apresentação das amostras e efeito residual.....	43
3.1.3 Análise do poder discriminativo.....	45
3.1.4 Diagnóstico da normalidade dos resíduos do modelo ANOVA.....	46
3.1.5 Diagnóstico da homoscedasticidade.....	49
3.1.6 Variação da expansibilidade.....	49
3.2 Escala de Ordenação x Escala Hedônica x Escala Autoajustável.....	51

4. CONCLUSÕES.....	53
5. REFERÊNCIAS BIBLIOGRÁFICAS.....	55
CAPÍTULO 3: ESCALA HEDÔNICA HÍBRIDA: VANTAGENS EM RELAÇÃO ÀS ESCALAS HEDÔNICA TRADICIONAL, AUTOAJUSTÁVEL E DE ORDENAÇÃO	57
Resumo.....	57
Abstract.....	58
1. INTRODUÇÃO.....	60
2. OBJETIVOS.....	62
3. MATERIAIS E MÉTODOS.....	62
3.1 Amostras.....	62
3.2 Provadores.....	63
3.3 Delineamento Experimental e Modelo Estatístico.....	63
3.4 Métodos sensoriais afetivos.....	64
3.5 Análise estatística.....	67
4. RESULTADOS E DISCUSSÃO.....	69
4.1 Escala Autoajustável x Escala Hedônica Estruturada x Escala Híbrida	69
4.1.1 Análise exploratória dos dados.....	69
4.1.2 Efeito de ordem de apresentação das amostras e efeito residual.....	72
4.1.3 Poder discriminativo	75
4.1.3.1 Uso do método “bootstrap” para análise do poder discriminativo.....	76
4.1.4 Diagnóstico da normalidade dos resíduos do modelo ANOVA	78
4.1.5 Diagnóstico da homoscedasticidade	79
4.1.6 Variação da expansibilidade	79
4.1.7 Diagnóstico da aditividade do modelo.....	80
4.1.8 Facilidade de uso.....	83
4.2 Análise não paramétrica dos dados de ordenação das 4 escalas.....	84
5. CONCLUSÕES.....	86
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	87
CAPÍTULO 4: DESEMPENHO DA ESCALA HEDÔNICA HÍBRIDA E AUTOAJUSTÁVEL NA GERAÇÃO DE MAPAS INTERNOS DE PREFERÊNCIA	89
Resumo.....	89

Abstract.....	90
1. INTRODUÇÃO.....	91
2. OBJETIVOS.....	94
3. MATERIAIS E MÉTODOS.....	94
3.1 Produtos.....	94
3.2 Consumidores.....	96
3.3 Delineamento experimental e modelo estatístico.....	97
3.4 Métodos sensoriais afetivos.....	98
3.5 Análise estatística.....	101
4. RESULTADOS E DISCUSSÃO.....	102
4.1 Caracterização dos consumidores segundo seus hábitos de consumo de vinho tinto seco.....	102
4.2 ANOVA e Mapa Interno de Preferência.....	105
4.2.1 Efeito da ordem de apresentação das amostras e efeito residual.....	107
4.2.2 Análise dos Mapas Internos de Preferência.....	108
4.2.3 Perfil de preferência e caracterização sócio-demográfica de cada segmento de consumidores.....	120
5. CONCLUSÕES.....	128
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	130
CONCLUSÕES GERAIS.....	133
APÊNDICE.....	135

RESUMO GERAL

Em testes sensoriais, a análise dos dados geralmente é realizada através de algum modelo ANOVA. Estes modelos pressupõem que as respostas experimentais sejam: i) independentes, ii) normalmente distribuídas, iii) homoscedásticas (variâncias iguais) e, iv) provenientes de uma mesma escala de medida (aditividade). Os principais problemas na análise de dados sensoriais através de modelos ANOVA referem-se aos dois últimos pressupostos. A homogeneidade das variâncias não pode ser assegurada devido à existência de pelo menos duas fontes potenciais de variabilidade dos dados, quais sejam: provadores e tratamentos. Por outro lado, a aditividade pode ser violada quando um provador utiliza faixas consistentemente mais (ou menos) amplas da escala para expressar a sua impressão sobre o produto. A maneira pessoal com que cada provador utiliza a escala para avaliar os produtos, chama-se de variação da expansibilidade entre provadores. Tanto a falta de homogeneidade das variâncias como a não aditividade do modelo, acarretam conseqüências sérias na obtenção do verdadeiro nível de significância para o efeito dos tratamentos, podendo afetar adversamente as comparações entre as médias dos tratamentos e comprometer seriamente tanto a interpretação dos resultados fornecidos pelo experimento como a validade do modelo ANOVA.

Em testes com consumidores, escalas tradicionais como a escala hedônica de 9 pontos freqüentemente apresentam a seguinte problemática: i) geram dados que freqüentemente não satisfazem os pressupostos estatísticos de normalidade, aditividade e homoscedasticidade exigidos nos modelos ANOVA, ii) oferecem pouca liberdade aos consumidores para expressarem suas percepções sensoriais, devido ao limitado número de categorias, iii) induzem efeitos numéricos e contextuais no julgamento dos provadores e, iv) os valores numéricos associados às suas categorias, embora numericamente possuam intervalos iguais, não refletem iguais diferenças em percepção.

Das metodologias utilizadas em testes sensoriais com consumidores, a escala hedônica de 9 pontos, é sem dúvida, a mais utilizada. Porém, em função da problemática anteriormente mencionada, surge a necessidade de serem pesquisadas escalas alternativas que possuam um melhor desempenho que a escala hedônica tradicional, tanto quando os dados são analisados através de modelos ANOVA como quando os mesmos são analisados através de métodos multivariados como Mapa Interno de Preferência - MDPREF. De um modo geral, o objetivo do presente trabalho foi pesquisar o desempenho de duas escalas alternativas em estudos com consumidores, quais sejam: escala autoajustável e escala hedônica híbrida, comparando-as com métodos afetivos tradicionais como a escala de ordenação e escala hedônica de 9 pontos. Para isso, três experimentos foram realizados conforme descrito a seguir:

O primeiro experimento foi realizado com o objetivo de se avaliar em condições reais de teste de consumidor, o desempenho da escala autoajustável em relação à escala hedônica de 9 pontos e escala de ordenação, utilizando-se os seguintes critérios: i) diferenças em expansibilidade entre provadores, ii) poder discriminativo e, iii) adequação dos dados coletados por cada escala aos pressupostos do modelo ANOVA. Três marcas comerciais de confeitos foram avaliadas por 288 consumidores. Os resultados obtidos através das escalas hedônica de 9 pontos e autoajustável foram analisados através de ANOVA e os resultados da escala de ordenação, através do teste de Friedman. Os valores de $pF_{amostra}$, $pF_{provador}$ e $QM_{resíduo}$ fornecidos pela ANOVA de cada escala, foram respectivamente utilizados para avaliar o poder discriminativo, a expansibilidade dos provadores e a variabilidade residual dos dados. Teste de Tukey foi também aplicado para análise do poder discriminativo de cada escala. A normalidade dos dados foi verificada através do cálculo dos coeficientes de assimetria e curtose, gráfico de probabilidade normal e teste de Kolmogorov-Smirnov. A homoscedasticidade, foi avaliada através de gráficos de dispersão e teste de Levene. Os resultados mostraram que a escala autoajustável foi efetiva para tratar o problema da expansibilidade entre provadores e da desigualdade das variâncias,

porém, os resíduos mostraram moderados desvios da normalidade. A escala hedônica de 9 pontos apresentou problemas de heteroscedasticidade. As escalas autoajustável e de ordenação apresentaram o menor e o maior poder discriminativo respectivamente. Apesar dos problemas detectados, as três escalas apresentaram as mesmas tendências de preferência dos produtos avaliados.

O segundo experimento foi realizado com o objetivo de se avaliar o desempenho da escala hedônica híbrida em estudos com consumidores, comparando-a à escala hedônica de 9 pontos, escala autoajustável, e escala de ordenação; através dos seguintes critérios: i) variabilidade das respostas sensoriais, ii) poder discriminativo, iii) adequação dos dados às suposições dos modelos ANOVA e, iv) facilidade de uso pelos consumidores. Cinco marcas de suco de laranja foram avaliadas por 80 consumidores, divididos em quatro grupos de 20 indivíduos cada. Todos os indivíduos avaliaram todas as amostras através de todas as escalas em 4 diferentes sessões de degustação. Um delineamento em quadrado latino 4x4, foi utilizado para controlar o efeito de ordem de apresentação das escalas e avaliar sem vícios a facilidade de uso das mesmas. Para cada escala, a ordem de apresentação das amostras e efeitos residuais ("carry-over") foram balanceados. Os resultados obtidos através das escalas hedônica tradicional, híbrida e autoajustável foram avaliados através de ANOVA. A normalidade dos dados foi verificada através do teste de Shapiro-Wilks, a homoscedasticidade através do teste de Brown-Forsythe e a aditividade, através do teste de Tukey para um grau de liberdade. Os valores de $pF_{amostra}$, $pF_{providor}$ e $QM_{resíduo}$ fornecidos pela ANOVA de cada escala, foram respectivamente utilizados para avaliar o poder discriminativo, a expansibilidade dos provedores e a variabilidade residual dos dados. O teste de REGWQ foi também aplicado para análise do poder discriminativo de cada escala. Os resultados obtidos através da escala de ordenação foram avaliados pelo teste de Friedman e, a facilidade de uso das escalas por testes de Cochran-Mantel-Haenszel. Os resultados sugeriram uma superioridade da escala hedônica híbrida sobre as escalas hedônica estruturada e autoajustável, tanto em função do poder discriminativo como da adequação dos dados às suposições de normalidade e homoscedasticidade. A

despite dos dados da escala autoajustável terem apresentado maior variabilidade e sérios desvios da normalidade dos resíduos, o poder discriminativo desta escala foi ligeiramente superior ao da escala hedônica estruturada. A escala de ordenação apresentou o menor poder discriminativo em relação às demais. As escalas hedônicas estruturada e híbrida foram consideradas significativamente ($p \leq 0,01$) mais fáceis de serem utilizadas que a autoajustável, não havendo diferença ($p \leq 0,05$) entre as duas primeiras.

Finalmente, o objetivo do terceiro experimento foi avaliar o desempenho das escalas hedônica estruturada, hedônica híbrida e autoajustável na construção de Mapas Internos de Preferência - MDPREF. Nesta pesquisa, a aceitação global de 8 marcas comerciais de vinho tinto, a maioria deles varietal Cabernet Sauvignon, foi avaliada por 112 consumidores. Foram utilizados delineamentos experimentais balanceados para ordem de apresentação das escalas, ordem de apresentação das amostras e efeitos residuais. Os dados foram analisados através de ANOVA e MDPREF. O critério de avaliação do desempenho da cada escala baseou-se no número de consumidores significativamente ajustados ($p \leq 0,05$) e no grau de segmentação dos produtos e dos consumidores produzidos pelo MDPREF. Os resultados sugeriram uma superioridade da escala híbrida sobre a escala hedônica tradicional e autoajustável. O MDPREF gerado pelos dados da escala híbrida produziu um maior número de dimensões significativas de preferência ($p \leq 0,05$), trazendo como decorrência, uma porcentagem de 79,5% consumidores significativamente ajustados ($p \leq 0,05$), enquanto a escala autoajustável ajustou 54,5% dos consumidores e a escala hedônica 51,8%. Em geral a escala hedônica de 9 pontos apresentou um desempenho inferior ao das demais escalas.

Os resultados do presente estudo sugerem fortemente que a escala hedônica híbrida é uma ferramenta válida e eficiente que pode ser utilizada na coleta de dados associados a estudos com consumidores, tanto quando eles forem analisados através de modelos normais para análise da variância como através da metodologia de Mapa Interno de Preferência.

GENERAL ABSTRACT

In sensory tests, the basic statistical tool for analyzing data is almost invariably some sort of analysis of variance models. These models presuppose that the experimental responses are: i) independent, ii) normally distributed, iii) homoscedastic (have equal variances) and, iv) scores are on the same scale of measurement (additivity). The main problems arising from the analysis of sensory data using ANOVA models are related to the last two assumptions. Homogeneity of error variance is not assured, especially as there are at least two potential sources of heterogeneity: treatments and assessors. On the other hand, the additivity could be violated if one assessor used a consistently larger (or smaller) portion of the scale range, scoring more (or less) expansively than other assessors to express his opinion of the product. The individual way in which each panelist uses the scale to evaluate the products is known as the differential expansiveness of scoring between assessors. Both the lack of homogeneity of the variances and the non-additivity of the model, result in serious consequences in obtaining a true level of significance for the effect of the treatments and may adversely affect the comparison of treatment means. The non-additivity can seriously affect and possibly invalidate the analysis of variance and the interpretation of the results that it provides.

In consumer tests, traditional scales such as the nine-point hedonic scale frequently present the following problems: i) they do not satisfy the statistical assumptions of independence, normality and homoscedasticity required by ANOVA models; ii) they give little freedom to the individuals to express their perceptions, due to the limited number of categories; iii) they induce numerical and contextual effects in the judgments by the panelists and, iv) the difference between numerical values associated with the categories do not reflect equivalent differences in perception.

Of the methodologies used in sensory tests with consumers, the 9-point hedonic scale is undoubtedly the most widely used. However, considering the previously mentioned problem, there is a need to investigate alternative scales providing better performance than the traditional hedonic scale, both when the data

are analyzed by ANOVA models and multivariate methods such as the Internal Preference Map – MDPREF. In general the objective of this research was to investigate the performance of two alternative scales in consumer studies, these being the self-adjusting scale and the hybrid hedonic scale, comparing them with traditional affective methods such as the ranking scale and the 9-point hedonic scale. With this objective three experiments were carried out as follows:

The first experiment was carried out with the objective of evaluating the performance of the self-adjusting scale as compared to the 9-point hedonic scale and ranking scale under real consumer test conditions, using the following criteria: i) differential expansiveness between assessors, ii) discriminating power and, iii) compliance of the data collected by each scale with the ANOVA assumptions. Three commercial brands of candy were evaluated by 288 consumers. The results obtained from the 9-point hedonic and self-adjusting scales were analyzed by ANOVA and those of the ranking test by Friedman's test. The values for pF_{sample} , pF_{assessor} and QM_{error} provided by ANOVA for each scale, were used respectively to evaluate the discriminating power, the expansiveness of scoring between assessors and the data variability. Tukey's test was also applied to analyze the discriminating power of each scale. Normal probability plots, Kolmogorov-Smirnov test and coefficients of skewness and kurtosis checked data normality. Homoscedasticity was evaluated by scatter plots and the Levene test. The results showed that the self-adjusting scale was effective to deal with differential assessor expansiveness and produced homogeneous variances, however the residuals showed moderate deviations from normality. The 9-point hedonic scale showed problems with heteroscedasticity. Rank and the self-adjusting scales showed the highest and the lowest discriminating powers, respectively. Despite the problems detected, the three scales presented the same tendencies for preference amongst the products tested.

The second experiment was carried out with the objective of evaluating the performance of the hybrid hedonic scale in consumer studies, comparing it with the 9-point hedonic, the self-adjusting and the ranking scales, using the following criteria: i) variability of sensory response, ii) discriminative power, iii) data adequacy to the

assumptions of ANOVA models and, iv) ease of use. Eighty consumers, divided into four groups of 20 individuals each, evaluated five brands of orange juice. All the individuals evaluated all the samples using all the scales, in 4 distinct tasting sessions. A 4 x 4 Latin square design was used to control the effect of the order of presentation of the scales and evaluate their ease of use without biases. For each scale the presentation order and carry-over were balanced. The results obtained using the traditional hedonic, hybrid hedonic and self-adjusting scales were evaluated using ANOVA. Data normality was evaluated using the Shapiro-Wilks test, homoscedasticity by the Brown-Forsythe's test and the Tukey's one degree of freedom test for non-additivity. The values for pF_{sample} , pF_{assessor} and QM_{error} , provided by ANOVA for each scale, were used respectively to evaluate discriminating power, expansiveness between assessors and data variability. The REGWF test was also applied to analyze the discriminative power of each scale. The results obtained from the ranking test were evaluated by Friedman's test and the ease of use of the scales by the Cochran-Mantel-Haenszel tests. The results indicated the superiority of the hybrid hedonic scale as compared to the structured hedonic and self-adjusting scales, both with respect to discriminative power and to data adequacy to the assumptions of normality and homoscedasticity. The self-adjusting scale presented a slightly greater discriminative power than the structured hedonic scale, despite the former having presented data with a greater variability and lack of normality of the residuals. Of all the methods, the ranking test presented the least discriminative power. The structured and hybrid hedonic scales were considered to be significantly ($p \leq 0.01$) easier to use than the self-adjusting scale, there being no difference ($p \leq 0.05$) between these first two scales.

Finally, the objective of the third experiment was to evaluate the performance of the nine-point hedonic, hybrid hedonic and self-adjusting scales in the segmentation of samples and consumers using Internal Preference Mapping methodology. One hundred and twelve consumers evaluated the overall acceptability of 8 commercial brands of red wine, the majority being Cabernet Sauvignon. The effects of presentation order –scales and samples- and carry over effects were balanced. The data were analyzed by ANOVA and MDPREF. Scale performance was

evaluated using as criteria: number of significant dimensions in the MDREF ($p \leq 0.05$), number of consumers significantly adjusted ($p \leq 0.05$) and the degree of segmentation of the products and consumers. The results suggested a superiority of the hybrid scale over the traditional hedonic and self-adjusting scales. The MDPREF generated by the hybrid scale data produced the greatest number of significant dimensions ($p=5\%$), yielding 79.5% of the consumers significantly adjusted ($p=5\%$), while the MDPREF generated by the self-adjusting scale adjusted 54.5% of the consumers and that of the hedonic scale, 51.8%. Overall, the 9-point hedonic scale showed the worst performance in relation to the other scales examined.

The results of this study strongly suggest that the hybrid hedonic scale is a valid and efficient tool for use in data collection associated with consumer studies, both when analyzed by normal models for the analysis of variance and by Internal Preference Mapping methodology.

INTRODUÇÃO GERAL

Na indústria alimentícia moderna, a avaliação das características sensoriais dos produtos, na forma como elas são percebidas pelos consumidores, é um componente essencial no desenvolvimento, manutenção, otimização, melhoria da qualidade e avaliação do potencial de mercado de um determinado produto alimentício (STONE & SIDEL, 1993; PIGGOT, 1995).

Dentro das diversas metodologias utilizadas pela Análise Sensorial, encontram-se os testes afetivos, cujo propósito principal, é a avaliação da preferência e/ou aceitação dos consumidores com relação a produtos e serviços (MEILGAARD *et al.*, 1991). Uma vez que a resposta dos consumidores é, via de regra, coletada através de escalas, aspectos como facilidade de uso, expansibilidade das respostas dos provadores, poder discriminativo, validade e confiabilidade das respostas coletadas, vícios numéricos e contextuais, problemas de análise estatística dos dados, entre outros, têm sido objeto de estudo de muitos cientistas que atuam na área de Análise Sensorial (PEARCE *et al.*, 1986; GIOVANNI & PANGBORN, 1983; LAWLESS & MALONE, 1986a, 1986b, TEPPER *et al.*, 1994; CHEN & RESURRECCION, 1996; HOUGH *et al.*, 1997).

Das limitações identificadas acima em metodologias de avaliações sensoriais que utilizam escalas, uma das mais citadas é a da variação da expansibilidade entre provadores, decorrente do fato de diversos provadores utilizarem diferentes porções da escala para expressarem a sua impressão sobre o produto (GAY & MEAD, 1992; MEAD & GAY, 1995). A variação da expansibilidade entre provadores constitui-se em problema particularmente grave no caso dos testes realizados com consumidores, porque produz variação

considerável nos efeitos principais analisados pelo experimento, quais sejam: diferenças entre tratamentos (MEAD & GAY, 1995; GAY & MEAD, 1992; GAY, 1988).

Para minimizar o problema da expansibilidade dos provadores em testes sensoriais, GAY (1988) e GAY & MEAD (1992), propuseram um novo tipo de escala chamado de "autoajustável". Nesta escala, o provador primeiro deve identificar as duas amostras que representam os extremos em relação ao estímulo que está sendo avaliado e alocá-las nas extremidades da escala, por exemplo, as amostras mais doce/menos doce, mais gostosa/menos gostosa, etc. As amostras restantes são então ordenadas e quantificadas na escala, como reflexo do grau relativo de diferença percebida entre o conjunto de amostras apresentadas. As vantagens da escala autoajustável, de acordo com seus autores são: facilidade de uso por provadores não treinados, a redução da variação devido à expansibilidade e ordenação preliminar das amostras anterior à sua quantificação. Quando utilizada em testes com consumidores, esta escala usa somente a essência da informação contida num conjunto de respostas sensoriais, quais sejam: a ordem de preferência das amostras e o grau relativo de diferença com relação à aceitação. Em função das vantagens observadas, os pesquisadores recomendaram o estudo do desempenho da escala autoajustável em pesquisas junto a consumidores. Entretanto, na bibliografia internacional, não se encontram facilmente estudos utilizando esta escala em testes afetivos.

Outras limitações citadas anteriormente e que causam grande impacto sobre o poder discriminativo de testes afetivos que utilizam escalas, notadamente aqueles que usam escalas de categoria como a tradicional escala hedônica estruturada de 9 pontos (PERYAM & PELGRIM, 1957) são: i) pouca liberdade para o consumidor expressar suas percepções sensoriais devido ao limitado número de categorias de respostas, ii) os valores numéricos atribuídos às suas

categorias, embora numericamente possuam intervalos iguais, não refletem iguais diferenças em percepção, iii) efeitos numéricos e contextuais são mais prováveis de acontecerem no uso desta escala, devido a diferenças na interpretação das diversas categorias e a erros de hábito, onde o provador tende a repetir as mesmas respostas em situações onde ele/ela precisa fazer várias avaliações consecutivas; iv) produzem erros de tendência central devido a que os provadores evitam o uso das categorias extremas, reduzindo o número de categorias a 7 e, trazendo em decorrência, a diminuição da sensibilidade da escala para detectar diferenças entre as amostras com alta ou baixa aceitação e, v) os dados coletados por esta escala freqüentemente não satisfazem os pressupostos estatísticos de normalidade e homoscedasticidade exigidos pelos modelos ANOVA (VILLANUEVA, *et al.* 1999; VILLANUEVA, *et al.* 2000; VILLANUEVA, *et al.* 2001; SCHUTZ & CARDELLO, 2001; GAY & MEAD, 1992; O'MAHONY, 1982; GIOVANI & PANGBORN, 1983; MCPHERSON & RANDALL, 1985).

Uma alternativa à escala hedônica tradicional e suas limitações seria o uso de escalas hedônicas não estruturadas ou lineares em testes com consumidores, conforme sugeriram GIOVANI & PANGBORN (1983). Entretanto, ainda que os trabalhos de GIOVANI & PANGBORN (1983), MCPHERSON & RANDALL (1985) entre outros, tenham mostrado um melhor desempenho da escala não estruturada em relação à escala hedônica de 9 pontos, oferecendo maior liberdade aos consumidores para expressar suas percepções sensoriais, diminuindo vícios numéricos e/ou contextuais e apresentando menores desvios da normalidade dos dados, o uso desta escala em estudos com consumidores é limitado, devido ao fato dos provadores terem maior dificuldade de entendê-la e utilizá-la (LAWLESS & MALONE, 1986a).

Uma possibilidade aparentemente inexplorada na literatura consultada, que possivelmente permitiria minimizar as limitações associadas tanto à tradicional escala hedônica de 9 pontos como à escala hedônica não estruturada, seria o uso de uma escala hedônica híbrida. A escala híbrida proposta nesta pesquisa é uma escala linear, ancorada na região central e nos extremos, com anotações verbais e numéricas. Os pontos equidistantes na linha horizontal servem para definir melhor o grau e a orientação do contínuo hedônico. As possíveis vantagens desta escala em relação à escala hedônica tradicional são: i) ela não se restringe a um limitado número de categorias, ou seja, o provador pode utilizar qualquer parte da escala para atribuir seus valores hedônicos, com possível aumento do poder discriminativo da escala, pelo fato dos indivíduos terem maior liberdade na escolha do valor numérico que expressará o seu grau de aceitabilidade do produto, ii) diminuição do erro psicológico de hábito, encorajando o provador a fazer sua avaliação de maneira mais consciente e criteriosa, iii) diminuição de efeitos contextuais favorecendo o seu uso em estudos interculturais e, iv) utilização de uma diversidade de análises estatísticas paramétricas e não paramétricas pelo fato desta escala gerar dados contínuos.

Assim em vista das limitações associadas às escalas afetivas tradicionalmente utilizadas em testes com consumidores, notadamente a escala hedônica de 9 pontos, o presente trabalho teve por objetivos:

1. Avaliar em condições reais de teste de consumidor, as vantagens e desvantagens da escala autoajustável, verificando o desempenho do método em relação a outros tipos de escala tradicionalmente utilizadas em pesquisas de consumidores, quais sejam: escala hedônica estruturada de 9 pontos e escala de ordenação, comparando os três métodos em relação ao poder discriminativo, à variabilidade das respostas sensoriais e à

adequação dos dados coletados às suposições estatísticas associadas à análise da variância (ANOVA).

2. Propor o uso de uma escala hedônica híbrida desenvolvida nesta pesquisa e avaliar o desempenho desta escala em relação à escala autoajustável, escala hedônica estruturada de 9 pontos e escala de ordenação, utilizando como critérios os seguintes parâmetros: i) variabilidade das respostas sensoriais, ii) poder discriminativo, iii) adequação dos dados coletados às suposições estatísticas para análise da variância: normalidade, homoscedasticidade, aditividade e, iv) facilidade de uso pelos consumidores, considerando-se aspectos como: facilidade de entendimento, rapidez para completar a ficha, menor restrição no uso e preferência dos consumidores com relação às escalas.

3. Avaliar o desempenho das escalas hedônica estruturada, hedônica híbrida e autoajustável, na identificação de segmentos pré-existentes de consumidores e de seus padrões de preferência através da metodologia de Mapa Interno de Preferência (MDPREF).

Para alcançar os objetivos propostos, três experimentos foram realizados e cada um deles encontra-se relatado na presente tese em formato de artigo independente. Assim, esta tese encontra-se dividida em quatro capítulos. O primeiro capítulo refere-se a uma revisão bibliográfica geral que procura introduzir ao leitor, conhecimentos e problemáticas básicas associadas ao uso de escalas afetivas em testes com produtos alimentícios. O segundo capítulo relata os resultados do primeiro experimento cujo objetivo foi testar o desempenho da escala autoajustável em estudos com consumidores, comparando-a com a escala hedônica de 9 pontos e escala de ordenação. Neste estudo, 288 consumidores avaliaram 3 marcas comerciais de doces. O terceiro capítulo relata os

resultados obtidos no segundo experimento onde cinco marcas comerciais de suco de laranja foram avaliadas por 80 consumidores e cujo objetivo foi comparar o desempenho das escalas hedônica híbrida e autoajustável em relação a duas escalas tradicionalmente utilizadas em estudos com consumidores: escala de ordenação e escala hedônica de 9 pontos. Finalmente, o quarto capítulo relata os resultados correspondentes ao terceiro experimento, realizado para comparar o desempenho da escala hedônica de 9 pontos, hedônica híbrida e autoajustável na geração de Mapas Internos de Preferência. Neste experimento, 8 marcas comerciais de vinho tinto seco foram avaliadas por 112 consumidores.

CAPÍTULO 1

REVISÃO BIBLIOGRÁFICA

1. ESCALAS AFETIVAS E SUAS PROBLEMÁTICAS

A função básica da Análise Sensorial é a obtenção de medidas subjetivas de características sensoriais de produtos alimentícios, tais como aparência, sabor, textura, aroma, etc. Assim, nesta área de pesquisa, o alimento é o elemento a ser medido e o homem, o instrumento de medição. Neste sentido, a informação que a Análise Sensorial produz é única, e diferencia-se claramente de outras fontes de informação onde, via de regra, métodos químicos e/ou instrumentais são utilizados para caracterizar o alimento (MEAD & GAY, 1995).

Dentro das diversas metodologias desenvolvidas pela Análise Sensorial, encontram-se os testes afetivos, cujo propósito principal, é a avaliação das respostas dos consumidores com relação à preferência e/ou aceitação dos produtos. Esses testes têm demonstrado ser altamente efetivos no desenvolvimento, melhoria da qualidade, otimização e avaliação do potencial do mercado de produtos alimentícios (MEILGAARD *et al.*, 1991).

Em testes afetivos, os métodos mais usados empregam escalas hedônicas para avaliar o grau de aceitação dos produtos. Uma escala internacionalmente aceita e largamente utilizada neste tipo de teste é a escala hedônica estruturada de 9 pontos (Figura 1). Esta escala consiste em 9 categorias verbais com termos variando entre “desgostei extremamente” (valor 1) e “gostei extremamente” (valor 9), e uma categoria neutra localizada no centro da escala associada ao termo “não gostei/nem desgostei” (valor 5) (PERYAM & PILGRIM, 1957). Vários problemas são apontados com relação a esta tradicional escala, quais sejam: i) pouca liberdade para o

consumidor expressar suas percepções sensoriais devido ao limitado número de categorias de respostas, ii) os valores numéricos atribuídos às categorias, embora numericamente possuam intervalos iguais, não refletem iguais diferenças em percepção, iii) efeitos numéricos e contextuais são mais prováveis de acontecerem no uso desta escala, devido às diferenças de interpretação das diversas categorias e a erros de hábito, onde o provador tende a repetir as mesmas respostas em situações onde várias avaliações consecutivas são necessárias, iv) produzem erros de tendência central devido a que os provadores evitam o uso das categorias extremas, reduzindo o número de categorias a 7 e trazendo, em decorrência, a diminuição da sensibilidade da escala para detectar diferenças entre as amostras com alta ou baixa aceitação e, iv) as respostas sensoriais freqüentemente apresentam problemas de não normalidade e heteroscedasticidade, pressupostos requeridos pelos modelos ANOVA usuais (O'MAHONY, 1982; GIOVANNI & PANGBORN, 1983; MCPHERSON & RANDALL, 1985; VIE *et al.*, 1991; GAY & MEAD, 1992; VILLANUEVA *et al.*, 1999; VILLANUEVA *et al.*, 2000; VILLANUEVA *et al.*, 2001; CURIA *et al.*, 2001; SCHUTZ & CARDELLO, 2001).

Nome: _____		Data: _____	
<p>Você está recebendo "x" amostras codificadas de _____. Avalie da esquerda para a direita cada uma das amostras e use a escala abaixo para indicar o quanto você gostou ou desgostou das mesmas.</p>			
<p>9 Gostei extremamente 8 Gostei muito 7 Gostei moderadamente 6 Gostei ligeiramente 5 Nem gostei / nem desgostei 4 Desgostei ligeiramente 3 Desgostei moderadamente 2 Desgostei muito. 1 Desgostei extremamente</p>			
AMOSTRA N ^o :		VALOR:	
_____		_____	
_____		_____	
Comentários _____			

Figura 1- Ficha de avaliação com escala hedônica estruturada de 9 pontos.

Outro problema apresentado por métodos sensoriais que utilizam escalas como instrumento de medida, é o da variação da expansibilidade entre provadores, decorrente do fato de diferentes provadores utilizarem diferentes porções da escala para expressarem a sua impressão sobre o produto. Essas diferenças intrínsecas no uso da escala de medida entre os provadores pode acontecer tanto em relação à localização das respostas sensoriais no interior da escala (diferenças em posição) como à faixa da escala utilizada (diferenças em dispersão). As diferenças em posição decorrem do fato de que alguns provadores utilizam somente a parte inferior da escala, tendendo a dar notas baixas enquanto outros, utilizam a parte superior, tendendo a dar notas altas; há ainda aqueles que utilizam a região central. Já as diferenças em dispersão, acontecem devido a que alguns provadores utilizam toda a escala para fazer suas avaliações enquanto outros se concentram sobre uma parte da escala (GAY & MEAD, 1992; MILLER, 1987; NAES, 1990; WILKINSON & YUKSEL, 1997). Às diferenças no uso efetivo de porções da escala chama-se de variação da expansibilidade entre provadores, constituindo-se em problema particularmente grave no caso dos testes realizados com consumidores, porque produz variação considerável nos efeitos principais analisados pelo experimento, quais sejam: diferenças entre tratamentos (GAY & MEAD, 1992; GAY, 1988).

Em testes sensoriais, a análise dos dados é habitualmente realizada através de modelos lineares para análise da variância (ANOVA). Estes modelos pressupõem que as respostas experimentais devam ser: i) independentes, ii) normalmente distribuídas, iii) de variâncias iguais (homoscedásticas) e, iv) provenientes de uma mesma escala de medida. GAY & MEAD (1992) afirmam que os principais problemas na análise de dados sensoriais são aqueles referentes aos dois últimos pressupostos. Desvios referentes à homogeneidade das variâncias ocorrem devido à existência de pelo menos duas fontes potenciais de variabilidade dos dados, quais sejam: provadores e tratamentos, acarretando conseqüências sérias na obtenção do verdadeiro nível de significância para o efeito dos tratamentos. Por outro lado, o quarto pressuposto (aditividade) pode ser violado pelo fato de que nas avaliações

sensoriais, cada provador tem sua impressão pessoal sobre o produto e utiliza faixas mais (ou menos) amplas da escala, atribuindo valores maiores (ou menores) em relação a outros provadores, ocasionando o problema da variação da expansibilidade.

Para minimizar o problema da expansibilidade dos provadores em testes sensoriais, WILLIAMS *et al.* (1979); GAY (1988) e GAY & MEAD (1992) propuseram um novo tipo de escala, ao qual os autores intitularam “escala autoajustável”. A escala autoajustável (Figura 2), é uma escala linear através da qual o provador primeiro deve identificar as duas amostras que representam os extremos em relação ao estímulo que está sendo avaliado e alocá-las nas extremidades da escala, por exemplo, as amostras mais doce/menos doce, mais gostosa/menos gostosa, etc. As amostras restantes são então ordenadas e quantificadas na escala, como reflexo do grau relativo de diferença percebida entre o conjunto de amostras apresentadas. Quando utilizada em testes com consumidores, esta escala busca somente utilizar a essência da informação contida num conjunto de respostas sensoriais, quais sejam: a ordem de preferência das amostras e o seu grau relativo de diferença.

Nome _____		Data _____	
<p>Você esta recebendo — amostras codificadas de —. Prove por favor as amostras da esquerda para direita. Escolha primeiro a amostra que você mais gostou e depois a que você menos gostou e anote os códigos nos extremos direito e esquerdo da escala abaixo, respectivamente. Avalie agora as outras amostras de forma comparativa com aquelas alocadas nos extremos, marcando um “x” e o código da amostra no ponto da escala que você achar que melhor represente o quanto você gostou ou desgostou de cada amostra.</p>			
X _____ X			
Menos gostoso		Mais gostoso	
Comentários: _____			

Figura 2- Ficha de avaliação sensorial utilizando a escala autoajustável.

As vantagens da escala autoajustável, de acordo com seus autores são a facilidade de uso por provadores não treinados, a diminuição da variação devido à expansibilidade e a ordenação preliminar das amostras anterior à sua quantificação. Já uma possível desvantagem no uso desta escala, seria o fato dos provadores

terem que experimentar todas as amostras antes de quantificá-las, o que tornaria sua avaliação trabalhosa, principalmente quando se requer que vários atributos do alimento sejam avaliados, tal como acontece na Análise Descritiva Quantitativa.

Na tentativa de contornar problemas como a não equivalência no grau de percepção entre os intervalos das categorias verbais e a falta de normalidade dos dados gerados por escalas estruturadas, GIOVANNI & PANGBORN (1983) e MCPHERSON & RANDALL (1985), propuseram a utilização de escalas lineares ou não estruturadas em estudos com consumidores. A escala não estruturada (Figura 3) é uma escala linear de um determinado comprimento (9 a 15cm), com âncoras verbais nos seu extremos, através do qual o provador é solicitado a expressar o grau de percepção de um determinado atributo do produto em avaliação, colocando uma marca na linha horizontal da escala. A distância entre essa marca e o extremo esquerdo da escala é medida com régua, e transformada em valor numérico proporcional ao grau de percepção do atributo avaliado. Os referidos autores concluíram que apesar da escala não estruturada apresentar ainda desvios da normalidade dos dados, mostrou um melhor desempenho com relação às escalas hedônicas estruturadas, oferecendo maior liberdade aos consumidores para expressarem suas percepções sensoriais e diminuindo efeitos numéricos e contextuais.

Nome: _____ Idade: _____ Data: _____
Você esta recebendo ___ amostras codificadas de _____. Por favor PROVE as amostras da esquerda para a direita e marque na escala correspondente um "x" e o código da amostra no lugar da escala que você achar que melhor represente o quanto você gostou ou desgostou de UM MODO GERAL de cada amostra.

Desgostei extremamente Gostei extremamente

Figura 3. Ficha de avaliação da aceitação utilizando a escala hedônica não estruturada.

Outra escala extensamente pesquisada em testes junto a consumidores pelos cientistas sensoriais é a escala de magnitude (Figura 4). Esta escala é considerada como sendo uma escala de razão e geralmente utiliza-se para medir percepções associadas às intensidades dos estímulos, de forma comparativa com uma referência. No uso da escala de magnitude, o provador recebe inicialmente uma amostra referência, com um valor de intensidade determinado por ele ou definido pelo experimentador, para um dado atributo da amostra (acidez, doçura, etc.). O provador é então instruído a avaliar todas as amostras subseqüentes de forma comparativa à amostra referência: se um estímulo da amostra referência foi especificado como 10 de intensidade e esse estímulo na amostra codificada for percebido como duas vezes mais intenso, o provador atribuirá a ele o valor de 20 (MEILGAARD *et al.*, 1991). Os resultados apresentados pelas diversas pesquisas utilizando esta escala são controversos. Ainda que autores como McDANIEL & SAWYER (1981) e LAVENKA & KAMEN (1994) tenham concluído em seus trabalhos que a escala de magnitude foi mais discriminativa que a escala hedônica de 9 pontos, GIOVANNI & PANGBORN (1983) e PANGBORN *et al.* (1989), entre outros, concluíram que esta escala pode ser inadequada para avaliar a aceitação de produtos a despeito dela ser apresentada utilizando referência fixa ou livre.

Nome: _____		Data: _____	
<p>Por favor, prove primeiramente a amostra referência "R" assinalando a ela o valor 10 de doçura. Em seguida avalie a doçura de cada amostra codificada em relação à amostra referência "R". Se a amostra codificada for 2 vezes mais doce que amostra R, dê à amostra codificada o valor 20, se for 2 vezes menos doce, dê o valor 5, e assim por diante.</p>			
AMOSTRA		DOÇURA	
R		10	
911		—	
732		—	
127		—	

Figura 4. Ficha de avaliação sensorial utilizando a escala de magnitude.

Recentemente, VILLANUEVA *et al.* (2000), VILLANUEVA *et al.* (2001) e VILLANUEVA *et al.* (2002), realizaram uma série de estudos com consumidores, com o objetivo de identificar escalas alternativas que minimizassem alguns dos problemas

Nome: _____ Data: _____

Você esta recebendo — amostras codificadas de —. Avalie por favor da esquerda para a direita, a APARÊNCIA das amostras e ordene-as em ordem decrescente de preferência, com relação a APARÊNCIA.

+ preferida _____ - preferida _____

Comentários: _____

Figura 6. Ficha de avaliação da preferência utilizando a escala de ordenação.

Visando contornar o problema das diferenças em posição e dispersão causados pelo uso de diferentes porções de escala pelos provadores de equipes sensoriais treinadas, vários autores propuseram métodos alternativos de análise estatística e geração de dados com menor variabilidade. NAES (1990); PRITCHETT (1992); BROCKHOFF & SKOVGAARD (1994) e WILKINSON & YUKSEL (1997), analisaram dados provenientes de equipes treinadas através de modelos lineares generalizados, mínimos quadrados ponderados e métodos de máxima verossimilhança.

Em vista dos problemas apresentados pela tradicional escala hedônica de 9 pontos, SCHUTZ & CARDELLO (2001), desenvolveram uma nova escala chamada de “escala de magnitude afetiva verbal” (“labeled affective magnitude” – LAM) (Figura 7) para uso com consumidores, com termos verbais consistentes com os da escala hedônica de 9 pontos. Os resultados desse estudo mostraram igual confiabilidade e sensibilidade da nova escala em relação à escala hedônica de 9 pontos, produzindo ainda uma discriminação ligeiramente maior entre os produtos de alta aceitação e distribuições normais dos dados. Quando comparada com a escala de magnitude tradicional, ambas as escalas produziram resultados semelhantes em termos das razões obtidas entre os atributos avaliados. Com relação à facilidade de uso, os consumidores reportaram que a escala de magnitude verbal foi tão fácil de usar

quanto a escala hedônica de 9 pontos e significativamente menos difícil ($p \leq 0,05$) do que a escala de magnitude.

Nome: _____ Data: _____

Você está recebendo ___ amostras de _____. Por favor prove cada uma das amostras da esquerda para a direita e marque na escala correspondente um "x" e o código da amostra no lugar da escala onde você acha que melhor representa o quanto você gostou ou desgostou da amostra.

—	Fantástico
—	Gostei extremamente
—	Gostei muito
—	Gostei moderadamente
—	Gostei ligeiramente
—	Nem gostei/nem desgostei
—	Desgostei ligeiramente
—	Desgostei moderadamente
—	Desgostei muito
—	Desgostei extremamente
—	Terrível

Figura 7. Ficha de avaliação sensorial da escala de magnitude afetiva verbal.

2. ESTUDOS COMPARATIVOS SOBRE DESEMPENHO DE ESCALAS

Uma vez que a resposta dos consumidores em testes afetivos é, via de regra, coletada através de escalas, aspectos como facilidade de uso, expansibilidade das respostas, poder discriminativo, validade e confiabilidade das respostas coletadas, vícios numéricos e contextuais, problemas de análise estatística dos dados sensoriais entre outros, têm sido objeto de pesquisa de muitos cientistas que atuam na área de Análise Sensorial.

MOSKOWITZ & SIDEL (1971), conduziram uma série de experimentos para comparar a escala de magnitude e a escala hedônica de 9 pontos em testes afetivos. Esses pesquisadores verificaram que ambos os procedimentos de avaliação foram igualmente sensíveis em detectar diferenças na aceitação de alimentos, porém cada procedimento fornecia informações adicionais: a escala de magnitude quantificava as razões da aceitabilidade entre os diversos itens e a escala hedônica fornecia categorias numéricas e verbais de aceitação. Os autores concluíram que uma escala pode ser usada para complementar à outra na avaliação da aceitação dos produtos alimentícios.

McDANIEL & SAWYER (1981), realizaram um estudo com consumidores de whisky utilizando dois tipos de teste (de laboratório e de uso doméstico) e dois tipos de escala (de magnitude e hedônica de 9 pontos) para comparar a eficiência tanto dos tipos de teste quanto dos métodos de escala. Para isso, dois grupos de consumidores avaliaram 6 amostras de whisky através das duas escalas. No teste de laboratório, participaram estudantes universitários e no teste de uso doméstico, voluntários adultos. Em ambos os testes, a escala de magnitude produziu um número maior de diferenças significativas ($p=5\%$) entre as amostras, sendo este número ainda maior para o teste de uso doméstico. Os autores concluíram que a escala de magnitude foi mais discriminativa do que a escala hedônica de 9 pontos, apresentando portanto melhor desempenho.

GIOVANNI & PANGBORN (1983), pesquisaram o desempenho de duas escalas: não estruturada de 10 cm e escala de magnitude, na geração de dados hedônicos e de intensidade. Ambas as escalas produziram resultados semelhantes na avaliação de bebidas, porém, a escala não estruturada mostrou maior poder discriminativo, melhor reprodutibilidade das respostas sensoriais, maior facilidade de uso pelos provadores e menor influência de efeitos numéricos e contextuais nas respostas sensoriais. Na escala de magnitude, os provadores utilizaram porções mais restritas da escala para expressar suas percepções, sendo esse efeito mais

acentuado no caso das respostas hedônicas. Os autores observaram que na escala de magnitude, as respostas de intensidade foram influenciadas pelo nível de intensidade do atributo observado na amostra referência e, por sua vez na geração das respostas hedônicas, os consumidores tiveram dificuldade em atribuir razões por causa da bipolaridade da escala.

Posteriormente, PANGBORN *et al.* (1989), realizaram 5 experimentos para comparar a intensidade percebida e o grau de aceitação do gosto amargo da cafeína e quinino em chocolate quente, utilizando escalas estruturadas, não estruturadas e de magnitude com referência fixa e referência livre. As escalas estruturadas e não estruturadas produziram resultados semelhantes com relação às estimativas da intensidade, indicando que ambas as escalas foram utilizadas pelos provadores de forma similar. Na avaliação da intensidade, o uso da escala de magnitude com referências fixa e livre, não alteraram a forma das funções de poder, já na avaliação da aceitação, mudanças foram observadas nas distribuições das respostas hedônicas em função da amostra referência. Os autores recomendaram a utilização de referência fixa, tanto pela maior facilidade de uso da escala como pela maior simplicidade na análise estatística e sugeriram que a escala de magnitude pode ser inadequada para avaliar a aceitação dos produtos independentemente dela ser apresentada com amostra referência fixa ou livre.

PEARCE *et al.* (1986), conduziram estudos para comparar a eficiência de três escalas: hedônica estruturada de 9 pontos, de magnitude unipolar e de magnitude bipolar, na coleta de respostas hedônicas para estímulos controlados. Um total de 533 indivíduos participaram do estudo, sendo 225 homens e 328 mulheres, os quais avaliaram 8 amostras de tecidos procedentes de diferentes fábricas americanas através das três escalas. Os resultados mostraram que as três escalas produziram respostas hedônicas bastante similares tanto no grau de aceitação quanto na magnitude das diferenças entre os estímulos avaliados. Nenhuma escala mostrou

clara superioridade sobre as demais em relação à confiabilidade, precisão ou poder discriminativo do teste.

LAWLESS & MALONE (1986a), avaliaram a sensibilidade relativa de quatro métodos de escala em testes de localização central com consumidores. As escalas foram: escala de categoria, escala não estruturada, escala de magnitude e escala híbrida tipo "scanner". Dois grupos de consumidores participaram deste estudo: um grupo feminino de estudantes colegiais com idade média de 21 anos, e um grupo heterogêneo de donas de casa de idade média de 38 anos. Em relação ao poder discriminativo para detecção de pequenas diferenças na intensidade dos atributos avaliados, as escalas de categoria, não estruturada e híbrida produziram resultados semelhantes. Por sua vez, a escala de magnitude foi tão eficiente quanto às demais para o grupo homogêneo de estudantes, entretanto, para o grupo heterogêneo de donas de casa, esta mostrou-se menos eficiente. Os autores relataram que os consumidores utilizaram as escalas com maior precisão, após adquirirem certa familiaridade com a avaliação dos produtos. Ambos os grupos de consumidores utilizaram amplas porções das escalas, apesar de existirem mínimas diferenças sensoriais entre os produtos avaliados.

Na seqüência, em um estudo semelhante ao anterior, LAWLESS & MALONE (1986b), utilizaram novamente os 4 tipos de escala - de categoria, não estruturada, de magnitude e híbrida tipo "scanner"- para avaliar a intensidade de atributos como aroma citral e linalol em dietil fitalato, brilho e cor em superfícies de madeira, aspereza em lixas comerciais e espessura em silicones. O objetivo desta pesquisa foi comparar o desempenho destas escalas com relação à capacidade para discriminar diferenças entre os produtos, à variabilidade das respostas geradas, à confiabilidade e facilidade de uso das mesmas. Esses autores verificaram que todos os métodos foram capazes de gerar diferenças altamente significativas ($p=5\%$) entre os produtos, porém, uma ligeira vantagem para a escala de categoria foi observada em quase todas as comparações. Eles reportaram ainda que com a escala de categoria e a

escala não estruturada, os consumidores tiveram maior facilidade de uso enquanto que com escala de magnitude tiveram maior dificuldade.

LAVENKA & KAMEM (1994), testaram 3 tipos de escala: escala hedônica de 9 pontos, escala de magnitude com referência fixa e escala de magnitude com referência livre, para determinar qual seria o melhor método de avaliação da aceitação global de produtos alimentícios. Uma lista impressa de 50 itens de alimentos foi avaliada por 236 recrutas do exército americano, os quais foram divididos em três grupos de acordo com as suas características demográficas. Desta forma, 71 indivíduos avaliaram os produtos através da escala hedônica de 9 pontos, 92 através da escala de magnitude com referência fixa e 72 através da escala de magnitude com referência livre. Essencialmente os mesmos resultados foram obtidos pelos três métodos. Foram observadas altas correlações de produto-momento de Pearson entre as respostas da escala hedônica de 9 pontos e a escala de magnitude com referência livre. Os autores reportaram que a escala de magnitude com referência livre poderia ser utilizada para coletar dados de maneira tão rápida e eficiente quanto as outras duas escalas, mesmo com provadores não treinados, e consideraram que estas escalas teriam a vantagem de produzir respostas menos viciadas em estudos inter-culturais por não possuírem categorias verbais.

TEPPER *et al.* (1994), realizaram um experimento com consumidores para avaliar a percepção e aceitação de 6 tipos de alimentos contendo diferentes níveis de gordura: leite, batata, frango, pudim, “snacks” e ovos. Os autores utilizaram dois métodos afetivos: teste de ordenação e teste de escala hedônica não estruturada. A eficiência dos dois métodos foi avaliada em função da capacidade dos mesmos em discriminar ($p=5\%$) os produtos contendo diferentes teores de gordura. O estudo mostrou que ambos os métodos foram eficientes e produziram padrões de resposta similares, ainda que os provadores tenham reportado que com o método de ordenação, a avaliação foi mais simples. A escala não estruturada teve a vantagem

de fornecer informações quantitativas sobre o grau com que os consumidores gostaram ou desgostaram do teor de gordura de cada produto.

CHEN & RESURRECCION (1996), conduziram estudos com o objetivo de determinar o número de categorias mais adequado na escala hedônica facial, para utilização em testes afetivos com crianças. Setenta e nove crianças de idade variando entre 3 e 6 anos, avaliaram o quanto gostavam ou desgostavam da cor e do sabor de duas marcas comerciais de leite: pasteurizado e UHT. Os resultados mostraram que crianças de idade variando entre 3 e 4 anos, utilizaram com maior eficiência a escala hedônica facial de três pontos, enquanto crianças com idade entre 4 e 5 anos, mostraram maior poder discriminativo utilizando escalas de 5 pontos e crianças entre 5 e 6 anos, escalas de 7 pontos.

HOUGH, *et al.* (1997), compararam a escala hedônica facial e a escala hedônica estruturada verbal de 9 pontos em testes de aceitação com crianças com idade média de 11 anos. As crianças avaliaram o quanto gostaram ou desgostaram de diversos atributos de leite achocolatado. Os resultados mostraram que ambas as escalas foram equivalentes para todos os atributos na sua capacidade de discriminação das amostras entre si.

Com o objetivo de avaliar o desempenho da escala autoajustável em testes com consumidores, VILLANUEVA *et al.* (2000), realizaram um estudo comparando a escala autoajustável com a tradicional escala hedônica, e teste de ordenação avaliando a expansibilidade entre provadores, o poder discriminativo das escalas e a adequação dos dados gerados aos pressupostos associados aos modelos ANOVA. Os autores constataram que a escala autoajustável resolveu problemas da expansibilidade entre provadores e produziu variâncias homogêneas entre as amostras, porém, esta escala apresentou problemas de não normalidade dos resíduos. A escala hedônica de 9 pontos, por sua vez, apresentou problemas de

heteroscedasticidade nas variâncias das amostras. As escalas de ordenação e autoajustável tiveram o melhor e pior poder discriminativo, respectivamente.

VILLANUEVA *et al.* (2001), realizaram um experimento para avaliar o desempenho da escala hedônica híbrida em estudos com consumidores, comparando-a à escala hedônica de 9 pontos, escala autoajustável, e escala de ordenação; utilizando os seguintes critérios: i) variabilidade das respostas sensoriais, ii) poder discriminativo, iii) adequação dos dados às suposições dos modelos ANOVA e, iv) facilidade de uso pelos consumidores. Cinco marcas comerciais de suco de laranja foram avaliadas por 80 consumidores. Os resultados encontrados pelos autores sugeriram uma superioridade da escala hedônica híbrida sobre as escalas hedônica tradicional e autoajustável, tanto em função do poder discriminativo, como em relação à adequação dos dados às suposições de normalidade e homoscedasticidade. Por sua vez, a escala autoajustável apresentou poder discriminativo ligeiramente maior do que a escala hedônica estruturada, a despeito dos dados da primeira apresentarem maior variabilidade e problemas de não normalidade dos resíduos. A escala de ordenação apresentou o menor poder discriminativo em relação às demais. As escalas hedônicas estruturada e híbrida foram consideradas significativamente mais fáceis de serem utilizadas que a escala autoajustável ($p \leq 0,05$), não havendo diferença entre as duas primeiras.

CURIA *et al.* (2001), pesquisaram o grau de ambigüidade no uso da escala hedônica de 9 pontos por consumidores argentinos quando esta escala foi traduzida para o espanhol. Três traduções diferentes foram testadas: uma tradução literal do inglês, uma tradução independente e uma tradução independente adaptada para uso com crianças. Participaram da pesquisa 288 indivíduos provenientes de duas cidades argentinas, entre crianças, adolescentes e adultos. Os testes consistiram em ordenar as frases de cada uma das três escalas começando de “desgostei muito” até “gostei muito”. Na maioria dos casos, a escala adaptada para crianças mostrou melhor desempenho, apresentando menor número de inversões das frases em relação às

outras duas escalas. Na escala adaptada, aproximadamente 30% dos indivíduos ordenaram as frases traduzidas em ordem diferente à versão em inglês; já no caso da tradução literal e independente, essas porcentagens foram maiores, ocorrendo duas ou mais inversões das frases. Os autores concluíram que a escala hedônica de 9 pontos em outros idiomas deve ser empregada com cuidado, assegurando que a população entenda as frases de maneira similar à escala originalmente proposta por PERYAM & PILGRIM (1957). Os autores também sugeriram que o uso de escalas numéricas ou não estruturadas poderia ser mais apropriado.

COX, *et al.* (2001), realizaram um estudo inter-cultural utilizando escalas hedônicas para verificar a existência de interações entre escala, cultura e sexo em consumidores australianos de origem europeia (n=61) e consumidores residentes na Austrália provenientes da Malásia (n=54). Um grupo de consumidores utilizou a escala hedônica verbal de 9 pontos e outro grupo, a escala não estruturada. Cada indivíduo era sorteado previamente antes de fazer parte de um dos grupos. Ambos os grupos utilizaram cabines informatizadas para fazerem suas avaliações sensoriais. Os resultados desse estudo mostraram que tanto os australianos quanto os malaios, utilizaram ambas as escalas de maneira similar. Porém, os resultados de análises não paramétricas sugeriram que a escala não estruturada encorajava o uso de uma maior porção da escala no grupo malaio e, portanto, nesta população o uso desta escala seria mais recomendável. Também, houve uma ligeira evidência de que ambos os grupos étnicos, confundiam de maneira diferente, uma ou mais categorias verbais da escala hedônica verbal, sugerindo que o uso da escala não estruturada pode ser preferível em estudos inter-culturais.

Em um recente estudo, VILLANUEVA *et al.* (2002), compararam o desempenho das escalas hedônica estruturada, hedônica híbrida e autoajustável, na geração de Mapas Internos de Preferência - MDPREF. Nesta pesquisa, a aceitação global de 8 marcas comerciais de vinho tinto, a maioria deles varietal Cabernet Sauvignon, foi avaliada por 112 consumidores. O critério de avaliação do

desempenho da cada escala baseou-se no número de consumidores significativamente ajustados ($p \leq 0,05$) e no grau de segmentação dos produtos e dos consumidores produzidos pelo MDPREF. Os resultados sugeriram uma superioridade da escala híbrida sobre a escala hedônica tradicional e autoajustável. O MDPREF gerado pelos dados da escala híbrida produziu um maior número de dimensões significativas de preferência ($p \leq 0,05$), trazendo como decorrência, uma porcentagem de 79,5% consumidores significativamente ajustados ($p \leq 0,05$), enquanto a escala autoajustável ajustou 54,5% dos consumidores e a escala hedônica 51,8%. Em geral a escala hedônica de 9 pontos apresentou um desempenho inferior ao das demais escalas.

3. A PROBLEMÁTICA DA ANÁLISE ESTATÍSTICA DE DADOS COLETADOS ATRAVÉS DE ESCALAS

O'MAHONY (1982) advertiu que, no que se refere a dados sensoriais, as ferramentas estatísticas devem ser utilizadas com muita cautela, já que a maioria dos testes estatísticos paramétricos assumem que os dados amostrais procedem de populações normalmente distribuídas. Da mesma forma, a maioria dos testes estatísticos para comparações de médias associados aos modelos ANOVA, tais como o da Mínima Diferença Significativa (teste t), Tukey, Scheffé, Fisher-Bonferroni e outros, são testes paramétricos que exigem normalidade nos dados. O autor também afirma que se não é feito um diagnóstico estatístico para verificação desses pressupostos e, se os dados produzidos pelas avaliações sensoriais não satisfizerem tais pressupostos, o que ocorre com muita frequência, pode-se chegar a conclusões incorretas ou resultados de pouca validade acerca dos tratamentos experimentais.

Segundo STONE & SIDEL (1993), O'MAHONY (1982) e O'MAHONY (1986), quando os dados procedem de escalas não estruturadas e escalas de magnitude, a maioria dos métodos estatísticos paramétricos podem ser empregados: médias, desvios padrões, análise da variância, análise de regressão, análise de correlação,

técnicas multivariadas, etc. Em complementação, para esses autores, os dados coletados por essas escalas são genuinamente quantitativos e podem também ser transformados em dados de ordenação, e alternativamente analisados através de métodos estatísticos não paramétricos.

Estudos têm mostrado que freqüentemente, os dados sensoriais produzem distribuições assimétricas e portanto não normais. Isto acontece quando as respostas sensoriais se concentram no extremo inferior ou superior da escala. Por exemplo, em testes de consumidores utilizando escalas hedônicas estruturadas, esse fato ocorre quando um dos produtos tem, de forma muito consensual, alta ou baixa aceitação em relação aos demais, o que resulta em freqüências altas de notas 9 ou de notas 1 para aquele produto (PEARCE *et al.*, 1986; WILKINSON & YUKSEL, 1997). De fato, GIOVANNI & PANGBORN (1983) observaram distribuições bimodais das respostas hedônicas ao utilizarem a escala não estruturada, infringindo claramente a suposição de normalidade dos dados.

Para determinar a verdadeira natureza da distribuição dos dados de aceitação dos alimentos, MCPHERSON & RANDALL (1985), avaliaram características como grau de assimetria e curtose das distribuições dos dados gerados por duas escalas: não estruturada de 10 cm e hedônica de 7 pontos. Para isso, a aceitação de vários tipos de vegetais foram avaliados por 99 mulheres voluntárias. Foram observadas distribuições com maior grau de curtose e assimetria nos alimentos mais preferidos, sendo que o grau de assimetria aumentou proporcionalmente ao grau de aceitação das amostras para ambas as escalas. Em geral, os resultados mostraram pouca justificativa para assumir normalidade nos dados, porém, a escala não estruturada mostrou um melhor desempenho que a escala hedônica de 7 pontos, oferecendo maior liberdade aos consumidores para expressarem suas percepções sensoriais, diminuindo efeitos contextuais devido à interpretação das diversas categorias e apresentando menores desvios da normalidade dos dados.

Em testes com consumidores, as escalas estruturadas, notadamente a escala hedônica de 7 ou 9 pontos são as mais utilizadas, por serem de fácil entendimento entre os provadores (CARDELLO & SCHUTZ, 1996). É esperado que os dados provenientes dessas escalas, violem mais fortemente os pressupostos de normalidade e homoscedasticidade implicados nos modelos ANOVA usuais, uma vez que essas escalas se limitam a um determinado número de categorias (VIE *et al.*, 1991, McPHERSON & RANDALL, 1985). Para contornar esse problema, CLONINGER & BALDWIN (1976), utilizaram o método de Guilford para normalização de dados provenientes de escalas de categoria. Atributos como sabor, aroma e suculência de vários produtos alimentícios foram avaliados usando escalas de categoria que variavam entre 5 e 15 pontos. Os resultados mostraram que a tendência dos provadores de evitarem o uso dos extremos da escala aumentou com o número de categorias, pois erros de tendência central foram observados nas avaliações de sabor na escala de 9 e 15 pontos. Após a normalização, a escala de 5 pontos apresentou distâncias aproximadamente iguais entre as categorias, o que não aconteceu nas escalas com maior número de categorias, levando os autores a concluir que a escala de 5 pontos produziria no teste F, níveis de significância mais exatos e no teste t, número de diferenças significativas mais confiáveis.

VIE *et al.* (1991) e CHAUHAN & O'MAHONY (1993), utilizaram um método alternativo de avaliação das preferências dos consumidores chamado de "Signal Detection R-index Analysis". Este método baseia-se na teoria de "Signal Detection" e fornece os valores dos índices R, que expressam a probabilidade de que um determinado produto alvo seja escolhido pelos consumidores amostrados quando o mesmo é apresentado em testes de comparações pareadas junto com outro produto. Segundo esses autores, esta metodologia se apresenta como uma alternativa à escala hedônica de 9 pontos e escala do ideal devido a que os valores dos índices R, ainda que conceitualmente simples, fornecem informações sofisticadas.

AGRADECIMENTO

Os autores agradecem à **Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)**, pelo apoio financeiro que fez possível a realização deste projeto de pesquisa.

4. REFERÊNCIAS BIBLIOGRÁFICAS

- BROCKHOFF, M.; SKOVGAARD, M. Modeling individual differences between assessors in sensory evaluations. *Food Quality and Preference*, v. 5, p. 215-224, 1994.
- CARDELLO, A. V.; SCHUTZ, H. G. Food appropriateness measures as an adjunct to consumer preference/acceptability evaluation. *Food Quality and Preference*, v.7, p. 239-249, 1996.
- CHAUHAN, J.; O'MAHONY, M. Use of signal detection ranking analysis to measure preference for commercial and "health modified cakes". *Journal of Sensory Studies*, v. 8, p. 69-75, 1993.
- CHEN, A. W.; RESURRECCION, V. A. Age appropriate hedonic scales to measure food preferences of young children. *Journal of Sensory Studies*, v. 11, p. 141-163, 1996.
- CLONINGER, M. R.; BALDWIN, R. E. Analysis of sensory rating scales. *Journal of Food Science*, v. 56, p. 1225-1228, 1976.
- COX, D. N.; CLARK, M. R.; MIALON, V. S. A cross-cultural methodological study of the uses of two common hedonic responses scales. *Food Quality and Preference*, v.12, p. 119-131, 2001.
- CURIA, A. V.; HOUGH G.; MARTÍNEZ, M. C.; MARGALEF, M. I. How Argentine consumers understand the Spanish translation of the 9-point hedonic scale. *Food Quality and Preference*, v.12, p. 217-221, 2001.
- GAY, C., A statistical assessment of taste-testing methods. Ph. D. Thesis, The University of Reading, Reading, UK, 1988, 276p.
- GAY, C.; MEAD, R. A statistical appraisal of the problem of sensory measurement. *Journal of Sensory Studies*, v. 7, p. 205-228, 1992.
- GIOVANNI, M. E.; PANGBORN, R. M. Measurement of taste intensity and degree of liking of beverages by graphic scales and magnitude estimation. *Journal of Food Science*, v. 48, p. 1175-1182, 1983.
- HOUGH, G.; SANCHEZ, R.; BARBIERI, T.; MARTINEZ, E. Sensory optimization of powdered chocolate milk formula. *Food Quality and Preference*, v. 8, p. 213-221, 1997.

- LAVENKA, N.; KAMEM, J. Magnitude estimation of food acceptance. *Journal of Food Science*, v. 59, p. 1322-1324, 1994.
- LAWLESS, H. T.; MALONE, G. J. The discriminative efficiency of common scaling methods. *Journal of Sensory Studies*, v. 1, p. 85-98, 1986a.
- LAWLESS, H. T.; MALONE, G. J. A comparison of rating scales: sensitivity, replicates and relative measurement. *Journal of Sensory Studies*, v. 1, p. 155-174, 1986b.
- MCDANIEL, M. R.; SAWYER, F. M. Preference testing of whiskey sour formulations: magnitude estimation versus the 9-point hedonic. *Journal of Food Science*, v. 46, p. 182-185, 1981.
- MCPHERSON, R. S; RANDALL, E. Line length measurement as a tool for food preference research. *Ecology of food and Nutrition*, v. 17, p. 149-156, 1985.
- MEAD, R.; GAY, C. Sequential design of sensory trials. *Food Quality and Preference*, v.6, p. 271-280, 1995.
- MEILGAARD, M.; CIVILLE, G. V.; CARR, B. T. *Sensory Evaluation Techniques*. 2nd Ed. CRC Press, Inc. Boca Raton, FL, USA, 1991, 281p.
- MILLER, A. J. Adjusting taste scores for variations in use of scales. *Journal of Sensory Studies*, v. 2, p. 231-242, 1987.
- MOSKOWITZ, H. R.; SIDEL J. L. Magnitude and hedonic scales of food acceptability. *Journal of Food Science*, v. 36, p. 677-680, 1971.
- NAES, T. Handling individual differences between assessors in sensory profiling. *Food Quality and Preference*, v. 2, p. 187-199, 1990.
- O'MAHONY, M. *Sensory Evaluation of Food Statistical Methods and Procedures*. Marcel Dekker, Inc. All Rights Reserved, New York, NY, USA, 1986.
- O'MAHONY, M. Some assumptions and difficulties with common statistics for sensory analysis. *Food Technology*, p. 75-82. 1982.
- PANGBORN, R. M.; GUINARD, J. X.; MEISELMAN, H. L. Evaluation of bitterness of caffeine in hot chocolate drink by category, graphic, and ratio scaling. *Journal of Sensory Studies*, v. 4, p. 31-53, 1989.
- PEARCE, J. H; KORTH, B.; WARREN, C. B. Evaluation of three scaling methods for hedonics. *Journal of Sensory Studies*, v. 1, p. 27-46, 1986.
- PERYAM, D. R. AND PILGRIM, F. J. Hedonic scale method of measuring food preferences. *Food Technology*, 11(9), *Suppl.* p. 9-14, 1957.
- PIGGOTT, J. R. Design questions in sensory and consumer science *Food Quality and Preference*, v.6, p. 217-220, 1995.
- PRITCHETT, M. P. A., Performance assessment of sensory panelists. *Journal of Sensory Studies*, v. 7, p. 229-252, 1992.

- SCHUTZ, H. G.; CARDELLO, A. V. A labeled affective magnitude (LAM) scale for assessing food liking/disliking. *Journal of Sensory Studies*, v. 16, p. 117-159, 2001.
- STONE, H.; SIDEL, J. *Sensory Evaluation Practices*. 2^a Ed. Academic Press, Inc. San Diego, CA, USA, 1993, 308p.
- TEPPER, B.J.; SHAFFER, S. E.; SHEARER, C. M. Sensory perception of fat in common foods using two scaling methods. *Food Quality and Preference*, v.5, p. 245-251, 1994.
- VIE, A.; GULLI, D.; O'MAHONY, M. Alternative hedonic measures. *Journal of Food Science*, v. 56, p. 1-5, 1991.
- VILLANUEVA, N. D. M, PETENATE A. J. & DA SILVA, M. A. A. P. Performance of three affective methods and diagnosis of the ANOVA model, *Food Quality and Preference*, Vol.11, N° 5, pp. 363-370, 2000.
- VILLANUEVA, N. D. M, PETENATE A. J. & DA SILVA, M. A. A. P. *Performance of the hybrid and self-adjusting hedonic scales in the generation of Internal Preference Maps*. In: 2002 Annual Meeting Technical Program and Food Expo. Anaheim, CA, USA, 2002. Abstract Book p. 265-266.
- VILLANUEVA, N. D. M., PETENATE A. J. & DA SILVA, M. A. A. P. *Self-adjusting and hybrid hedonic scales: Advantages and disadvantages with respect to traditional affective methods*. In: The 4th Pangborn Sensory Science Symposium. Dijon, France, 2001. Conference Program Abstracts, p. 136.
- VILLANUEVA, N. D. M.; DA SILVA, M. A. A. P.; PETENATE, A. J. *Discriminating power and reliability of three scales for consumer data collection: Diagnosis of the ANOVA model*. In: Annual Meeting Technical Program and Food Expo. Chicago, 1999. Abstract Book p. 199.
- WILKINSON, C.; YUKSEL, D., Modeling differences between panelists in use of measurement scales. *Journal of Sensory Studies*, v. 12, p. 55-68, 1997.
- WILLIAMS, A. A.; SPENCER, H. W.; CARTER, C. S.; ARNOLD, G. M.; LANGRON, S. *Manual of Sensory Difference Testing of Wines*, Long Ashton Research Station, UK, 1979, 70p.

CAPÍTULO 2

AVALIAÇÃO DO DESEMPENHO DE TRÊS MÉTODOS DE ESCALA E DIAGNÓSTICO DO MODELO ANOVA USUAL ADITIVO EM ESTUDOS COM CONSUMIDORES

Resumo

Os dados provenientes de estudos de consumidores, usualmente coletados através de escalas hedônicas tradicionais, com frequência não satisfazem os pressupostos estatísticos de independência, normalidade e homoscedasticidade. Em complementação, métodos sensoriais que utilizam escalas como instrumento de medida apresentam o problema da variação da expansibilidade entre provadores. A escala autoajustável foi desenvolvida para minimizar alguns destes problemas, porém, ainda não foram reportados resultados sobre seu desempenho em condições reais de teste com consumidor. Este estudo comparou o desempenho da escala hedônica de 9 pontos, escala autoajustável e escala de ordenação com relação às diferenças de expansibilidade entre provadores, poder discriminativo e adequação dos dados coletados por cada escala aos pressupostos do modelo ANOVA. Três marcas comerciais de confeitos foram avaliadas por 288 consumidores. A normalidade dos dados foi avaliada através do cálculo dos coeficientes de assimetria e curtose, gráfico de probabilidade normal e o teste de Kolmogorov-Smirnov. A homoscedasticidade foi avaliada através de gráficos de dispersão e testes de Levene. A escala autoajustável foi efetiva para tratar o problema da expansibilidade e homoscedasticidade, porém, os dados gerados por esta escala mostraram moderadas distorções da normalidade. A escala hedônica apresentou problemas de heteroscedasticidade. As escalas autoajustável e de ordenação apresentaram o menor e o maior poder discriminativo, respectivamente. Apesar dos problemas detectados, as três escalas apresentaram as mesmas tendências de preferência dos produtos avaliados.

Abstract

Consumer data arising from traditional scales frequently do not comply with ANOVA assumptions of normality, independence and homoscedasticity. Differential expansiveness scoring between assessors is also a limitation for the use of scales. The self-adjusting scale was designed to minimize some of these problems but its performance with consumers has not been reported. This study compared the performance of the 9-point hedonic scale, the self-adjusting scale and the ranking test with respect to: differential expansiveness between assessors, discriminating power and compliance of data with ANOVA assumptions. Candies were evaluated by 288 consumers. Data normality was checked by normal probability plots, Kolmogorov-Smirnov test and coefficients of skewness and kurtosis. Homoscedasticity assumption was evaluated by Levene test and scatter plots. The self-adjusting scale was effective to deal with differential assessor expansiveness, produced homogeneous variances but its data showed moderate deviations from normality. Hedonic scale data showed lack of homoscedasticity. Rank and self-adjusting scale showed the highest and the lowest discriminating power respectively. Despite the problems detected, the three scales presented the same tendencies for preference amongst the products tested.

1. INTRODUÇÃO

Na indústria alimentícia moderna, a avaliação das características sensoriais dos alimentos, da forma como elas são percebidas pelos consumidores, é um componente essencial no processo de desenvolvimento, otimização, melhoria da qualidade e avaliação do potencial de mercado dos produtos (SIDEL & STONE, 1993; PIGOTT, 1995).

Dentro das diversas metodologias utilizadas pela Análise Sensorial, encontram-se os testes afetivos, cujo propósito principal, é a avaliação da preferência e/ou aceitação dos consumidores com relação a produtos e serviços (MEILGAARD *et al.*, 1987). Uma vez que a resposta dos consumidores é via de regra coletada através de escalas, estas têm sido objeto de estudo de muitos cientistas que atuam na área

de Análise Sensorial. Vários estudos sobre desempenho de escalas têm avaliado a eficiência, o poder discriminativo e a facilidade de uso das mesmas por consumidores (PEARCE *et. al.*, 1986; GIOVANNI & PANGBORN, 1983; LAWLESS & MALONE, 1986a, 1986b; TEPPER *et. al.*, 1994; CHEN & RESURRECCION, 1996; HOUGH *et. al.*, 1997).

Uma das limitações identificadas em metodologias de avaliações sensoriais que utilizam escalas, é o da variação da expansibilidade entre provadores, decorrente do fato de diversos provadores utilizarem diferentes porções da escala para expressarem a sua impressão sobre o produto (GAY, 1988; GAY & MEAD, 1992; MEAD & GAY, 1995). A variação da expansibilidade entre provadores constitui-se em problema particularmente grave no caso dos testes realizados com consumidores, porque produz variação considerável nos efeitos principais analisados pelo experimento, quais sejam: diferenças entre tratamentos (GAY & MEAD, 1992; MEAD & GAY, 1995).

Além da problemática da variação da expansibilidade entre provadores, uma outra limitação que deve ser tratada com maior rigor pela Análise Sensorial, refere-se ao fato de que a maioria dos testes estatísticos aplicados a medidas sensoriais, assumem que os dados amostrais procedem de populações normalmente distribuídas. Em particular, os modelos lineares normais para análise da variância, teste muito utilizado em medidas sensoriais, partem dos pressupostos estatísticos de independência, normalidade e homoscedasticidade das respostas experimentais. Da mesma forma, a maioria dos testes estatísticos utilizados para comparações de médias, tais como o da Mínima Diferença Significativa (teste t), Duncan, Tukey, Scheffé, e outros, são testes paramétricos que exigem normalidade nos dados.

Infelizmente, o que se observa com razoável freqüência, é que os dados provenientes das avaliações sensoriais não satisfazem os pressupostos estatísticos de independência, normalidade e homoscedasticidade. Em particular, dados provenientes de escalas hedônicas estruturadas, violam a suposição de normalidade e homoscedasticidade, pois, mesmo que valores numéricos sejam atribuídos às

categorias, reconhecidamente, eles não refletem iguais diferenças em percepção, embora numericamente os intervalos sejam iguais. (O'MAHONY, 1982; VIE *et al.*, 1991; GAY & MEAD, 1992).

PEARCE *et al.* (1986) e WILKINSON & YUKSEL (1997), afirmam que com frequência, os dados sensoriais produzem distribuições assimétricas e, portanto, não normais. Em testes de consumidores utilizando escalas hedônicas estruturadas, este fato ocorre quando um dos produtos tem, de forma muito consensual, alta ou baixa aceitação com relação aos demais, o que resulta em frequências altas de notas entre 7 e 9 ou entre 1 e 3 para aquele produto. De fato, GIOVANNI & PANGBORN (1983), observaram distribuições bimodais das respostas hedônicas ao utilizarem a escala não estruturada, o que infringe claramente a suposição de normalidade dos dados.

Visando contornar problemas relativos à independência, homoscedasticidade e normalidade de dados sensoriais, vários autores têm proposto métodos alternativos de geração e análise estatística desses dados (NAES, 1990; PRITCHETT, 1992; WILKINSON & YUKSEL, 1997).

Para minimizar o problema da expansibilidade dos provadores em testes sensoriais, WILLIAMS *et al.* (1979), GAY (1988) e GAY & MEAD (1992), propuseram um novo tipo de escala chamado de "autoajustável", que procura utilizar somente a essência da informação contida num conjunto de respostas sensoriais: a ordem de preferência das amostras e seu grau relativo de diferença. Os autores realizaram estudos com equipes treinadas para verificar a eficácia desta nova escala, e relataram maior poder discriminativo e coeficientes de expansibilidade substancialmente menores para a escala autoajustável em relação à escala de magnitude. Em função das vantagens observadas e também pela sua facilidade de uso por parte dos provadores, os pesquisadores recomendaram o estudo do desempenho da escala autoajustável em pesquisas junto a consumidores.

Uma vez que a utilização da escala autoajustável apresenta-se hoje como uma possibilidade a ser investigada, em vista da novidade do método e da

interpretabilidade da informação que se espera obter, o presente trabalho teve por objetivos:

1. Avaliar em condições reais de teste de consumidor, as vantagens e desvantagens da escala autoajustável, verificando o desempenho do método em relação a outros tipos de escala tradicionalmente utilizadas em pesquisas de consumidores, quais sejam: escala hedônica estruturada de 9 pontos e escala de ordenação.
2. Comparar os três métodos listados acima com relação ao poder discriminativo, variabilidade das respostas sensoriais e adequação dos dados coletados às suposições estatísticas para análise da variância (ANOVA).

2. MATERIAIS E MÉTODOS

2.1. Produtos

Foram avaliados três diferentes formulações comerciais de confeito tipo bala. A aceitação de cada amostra foi avaliada em relação à aparência e impressão global após degustação.

2.2. Provadores

Participaram do estudo 288 consumidores habituais de confeitos, sendo 84 homens e 204 mulheres, de idade variando entre 14 e 18 anos, provenientes de escolas de segundo grau da cidade de Campinas.

2.3. Delineamento Experimental e Modelo Estatístico

O delineamento experimental utilizado para apresentação das amostras aos consumidores e geração de respostas, foi o delineamento tipo "cross-over", balanceado para efeitos de ordem de apresentação de amostras e efeitos residuais de primeira ordem. No presente trabalho, este delineamento foi construído utilizando

o algoritmo apresentado em MACFIE *et al.* (1989), uma vez que esses autores só apresentam tabelas para $n \geq 4$ amostras. Assim sendo, os resultados experimentais do presente estudo foram analisados mediante o modelo estatístico para análise da variância associado a este tipo de delineamentos, o qual explicita-se a seguir,

$$Y_{ij} = \mu + p_i + o_j + a_{d(i,j)} + r_{d(i,j-1)} + \varepsilon_{ij}, \quad i=1, \dots, 96; j=1, \dots, 3$$

Onde,

Y_{ij} : resposta do i -ésimo provador na avaliação da amostra $d_{(i,j)}$ apresentada na j -ésima ordem

μ : média geral

p_i : efeito do i -ésimo provador

o_j : efeito da j -ésima ordem de apresentação da amostra

$a_{d(i,j)}$: efeito direto da amostra $d_{(i,j)}$, avaliada pelo i -ésimo provador na j -ésima ordem de apresentação

$r_{d(i,j-1)}$: efeito residual de primeira ordem da amostra $d_{(i,j-1)}$, avaliada pelo i -ésimo provador na $(j-1)$ -ésima ordem de apresentação, onde $j \geq 2$ e $r_{d(i,0)} = 0$

ε_{ij} : perturbações aleatórias normais independentes e identicamente distribuídas com $E(\varepsilon_{ij}) = 0$ e $Var(\varepsilon_{ij}) = \sigma^2$

2.4 Métodos Sensoriais Afetivos

Para realização dos testes foram utilizados três diferentes métodos afetivos: escala hedônica estruturada de 9 pontos (Figura 1); escala de ordenação, através da qual o provador classificou as amostras em ordem decrescente de preferência (Figura 2) e escala autoajustável (Figura 3). Os 288 consumidores foram divididos aleatoriamente em três grupos distintos de 96 indivíduos cada. Desta forma, 96

indivíduos avaliaram as três amostras através da escala hedônica de 9 pontos, 96 através da escala de ordenação e 96, através da escala autoajustável.

Os testes sensoriais foram realizados no Laboratório de Análise Sensorial da Faculdade de Engenharia de Alimentos da UNICAMP, o qual dispunha de 7 cabines individuais. Os provadores foram conduzidos às cabines onde após se acomodarem, receberam as instruções sobre o uso da escala, a natureza dos produtos e os atributos a serem avaliados. No caso das escalas de ordenação e autoajustável, as três amostras foram avaliadas comparativamente, já na escala hedônica de 9 pontos, as 3 amostras foram avaliadas monadicamente.

Aproximadamente 10g de cada amostra foram servidos em pratos de fundo branco, codificados com números aleatórios de três dígitos e dispostos sob luz fluorescente branca. Os provadores avaliaram primeiro o quanto gostaram ou desgostaram da aparência, e após a degustação, avaliaram o quanto gostaram ou desgostaram de cada amostra de uma forma global. Foi oferecido juntamente com as amostras, um copo de água para enxágüe da boca entre uma amostra e outra. Os consumidores que utilizaram a escala hedônica para suas avaliações, foram também solicitados a descreverem o que mais gostaram e menos gostaram de cada amostra. Nas demais escalas, os provadores fizeram seus comentários sem que fossem explicitamente solicitados.

Nome: _____ Data: _____

Você esta recebendo 3 amostras codificadas de confeito. Avalie da esquerda para a direita a aparência de cada uma das amostras e use a escala abaixo para indicar o quanto você gostou ou desgostou das mesmas.

9 Gostei extremamente
 8 Gostei muito
 7 Gostei moderadamente
 6 Gostei ligeiramente
 5 Nem gostei / nem desgostei
 4 Desgostei ligeiramente
 3 Desgostei moderadamente
 2 Desgostei muito.
 1 Desgostei extremamente

AMOSTRA N^o: _____ VALOR: _____

Comentários: _____

Figura 1. Ficha de avaliação da aceitação utilizando a escala hedônica de 9 pontos.

Nome: _____ Data: _____

Você esta recebendo 3 amostras codificadas de confeito. Avalie por favor, da esquerda para a direita, a APARÊNCIA das amostras e ordene-as em ordem decrescente de preferência, com relação à APARÊNCIA.

+ preferida _____ - preferida

Comentários: _____

Figura 2. Ficha de avaliação da preferência utilizando a escala de ordenação.

Nome: _____ Data: _____

Você esta recebendo 3 amostras codificadas de confeito. Avalie por favor, a APARÊNCIA das amostras da esquerda para direita. Escolha primeiro a amostra que você mais gostou e depois a que você menos gostou e anote os códigos nos extremos direito e esquerdo da escala abaixo, respectivamente. Avalie agora as outras amostras de forma comparativa com aquelas alocadas nos extremos, marcando um "x" e o código da amostra no ponto da escala que você achar que melhor represente o quanto você gostou ou desgostou de cada amostra.

x _____ x
 Menos gostoso Mais gostoso

Comentários: _____

Figura 3. Ficha de avaliação sensorial utilizando a escala autoajustável.

2.5. Análise Estatística

Inicialmente para cada escala, procedeu-se uma análise exploratória dos dados referente a cada amostra através do cálculo de estatísticas básicas como a média, mediana, mínimo, máximo e desvio padrão. Essas análises foram complementadas com histogramas de frequências para cada amostra em cada atributo avaliado, calculando-se para cada histograma os seus coeficientes de assimetria e curtose. Esses procedimentos foram realizados com o intuito de se visualizar a variabilidade, a tendência central e a forma da distribuição dos dados.

Os resultados obtidos através das escalas hedônica e autoajustável foram analisados através de ANOVA ($p=5\%$) para delineamentos tipo “cross-over”, considerando-se no modelo os efeitos: provador, ordem de apresentação, amostras e efeito residual. Pelo fato de todos os provadores terem experimentado todas as amostras, para análise dos dados gerados através da escala autoajustável, não foi necessária a utilização do método iterativo de duas etapas proposto por GAY (1988), utilizando apenas o procedimento de análise da variância usual, conforme sugerido por WILLIAMS *et al.* (1979). Para avaliar as diferenças entre as amostras foram realizadas comparações pareadas das médias, aplicando-se o teste de Tukey (MONTGOMERY, 1984).

Os resultados da escala de ordenação foram analisados utilizando-se o teste não paramétrico de Friedman. O mesmo teste foi aplicado aos dados de ordenação da escala hedônica e autoajustável, para comparar a eficiência das três escalas (STONE & SIDEL, 1993). Para detecção das diferenças das preferências entre as amostras, foram feitas comparações pareadas dos totais de ordenação, aplicando-se também o teste não paramétrico de Friedman, cujos valores críticos para um grande número de provadores encontram-se tabelados em BASKER (1988).

O parâmetro utilizado para se comparar o poder discriminativo das escalas hedônica de 9 pontos e autoajustável foi o nível de significância do efeito devido à amostra, obtido na ANOVA ($p_{amostra}$).

O diagnóstico da adequação dos resultados das escalas hedônica de 9 pontos e autoajustável aos modelos ANOVA, foram realizados através de testes de Kolmogorov-Smirnov e gráficos da probabilidade normal dos resíduos, para verificação da normalidade. Testes de Levene (LEVENE, 1960) e gráficos de dispersão dos desvios padrões em função das médias das amostras foram utilizados para verificação da homoscedasticidade.

A variação da expansibilidade entre provadores nas escalas hedônica de 9 pontos e autoajustável foi avaliada através do valor de pF_{provador} fornecido pela ANOVA e teste de Levene, para homogeneidade das variâncias dos provadores. A variabilidade das respostas foi avaliada através do erro experimental ($QM_{\text{resíduo}}$) e dos desvios padrões das amostras.

Para análise dos dados foram utilizados os programas estatísticos SAS (Statistical Analysis System), Versão 8.0 e Statistica, Versão 5.5.

3. RESULTADOS

3.1 Escala Autoajustável x Escala Hedônica

3.1.1 Análise exploratória dos dados: Estatísticas básicas referentes aos dados coletados através das escalas hedônica e autoajustável são mostradas nas Tabelas 1 e 2, onde se observa que as médias de aceitação das amostras apresentam uma mesma tendência para ambas as escalas: na avaliação da aparência, em ambas as escalas, a amostra 2 foi a mais aceita e a 3 a menos aceita, enquanto que na avaliação global, a amostra 1 foi a mais aceita e a amostra 3 a menos aceita.

Os valores da mediana na escala hedônica (Tabelas 1 e 2), encontram-se bastante próximos aos da média, enquanto que na escala autoajustável, para as amostras de maior e menor aceitação, há uma grande diferença entre a média e a mediana, indicando que no mínimo 50% dos provadores fizeram suas avaliações num dos extremos das escalas. Esses resultados indicam uma razoável assimetria nos

dados obtidos através da escala auto-ajustável sugerindo cuidado no uso da ANOVA para análise dos mesmos.

Tabela 1. Estatísticas básicas dos dados da avaliação da aparência de confeitados coletados através da escala hedônica de 9 pontos e da escala autoajustável (n=96 provadores)

Escaia	Amostra	Média ⁽¹⁾	Mediana	Mínimo	Máximo	Desvio Padrão
Hedônica ⁽²⁾	1	6,4 ^a	6,0	2	9	1,50
	2	7,3 ^b	7,0	4	9	1,23
	3	5,0 ^c	5,0	1	9	2,09
Autoajustável ⁽³⁾	1	6,0 ^a	6,5	0	10	3,61
	2	7,2 ^a	10,0	0	10	3,64
	3	2,4 ^b	0,0	0	10	3,83

- (1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ($p=5\%$). Esses resultados foram obtidos através do teste de Tukey.
 (2) 1= desgostei extremamente, 9= gostei extremamente.
 (3) 0= menos gostoso, 10= mais gostoso

Tabela 2. Estatísticas básicas dos dados da avaliação da aceitação global de confeitados coletados através da escala hedônica de 9 pontos e da escala autoajustável (n=96 provadores)

Escaia	Amostra	Média ⁽¹⁾	Mediana	Mínimo	Máximo	Desvio Padrão
Hedônica ⁽²⁾	1	7,5 ^a	8,0	4	9	1,33
	2	6,6 ^b	7,0	1	9	2,15
	3	4,8 ^c	5,0	1	9	2,33
Autoajustável ⁽³⁾	1	7,8 ^a	10,0	0	10	3,37
	2	6,1 ^b	6,4	0	10	3,48
	3	1,8 ^c	0,0	0	10	3,51

- (1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ($p=5\%$). Esses resultados foram obtidos através do teste de Tukey.
 (2) 1= desgostei extremamente, 9= gostei extremamente.
 (3) 0= menos gostoso, 10= mais gostoso

As Figuras de 4 a 7 mostram as distribuições dos dados gerados através das escalas hedônica e autoajustável. No caso da escala hedônica, observa-se uma moderada assimetria à esquerda para as amostras mais aceitas (amostra 1 e 2) e um razoável achatamento para a amostra de menor aceitação (amostra 3). No caso da escala autoajustável os problemas de assimetria e achatamento ocorreram de forma mais grave que na escala hedônica (Figuras 6 e 7).

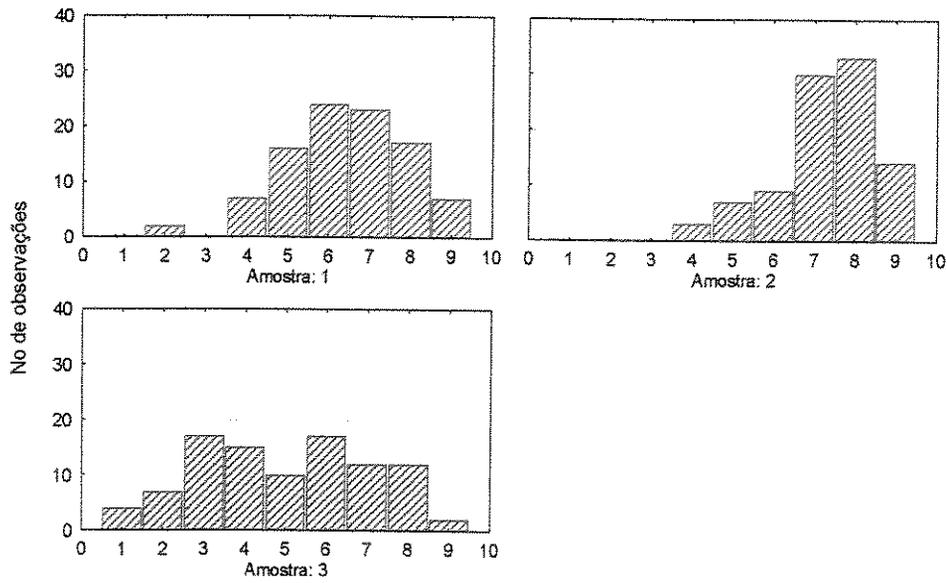


Figura 4. Histogramas de freqüências dos dados de aceitação da aparência gerados através da escala hedônica, para cada amostra.

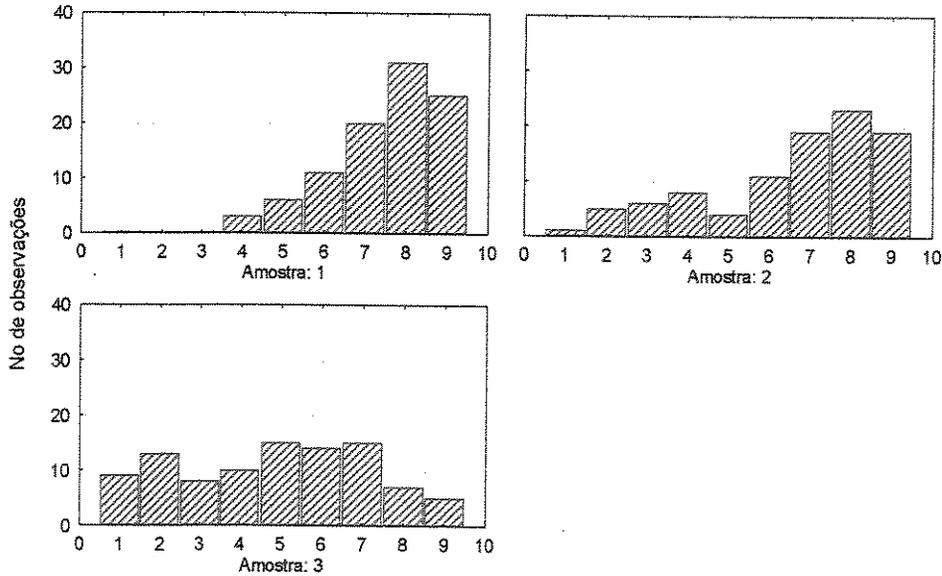


Figura 5. Histogramas de freqüências dos dados de aceitação global gerados através da escala hedônica, para cada amostra.

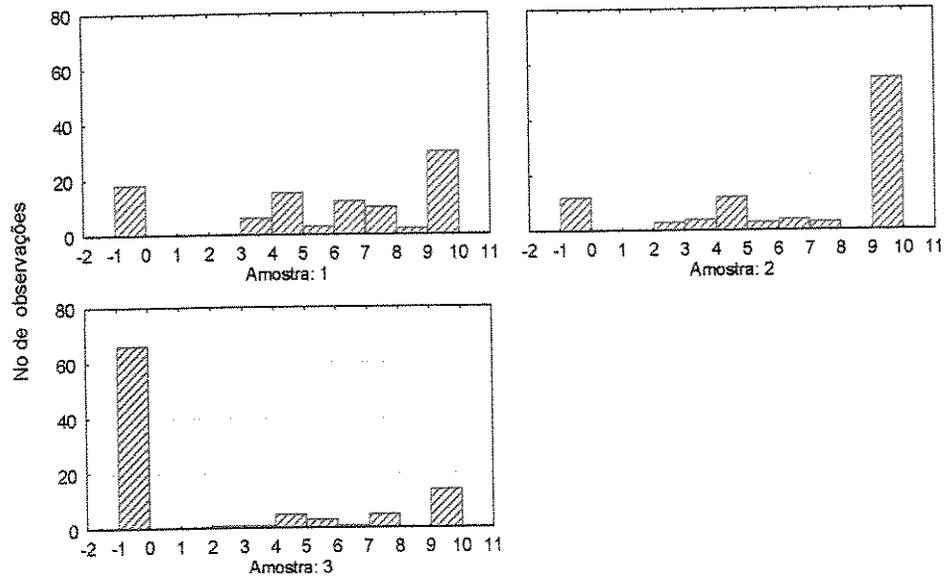


Figura 6. Histogramas de freqüências dos dados de aceitação da aparência gerados através da escala autoajustável, para cada amostra.

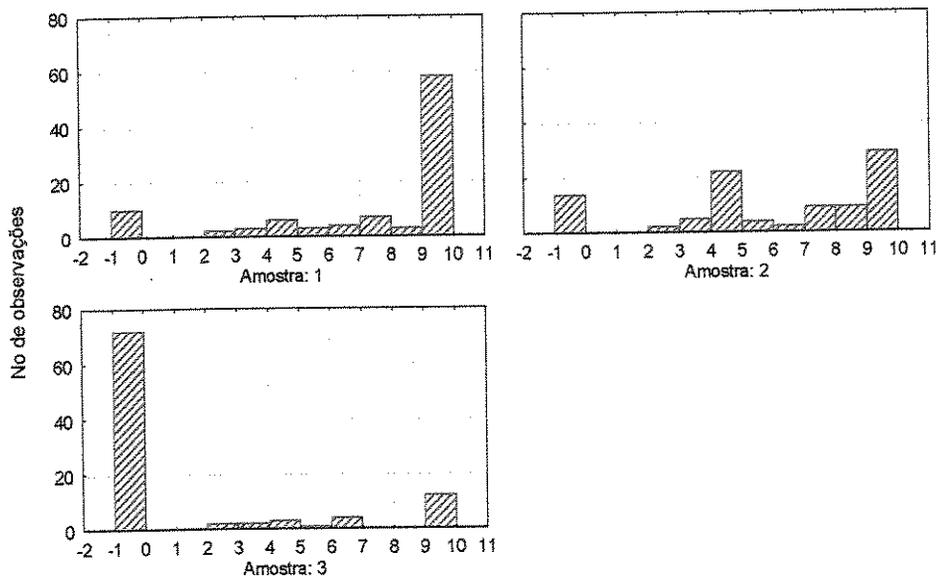


Figura 7. Histogramas de freqüências dos dados de aceitação global gerados através da escala autoajustável, para cada amostra.

Quando os coeficientes de assimetria e curtose foram calculados para se avaliar o grau de afastamento da simetria e de achatamento/apontamento das distribuições (Tabelas 3 e 4) em relação à distribuição normal, confirmou-se maiores desvios para a escala auto-ajustável que para a escala hedônica.

Tabela 3. Coeficientes de assimetria e curtose dos dados da aparência de confeitos coletados através da e escala hedônica de 9 pontos e da escala autoajustável (n=96 provadores).

Escala	Amostra	Coef. de assimetria ⁽¹⁾	Coef. de curtose ⁽²⁾
Hedônica	1	-0,40	0,21
	2	-0,77	0,32
	3	-0,02	-1,00
Autoajustável	1	-0,47	-1,02
	2	-0,89	-0,66
	3	1,17	-0,34

(1) valores < 0 indicam assimetria à esquerda, > 0 indicam assimetria à direita e =0 indicam normalidade dos dados. (2) valores < 0 indicam achatamento, > 0 indicam distribuição em ponta e =0 indicam normalidade dos dados.

Os coeficientes de assimetria das distribuições dos dados gerados pela escala hedônica variaram entre -0,02 e -0,77 para a aparência e entre -0,08 e -0,85 para a aceitação global (Tabelas 3 e 4), indicando grau de assimetria à esquerda. Essa assimetria foi maior para as amostras mais aceitas. Em relação à curtose, estes valores variaram entre -1,00 e 0,32 (Tabela 3), para aparência e -1,04 e 0,05 (Tabela 4) para avaliação global, mostrando que quanto menor a aceitação, mais achatada foi a distribuição.

No caso das distribuições geradas pela escala autoajustável, os coeficientes de assimetria dos dados de aparência variaram entre -0,89 e 1,17 (Tabela 3) e para avaliação global entre -1,33 e 1,63 (Tabela 4), mostrando alto grau de assimetria pela direita para a amostra mais aceita e pela esquerda, para a menos aceita. Já os coeficientes de curtose da escala autoajustável para os valores de aceitação da aparência, variaram entre -1,02 e -0,34 (Tabela 3) e para aceitação global entre -1,01 e 1,05 (Tabela 4), mostrando tanto distribuições achatadas, quanto pontiagudas. Assim os dados obtidos através da escala autoajustável mostram maior

restrição ao uso da ANOVA para sua análise, do que aqueles gerados através da escala hedônica.

Tabela 4. Coeficientes de assimetria e curtose dos dados da avaliação global de confeitos coletados através da e escala hedônica de 9 pontos e da escala autoajustável (n=96 provadores)

Escala	Amostra	Coef. de assimetria⁽¹⁾	Coef. de curtose⁽²⁾
Hedônica	1	-0,82	0,05
	2	-0,85	-0,29
	3	-0,08	-1,04
Autoajustável	1	-1,33	0,45
	2	-0,44	-1,01
	3	1,63	1,05

(1) valores < 0 indicam assimetria à esquerda, > 0 indicam assimetria à direita e =0 indicam normalidade dos dados. (2) valores < 0 indicam achatamento, > 0 indicam distribuição em ponta e =0 indicam normalidade dos dados.

Os valores dos desvios padrões mostram claramente que houve maior variabilidade nos dados produzidos pela escala autoajustável (Tabelas 1 e 2). Assim, se por um lado esperava-se aumento do poder discriminativo da escala autoajustável em relação à escala hedônica devido a que a primeira elimina o problema da expansibilidade entre os provadores, por outro lado os resultados mostram que a escala autoajustável produziu maior variabilidade, o que pode ter concorrido na diminuição da sensibilidade do experimento para detectar diferenças significativas entre as amostras.

3.1.2 Efeito da ordem de apresentação das amostras e efeito residual

Na análise da variância dos dados obtidos através das escalas hedônica tradicional e autoajustável, não foram detectados efeitos de ordem de apresentação das amostras para nenhuma das duas escalas, tanto para os dados da aparência como para os dados de aceitação global ($p \geq 0,16$). Porém, os resultados da ANOVA revelaram efeitos residuais de primeira ordem significativos para os dados obtidos através da escala autoajustável, tanto na avaliação da aparência ($p = 0,004$) como na avaliação da aceitação global ($p = 0,002$). Já na análise dos dados coletados através

da escala hedônica de 9 pontos, não foram detectados efeitos residuais significativos ($p \geq 0,50$) para nenhum dos atributos avaliados.

A Tabela 5, apresenta para ambas as escalas, os valores hedônicos médios de aceitação da aparência e de aceitação global, atribuídos pelos provadores após terem avaliado cada uma das três amostras. Na avaliação da aparência, observa-se que os provadores tenderam a dar notas maiores para a amostra que precedeu à pior amostra (amostra 3) e notas menores, para a amostra que precedeu à melhor amostra (amostra 2). Uma tendência semelhante foi observada na avaliação da aceitação global, onde também os provadores tenderam a dar notas maiores após a avaliação da pior amostra (amostra 3) e notas menores após a avaliação da melhor amostra (amostra 1). Isto significa que as avaliações sensoriais da aparência e da aceitação global, foram bastante influenciadas pelo efeito de contraste entre as amostras (MEILGAARD *et al.*, 1987). Uma possível explicação para esses resultados, seria o fato da escala autoajustável dar ênfase à identificação das amostras extremas que funcionam como âncoras ou referências, maximizando o efeito residual/contraste; notadamente, no caso de se avaliar apenas três amostras. Neste caso, o efeito residual provocado pela ancoragem seria ainda maior, visto que a terceira amostra será avaliada inteiramente em função das amostras extremas.

Pelo exposto, não apenas o número ou o tipo de amostras influenciam sobre o efeito residual, como também o tipo de escalonamento utilizado para avaliação das amostras. Portanto, um aspecto importante na geração de dados confiáveis é a escolha da escala a ser utilizada. Esta escolha deve fundamentar-se em conhecimentos prévios acerca de aspectos psicométricos da escala, assim como das suas vantagens e limitações. Caso contrário, podem ser obtidos dados viciados e tiradas conclusões pouco válidas sobre os resultados experimentais.

Tabela 5. Médias⁽¹⁾ de aceitação da aparência e aceitação global dos valores hedônicos atribuídos pelos provadores após avaliarem cada uma das três amostras, referentes aos dados coletados através das escalas hedônica de 9 pontos e autoajustável (n=96 provadores).

Atributo	Amostra	Escala hedônica de 9 pontos ⁽²⁾	Escala Autoajustável ⁽³⁾
Aparência	1	6,09 a	6,15 a
	2	5,86 a	3,42 b
	3	6,44 a	5,54 a
Aceitação global	1	5,26 a	4,85 ab
	2	6,31 a	3,61 b
	3	6,74 a	6,05 a

(1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ($p \leq 0,05$). Esses resultados foram obtidos através do teste de Tukey.

(2) 1= desgostei extremamente, 9= gostei extremamente.

(3) 0= menos gostoso, 10= mais gostoso.

3.1.3 Análise do poder discriminativo: os resultados da análise da variância dos dados de aceitação com relação à aparência e avaliação global para ambas as escalas, são mostradas na Tabela 6.

Tabela 6. Valores de QM_{erro} , pF_{amostra} e pF_{provador} obtidos da ANOVA dos valores de aceitação gerados pela escala hedônica e autoajustável (n=96 provadores).

Escala	Atributo	QM_{erro}	pF_{amostra}	$pF_{\text{provador}}^{(1)}$
Hedônica	Aparência	2,863	$p=0,0001$	$p=0,7690$
	Aceitação global	3,540	$p=0,0001$	$p=0,0761$
Autoajustável	Aparência	18,872	$p=0,0001$	$p=1,000$
	Aceitação global	16,489	$p=0,0001$	$p=1,000$

(1) $pF_{\text{provador}} \leq 0,05$ indica problema de expansibilidade entre provadores.

Os valores de pF_{amostra} em ambas as escalas foram altamente significativos ($p < 0,001$) sugerindo alto poder discriminativo para ambas as escalas. O teste de Tukey para comparação de médias (Tabelas 1 e 2), mostrou que a escala hedônica discriminou todas as amostras entre si ($p < 0,05$) com relação aos dois atributos julgados. O poder discriminativo da escala autoajustável foi ligeiramente inferior ao da escala hedônica uma vez que, na escala autoajustável, não foram detectadas diferenças entre as amostras 1 e 2, para o atributo aparência. Um menor poder discriminativo da escala autoajustável com relação à escala hedônica pode ser

justificado em função dos altos valores de QM_{erro} produzidos pela ANOVA dos dados gerados pela escala autoajustável (Tabela 6).

3.1.4. Diagnóstico de normalidade dos resíduos do modelo ANOVA: os gráficos para diagnóstico da normalidade dos resíduos do modelo ANOVA dos dados da aparência e avaliação global obtidos através da escala hedônica (Figuras 8 e 9) e da escala autoajustável (Figuras 10 e 11), não parecem mostrar grandes desvios da normalidade para nenhum dos atributos. No caso da escala hedônica, o teste de Kolmogorov-Smirnov (Tabela 7) não rejeitou a hipótese de normalidade dos resíduos para nenhum dos atributos julgados ($p \geq 0,15$). Já na escala autoajustável, foram detectados moderados desvios da normalidade dos resíduos para a avaliação da aparência ($p=0,015$). Porém, para a avaliação global, não foi rejeitada a hipótese de normalidade dos resíduos ($p=0,096$).

Em geral, nos modelos de análise de variância com efeitos fixos, o teste F é bastante robusto para ligeiros e moderados desvios da normalidade, principalmente, quando o número de graus de liberdade do erro é grande (SCHEFFÉ, 1959). Uma distribuição dos resíduos com caudas consideravelmente mais pesadas (curtose negativa) ou leves (curtose positiva) que as da distribuição normal, causam problemas mais sérios que uma distribuição assimétrica. Se a curtose é positiva, o valor de F tenderá a ser tão pequeno que não se poderá rejeitar a hipótese nula mesmo ela sendo falsa. Se a curtose é negativa, o valor de F tenderá a ser grande, levando à rejeição da hipótese nula mesmo ela sendo verdadeira. A assimetria dos resíduos, em geral não tem um efeito medível no valor da estatística F. O poder do teste e os níveis de significância obtidos no teste F, são apenas aproximados quando o pressuposto da normalidade dos resíduos é violada (BOX & ANDERSEN, 1955).

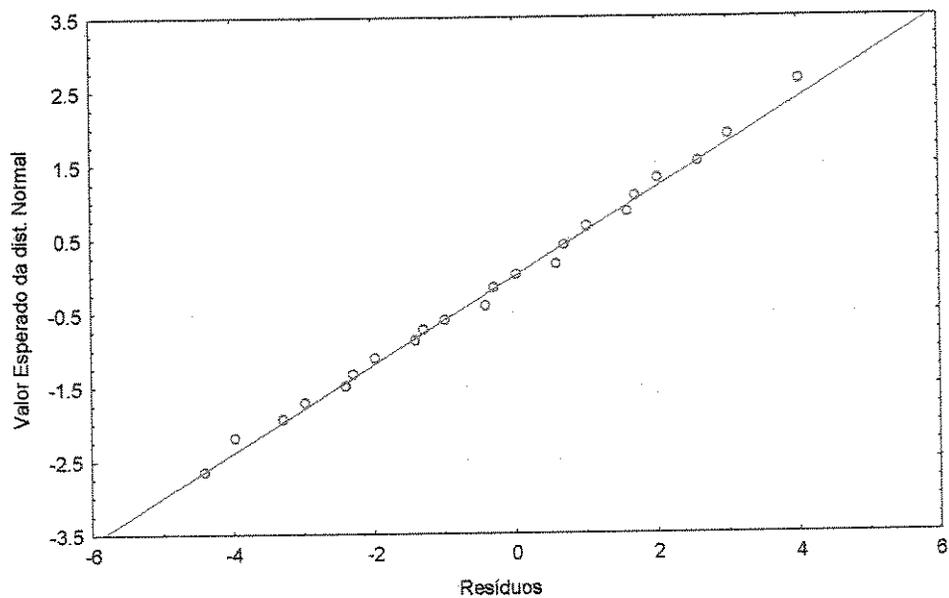


Figura 8. Gráfico da probabilidade normal dos resíduos do modelo ANOVA dos dados de aparência utilizando escala hedônica.

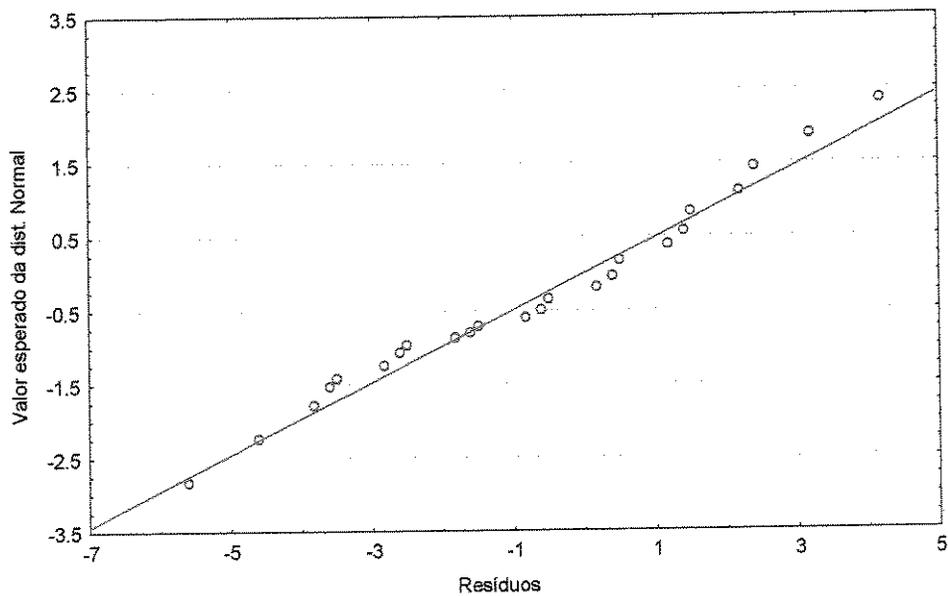


Figura 9. Gráfico da probabilidade normal dos resíduos do modelo ANOVA dos dados de avaliação global utilizando escala hedônica.

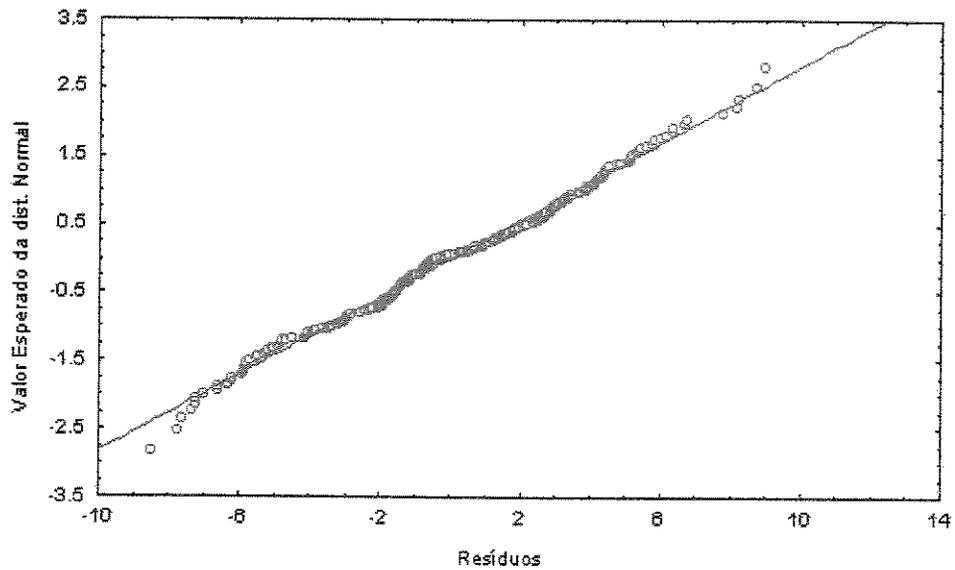


Figura 10. Gráfico da probabilidade normal dos resíduos do modelo ANOVA dos dados de aparência utilizando escala autoajustável.

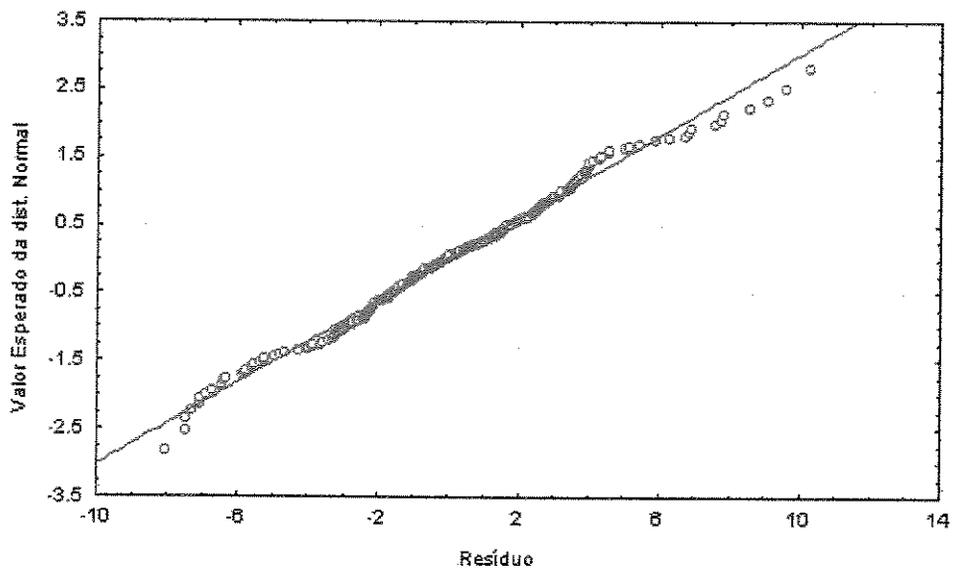


Figura 11. Gráfico da probabilidade normal dos resíduos do modelo ANOVA dos dados de avaliação global utilizando escala autoajustável.

3.1.5 Diagnóstico da homoscedasticidade: os gráficos de dispersão dos desvios padrões das amostras em função de suas respectivas médias de aceitação, obtidas através da escala hedônica, sugeriram correlações negativas entre a variabilidade dos dados e o grau de aceitação dos produtos (Figuras 12 e 13). De fato, as Tabelas 1 e 2 mostram claramente que enquanto os desvios padrões correspondentes às avaliações realizadas através da escala autoajustável foram homogêneos, os desvios associados à escala hedônica foram menores para as amostras de maior aceitação. Os resultados do teste de Levene (Tabela 7) confirmam que enquanto a hipótese de homogeneidade das variâncias foi satisfeita para a escala autoajustável ($p=0,73$), o contrário aconteceu para a escala hedônica, onde se obtiveram valores altamente significativos ($p<0,001$).

Se o pressuposto de homogeneidade das variâncias não é satisfeita, o teste F é ligeiramente afetada nos modelos de análise da variância com efeitos fixos. Porém, o problema é mais sério quando as variâncias tendem a ser função das médias, ou em casos onde uma das variâncias é bastante maior em relação às demais. Nestes casos, os níveis de significância produzidos pelo teste F, os intervalos de confiança para as médias e o poder do teste podem ser bastante afetados (SCHEFFÉ, 1959, MONTGOMERY, 1984).

3.1.6 Variação da expansibilidade: os valores de pF_{provedor} obtidos na ANOVA dos dados da escala hedônica (Tabela 6), indicam que não houve diferenças nas médias dos valores hedônicos dos provedores na avaliação da aparência das amostras ($p=5\%$). Já no caso da avaliação da aceitação global, foi atingido um nível de 7,61% de significância para o efeito provedor, indicando ligeiras diferenças em posição para os valores hedônicos dos provedores. Na escala autoajustável, resultaram valores de pF_{provedor} próximos a 1, indicando que não houve diferenças nas médias dos provedores.

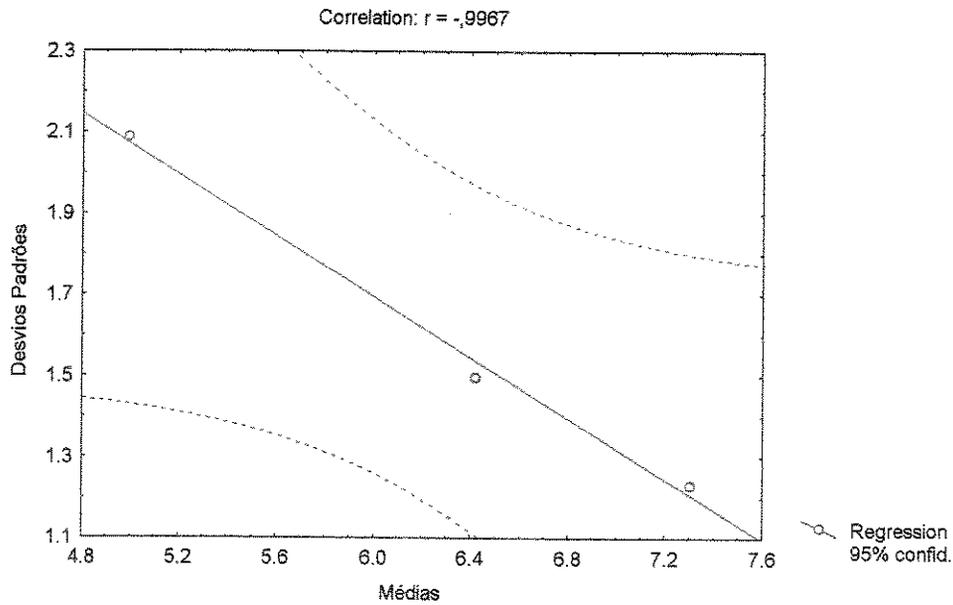


Figura 12. Gráfico de dispersão das médias de aceitação da aparência das amostras vs. os desvios padrões, correspondente aos dados coletados através da escala hedônica de 9 pontos.

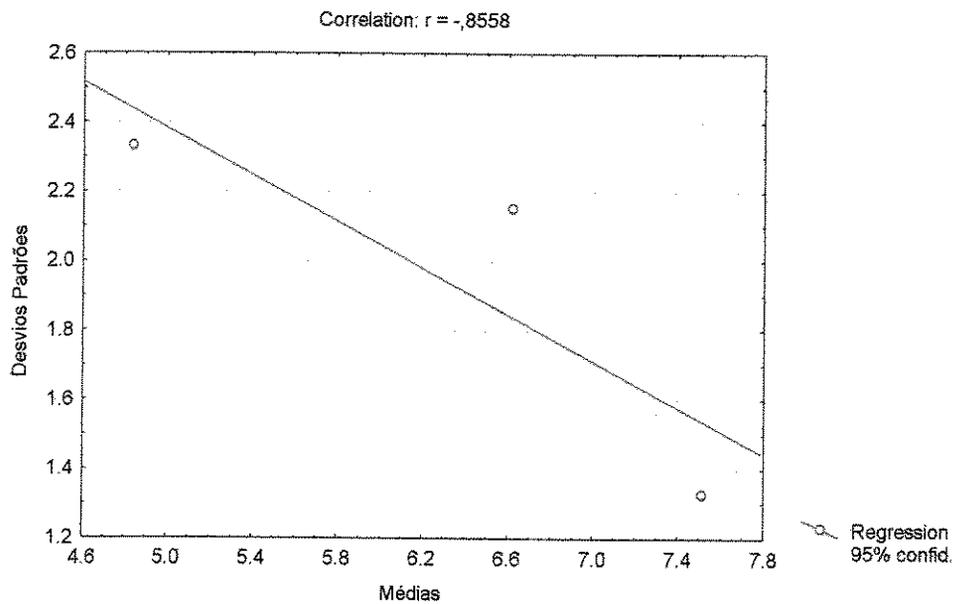


Figura 13. Gráfico de dispersão das médias de aceitação global das amostras vs. os desvios padrões, correspondente aos dados coletados através da escala hedônica de 9 pontos.

Também, os resultados do teste de Levene (Tabela 7) confirmam que enquanto a hipótese de homogeneidade das variâncias dos provadores foi satisfeita para a escala autoajustável ($p \approx 1,00$), o contrário aconteceu para a escala hedônica, onde se obtiveram valores altamente significativos ($p < 0,001$). Desta forma, verificaram-se na escala hedônica, diferenças em dispersão (variâncias) dos valores hedônicos correspondentes aos provadores, ocasionado pelas diferenças individuais no uso efetivo da amplitude da escala, confirmando-se a existência de variação da expansibilidade entre provadores. Como era esperado, na escala autoajustável, não se observaram diferenças em posição nem em dispersão dos dados sensoriais correspondentes aos provadores, uma vez que nesta escala, o provador utiliza o comprimento total da escala para fazer suas avaliações, eliminando a priori o problema da variação da expansibilidade entre provadores.

Tabela 7. Tabela dos valores dos níveis de significância p do teste de normalidade e homogeneidade para cada atributo avaliado em cada escala.

Atributo	Teste de Kolmogorov-Smirnov ⁽¹⁾		Teste de Levene ⁽¹⁾ (amostras)		Teste de Levene ⁽¹⁾ (provadores)	
	Escala Hedônica	Escala Autoajustável	Escala Hedônica	Escala Autoajustável	Escala Hedônica	Escala Autoajustável
Aparência	p= 0,150	p= 0,015	p= 0,0001	p= 0,7333	p= 0,0001	p≈ 1,000
Avaliação global	p= 0,150	p= 0,096	p= 0,0001	p= 0,9203	p= 0,0001	p≈ 1,000

(1) Se $p \geq 0,05$, não é rejeitada a hipótese de normalidade e homoscedasticidade.

3.2 Escala de Ordenação x Escala Hedônica x Escala Autoajustável

O teste não paramétrico de Friedman aplicado aos dados de preferência gerados pela escala de ordenação e aos dados de aceitação coletados através da escala hedônica e autoajustável, produziram os resultados mostrados na Tabela 8.

Tabela 8. Valores dos totais de ordenação para cada amostra em cada escala e atributo avaliado.

Escala	Amostra	Aparência			Aceitação global		
		Total de ordenação ⁽¹⁾	p ⁽²⁾ =0,05	p ⁽²⁾ =0,01	Total de ordenação ⁽¹⁾	p ⁽²⁾ =0,05	p ⁽²⁾ =0,01
De ordenação	1	190	a	a	254	a	a
	2	254	b	b	197	b	b
	3	132	c	c	125	c	c
Hedônica de 9 pontos	1	196,5	a	a	236	a	a
	2	239,5	b	b	201	b	a
	3	140	c	c	139	c	b
Autoajustável	1	202	a	a	239	a	a
	2	234	a	a	205	b	a
	3	140	b	b	132	c	b

- (1) Totais com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si. Esses resultados foram obtidos através do teste não paramétrico de Friedman.
 (2) p= Nível de significância.

A Tabela 8 mostra que a escala de ordenação apresentou diferenças significativas a 1% de significância na comparação entre todos os pares de amostras, tanto na avaliação da aparência, quanto na avaliação da aceitação global. No caso da escala hedônica foram detectadas diferenças significativas entre cada par de amostras ao nível de 5% de significância em ambos os atributos, sendo que ao nível de 1% não houve diferença significativa entre as amostras mais aceitas (amostras 1 e 2) na avaliação da aceitação global das mesmas. No que se refere à escala autoajustável, não foram detectadas diferenças entre as amostras mais aceitas ao nível de 5% de significância na avaliação da aparência e, ao nível de 1%, na avaliação da aceitação global. Mais uma vez, os resultados deste estudo revelam que se por um lado, a escala autoajustável resolve o problema da expansibilidade entre os provadores, por outro lado, a utilização desta escala produz um aumento da variabilidade dos dados inclusive em termos de ordenação. Uma vez que em ambas as escalas os consumidores foram solicitados a ordenarem as amostras, o menor poder discriminativo da escala autoajustável com relação à escala de ordenação pode ser proveniente de uma maior dificuldade do consumidor em entender o procedimento solicitado e, portanto, utilizá-la. GAY & MEAD (1992) também observaram baixa eficiência da escala autoajustável quando a mesma foi utilizada apenas para

discriminar três amostras por provador (bloco) e sugeriram que neste caso, o método baseado na análise das ordens proposto por PETTITT (1982), poderia ser uma melhor alternativa para análise desses dados.

4. CONCLUSÕES

As três escalas apresentaram as mesmas tendências de ordenação das amostras em função das preferências dos consumidores. Entretanto, a escala autoajustável mostrou-se ligeiramente menos eficiente em relação ao poder discriminativo.

No caso de se utilizar um delineamento em blocos completos para apresentação das amostras aos provadores, a escala autoajustável resolveu o problema da expansibilidade entre os provadores, porém os resultados do presente estudo mostram que ela aumentou a variabilidade das respostas dos consumidores, diminuindo, conseqüentemente, o poder discriminativo do teste.

A escala hedônica de 9 pontos apresentou problemas de heteroscedasticidade no diagnóstico do modelo ANOVA. Os resultados de uma ANOVA onde não são satisfeitos os pressupostos estatísticos subjacentes são apenas aproximados, uma vez que nestes casos se desconhecem os níveis exatos de significância para os efeitos dos tratamentos. Assim, na análise de dados provenientes da escala hedônica, não seria prudente estabelecer níveis de significância (1%, 5%, etc.) rigorosos para os testes. Também, para avaliar os efeitos do afastamento dos pressupostos sobre os níveis de significância obtidos na análise dos dados através de modelos normais aditivos, torna-se fundamental, antes de qualquer conclusão, realizar o diagnóstico do modelo ANOVA através da análise dos resíduos.

Os dados gerados pela escala autoajustável neste experimento, mostraram moderados desvios da normalidade dos resíduos para a avaliação da aparência. Já na avaliação da aceitação global, não foram detectados desvios da normalidade.

Também, não foram detectados problemas de heteroscedasticidade das amostras, indicando que a ANOVA usual poderia ser uma análise estatística adequada para eles. Por outro lado, a aproximação heurística proposta por GAY (1988), para análise de dados gerados por esta escala, não seria uma alternativa apropriada em casos como os de nosso estudo, onde no delineamento experimental utilizado para apresentação das amostras, todos os provadores (blocos) recebem todas as amostras.

É evidente que a confirmação dos resultados obtidos nesta pesquisa em relação à escala autoajustável se faz necessária e só pode ser obtida pela realização de um maior número de experimentos independentes, envolvendo maior número de amostras, diferentes produtos e diferentes grupos de consumidores. Entretanto, os resultados do presente estudo indicam a necessidade de métodos alternativos para a análise de dados coletados através das escalas hedônica e autoajustável. Esses métodos não devem basear-se em pressupostos de normalidade e homoscedasticidade, devendo ao invés utilizar recursos como modelos lineares generalizados, análise de dados categóricos, testes não paramétricos, dentre outros ou; alternativamente, utilizar transformação de dados para normalização e/ou estabilização da variância.

Ainda outra sugestão, seria a realização de testes de hipótese para modelos de aleatorização, visto que os dados gerados nesta pesquisa foram obtidos através de experimentos planejados utilizando delineamentos tipo "cross-over". Assim, níveis exatos de significância poderiam ser obtidos utilizando testes de aleatorização. Esses testes não se baseiam em pressupostos de normalidade, homogeneidade das variâncias ou independência.

AGRADECIMENTO

Os autores agradecem à **Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)**, pelo apoio financeiro que fez possível a realização deste projeto de pesquisa.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- BASKER, D. Critical values of differences among rank sums for multiple comparisons. *Food Technology*. p. 79-84, 1988.
- BOX, G. E. P; ANDERSEN, S. L. Permutation theory in the derivation of robust criteria and the study of departures from assumption. *J. Royal Statistical Society*, series B, Vol. 17, p. 1-34, 1955.
- CHEN, A. W.; RESURRECCION, V. A. Age appropriate hedonic scales to measure food preferences of young children. *Journal of Sensory Studies*, v. 11, p. 141-163, 1996.
- GAY, C. *A statistical assessment of taste-testing methods*. Ph. D. Thesis, The University of Reading, 1988, 276p.
- GAY, C.; MEAD, R. A statistical appraisal of the problem of sensory measurement. *Journal of Sensory Studies*, v. 7, p. 205-228, 1992.
- GIOVANNI, M. E.; PANGBORN, R. M. Measurement of taste intensity and degree of liking of beverages by graphic scales and magnitude estimation. *Journal of Food Science*, v. 48, p. 1175-1182, 1983.
- HOUGH, G.; SANCHEZ, R.; BARBIERI, T.; MARTINEZ, E. Sensory optimization of powdered chocolate milk formula. *Food Quality and Preference*, v. 8, p. 213-221, 1997.
- LAWLESS, H. T.; MALONE, G. J. The discriminative efficiency of common scaling methods. *Journal of Sensory Studies*, v. 1, p. 85-98, 1986a.
- LAWLESS, H. T.; MALONE, G. J. A comparison of rating scales: sensitivity, replicates and relative measurement. *Journal of Sensory Studies*, v. 1, p. 155-174, 1986b.
- LEVENE, H. Robust Tests for the Equality of Variance, *Contributions to Probability and Statistics*, ed. I. Olkin, Palo Alto, CA: Stanford University Press, p. 278 -292, 1960.
- MACFIE, H. J. ; BRATCHELL, N.; GREENHOFF, K.; VALLIS, L. V. Designs to balance the effect of order of presentation and first-order carry-over effects in hall tests. *Journal of Sensory Studies*, v. 4, p. 129-148, 1989.
- MCTIGUE, M. C.; KOEHLER, H. H.; SILBERNAGEL, M. J. Comparison of four sensory evaluation methods for assessing cooked dry bean flavor. *Journal Food Science*, v. 5, p. 1278-1283, 1989.

- MEAD, R.; GAY, C. Sequential design of sensory trials. *Food Quality and Preference*, v.6, p. 271-280, 1995.
- MEILGAARD, M.; CIVILLE, G. V.; CARR, B. T. *Sensory Evaluation Techniques*. CRC Press, Inc. Boca Raton, FL, USA, 1987, 281p.
- MONTGOMERY, D. C. *Design and Analysis of Experiments*, 2^a Ed. John Wiley & Sons, New York, USA, 1984, 538p.
- NAES, T. Handling individual differences between assessors in sensory profiling. *Food Quality and Preference*, v. 2, p. 187-199, 1990.
- O'MAHONY, M. Some assumptions and difficulties with common statistics for sensory analysis. *Food Technology*, p. 75-82, 1982.
- PETTITT, A. N. Inference for the linear model using a likelihood based on ranks. *J. R. Statistic. Soc. B* 44, p. 234-243, 1982.
- PEARCE, J. H; KORTH, B.; WARREN, C. B. Evaluation of three scaling methods for hedonics. *Journal of Sensory Studies*, v. 1, p. 27-46, 1986.
- PIGGOT, J. R. Design questions in sensory and consumer science *Food Quality and Preference*, v.6, p. 217-220, 1995.
- PRITCHETT, M. P. A., Performance assessment of sensory panelists. *Journal of Sensory Studies*, v. 7, p. 229-252, 1992.
- SCHEFFÉ, H. *The Analysis of Variance*. John Wiley & Sons, Inc. New York, 1959.
- SIDEL J. L.; STONE, H., The role of sensory evaluation in the food industry. *Food Quality and Preference*, v.4, p. 65-73, 1993.
- STONE, H.; SIDEL, J. *Sensory Evaluation Practices*. 2^a Ed. Academic Press, Inc. San Diego, 1993.
- TEPPER, B.J.; SHAFFER, S. E.; SHEARER, C. M. Sensory perception of fat in common foods using two scaling methods. *Food Quality and Preference*, v.5, p. 245-251, 1994.
- VIE, A.; GULLI, D.; O'MAHONY, M. Alternative hedonic measures. *Journal of Food Science*, v. 56, p. 1-5, 1991.
- WILKINSON, C.; YUKSEL, D., Modeling differences between panelists in use of measurement scales. *Journal of Sensory Studies*, v. 12, p. 55-68, 1997.
- WILLIAMS, A. A.; SPENCER, H. W.; CARTER, C. S.; ARNOLD, G. M.; LANGRON, S. *Manual of Sensory Difference Testing of Wines*, Long Ashton Research Station, UK, 1979, 70p.

CAPÍTULO 3

ESCALA HEDÔNICA HÍBRIDA: VANTAGENS EM RELAÇÃO ÀS ESCALAS HEDÔNICA TRADICIONAL, AUTOAJUSTÁVEL E DE ORDENAÇÃO

Resumo

Em testes com consumidores, escalas tradicionais como a escala hedônica de 9 pontos freqüentemente apresentam a seguinte problemática: i) não satisfazem os pressupostos estatísticos de independência, normalidade e homoscedasticidade exigidos nos modelos ANOVA, ii) dão pouca liberdade para os indivíduos expressarem suas percepções, devido ao limitado número de categorias, iii) induzem efeitos numéricos e contextuais no julgamento dos provadores e, iv) os valores numéricos associados às suas categorias, embora numericamente possuam intervalos iguais, não refletem iguais diferenças em percepção. Algumas pesquisas envolvendo escalas hedônicas não estruturadas, mostram um bom desempenho destas, porém, mencionam dificuldade de uso por consumidores. Neste contexto surge a proposta de uma escala alternativa intitulada "escala hedônica híbrida", elaborada pela combinação da escala estruturada e não estruturada. Pelo exposto, o objetivo deste trabalho foi avaliar o uso da escala hedônica híbrida em estudos com consumidores comparando-a à escala hedônica de 9 pontos, escala autoajustável, e escala de ordenação; através dos seguintes critérios: i) variabilidade das respostas sensoriais, ii) poder discriminativo, iii) adequação dos dados às suposições dos modelos ANOVA e, iv) facilidade de uso. Cinco marcas de suco de laranja foram avaliadas por 80 consumidores, divididos em quatro grupos de 20 indivíduos cada. Todos os indivíduos avaliaram todas as amostras através de todas as escalas em 4 diferentes sessões de degustação. Um delineamento em quadrado latino 4x4 foi utilizado para controlar o efeito de ordem de apresentação das escalas e, avaliar sem vícios, a facilidade de uso das mesmas. Para cada escala, a ordem de apresentação das amostras foi balanceada. Os resultados obtidos através das escalas hedônica

tradicional, híbrida e autoajustável foram avaliados através de ANOVA, considerando-se os efeitos devido a provador, amostra, ordem de apresentação das amostras e efeito residual. A normalidade dos resíduos foi verificada através do teste de Shapiro-Wilks e a homoscedasticidade através do teste de Brown-Forsythe. Os valores de $pF_{amostra}$, $pF_{provador}$ e $QM_{resíduo}$ fornecidos pela ANOVA de cada escala, foram respectivamente utilizados para avaliar o poder discriminativo, a expansibilidade dos provadores e a variabilidade dos dados. O teste de REGWF foi também aplicado para análise do poder discriminativo de cada escala. Os resultados obtidos utilizando a escala de ordenação foram analisados através do teste de Friedman e, a facilidade de uso das escalas pelo teste de Cochran-Mantel-Haenszel. Os resultados sugeriram uma superioridade da escala hedônica híbrida sobre as escalas hedônica estruturada e autoajustável, tanto em função do poder discriminativo como em relação à adequação dos dados às suposições de normalidade e homoscedasticidade. A escala autoajustável apresentou poder discriminativo ligeiramente maior do que a escala hedônica estruturada a despeito dos dados da primeira apresentarem maior variabilidade e não normalidade dos resíduos. A escala hedônica de 9 pontos apresentou problemas de heteroscedasticidade nas variâncias das amostras. A escala de ordenação apresentou o menor poder discriminativo em relação às demais. As escalas hedônicas estruturada e híbrida foram consideradas significativamente ($p \leq 0,01$) mais fáceis de serem utilizadas que a autoajustável, não havendo diferença ($p \leq 0,05$) entre as duas primeiras.

Abstract

In consumer tests, traditional scales such as the nine-point structured hedonic scale frequently present the following problems: i) they do not satisfy the statistical assumptions of independence, normality and homoscedasticity required by ANOVA models; ii) they give little freedom to the individuals to express their perceptions, due to the limited number of categories; iii) they induce numerical and contextual effects in the judgments by the panelists and, iv) the difference between numerical values associated with the categories do not reflect equivalent differences in perception.

Some studies involving unstructured hedonic scales have shown good performance of these scales, mentioning although difficulty in their use with consumers. In this context the proposal of a different scale, known as the “hybrid hedonic scale” has arisen. It represents a combination of structured and unstructured scales. Thus, the objective of this study was to evaluate the use of the hybrid hedonic scale in consumer studies, comparing it with the nine-point hedonic scale, self-adjusting scale and ranking test by the following criteria: i) variability of sensory response, ii) discriminative power, iii) data adequacy to the assumptions of ANOVA models and, iv) ease of use. Eighty consumers, divided into four groups of 20 individuals each, evaluated five brands of orange juice. All the individuals evaluated all the samples using all the scales, in 4 distinct tasting sessions. A 4 x 4 Latin square design was used to control the effect of the order of presentation of the scales and evaluate their ease of use without biases. For each scale, the presentation order was balanced. The results obtained with the traditional, hybrid and self-adjusting hedonic scales were evaluated using ANOVA (judges, sample, sample presentation order, carry-over). Data normality was evaluated using the Shapiro-Wilks test and homoscedasticity by Brown-Forsythe’s test. The values for pF_{sample} , pF_{judge} and MS_{error} provided by ANOVA for each scale were used to evaluate, respectively, discriminative power, differential expansiveness between assessors and data random variability. The REGWF test and Bootstrap method was also applied to analyze the discriminative power of each scale. Ranking test results were evaluated by Friedman’s test and the ease of use of the scales by the Cochran-Mantel-Haenszel tests. The results indicated the superiority of the hybrid hedonic scale as compared to the structured and self-adjusting scales both with respect to the discriminatory power and to data adequacy to the assumptions of normality and homoscedasticity. The self-adjusting scale presented a slightly greater discriminative power than the structured hedonic scale, despite the former having presented data with a greater random variability and lack of normality of the residuals. Of all the methods, the ranking test presented the least discriminative power. The structured and hybrid hedonic scales were considered to be significantly ($p \leq 0.01$) easier to use than the self-adjusting scale, there being no difference between these first two scales.

1. INTRODUÇÃO

Dados sensoriais coletados junto a consumidores através de escalas tradicionais como a escala hedônica de 9 pontos (PERYAM & PILGRIM, 1957), freqüentemente não satisfazem os pressupostos estatísticos de independência, normalidade e homoscedasticidade exigidos nos modelos para análise da variância dos dados (VILLANUEVA *et al.*, 2000; GAY & MEAD, 1992; O'MAHONY, 1982). Problemas de variação da expansibilidade entre provadores têm também sido reportados como uma limitação adicional ao uso das escalas afetivas tradicionais (GAY, 1988; GAY & MEAD, 1992; MEAD & GAY, 1995). Com o objetivo de minimizar alguns dos problemas anteriormente citados, GAY (1988) testou com provadores treinados, uma escala intitulada de "escala autoajustável", anteriormente proposta por WILLIAMS *et al.* (1979).

Com o objetivo de avaliar o desempenho desta escala em testes com consumidores, VILLANUEVA *et al.* (2000), realizaram um estudo comparando a escala autoajustável com a tradicional escala hedônica e escala de ordenação em relação à expansibilidade entre provadores, poder discriminativo e adequação às hipóteses associadas aos modelos ANOVA. Os autores constataram que a escala autoajustável resolveu problemas da expansibilidade entre provadores e produziu variâncias homogêneas entre as amostras, porém os dados gerados através desta escala mostraram moderadas distorções da normalidade. A escala hedônica de 9 pontos, por sua vez, apresentou problemas de heteroscedasticidade nas variâncias das amostras. As escalas de ordenação e autoajustável tiveram o melhor e pior poder discriminativo, respectivamente.

Em vista do desempenho pouco vantajoso da escala autoajustável em relação à escala hedônica tradicional, os referidos autores apontaram a necessidade de se continuar desenvolvendo novas escalas, mais adequadas que a tradicional escala hedônica de 9 pontos, para a obtenção de dados afetivos.

Alguns problemas apontados em relação à tradicional escala hedônica estruturada de 9 pontos são: i) pouca liberdade para expressar as percepções

sensoriais devido ao limitado número de categorias de respostas, ii) os valores numéricos atribuídos às categorias, embora numericamente possuam intervalos iguais, não refletem iguais diferenças em percepção, iii) efeitos numéricos e contextuais são mais prováveis de acontecerem no uso desta escala, devido às diferenças na interpretação das diversas categorias e a erros de hábito, onde o provador tende a repetir as mesmas respostas em situações onde várias avaliações consecutivas são necessárias (GIOVANNI & PANGBORN, 1983; MCPHERSON & RANDALL, 1985). Como alternativa para se minimizar esses problemas, surgiram as escalas hedônicas não estruturadas ou lineares. Entretanto, ainda que os trabalhos de GIOVANNI & PANGBORN (1983), MCPHERSON & RANDALL (1985) entre outros, tenham mostrado melhor desempenho desta escala em relação à escala hedônica de 9 pontos, oferecendo maior liberdade aos consumidores para expressar as percepções sensoriais, diminuindo efeitos contextuais e apresentando menores desvios da normalidade dos dados, o uso da escala não estruturada em estudos com consumidores é limitado, devido ao fato dos provadores terem maior dificuldade de entender e utilizar esta escala (LAWLESS & MALONE, 1986a). Neste contexto surge a proposta de uma escala intermediária, intitulada "escala hedônica híbrida", a qual consiste na combinação da escala estruturada e não estruturada.

A escala hedônica híbrida é uma escala linear, resultado da combinação da escala estruturada e não estruturada. Para ser mais facilmente utilizada pelos consumidores do que a escala não estruturada, esta escala deve ser ancorada na região central e nos extremos, com anotações verbais. As demais porções do seu contínuo devem ser marcadas com pontos equidistantes para melhor definir os graus e a orientação do contínuo hedônico. As possíveis vantagens desta escala em relação à escala hedônica tradicional são: i) ela não se restringe a um limitado número de categorias, ou seja, o provador pode utilizar qualquer parte da escala para atribuir seus valores hedônicos, com possível aumento do poder discriminativo da escala pelo fato dos provadores terem maior liberdade na escolha do seu grau de aceitabilidade, ii) diminuição do erro psicológico de hábito, encorajando ao provador a fazer sua avaliação de maneira mais consciente e criteriosa, iii) diminuição de efeitos numéricos e contextuais, favorecendo o seu uso em estudos inter-culturais, iv) pelo

fato desta escala gerar dados contínuos, é possível que seus dados permitam a utilização mais segura de uma diversidade de análises estatísticas paramétricas e não paramétricas (O'MAHONY, 1982). Apesar de todas essas possíveis vantagens, na literatura corrente da área de Análise Sensorial, quase não se encontram estudos que avaliem o desempenho, vantagens e desvantagens desta escala em relação às demais escalas afetivas.

2. OBJETIVOS

Propor o uso da escala hedônica híbrida em teste de consumidor e avaliar o desempenho desta escala em relação à escala autoajustável, escala hedônica estruturada de 9 pontos e teste de ordenação, utilizando como critérios os seguintes parâmetros: i) variabilidade das respostas sensoriais, ii) poder discriminativo, iii) adequação dos dados coletados às suposições estatísticas para análise da variância: independência, normalidade, homoscedasticidade e, iv) facilidade de uso pelos consumidores, considerando-se aspectos como: facilidade de entendimento, menor restrição no uso e preferência dos consumidores com relação às escalas.

3. MATERIAIS E MÉTODOS

3.1 Amostras

Foram avaliadas cinco marcas de suco de laranja de maior comercialização no mercado consumidor brasileiro, sendo duas marcas de suco reconstituído (amostras 1 e 5) e três de suco integral pasteurizado (amostras 2, 3 e 4). As duas amostras de suco reconstituído eram acondicionadas em embalagens Tetra Brik[®] (longa vida), por processo asséptico e armazenadas à temperatura ambiente (amostras 1 e 5). Duas dentre as três marcas de suco integral encontravam-se acondicionadas em embalagens Tetra Rex[®], as quais foram armazenadas sob refrigeração, a aproximadamente 10 °C (amostras 3 e 4). A última amostra encontrava-se acondicionada em embalagem de material similar àquele da embalagem Tetra Brik[®] e armazenada à temperatura ambiente, porém, seu processo de enchimento foi semi-asséptico (amostra 2).

3.2 Provadores

Os provadores foram recrutados junto aos alunos do Colégio Técnico da Universidade Estadual de Campinas - UNICAMP. Das 115 fichas de recrutamento respondidas, foram selecionados 80 consumidores habituais de suco de laranja, 44 homens e 36 mulheres de idade variando entre 16 e 35 anos, tendo-se como critério de seleção uma frequência de consumo mínima de 1 a 2 copos de suco de laranja por semana e gostar de suco de laranja em um grau igual ou superior a "gostar moderadamente" na escala hedônica (1= desgostei extremamente, 9= gostei extremamente).

3.3 Delineamento Experimental e Modelo Estatístico

Para avaliar sem vícios a facilidade de uso das escalas, os 80 provadores foram divididos aleatoriamente em quatro grupos de 20 indivíduos cada. Assim, para cada grupo, a ordem de apresentação das escalas foi balanceada obedecendo-se a um delineamento em quadrado latino 4x4 (Tabela 1), utilizado para controlar o efeito de ordem de apresentação das escalas. Desta forma, em cada grupo, todos os consumidores avaliaram todas as amostras, utilizando as quatro escalas, em quatro sessões distintas de avaliação.

Tabela 1. Delineamento para balancear a ordem de apresentação das escalas (n= 80 provadores, divididos em grupos de 20 indivíduos cada).

GRUPO DE PROVADORES	ORDEM DE APRESENTAÇÃO DAS ESCALAS			
	1 ^a	2 ^a	3 ^a	4 ^a
1 ^o	EHE	EHH	EA	EO
2 ^o	EHH	EA	EO	EHE
3 ^o	EA	EO	EHE	EHH
4 ^o	EO	EHE	EHH	EA

EHE= Escala hedônica estruturada

EHH= Escala hedônica híbrida

EA= Escala autoajustável

EO= Escala de ordenação

Para cada escala, o delineamento experimental utilizado para apresentação das amostras aos consumidores e geração de respostas, foi o delineamento tipo

“cross-over”, balanceado para controlar efeitos de ordem de apresentação das amostras e efeitos residuais de primeira ordem. Este delineamento encontra-se tabelado em MACFIE & WAKELING (1995). Desta forma, os resultados experimentais do presente estudo foram analisados mediante o modelo estatístico para análise da variância associado a este tipo de delineamentos, o qual explicita-se a seguir

$$Y_{ij} = \mu + p_i + o_j + a_{d(i,j)} + r_{d(i,j-1)} + \varepsilon_{ij}, \quad i=1, \dots, 80; \quad j=1, \dots, 5$$

Onde,

Y_{ij} : resposta do i -ésimo provador na avaliação da amostra $d_{(i,j)}$ apresentada na j -ésima ordem

μ : média geral

p_i : efeito do i -ésimo provador

o_j : efeito da j -ésima ordem de apresentação da amostra

$a_{d(i,j)}$: efeito direto da amostra $d_{(i,j)}$, avaliada pelo i -ésimo provador na j -ésima ordem de apresentação

$r_{d(i,j-1)}$: efeito residual de primeira ordem da amostra $d_{(i,j-1)}$, avaliada pelo i -ésimo provador na $(j-1)$ -ésima ordem de apresentação, onde $j \geq 2$ e $r_{d(i,0)} = 0$

ε_{ij} : perturbações aleatórias normais independentes e identicamente distribuídas com $E(\varepsilon_{ij})=0$ e $Var(\varepsilon_{ij})=\sigma^2$

3.4 Métodos sensoriais afetivos

Os testes sensoriais foram realizados em 4 sessões diferentes de avaliação no Laboratório de Análise Sensorial do Colégio Técnico da Universidade Estadual de Campinas - UNICAMP. As amostras foram avaliadas por quatro métodos afetivos:

escala hedônica de 9 pontos (Figura 1), escala autoajustável (Figura 2), escala hedônica híbrida (Figura 3), e escala de ordenação (Figura 4).

Em cada sessão, cada grupo de provadores avaliou as 5 amostras, utilizando uma escala diferente. Os provadores foram conduzidos às cabines individuais, onde após se acomodarem, receberam as instruções sobre o uso da escala específica a ser utilizada, a natureza dos produtos e o tipo de avaliação a ser realizada. No caso das escalas hedônica estruturada e hedônica híbrida, as 5 amostras foram avaliadas monadicamente; já nas escalas de ordenação e autoajustável, as 5 amostras foram avaliadas comparativamente. Após a primeira sessão, os provadores compareceram às demais até completarem suas avaliações através das 4 escalas. Apenas uma sessão de avaliação por dia era realizada por cada provador, para evitar o cansaço e problemas de fadiga sensorial.

Aproximadamente 35ml de cada amostra foram servidas sob luz branca à temperatura de refrigeração entre 10 e 15 °C, em béqueres de vidro de 50 ml, codificados com números aleatórios de três dígitos. Os provadores foram solicitados a provarem o produto e, após a degustação, avaliarem o quanto gostaram ou desgostaram de cada amostra. Depois de terminada a última das 4 sessões de avaliação, foi apresentado a cada provador, um questionário onde eles/elas responderam perguntas concernentes à facilidade de uso da escala hedônica estruturada, hedônica híbrida e autoajustável. Foram avaliados, através de escalas estruturadas verbais, três atributos: a facilidade de uso (1= extremamente difícil, 9= extremamente fácil), a liberdade oferecida aos provadores para expressarem suas percepções sensoriais (0=nenhuma liberdade, 6= total liberdade) e a preferência dos consumidores com relação às escalas. Também lhes foi perguntado o motivo de sua preferência pela referida escala. O questionário utilizado para avaliação da facilidade de uso das escalas encontra-se apresentado no Apêndice.

Nome: _____ Idade: _____ Data: _____

Você esta recebendo 5 amostras codificadas de suco de laranja. Por favor, PROVE da esquerda para a direita cada uma das amostras avaliando-as de UM MODO GERAL. Utilizando a escala abaixo indique o quanto você gostou ou desgostou das mesmas.

- 9 Gostei extremamente
- 8 Gostei muito
- 7 Gostei moderadamente
- 6 Gostei ligeiramente
- 5 Nem gostei / nem desgostei
- 4 Desgostei ligeiramente
- 3 Desgostei moderadamente
- 2 Desgostei muito
- 1 Desgostei extremamente

Figura 1. Ficha de avaliação da aceitação utilizando a escala hedônica estruturada de 9 pontos.

Nome: _____ Idade: _____ Data: _____

Você esta recebendo 5 amostras codificadas de suco de laranja. Por favor, PROVE as amostras da esquerda para a direita e marque na escala correspondente um "x" e o código da amostra no lugar da escala que você achar que melhor represente o quanto você gostou ou desgostou de UM MODO GERAL de cada amostra.

Amostra N° _____

Figura 2. Ficha de avaliação da aceitação utilizando a escala hedônica híbrida.

Nome: _____ Idade: _____ Data: _____

Você esta recebendo 5 amostras codificadas de suco de laranja. PROVE agora as amostras da esquerda para a direita avaliando-as de UM MODO GERAL. Escolha primeiro a amostra que você mais gostou e depois a que você menos gostou e anote os códigos dessas amostras respectivamente nos extremos direito e esquerdo da escala abaixo. Avalie agora as outras amostras de forma comparativa com aquelas alocadas nos extremos, marcando um "x" e o código da amostra no ponto da escala que você achar que melhor represente o quanto você gostou ou desgostou de cada amostra após degustá-las.

Figura 3. Ficha de avaliação sensorial utilizando a escala autoajustável.

ESCALA DE ORDENAÇÃO				
Nome: _____	Idade: _____	Data: _____		
<p>Você está recebendo 5 amostras codificadas de suco de laranja. Agora PROVE as amostras da esquerda para a direita avaliando-as de um MODO GERAL e ordene-as em ordem decrescente de preferência.</p>				
+ preferida	_____	_____	_____	_____ - preferida

Figura 4. Ficha de avaliação sensorial utilizando a escala de ordenação.

3.5 Análise estatística

A análise da variância dos dados foi realizada utilizando o PROC GLM do SAS (SAS INSTITUTE INC., 2000). Os resultados obtidos através das escalas hedônica tradicional, hedônica híbrida e autoajustável, foram avaliados através de modelos ANOVA ($p= 5\%$) para delineamentos tipo “cross-over”, considerando-se no modelo estatístico os seguintes efeitos: provador, ordem de apresentação, amostras e efeito residual. Também, histogramas de frequência foram obtidos para complementar esta análise.

Para avaliar as diferenças entre as amostras, foram efetuadas comparações pareadas das médias aplicando-se o teste de Ryan-Einot-Gabriel-Welsch (REGWQ). Este teste baseado na amplitude estudentizada é considerado um dos testes mais poderosos entre os testes de múltiplas etapas (“Multiple-Stage Tests”). O procedimento do teste de REGWQ, encontra-se detalhado em RAMSEY (1978).

O diagnóstico da adequação dos resultados aos modelos de análise de variância, foi realizado através da análise dos resíduos e testes para verificação das suposições do modelo. A hipótese de normalidade foi avaliada através do teste de Shapiro-Wilks, considerado um dos testes mais poderosos e efetivos inclusive para tamanhos pequenos de amostras (PEARSON, D'AGOSTINO & BOWMAN, 1977). A homoscedasticidade, foi avaliada através do teste de Brown-Forsythe o qual é considerado robusto para desvios da normalidade (BROWN & FORSYTHE, 1974). A

variação da expansibilidade entre provadores foi avaliada através do valor de pF_{provador} fornecido pela ANOVA e pelo teste de Brown-Forsythe para avaliação da homogeneidade das variâncias dos provadores. Finalmente, o diagnóstico da aditividade do modelo foi realizado através de gráficos de dispersão dos resíduos vs. os valores preditos pelo modelo ANOVA e, do método da aproximação geral do teste de Tukey de um grau de liberdade para não aditividade transformável, proposto por BOX, HUNTER & HUNTER (1978).

A partir dos valores de pF_{amostra} e QM_{residuo} fornecidos pela análise de variância dos dados sensoriais correspondentes a cada escala, foram respectivamente explorados o poder discriminativo e a variabilidade residual das respostas sensoriais.

Para um estudo mais aprofundado do poder discriminativo das escalas, foram gerados para cada escala, valores da estatística F_{amostra} através do método de re-amostragem “bootstrap” descrito em EFRON & TIBSHIRANI (1986). Para isso, de cada conjunto de dados originais, foram obtidas $B= 300$ amostras “bootstrap” independentes, utilizando um programa computacional desenvolvido em linguagem SAS/GLM (SAS INSTITUTE INC., 2000). A obtenção de cada amostra “bootstrap” foi realizada através de amostragens aleatórias com reposição dos dados originais, utilizando no programa computacional um gerador de números aleatórios para selecionar os provadores. Assim, cada amostra “bootstrap” consistia de 80 provadores selecionados aleatoriamente de acordo com o delineamento experimental original, onde cada provador poderia ter sido selecionado mais de uma vez, devido à amostragem com reposição. Após a obtenção das 300 amostras “bootstrap” para cada escala, valores de F_{amostra} foram calculadas através da ANOVA dos dados gerados por cada amostra “bootstrap”. Em seguida, a partir desses valores, estatísticas básicas como a média e o desvio padrão foram calculadas, junto com os histogramas “bootstrap”.

Os resultados do questionário formulado para avaliar a facilidade de uso de cada escala, foram analisados através de testes de Cochran-Mantel-Haenszel,

atribuindo-se valores numéricos a cada categoria de resposta e controlando-se o efeito dos provadores, para se detectar a existência de diferenças significativas ($p \leq 0,05$) entre os atributos avaliados: facilidade de uso, liberdade dos consumidores expressarem suas percepções sensoriais e preferência dos mesmos com relação a uma das escalas.

4. RESULTADOS E DISCUSSÃO

4.1 Escala Autoajustável x Escala Hedônica Estruturada x Escala Híbrida

4.1.1 **Análise Exploratória dos dados:** estatísticas básicas referentes aos dados coletados através das escalas hedônica de 9 pontos, autoajustável e híbrida são apresentadas na Tabela 2, onde se observa que as médias de aceitação das amostras mostram, em geral, uma mesma tendência nas três escalas: a amostra 3 foi a mais aceita enquanto as amostras 1 e 4 foram as menos aceitas.

Tabela 2. Estatísticas básicas dos dados de aceitação global de suco de laranja coletados através da escala hedônica estruturada, escala autoajustável e escala híbrida (n= 80 provadores)

Escala	Amostra	Média ⁽¹⁾	p =5%	p =1%	Mediana	Diferença Média-Mediana	Desvio Padrão
Hedônica de 9 pontos ⁽²⁾	1	4,4	b	b	5,0	0,6	2,19
	2	5,0	ab	ab	5,0	0,0	1,80
	3	5,6	a	a	6,0	0,4	2,15
	4	4,8	b	ab	5,0	0,2	1,98
	5	5,2	ab	ab	5,0	0,2	1,99
Autoajustável ⁽³⁾	1	4,1	b	b	3,4	0,7	3,56
	2	4,9	ab	ab	4,7	0,2	3,60
	3	6,0	a	a	6,2	0,2	3,39
	4	3,8	b	b	3,3	0,5	3,42
	5	4,8	ab	ab	4,6	0,2	3,92
Hedônica híbrida ⁽⁴⁾	1	4,3	b	b	4,5	0,2	2,34
	2	4,7	b	b	4,5	0,2	1,98
	3	5,7	a	a	6,0	0,3	2,31
	4	4,4	b	b	4,2	0,2	2,28
	5	4,8	b	ab	5,0	0,2	2,44

(1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ($p \leq 0,05$). Esses resultados foram obtidos através do teste de REGWQ.

(2) 1= desgostei extremamente, 9= gostei extremamente.

(3) 0= menos gostoso, 10= mais gostoso.

(4) 0=desgostei extremamente, 10= gostei extremamente.

Os valores da mediana na escala hedônica híbrida (Tabela 2), encontram-se bastante próximos aos da média, sugerindo simetria nas distribuições dos valores hedônicos referentes a cada amostra e mostrando proximidade das mesmas à distribuição normal. Na escala autoajustável e hedônica tradicional parece haver maior diferença entre as médias e as medianas, indicando um certo grau de assimetria nas distribuições dos valores hedônicos das amostras, principalmente quando se trata das amostras mais aceitas (amostra 3) e menos aceitas (amostras 1 e 4). Esta ocorrência em relação aos dados gerados pela escala autoajustável, também foi observada por VILLANUEVA *et al.* (2000). Na escala híbrida e autoajustável, as medianas variaram entre si, notadamente no caso da escala autoajustável. O mesmo não ocorre com a escala hedônica estruturada, sugerindo que quando os provadores avaliaram as amostras através das escalas hedônica híbrida e autoajustável, os mesmos utilizaram melhor o contínuo psicológico para discriminar as amostras do que no caso da escala hedônica tradicional.

As Figuras 5 e 6 mostram as distribuições dos dados gerados através das escalas hedônica e híbrida as quais, embora similares, apresentam no caso da escala híbrida, formas mais regulares e próximas à distribuição normal. Entretanto, os dados coletados pela escala autoajustável geraram distribuições bastante irregulares, com modas em um ou ambos os extremos da escala (Figura 7), polarizando os resultados para todas as amostras. Esses resultados sugerem que, se por um lado, a escala autoajustável elimina o problema da presença de efeitos extremos, causados pela relutância dos provadores em utilizarem um ou ambos os extremos da escala, por outro lado, ela gera o problema de compressão devido ao aglomeramento dos dados nos extremos da escala, produzindo distribuições bastante assimétricas, principalmente quando se trata das amostras menos ou mais aceitas. Para estas amostras, os dados concentram-se nos extremos da escala, gerando distribuições comprimidas à direita ou à esquerda. Os modelos paramétricos para análise da variância, pressupõem que as observações experimentais correspondentes aos tratamentos, procedem de populações normais (BOX, HUNTER & HUNTER, 1978). Assim, devido à sua natureza psicométrica, os dados gerados pela escala

autoajustável tendem a não satisfazer este pressuposto, conforme pode ser visualizado na Figura 7.

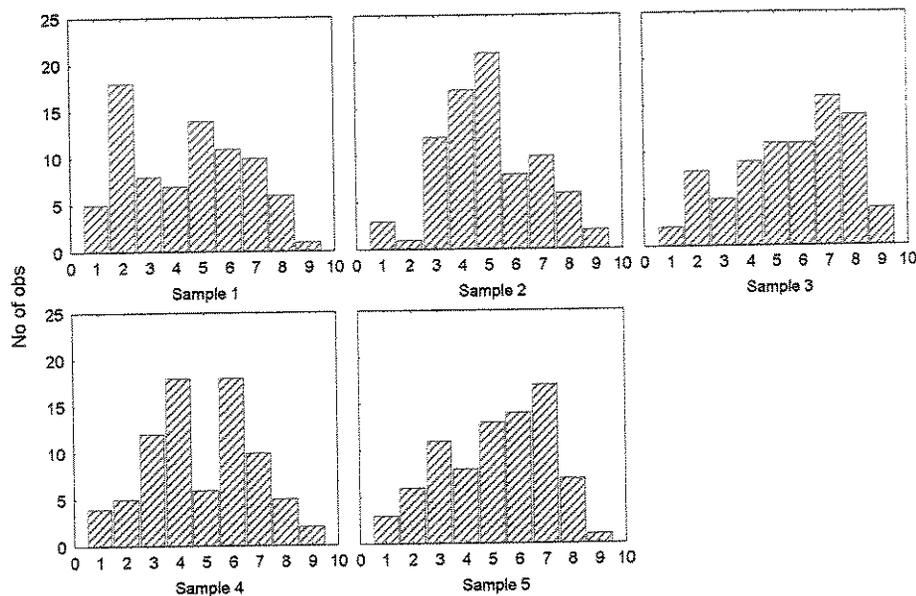


Figura 5. Histogramas de freqüências dos dados de aceitação global das amostras, gerados através da escala hedônica de 9 pontos.

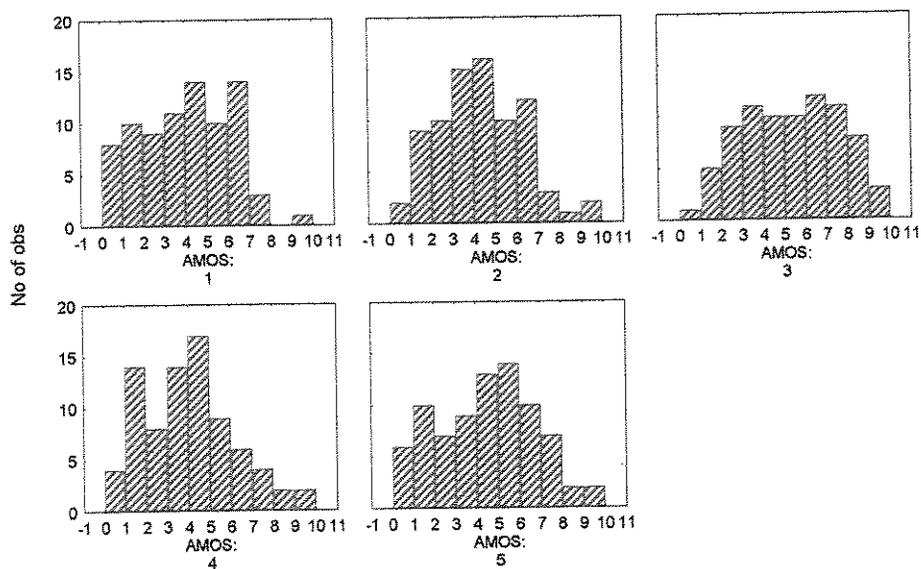


Figura 6. Histogramas de freqüências dos dados da aceitação global das amostras, gerados através da escala hedônica híbrida.

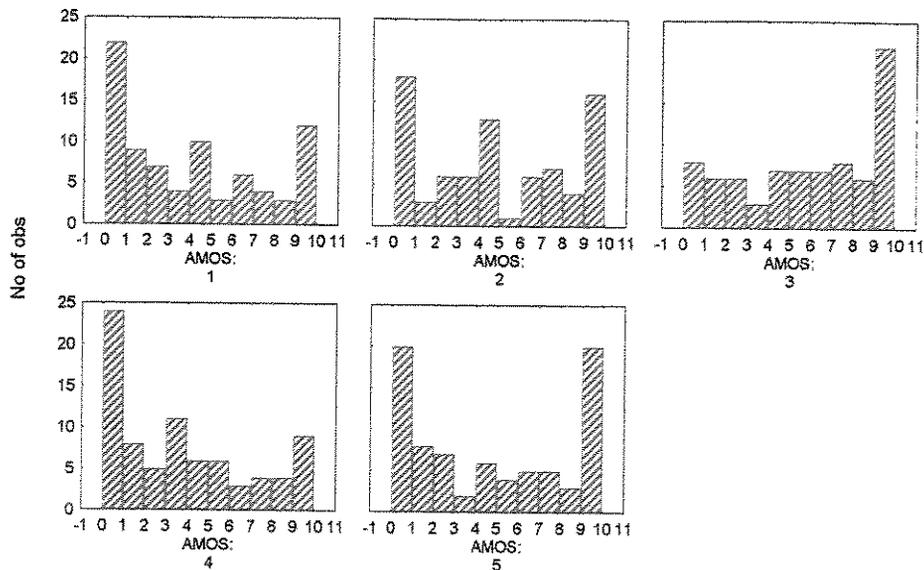


Figura 7. Histogramas de freqüências dos dados de aceitação global das amostras, gerados através da escala autoajustável.

Os valores dos desvios padrões (Tabela 2) e do QM_{erro} (Tabela 4), indicam maior variabilidade nos dados produzidos pela escala autoajustável. Esses resultados confirmam aqueles já observados por VILLANUEVA *et al.* (2000) de que a eliminação das diferenças na expansibilidade entre os provadores, traz como decorrência o aumento da variabilidade das respostas sensoriais, causado pela polarização dos valores atribuídos para todas as amostras devido ao uso obrigatório dos extremos da escala. Assim, o fato dos provadores utilizarem os extremos da escala, maximiza a diferença afetiva entre as amostras, porém produz um aumento substancial da variabilidade dos dados como se observa nitidamente na Tabela 2 e Figura 7.

4.1.2 Efeito da ordem de apresentação das amostras e efeito residual

Os resultados da ANOVA dos dados obtidos através das escalas hedônica tradicional, hedônica híbrida e autoajustável, revelaram nas duas primeiras escalas, efeitos de ordem de apresentação das amostras significativos, com valores de pF_{ordem} de 0,0001 e 0,02 respectivamente.

A Tabela 3, apresenta a médias dos valores hedônicos para cada ordem de apresentação das amostras em cada escala, na qual observa-se uma tendência decrescente das médias hedônicas em função da ordem de apresentação. Pode-se observar nesta tabela, que no caso da escala hedônica de 9 pontos, houve diferença significativa ($p \leq 0,05$) tanto entre a média da primeira ordem e as médias da terceira e quarta ordem; como entre as médias da segunda ordem em relação à terceira e quarta ordem. No caso da escala híbrida, só houve efeito da primeira ordem de apresentação, observando-se uma queda importante nos valores hedônicos para as demais ordens. Assim, pode-se concluir que não apenas o número ou o tipo de amostras tem influência no efeito de ordem de apresentação, como também o tipo de escala com o qual o provador faz sua avaliação. No caso da escala hedônica tradicional, a diferença de uma unidade entre cada categoria do contínuo afetivo pode concorrer para a maximização do efeito de ordem de apresentação das amostras, o que não ocorreu com as outras duas escalas que, por serem lineares, oferecem o intervalo todo para o provador fazer sua avaliação, possibilitando que variações menores dos valores hedônicos entre cada ordem sejam expressas.

Na ANOVA dos dados da escala autoajustável não foram detectados efeitos de ordem de apresentação ($p = 0.13$). Porém, observa-se na Tabela 3 uma clara tendência decrescente das médias em função da ordem de apresentação das amostras, observando-se uma queda importante nos valores hedônicos das demais ordens em relação à primeira ordem. A não significância desse efeito, provavelmente se deva à grande variabilidade dos dados produzidos por esta escala, refletidos nos tamanhos dos desvios padrões para cada ordem de apresentação (Tabela 3). Aparentemente, a despeito de sua característica psicométrica singular, a escala autoajustável parece não eliminar completamente o efeito de ordem de apresentação das amostras dos testes afetivos.

Tabela 3. Médias e desvios padrões dos valores hedônicos dos dados de aceitação global para cada ordem de apresentação das amostras, coletados através da escala hedônica de 9 pontos, escala autoajustável e escala híbrida (n= 80 provadores)

Ordem de apresentação das amostras	Escala hedônica de 9 pontos ⁽²⁾		Escala Autoajustável ⁽³⁾		Escala hedônica híbrida ⁽⁴⁾	
	Média ⁽¹⁾	Desvio Padrão	Média ⁽¹⁾	Desvio Padrão	Média ⁽¹⁾	Desvio Padrão
1 ^a	5,7 a	1,96	6,0 a	3,80	5,5 a	2,01
2 ^a	5,4 ab	1,80	4,5 a	3,66	4,7 b	2,35
3 ^a	4,6 c	2,08	4,1 a	3,41	4,6 b	2,26
4 ^a	4,5 c	2,13	4,6 a	3,70	4,5 b	2,51
5 ^a	4,8 bc	2,04	4,4 a	3,44	4,3 b	2,32

(1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ($p \leq 0,05$). Esses resultados foram obtidos através do teste de REGWQ.

(2) 1= desgostei extremamente, 9= gostei extremamente.

(3) 0= menos gostoso, 10= mais gostoso.

(4) 0=desgostei extremamente, 10= gostei extremamente.

Os resultados da ANOVA revelaram efeitos residuais significativos para os dados obtidos através da escala hedônica de 9 pontos ($p = 0,0017$). No caso das escalas híbrida e autoajustável, não foram detectados efeitos residuais significativos das amostras ($p \geq 0,0813$).

A Tabela 4, apresenta para as três escalas, os valores hedônicos médios de aceitação global, atribuídos pelos provadores após terem avaliado cada uma das cinco amostras. Observa-se no caso da escala hedônica tradicional, que os provadores tenderam a dar notas significativamente maiores para amostras que precederam à pior amostra (amostra 1) e notas menores, para amostras que precederam à melhor amostra (amostra 3).

No caso da escala hedônica tradicional, a diferença de uma unidade entre cada categoria do contínuo afetivo pode concorrer para a maximização do efeito residual das amostras, o que não ocorreu com as outras duas escalas que, por serem lineares, oferecem o intervalo todo para o provador fazer sua avaliação, possibilitando que variações menores dos valores hedônicos nos efeitos residuais de cada amostra sejam expressas. Outro fator que poderia contribuir para a presença desses efeitos, seriam as expressões verbais da escala que induziriam os provadores a utilizarem a escala mais em função dos termos verbais do que em função do contínuo hedônico.

Tabela 4. Médias e desvios padrões dos valores hedônicos dos dados de aceitação global para cada efeito residual das amostras, coletados através da escala hedônica de 9 pontos, escala autoajustável e escala híbrida (n= 80 provadores)

Efeito residual das amostras	Escala hedônica de 9 pontos ⁽²⁾		Escala Autoajustável ⁽³⁾		Escala hedônica híbrida ⁽⁴⁾	
	Média ⁽¹⁾	Desvio Padrão	Média ⁽¹⁾	Desvio Padrão	Média ⁽¹⁾	Desvio Padrão
1 ^a	5,7 a	1,80	4,5 a	3,57	5,2 a	1,96
2 ^a	4,9 abc	1,81	5,1 a	3,58	4,8 a	2,48
3 ^a	4,4 c	2,41	3,3 a	3,29	3,7 a	2,22
4 ^a	4,6 bc	1,96	4,7 a	3,54	4,8 a	2,56
5 ^a	5,2 ab	2,02	5,3 a	3,80	5,0 a	2,22

(1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ($p \leq 0,05$). Esses resultados foram obtidos através do teste de REGWQ.

(2) 1= desgostei extremamente, 9= gostei extremamente.

(3) 0= menos gostoso, 10= mais gostoso.

(4) 0=desgostei extremamente, 10= gostei extremamente.

4.1.3 Poder discriminativo

Os resultados da análise da variância dos dados de aceitação global para as três escalas são mostradas na Tabela 5.

Tabela 5. Valores de QM_{erro} , pF_{amostra} , e pF_{provedor} , obtidos da ANOVA dos dados de aceitação global gerados pelas escalas hedônica de 9 pontos, autoajustável e híbrida (n= 80 provadores).

Escala	QM_{erro}	pF_{amostra}	$pF_{\text{provedor}}^{(1)}$
Hedônica de 9 pontos	3,247	$p = 0,0205$	$p = 0,0001$
Autoajustável	15,152	$p = 0,0075$	$p = 1,0000$
Hedônica híbrida	4,193	$p = 0,0018$	$p = 0,0001$

(1) $pF_{\text{provedor}} \leq 0,05$ indica problema de expansibilidade entre provedores.

Os valores de pF_{amostra} na escalas hedônica de 9 pontos, autoajustável e híbrida foram significativos ao nível de 2%, 1% e 0,2%, respectivamente. Isto mostra que a escala hedônica híbrida apresentou maior evidência de efeitos significativos de amostras no teste F do que as escalas tradicional e autoajustável.

O teste de REGWQ para comparação de médias (Tabela 2), mostrou que a escala hedônica híbrida discriminou melhor a amostra mais aceita das demais amostras, detectando 4 diferenças significativas ($p = 5\%$), enquanto que as escalas autoajustável e hedônica de 9 pontos detectaram apenas 2 diferenças. Já ao nível de

significância de 1%, a escala híbrida detectou 3 diferenças significativas, a escala autoajustável 2 diferenças e, a hedônica tradicional apenas 1. Neste sentido, o poder discriminativo da escala hedônica de 9 pontos foi ligeiramente inferior ao da escala autoajustável, e esta por sua vez inferior à escala híbrida. Esta mesma conclusão é obtida através da análise das medianas (Tabela 2), sugerindo maior poder discriminativo da escala hedônica híbrida sobre as demais escalas.

Cabe observar que, apesar dos dados gerados pela escala autoajustável apresentarem maior variabilidade residual, esta escala teve poder discriminativo ligeiramente maior do que a escala hedônica de 9 pontos. Isto pode ser atribuído ao fato de que no momento da avaliação sensorial utilizando esta escala, o provador compara as amostras, ordena-as e somente depois atribui seus valores hedônicos. Quando se trata das melhores e piores amostras, a distância entre as médias de seus valores hedônicos correspondentes é maximizada, devido à ênfase da escala autoajustável, na identificação dessas amostras no momento da avaliação sensorial. Já com a escala hedônica estruturada, possíveis explicações para seu menor poder discriminativo seriam tanto a pouca liberdade dos indivíduos para expressarem suas percepções sensoriais por causa do limitado número de categorias como ao fato que os provadores podem dar os mesmos valores hedônicos para todas as amostras, ocorrendo erros de hábito.

4.1.3.1 Uso do método “bootstrap” para análise do poder discriminativo

Na Tabela 6 são apresentadas estatísticas básicas referentes aos valores de $F_{amostra}$, obtidos através da ANOVA dos dados de $B= 300$ amostras “bootstrap” independentes, geradas a partir dos dados originais coletados através das escalas hedônica, autoajustável e híbrida. Observa-se nesta tabela que a escala hedônica híbrida produziu valores mais elevados de $F_{amostra}$ do que a escala hedônica tradicional e autoajustável, apresentando um valor médio de 7,0. No caso das escalas tradicional e autoajustável, estas médias foram de 5,2 e 5,4 respectivamente. O teste de REGWQ para comparação de médias detectou diferenças significativas ($p= 5\%$) entre a média dos valores de $F_{amostra}$ da escala híbrida e as da escala hedônica e

autoajustável. Também, a escala híbrida apresentou o coeficiente de variação mais baixo (26%), indicando que a variabilidade dos valores de $F_{amostra}$ nesta escala foi bem inferior quando comparado aos coeficientes da escala tradicional (49,6%), e autoajustável (36,8%). O valor elevado do coeficiente de variação da escala hedônica de 9 pontos, indica que o valor de $F_{amostra}$ para esta escala foi bastante influenciado pelos valores hedônicos atribuídos às amostras pelos diferentes subgrupos de consumidores retirados nas sucessivas re-amostragens do grupo original.

Tabela 6. Estatísticas básicas dos valores de $F_{amostra}$ correspondentes às 300 amostras “bootstrap” obtidas a partir dos dados de aceitação global coletados através da escala hedônica de 9 pontos, escala autoajustável e escala híbrida.

Escala	Média⁽¹⁾ $F_{amostra}$	Mediana $F_{amostra}$	Desvio Padrão	CV (%)
Hedônica de 9 pontos	5,2 b	5,0	2,58	49,6
Autoajustável	5,4 b	5,2	2,00	36,8
Híbrida	7,0 a	6,8	1,81	26,0

(1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ($p \leq 0,05$). Esses resultados foram obtidos através do teste de REGWQ.

Em complementação, na Figura 8, apresenta-se os histogramas dos valores de $F_{amostra}$ correspondentes às 300 amostras “bootstrap” geradas para cada escala. Observa-se nos histogramas que, enquanto na escala hedônica de 9 pontos e autoajustável, grande parte dos valores de $F_{amostra}$ variam entre 2 e 8, os valores da escala híbrida variam entre 4 e 10, mostrando claramente que a escala híbrida produz valores mais elevados de $F_{amostra}$. Dado que no nosso experimento, o valor crítico para comparar as médias das amostras ao nível de 5% de significância na ANOVA é: $F_{4,308,0.05} = 2,4$; foram calculados nos histogramas “bootstrap” de cada escala, as porcentagens de valores de $F_{amostra}$ inferiores a 2,4, obtendo-se no caso da escala hedônica tradicional, que a mesma não rejeitaria a hipótese nula de igualdade entre as médias de aceitação em 14% dos casos. No caso das escalas autoajustável e híbrida, estas porcentagens foram de 3% e 0%, respectivamente, mostrando um poder discriminativo razoavelmente superior destas duas escalas em relação à escala hedônica tradicional.

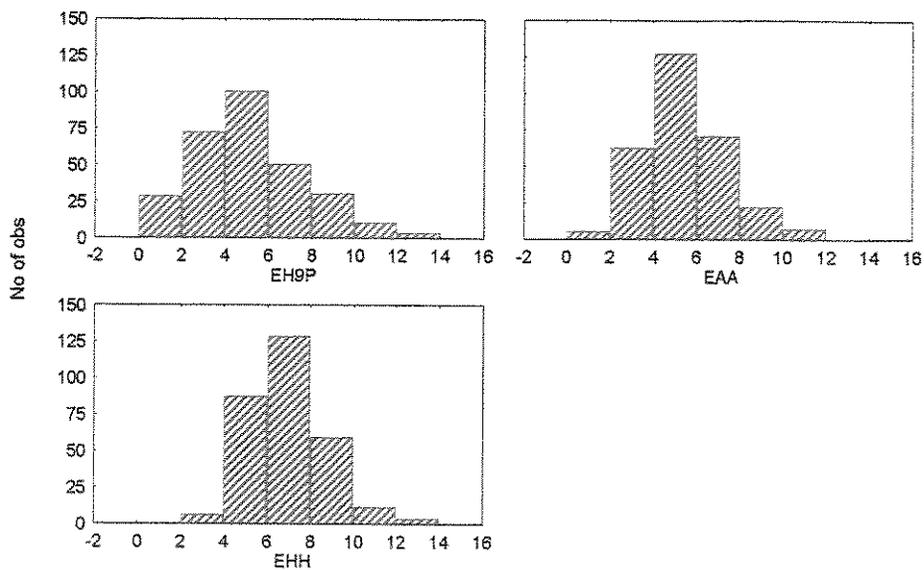


Figura 8. Histograma de freqüências dos valores de $F_{amostra}$ correspondentes às 300 amostras “bootstrap” dos dados de aceitação global gerados através da escala hedônica de 9 pontos, escala autoajustável e escala hedônica híbrida, respectivamente.

4.1.4 Diagnóstico de normalidade dos resíduos do modelo ANOVA

Os histogramas dos resíduos obtidos através da análise da variância dos valores de aceitação da escala hedônica, autoajustável e híbrida e os coeficientes de assimetria e curtose destas distribuições (Tabela 7) mostram que a escala hedônica de 9 pontos apresenta uma distribuição dos resíduos simétrica e ligeiramente achatada. No caso da escala autoajustável, a distribuição dos resíduos é ligeiramente assimétrica à direita, porém, bastante achatada, sugerindo um forte afastamento da normalidade. A distribuição dos resíduos da escala híbrida apresentou-se simétrica e mesocúrtica, mostrando características muito semelhantes às da distribuição normal.

A não normalidade dos resíduos para a escala autoajustável foi confirmada pelo teste de Shapiro-Wilks (Tabela 8), que rejeitou a hipótese de normalidade ($p=0,0001$). No caso das escalas hedônica de 9 pontos e híbrida, não foram

rejeitadas a hipótese de normalidade dos resíduos, sendo este pressuposto satisfeito para ambas as escalas ($p= 5\%$).

Tabela 7. Coeficientes de assimetria e curtose da distribuição dos resíduos do modelo ANOVA, calculados para cada escala.

Escala	Coef. de assimetria ⁽¹⁾	Coef. de curtose ⁽²⁾
Hedônica de 9 pontos	0,03	-0,45
Autoajustável	0,10	-1,01
Hedônica híbrida	0,02	0,18

(1) valores < 0 indicam assimetria à esquerda, > 0 indicam assimetria à direita e $= 0$ indicam normalidade dos resíduos.

(2) valores < 0 indicam achatamento, > 0 indicam distribuição em ponta e $= 0$ indicam normalidade dos resíduos.

4.1.5 Diagnóstico da homoscedasticidade

Os resultados do teste de Brown-Forsythe (Tabela 8) mostraram que a hipótese de homogeneidade das variâncias das amostras foi satisfeita pelas escalas hedônica híbrida e autoajustável ($p= 5\%$). Porém, foram detectadas problemas de heteroscedasticidade das variâncias na ANOVA dos dados coletados através da escala hedônica de 9 pontos ($p= 0,03$).

Tabela 8. Valores dos níveis de significância (p) do teste de normalidade, homogeneidade e aditividade para validação do modelo ANOVA para os dados de aceitação global gerados através de cada escala.

Escala	Teste de Shapiro-Wilks ⁽¹⁾	Teste de Brown-Forsythe ⁽²⁾ (amostras)	Teste de Brown-Forsythe ⁽²⁾ (provedores)	Teste de Tukey ⁽³⁾
Hedônica	0,1941	0,0302	$p= 0,9989$	0,474
Autoajustável	0,0001	0,1452	$p\approx 1,0000$	0,975
Híbrida	0,7054	0,3200	$p= 0,6611$	0,268

(1) Se $p\geq 0,05$, não é rejeitada a hipótese de normalidade dos resíduos.

(2) Se $p\geq 0,05$, não é rejeitada a hipótese de homoscedasticidade .

(3) Se $p\geq 0,05$, não é rejeitada a hipótese de linearidade do modelo.

4.1.6 Variação da expansibilidade

Os valores altamente significativos de pF_{provedor} obtidos na ANOVA dos dados coletados através das escalas hedônica tradicional e hedônica híbrida (Tabela 5), indicam que houve diferenças entre as médias dos valores hedônicos dos provedores

na avaliação da aceitação global das amostras ($p= 5\%$). Já na escala autoajustável, não foram detectadas diferenças significativas entre as médias dos diferentes provadores, apresentando a ANOVA, um valor de pF_{provador} próximo de 1. Também, os resultados do teste de Brown-Forsythe (Tabela 8) indicam que a hipótese de homogeneidade das variâncias dos provadores foi satisfeita para as três escalas ($p= 5\%$).

Os resultados acima evidenciaram a existência de variação da expansibilidade entre provadores para as escalas hedônica tradicional e híbrida, detectando-se diferenças em posição (médias), porém, não em dispersão (variâncias) em relação aos valores hedônicos atribuídos pelos provadores. Isto significa que apesar dos provadores terem utilizado diferentes porções da escala para expressarem suas percepções sensoriais, não houve diferenças significativas no tamanho efetivo dessas porções por eles/elas utilizadas. Conseqüentemente, a variação da expansibilidade ocorrida não ocasionou problemas de heteroscedasticidade nas variâncias dos provadores. Como era esperado, na escala autoajustável, não se observaram diferenças em posição nem em dispersão dos dados sensoriais correspondentes aos provadores, pois nesta escala, o provador é forçado a utilizar os extremos para fazer suas avaliações, resolvendo a priori o problema da variação da expansibilidade entre provadores.

4.1.7 Diagnóstico da aditividade do modelo

A presença de efeitos de interação entre blocos e tratamentos pode invalidar seriamente o modelo linear implicado na análise da variância dos dados. Desta forma, quando o modelo não é aditivo, dois tipos de interações podem ocorrer: i) interações transformáveis, que podem ser eliminadas através de transformações adequadas dos dados originais e, ii) interações não transformáveis. Em geral a presença da interação infla o QM_{erro} podendo afetar adversamente as comparações entre as médias dos tratamentos, causando um impacto dramático na interpretação dos resultados.

No presente estudo, o diagnóstico da aditividade do modelo ANOVA foi realizado tanto através do teste de Tukey de um grau de liberdade para não aditividade transformável (Tabela 8) como através de gráficos de dispersão dos resíduos versus os valores hedônicos preditos pelo modelo (Figuras 9, 10 e 11). Não foram detectados problemas de não aditividade dos modelos ANOVA utilizados para análise dos dados em nenhuma das três escalas. Porém, o gráfico de dispersão na Figura 9, mostra os efeitos da categorização da escala hedônica de 9 pontos mediante granularidade dos pontos. No caso da escala autoajustável (Figura 10), observar-se que o gráfico apresentou duas delimitações onde os resíduos aglomeram-se, refletindo o efeito de compressão provocado pelo uso desta escala. Esse aglomeramento produziu uma distribuição dos resíduos com caudas mais pesadas do que as da distribuição normal, ocasionando na mesma um forte afastamento da normalidade. Já no caso da escala hedônica híbrida (Figura 11), observou-se uma dispersão bastante aleatória dos resíduos, o que sugere que o modelo ANOVA para esta escala é adequado.

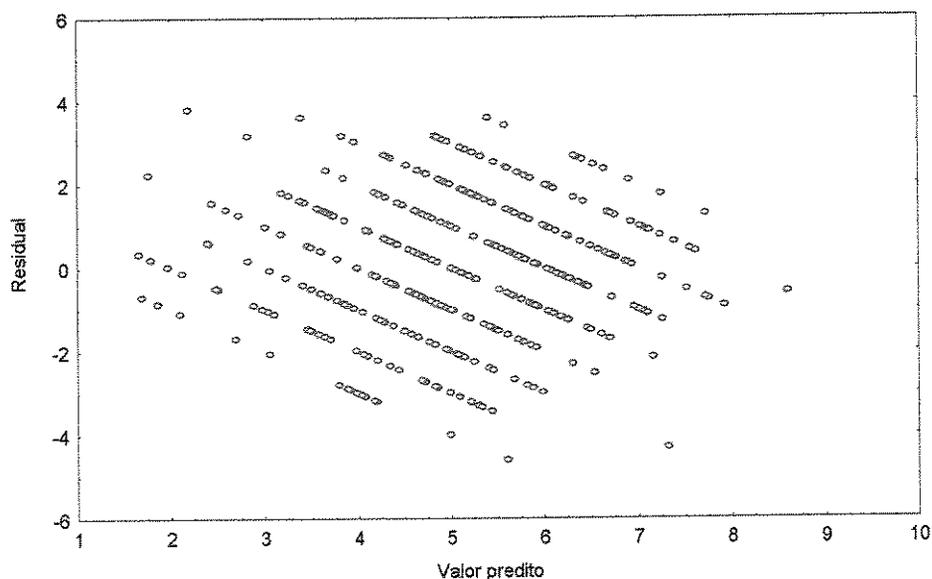


Figura 9. Diagrama de dispersão dos resíduos versus os valores ajustados pelo modelo ANOVA dos dados coletados pela escala hedônica de 9 pontos.

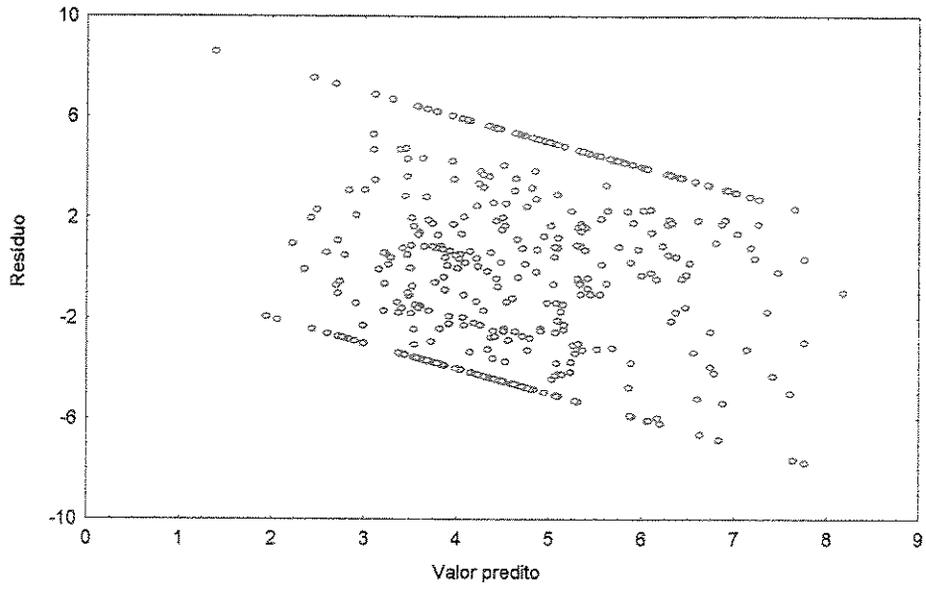


Figura 10. Diagrama de dispersão dos resíduos versus os valores ajustados pelo modelo ANOVA dos dados coletados pela escala autoajustável.

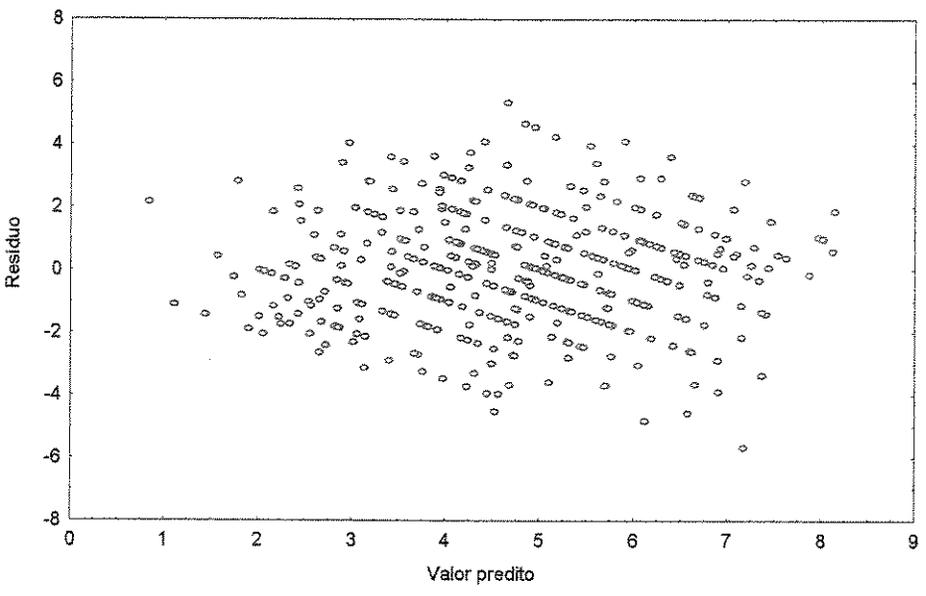


Figura 11. Diagrama de dispersão dos resíduos versus os valores ajustados pelo modelo ANOVA dos dados coletados pela escala hedônica híbrida.

4.1.8 Facilidade de uso

Os resultados referentes à facilidade de uso das escalas encontram-se na Tabela 9.

Tabela 9. Valores médios dos atributos “facilidade de uso”, “liberdade para expressarem suas percepções sensoriais” e “preferência de uso das escalas”.

Escala	Facilidade de uso^{(1), (2)}	Liberdade^{(1), (3)}	Preferência^{(1), (4)} (Somatória)
Hedônica estruturada	7,3 a	3,5 a	170 a
Autoajustável	6,4 b	3,9 a	147 a
Hedônica híbrida	7,0 a	3,8 a	163 a

(1) Valores com letras em comum numa mesma coluna, indicam atributos em que as escalas não diferem significativamente entre si ($p= 5\%$); (2) 1= extremamente difícil, 5= nem fácil/nem difícil, 9= extremamente fácil; (3) 1= nenhuma liberdade, 6= total liberdade; (4) 1= menos preferido, 3= mais preferido.

Na opinião dos consumidores, as escalas estruturada e híbrida foram consideradas significativamente ($p= 5\%$) mais fáceis de serem utilizadas e entendidas do que a escala autoajustável, tendo sido consideradas escalas moderadamente fáceis de utilizar, enquanto que a escala autoajustável foi considerada ligeiramente fácil de utilizar.

No atributo “liberdade para expressarem suas percepções sensoriais” e na avaliação da preferência entre as escalas, as três não diferiram significativamente entre si ($p= 5\%$), sendo consideradas escalas que oferecem muita liberdade de expressão das percepções sensoriais. O grupo que preferiu a escala hedônica de 9 pontos, em sua maioria apontou como razão de sua preferência, a facilidade de uso da escala devido à simplicidade dos termos. O grupo que preferiu a escala autoajustável apontou como vantagem, tanto a liberdade que a mesma oferece para expressar suas percepções sem obrigar ao consumidor a ficar restrito a uma opinião pré-estabelecida, como a possibilidade de fazer comparações diretas entre as amostras, dando-lhes uma visão completa das mesmas antes de expressarem seus julgamentos. Os consumidores que preferiram a escala híbrida relataram que a mesma oferece melhor noção do valor hedônico que a amostra merece, proporcionando maior precisão no momento de avaliar as amostras, o que ajudaria

numa melhor discriminação do produto. Outros aspectos positivos apontados com relação à escala híbrida foram: a familiaridade com uma escala de 0 a 10, sistema também utilizado nas instituições de ensino brasileiro para avaliação de alunos, e a facilidade de uso da mesma.

Os consumidores também mencionaram aspectos negativos em relação à escala hedônica de 9 pontos: eles/elas consideraram esta escala impositiva por forçar o indivíduo a enquadrar sua resposta a uma das categorias de expressão verbal, o que nem sempre corresponde ao que eles/elas acharam do produto. Na escala autoajustável eles/elas manifestaram seu desagrado por ter que usar os extremos da escala e não terem a possibilidade de empate entre as amostras. A escala híbrida desagradou aos provadores por obrigá-los a pensar mais antes de escolher o valor de aceitação de cada amostra.

4.2 Análise não paramétrica dos dados de ordenação das 4 escalas

O teste não paramétrico de Friedman aplicado aos dados de preferência gerados pela escala de ordenação e aos dados de ordenação das escalas hedônica estruturada, autoajustável e hedônica híbrida, produziram os resultados mostrados na Tabela 10.

Testes de comparações pareadas das preferências entre as amostras foram realizadas através do teste não paramétrico de Friedman. A Tabela 10 mostra que a escala de ordenação não detectou nenhuma diferença significativa entre as amostras ao nível de 5% de significância; a escala hedônica de 9 pontos detectou 1 diferença; a escala autoajustável 2 diferenças e, a escala hedônica híbrida, 3 diferenças. Assim, também neste tipo de análise de dados, a escala hedônica híbrida mostrou maior poder discriminativo. No caso da escala de ordenação, como não há uma razoável diferença entre as amostras, a tendência nas preferências também não foi clara, gerando totais de ordenação similares e, portanto, valores de p não significativos ao nível de 5%.

Tabela 10. Valores dos totais de ordenação para cada amostra em cada escala.

Escala	Amostra	Total de ordenação ⁽¹⁾	
Ordenação	1	220	a
	2	231	a
	3	268	a
	4	253	a
	5	228	a
Hedônica de 9 pontos	1	203	b
	2	235	ab
	3	277	a
	4	231	ab
	5	253	ab
Autoajustável	1	217	b
	2	247	ab
	3	284	a
	4	211	b
	5	241	ab
Hedônica Híbrida	1	218	b
	2	231,5	b
	3	295,5	a
	4	209	b
	5	246	ab

(1) Totais com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ($p= 5\%$) com relação à preferência dos consumidores. Esses resultados foram obtidos utilizando o teste de Friedman.

Estes resultados confirmam o trabalho de BAKER (1962), que afirma que o teste de ordenação é mais eficiente quando as diferenças entre as amostras são bastante perceptíveis. Também, HARRISON & ELDER (1950), consideram este método menos eficiente que os métodos de avaliação hedônica da aceitação. Segundo LAND & SHEPHERD (1988), o teste de ordenação é útil para fazer uma seleção preliminar dos produtos, pelo fato de ser um método rápido e fácil de ser realizado por provadores não treinados. No caso de testes que incluem poucas amostras, em particular para 3 amostras, os trabalhos de GAY & MEAD (1992) e VILLANUEVA *et al.* (2000) mostram que a escala de ordenação pode ser uma alternativa mais eficiente do que a escala de magnitude, autoajustável e hedônica tradicional.

5. CONCLUSÕES

Os resultados do presente estudo sugerem fortemente uma superioridade da escala hedônica híbrida sobre as escalas hedônica tradicional e autoajustável, tanto em função do poder discriminativo como em relação à adequação dos dados coletados às suposições estatísticas de normalidade e homoscedasticidade implicadas no uso dos modelos ANOVA. O poder discriminativo da escala hedônica híbrida foi superior ao das demais escalas tanto na análise dos dados de aceitação realizada através de ANOVA quanto na análise dos dados de preferência realizados através do teste não paramétrico de Friedman. A escala autoajustável apresentou poder discriminativo ligeiramente maior do que a escala hedônica de 9 pontos a despeito dos dados por ela gerados apresentarem maior variabilidade. Os resultados obtidos através do método “bootstrap”, confirmam a superioridade do poder discriminativo da escala hedônica híbrida em relação à escala estruturada e autoajustável. A escala de ordenação mostrou baixo poder discriminativo em relação às demais escalas.

No diagnóstico do modelo ANOVA, a escala hedônica de 9 pontos apresentou problemas de heteroscedasticidade das amostras e a escala autoajustável problemas de não normalidade dos resíduos. Por esse motivo, os níveis de significância associados aos valores de $F_{amostra}$ para essas escalas são apenas aproximados. Esses resultados confirmam aqueles reportados por VILLANUEVA *et al.* (2000) e sugerem que para dados provenientes dessas duas escalas, não seria prudente estabelecer níveis de significância rígidos para testar o efeito de amostra, tomando-se fundamental antes de qualquer conclusão, realizar o diagnóstico da adequação do modelo ANOVA através da análise dos resíduos. Em complementação, os resultados produzidos pelos modelos ANOVA dos dados provenientes da escala híbrida terão a grande vantagem de fornecer interpretações claras e conclusões válidas sobre os efeitos dos tratamentos experimentais, pelo fato dos pressupostos implicados na análise ANOVA serem adequadamente satisfeitos.

Com relação à facilidade de uso, as escalas hedônicas estruturada e híbrida foram consideradas pelos consumidores como sendo significativamente ($p < 0,01$) mais fáceis de usar que a autoajustável. Porém, em relação à preferência de uso das escalas, não houve uma preferência significativa entre as três escalas ($p = 0,167$).

Neste trabalho observou-se uma forte evidência de que a distribuição dos dados de aceitação depende do tipo de escala utilizada para a realização das avaliações hedônicas. No caso da escala hedônica híbrida, os resultados deste trabalho mostram fortes indícios de que ela produz dados de aceitação normalmente distribuídos, enquanto que as distribuições dos dados sensoriais produzidos pela escala autoajustável apresentaram desvios severos da normalidade.

AGRADECIMENTO

Os autores agradecem à **Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)**, pelo apoio financeiro que fez possível a realização deste trabalho de pesquisa.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- BAKER, R. A. Subjective panel testing. *Ind. Quality Control*, v. 19, p. 22-28, 1962.
- BOX, G. E. P.; HUNTER, W. G.; HUNTER, J. S. *Statistics for Experimenters*, John Wiley & Sons, New York, USA, 1978, 653p.
- BROWN, M.B.; FORSYTHE, A.B. Robust Tests for Equality of Variances. *Journal of the American Statistical Association*, 69, 364 -367, 1974.
- EFRON, B; TIBSHIRANI, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, v.1, p. 54-77, 1986.
- GAY, C., A statistical assessment of taste-testing methods. Ph. D. Thesis, The University of Reading, 1988.
- GAY, C.; MEAD, R. A statistical appraisal of the problem of sensory measurement. *Journal of Sensory Studies*, v. 7, p. 205-228, 1992.
- GIOVANNI, M. E.; PANGBORN, R. M. Measurement of taste intensity and degree of

- liking of beverages by graphic scales and magnitude estimation. *Journal of Food Science*, v. 48, p. 1175-1182, 1983.
- HARRISON, S.; ELDER, L. W. Some applications of statistics to laboratory taste testing. *Food Technology*, v.4, p. 434-439, 1950.
- LAND D. G.; SHEPHERD R. Scaling and ranking methods In *Sensory Analysis of Foods*, ed. J. R. Piggott. Elsevier Applied Science, London, pp. 155-186, 1988.
- LAWLESS, H. T.; MALONE, G. J. The discriminative efficiency of common scaling methods. *Journal of Sensory Studies*, v. 1, p. 85-98, 1986a.
- MACFIE, H. J.; WAKELING, I. N. Designing consumer trials balanced first and higher orders of carry-over effects when only a subset of k samples from t may be tested. *Food Quality and Preference*, v. 6, p. 299-308, 1995.
- MCPHERSON, R. S; RANDALL, E. Line length measurement as a tool for food preference research. *Ecology of food and Nutrition*, v. 17, p. 149-156, 1985.
- MEAD, R.; GAY, C. Sequential design of sensory trials. *Food Quality and Preference*, v.6, p. 271-280, 1995.
- O'MAHONY, M. Some assumptions and difficulties with common statistics for sensory analysis. *Food Technology*, p. 75-82. 1982.
- PEARSON, E. S.; D'AGOSTINO, R. B.; BOWMAN, K. D. Tests for departure from normality: Comparison of powers. *Biometrika*, 64, p. 231-46, 1977.
- PERYAM, D. R. AND PILGRIM, F. J. Hedonic scale method of measuring food preferences. *Food Technology*, 11(9), *Suppl.* p. 9-14, 1957.
- RAMSEY, P.H. Power Differences Between Pairwise Multiple Comparisons. *Journal of the American Statistical Association*, 73, p. 363, 1978.
- SAS INSTITUTE INC. *SAS/IML® Software: Usage and Reference, Version 8.0*, SAS Institute Inc., Cary, NC, USA, 2000.
- VILLANUEVA, N. D. M, PETENATE A. J.; Da SILVA, M. A. A. P. Performance of three affective methods and diagnosis of the ANOVA model. *Food Quality and Preference*. Vol. 11, pp. 363-370, 2000.
- WILLIAMS, A. A. *et al. Manual of Sensory Difference Testing of Wines*, Long Ashton Research Station, 1979, 70 p.

CAPÍTULO 4

DESEMPENHO DA ESCALA HEDÔNICA HÍBRIDA E AUTOAJUSTÁVEL NA GERAÇÃO DE MAPAS INTERNOS DE PREFERÊNCIA

Resumo

A avaliação das características sensoriais dos alimentos, na forma como elas são percebidas pelos consumidores, é essencial para o desenvolvimento, otimização, melhoria da qualidade e avaliação do potencial de mercado dos produtos alimentícios. Por isso, novas técnicas estatísticas que permitem identificar a preferência individual de cada consumidor com relação aos produtos avaliados, bem como detectar segmentos de consumidores com diferentes padrões de preferência e consumo têm sido desenvolvidas, sendo uma das mais utilizadas o "Mapa Interno de Preferência (MDPREF)". Na coleta de dados realizados em estudos com consumidores, a escala hedônica de 9 pontos tem sido a mais utilizada. Porém, recentes pesquisas mostraram que esta escala tem baixo poder discriminativo, reduzindo a capacidade do MDPREF em segmentar amostras e consumidores. Assim, para estudos com consumidores, faz-se necessária a pesquisa de escalas alternativas que possuam maior poder discriminativo que a escala hedônica tradicional. O objetivo desta pesquisa foi avaliar o desempenho da escala hedônica estruturada, hedônica híbrida e autoajustável no estudo das preferências dos consumidores através do MDPREF. A aceitação global de 8 marcas comerciais de vinho tinto, a maioria deles varietal Cabernet Sauvignon, foi avaliada por 112 consumidores. Foram utilizados delineamentos experimentais balanceados para ordem de apresentação das escalas, ordem de apresentação das amostras e efeitos residuais. Os dados foram analisados através de ANOVA, teste de REGWQ, MDPREF e teste de Permutação. Os critérios de avaliação do desempenho da cada escala foram baseados no número de dimensões significativas nos mapas, número de consumidores significativamente ajustados ($p \leq 0,05$) e no grau de segmentação dos produtos e dos consumidores. Os resultados obtidos, sugeriram uma

superioridade da escala híbrida sobre a escala hedônica tradicional e autoajustável. O MDPREF gerado pelos dados da escala híbrida produziu um maior número de dimensões significativas de preferência ($p \leq 0,05$), trazendo como decorrência, uma porcentagem de 79,5% dos consumidores significativamente ajustados ($p \leq 0,05$), enquanto que o MDPREF gerado pela escala autoajustável ajustou 54,5% dos consumidores e o da escala hedônica, 51,8%. Em geral, a escala hedônica de 9 pontos apresentou um desempenho inferior ao das demais escalas. Os resultados do presente estudo sugerem fortemente que a escala hedônica híbrida é uma ferramenta válida e eficiente a ser utilizada na geração de Mapas Internos de Preferência associados a estudos com consumidores.

Abstract

Evaluating the sensorial characteristics of food as they are perceived by consumers is essential for the development, optimization, quality improvement and assessment of a potential market for food products. For this reason, new statistic techniques have been developed, which permit to identify each consumer's individual preference regarding the products evaluated, as well as detect segments of consumers with different preference patterns. Among such techniques, the "Internal Preference Map (MDPREF)" is one of the most used. To generate MDPREF, the 9-point hedonic scale has been widely employed for data collection. However, several researchers believe that this scale presents limitations, which decrease its discriminative power, reducing the MDPREF capacity to segment samples and consumers. Thus, in consumer studies, it is important to research alternative scales, presenting better performance than the traditional hedonic scale in the generation of MDPREF. The objective of this study was to evaluate the performance of the 9-point hedonic, hybrid and self-adjusting scales in the segmentation of samples and consumers using MDPREF methodology. The overall acceptability of 8 Cabernet Sauvignon red wines was evaluated by 112 consumers. The effects of presentation order –scales and samples– and "carry-over" were balanced. The data were analyzed by ANOVA, REGWQ test, MDPREF and permutation tests. Scale performance was judged using as criteria:

number of significant dimensions ($p \leq 0.05$) in the MDPREF, number of consumers significantly adjusted ($p \leq 0.05$) and the segmentation of products and consumers. The results obtained suggested a superiority of the hybrid scale over the traditional hedonic and self-adjusting scales. The MDPREF generated by the hybrid scale data produced the greatest number of significant dimensions ($p=5\%$), yielding 79.5% of the consumers significantly adjusted ($p=5\%$), while the MDPREF generated by the self-adjusting scale adjusted 54.5% of the consumers and that of the hedonic scale, 51.8%. Overall, the 9-point hedonic scale showed the worst performance in relation to the other researched scales. The results of the current study strongly suggest that the hybrid hedonic scale is a valid and efficient tool for using in the generation of Internal Preference Maps in consumer studies.

1. INTRODUÇÃO

Na indústria alimentícia moderna, a avaliação das características sensoriais dos produtos, na forma como elas são percebidas pelos consumidores, é um componente essencial no desenvolvimento, manutenção, otimização, melhoria da qualidade e avaliação do potencial de mercado de um determinado produto alimentício (SIDEL & STONE, 1993; PIGGOT, 1995).

Das metodologias utilizadas em testes sensoriais com consumidores, a escala hedônica estruturada de 9 pontos (PERYAM & PELGRIM, 1957), é sem dúvida a mais utilizada. Entretanto, a despeito de sua popularidade entre analistas sensoriais, esta escala apresenta uma série de limitações, como: i) oferece pouca liberdade para os indivíduos expressarem suas percepções sensoriais devido ao limitado número de categorias de respostas, ii) os valores numéricos atribuídos às categorias, embora numericamente possuam intervalos iguais, não refletem iguais diferenças em percepção, iii) efeitos numéricos e contextuais são mais prováveis de acontecerem no uso desta escala, devido a diferenças na interpretação das diversas categorias e a erros de hábito, onde o provador tende a repetir as mesmas respostas em situações onde ele/ela precisa fazer várias avaliações consecutivas e; iv) os dados coletados por esta escala freqüentemente não satisfazem os pressupostos

estatísticos de normalidade e homoscedasticidade exigidos pelos modelos ANOVA (VILLANUEVA *et al.*, 2000; GAY & MEAD, 1992; O'MAHONY, 1982; GIOVANNI & PANGBORN, 1983; MCPHERSON & RANDALL, 1985).

Recentemente, VILLANUEVA *et al.* (2001) realizaram uma série de estudos com consumidores, com o objetivo de identificarem escalas alternativas que minimizassem alguns dos problemas apresentados pela escala hedônica tradicional. Os autores avaliaram, de forma comparativa, o desempenho da escala hedônica de 9 pontos e da escala de ordenação, com o das escalas autoajustável e hedônica híbrida; as duas últimas ainda muito pouco exploradas em testes com consumidores. Nesta pesquisa os autores exploraram as referidas escalas, em situação onde os consumidores avaliaram cinco amostras de suco de laranja em uma única sessão de degustação. Os resultados encontrados por VILLANUEVA *et al.* (2001), sugeriram uma superioridade da escala hedônica híbrida sobre as escalas hedônica tradicional e autoajustável, tanto em função do poder discriminativo, como em relação à adequação dos dados às suposições de normalidade e homoscedasticidade. Por sua vez, a escala autoajustável apresentou poder discriminativo ligeiramente maior do que a escala hedônica estruturada, a despeito dos dados da primeira apresentarem maior variabilidade aleatória e problemas de não normalidade dos resíduos. Assim, os resultados encontrados por VILLANUEVA *et al.* (2001) apontaram a vantagem do uso de escalas alternativas, de forma a transpor-se algumas limitações associadas à tradicional escala hedônica; limitações estas que concorrem para a redução do seu poder discriminativo. Infelizmente, problemas de não normalidade dos resíduos e heteroscedasticidade das variâncias não são as únicas limitações encontradas por analistas sensoriais em testes com consumidores.

Na análise dos dados de aceitação através de modelos ANOVA usuais, a hipótese de igualdade de aceitação entre as amostras é testada através de testes de médias como Tukey, Scheffé, etc. Entretanto, cabe observar que as médias podem ser afetadas por valores extremos, por distribuições assimétricas ou por distribuições multi-modais dos dados. Em testes com consumidores, esta limitação é grave porque impede a detecção de segmentos de indivíduos em função das amostras de sua

preferência. Por exemplo, se em um teste de aceitação 50% dos consumidores mostrarem alta aceitação por uma determinada amostra, e os demais consumidores mostrarem baixa aceitação por esta amostra, a média dos dados coletados resultará num valor intermediário de aceitação, o que não representará a verdadeira opinião de nenhum dos dois grupos de indivíduos. Assim, na análise ANOVA e demais testes de médias a ela associados -Tukey, Bonferroni, etc.-, informações sobre preferências individuais dos consumidores ou de grupos de consumidores são perdidas.

No entanto, a detecção de preferências individuais tem sido fundamental para o moderno profissional da área de marketing, que em função da atual competitividade dos mercados, procura identificar consumidores potenciais e dirigir a otimização e venda de produtos para mercados específicos.

Pelos motivos expostos, novas técnicas estatísticas que permitem identificar a preferência individual de cada consumidor com relação aos produtos avaliados têm sido desenvolvidas. Dentre elas, a técnica intitulada "Mapa Interno de Preferência (MDPREF)" tem sido uma das mais utilizadas (GREENHOFF & MACFIE, 1994; FALAHEE & MACRAE, 1995; DAILLANT-SPINLER *et al.*, 1996; ARDITTI, 1997; HELGESEN *et al.*, 1997; HEYD & DANZART, 1998; MONTELEONE *et al.*, 1998; GUINARD, UOTANI & SCHLICH, 2001).

O Mapa Interno de Preferência (MDPREF) dá um tratamento multidimensional aos dados afetivos, semelhante a uma Análise de Componentes Principais (CHANG & CARROL, 1968; GREENHOFF & MACFIE, 1994). Numa primeira etapa, a maior variação dentro dos dados hedônicos é identificada, extraída como preferência e alocada na primeira dimensão. Em seguida, uma segunda dimensão de preferência, ortogonal à primeira, é extraída, e assim sucessivamente até que toda a variação dos dados seja explicada. Desta forma, o MDPREF gera um gráfico onde os produtos são representados como pontos e cada indivíduo é representado por um vetor, sendo que os pontos mais próximos de um conjunto de vetores correspondem aos produtos de maior preferência daquele segmento de consumidores. As grandes vantagens do MDPREF sobre a ANOVA e tradicionais testes de médias são: a identificação da

preferência individual de cada consumidor em relação aos produtos avaliados, e a identificação de grupos de consumidores segmentados em função de suas preferências, possibilitando o estudo de cada segmento em função de sexo, hábitos e padrões de consumo, renda, nível educacional, etc. (GREENHOFF & MACFIE, 1994; FALAHEE & MACRAE, 1995; HELGESEN *et al.*, 1997). Este tipo de conhecimento tem sido muito útil para os profissionais de marketing das empresas alimentícias e por este motivo, desde sua proposta inicial por CHANG & CARROL (1968) e de estudos posteriores realizados por GREENHOFF & MACFIE (1994), o uso do MDPREF cresceu muito rapidamente neste setor.

Na literatura científica, o MDPREF também é largamente utilizado e, via de regra, associado à escala hedônica estruturada de 7 ou 9 pontos, para coleta de dados afetivos (MONTELONE *et al.*, 1998; HEYD & DANZART, 1998; YACKINOUS, WEE & GUINARD, 1999; MURRAY & DELAHUNTY, 2000; GUINARD *et al.*, 2001). É surpreendente que poucos cientistas tenham explorado métodos alternativos de coleta de respostas hedônicas junto aos consumidores, os quais produzissem uma melhoria no desempenho do MDPREF em relação à segmentação das amostras e dos indivíduos.

2. OBJETIVOS

1. Avaliar o desempenho das escalas hedônica estruturada, hedônica híbrida e autoajustável, na identificação de segmentos pré-existent de consumidores e seus padrões de preferência através da técnica de Mapa Interno de Preferência, tendo como critério de julgamento o grau de segmentação de amostras e dos consumidores.

3. MATERIAIS E MÉTODOS

3.1 Produtos

Foram avaliadas 8 marcas comerciais nacionais e importadas de vinho tinto, sendo 6 (seis) de varietal Cabernet Sauvignon, 1 (uma) de varietal Cabernet e 1

(uma) de vinho “Assemblage” elaborado com uma mistura de uvas Cabernet Franc e Merlot. Todas as marcas foram escolhidas em função de serem largamente comercializadas no mercado consumidor brasileiro. Os vinhos importados foram procedentes da Argentina e do Chile.

A Tabela 1 lista a linha varietal de cada marca de vinho utilizado na presente pesquisa, bem como sua safra, região de procedência, graduação alcoólica, preço e nota média de avaliação sensorial.

Tabela 1. Linhas varietais, safra, procedência, graduação alcoólica, preço e notas de avaliação das marcas de vinho tintos analisadas.

Amostra	Varietal	Safra	Procedência	Grau alcoólico (% Vol)	Preço⁽¹⁾ (R\$)	Nota de Avaliação⁽²⁾
1	Cabernet Sauvignon (Reserva)	1997	RS, Brasil	11,3	19,90	69,9
2	Cabernet Sauvignon (Reserva)	2000	RS, Brasil	11,5	15,15	70,6
3	Cabernet Sauvignon (Reserva)	2000	RS, Brasil	11,5	10,90	71,0
4	Cabernet Sauvignon	1999	Mendoza, Argentina	12,5	12,30	78,5
5	Cabernet Franc/Merlot	Sem informação	RS, Brasil	10,5	6,60	60,1
6	Cabernet	2000	RS, Brasil	11,5	6,90	67,0
7	Cabernet Sauvignon (Reserva)	1998	RS, Brasil	12	19,95	74,8
8	Cabernet Sauvignon (Reserva)	1999	Lontué, Chile	12,5	12,50	82,0

(1) Data de compra: junho e julho de 2001.

(2) Notas médias de avaliação obtidas a partir de testes cegos realizados por sete avaliadores especializados da EMBRAPA. Esses testes, liderados pelo Dr. Alberto Miele, foram realizados utilizando a metodologia da Associação Internacional de Enólogos e uma escala de 0 a 100.

3.2 Consumidores

Os consumidores foram recrutados dentro da comunidade universitária da Universidade Estadual de Campinas-UNICAMP, Campinas, SP, e do Instituto de Tecnologia de Alimentos (ITAL), Campinas, SP. Assim, professores, pesquisadores, técnicos e alunos de pós-graduação foram convidados a participarem do teste, através do preenchimento de uma ficha de recrutamento (Figura 1). Das 150 fichas respondidas, foram selecionados 112 consumidores habituais de vinho tinto, sendo 48 mulheres e 64 homens, de idade variando entre 20 e 55 anos, tendo-se como critério de seleção: frequência de consumo mínimo de até 1 copo de vinho tinto seco por semana, durante o inverno, e gostar de vinho tinto em um grau igual ou superior a “gostar moderadamente” na escala hedônica (Figura 1).

Análise Sensorial de Vinho Tinto Seco	
Nome: _____	Idade: _____ Curso: _____
1. Você gosta de vinho tinto seco?	Sim <input type="checkbox"/> Não <input type="checkbox"/>
2. Indique na escala abaixo, o quanto você gosta ou desgosta de vinho tinto seco :	
Gosto extremamente	<input type="checkbox"/>
Gosto muito	<input type="checkbox"/>
Gosto moderadamente	<input type="checkbox"/>
Gosto ligeiramente	<input type="checkbox"/>
Nem gosto/nem desgosto	<input type="checkbox"/>
Desgosto ligeiramente	<input type="checkbox"/>
Desgosto moderadamente	<input type="checkbox"/>
Desgosto muito	<input type="checkbox"/>
Desgosto extremamente	<input type="checkbox"/>
3. Indique na escala abaixo, quanto de vinho tinto seco você consome em média durante os meses de inverno:	
Quase sempre (quase todo dia)	<input type="checkbox"/>
Muito (3 a 4 copos/ semana)	<input type="checkbox"/>
Moderado (1 a 2 copos/ semana)	<input type="checkbox"/>
Pouco (até 1 copo por semana)	<input type="checkbox"/>
Quase nunca (menos de 1 copo/mês)	<input type="checkbox"/>

Figura 1. Ficha para recrutamento de consumidores de vinho tinto seco.

3.3 Delineamento experimental e modelo estatístico

Para avaliar sem vícios o desempenho das escalas na obtenção dos mapas de preferência, os 112 consumidores foram aleatoriamente divididos em três grupos contendo 40, 40 e 32 indivíduos respectivamente. Assim para cada grupo, a ordem de utilização das escalas foi balanceada seguindo-se um delineamento em quadrado latino 3x3 (Tabela 2), utilizado para controlar o efeito de ordem de apresentação das escalas. Desta forma, em cada grupo, todos os consumidores avaliaram todas as amostras, utilizando as três escalas, em três sessões distintas de degustação.

Tabela 2. Delineamento utilizado para balancear a ordem de apresentação das escalas aos consumidores (n=112 consumidores divididos em grupos de 40, 40 e 32 indivíduos cada).

GRUPO DE CONSUMIDORES	ORDEM DE APRESENTAÇÃO DAS ESCALAS		
	1ª Sessão	2ª Sessão	3ª Sessão
1º (40 indivíduos)	EHH	EHE	EAA
2º (40 indivíduos)	EHE	EAA	EHH
3º (32 indivíduos)	EAA	EHH	EHE

EHE= Escala hedônica estruturada; EHH= Escala hedônica híbrida; EAA= Escala autoajustável

Para cada escala, o delineamento experimental utilizado para apresentação das amostras aos consumidores e geração de respostas, foi o delineamento tipo "cross-over", balanceado para controlar efeitos de ordem de apresentação das amostras e efeitos residuais de primeira ordem. Este delineamento encontra-se tabelado em MACFIE & WAKELING (1995). Desta forma, os resultados experimentais foram analisados mediante o modelo estatístico para análise da variância associado a este tipo de delineamentos, o qual explicita-se a seguir

$$Y_{ij} = \mu + p_i + o_j + a_{d(i,j)} + r_{d(i,j-1)} + \varepsilon_{ij}, \quad i=1, \dots, 112; j=1, \dots, 8$$

Onde,

Y_{ij} : resposta do i -ésimo provador na avaliação da amostra $d_{(i,j)}$ apresentada na j -ésima ordem

μ : média geral

p_i : efeito do i -ésimo provador

o_j : efeito da j -ésima ordem de apresentação da amostra

$a_{d(i,j)}$: efeito direto da amostra $d_{(i,j)}$, avaliada pelo i -ésimo provador na j -ésima ordem de apresentação

$r_{d(i,j-1)}$: efeito residual de primeira ordem da amostra $d_{(i,j-1)}$, avaliada pelo i -ésimo provador na $(j-1)$ -ésima ordem de apresentação, onde $j \geq 2$ e $r_{d(i,0)} = 0$

ε_{ij} : perturbações aleatórias normais independentes e identicamente distribuídas com $E(\varepsilon_{ij}) = 0$ e $Var(\varepsilon_{ij}) = \sigma^2$

3.4 Métodos sensoriais afetivos

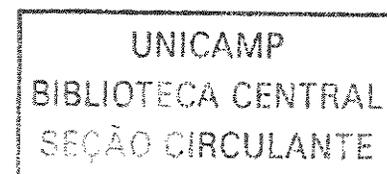
Os testes sensoriais foram realizados no Laboratório de Análise Sensorial do Departamento de Planejamento Alimentar e Nutrição da Universidade Estadual de Campinas, Campinas, SP, e no Laboratório de Análise Sensorial do Centro de Tecnologia de Carnes do Instituto de Tecnologia de Alimentos (ITAL), Campinas, SP. As amostras foram avaliadas utilizando-se três métodos afetivos: escala hedônica estruturada de 9 pontos, escala hedônica híbrida e escala autoajustável (Figuras 2, 3 e 4, respectivamente).

Três sessões de avaliações sensoriais distintas foram conduzidas; uma para cada escala. Em cada sessão, os indivíduos foram conduzidos às cabines individuais

de degustação, onde após se acomodarem, receberam instruções sobre a escala específica a ser utilizada, a natureza dos produtos e o tipo de avaliação a ser realizada. No caso das escalas hedônica estruturada e hedônica híbrida, as 8 amostras foram avaliadas monádicamente; já na escala autoajustável, as 8 amostras foram avaliadas comparativamente. Após a primeira sessão, os provadores compareceram às demais até completarem suas avaliações através das 3 escalas. Apenas uma única sessão de avaliação por dia era realizada por cada provador, para evitar o cansaço e problemas de fadiga sensorial.

Aproximadamente 30 ml de cada amostra foram servidos em cálices de vidro tipo tulipa, codificados com números aleatórios de 3 dígitos, e tampados com vidros de relógio. As amostras foram servidas à temperatura entre 16 e 18 °C, suficiente para proporcionar a sensação de frescor e permitir a detecção de odores no vinho (AMERINE & ROESSLER, 1983). Os consumidores foram solicitados a provarem o produto e, após a degustação, avaliarem de forma global, quanto gostaram ou desgostaram de cada amostra através da escala especificada naquela sessão de avaliação.

Após o término da terceira e última sessão, cada provador preencheu um questionário com perguntas concernentes a hábitos, grau de conhecimento e frequência de consumo de vinho tinto, cujos resultados serviram para delinear o perfil dos consumidores que participaram da pesquisa. O questionário utilizado para tal finalidade, apresenta-se no Apêndice.



ESCALA HEDÔNICA ESTRUTURADA

Nome: _____ Idade: _____ Data: _____

Você esta recebendo 8 amostras codificadas de vinho tinto seco. PROVE por favor, da esquerda para a direita cada uma das amostras avaliando-as de UM MODO GERAL. Utilizando a escala abaixo indique o quanto você gostou ou desgostou das mesmas.

- 9 Gostei extremamente
- 8 Gostei muito
- 7 Gostei moderadamente
- 6 Gostei ligeiramente
- 5 Nem gostei / nem desgostei
- 4 Desgostei ligeiramente
- 3 Desgostei moderadamente
- 2 Desgostei muito
- 1 Desgostei extremamente

Figura 2. Ficha de avaliação da aceitação utilizando a escala hedônica estruturada de 9 pontos.

ESCALA HEDÔNICA HÍBRIDA

Nome: _____ Idade: _____ Data: _____

Você esta recebendo 8 amostras codificadas de vinho tinto seco. Por favor, PROVE as amostras da esquerda para a direita e marque na escala correspondente um "x" e o código da amostra no lugar da escala que você achar que melhor represente o quanto você gostou ou desgostou de UM MODO GERAL de cada amostra.

Amostra N° _____

0 ————— 5 ————— 10
 Desgostei extremamente Nem gostei nem desgostei Gostei extremamente

Figura 3. Ficha de avaliação da aceitação utilizando a escala hedônica híbrida.

ESCALA AUTOAJUSTÁVEL

Nome: _____ Idade: _____ Data: _____

Você esta recebendo 8 amostras codificadas de vinho tinto seco. PROVE agora as amostras da esquerda para a direita avaliando-as de UM MODO GERAL. Escolha primeiro a amostra que você mais gostou e depois a que você menos gostou e anote os códigos dessas amostras respectivamente nos extremos direito e esquerdo da escala abaixo. Avalie agora as outras amostras de forma comparativa com aquelas alocadas nos extremos, marcando um "x" e o código da amostra no ponto da escala que você achar que melhor represente o quanto você gostou ou desgostou de cada amostra após degustá-las.

X ————— X
 Menos gostoso Mais gostoso

Figura 4. Ficha de avaliação sensorial utilizando a escala autoajustável.

3.5 Análise estatística

Os resultados de aceitação dos vinhos obtidos a partir das respostas dos 112 consumidores nas três escalas, foram analisados inicialmente através de ANOVA ($p=5\%$) para delineamentos tipo “cross-over”, considerando-se no modelo estatístico os seguintes efeitos: provador, ordem de apresentação, efeito residual e amostras. Diferenças de aceitação entre as amostras foram verificadas através de comparações pareadas das médias, aplicando-se o teste de REGWQ ao nível de 5% de significância.

Para cada escala, a representação gráfica das diferenças na aceitação das amostras e a identificação de cada indivíduo no espaço das preferências, foram obtidas através da metodologia tradicional do Mapa Interno de Preferência (GREENHOFF & MACFIE, 1994). Os Mapas de Preferência modernos representam as amostras através de elipses de nível de confiança de 95%. A obtenção das elipses de confiança de 95% no espaço de preferências foi realizada através de métodos “bootstrap” (EFRON & TIBSHIRANI, 1986). Para isso, foram retiradas $B=200$ amostras “bootstrap” com reposição, recalculando-se o modelo de preferência para cada amostra “bootstrap” de forma a se obter a variabilidade das posições das amostras no mapa. Este procedimento parte do pressuposto de que as 200 posições das amostras, obtidas a partir das re-amostragens, procedem de uma distribuição normal bivariada (ALTMAN, 1978).

O estabelecimento da dimensionalidade do espaço das preferências e a significância do ajuste dos dados de cada indivíduo pelo modelo de preferência, foram realizados através do teste de permutação (EDGINTON, 1987). Este procedimento rearranja aleatoriamente as respostas de cada indivíduo, produzindo novos conjuntos de dados permutados. Esses conjuntos são obtidos atribuindo aleatoriamente às amostras, as respostas hedônicas de cada indivíduo. Para isso, foram realizadas 500 permutações aleatórias, recalculando os modelos de preferência para cada permutação com a finalidade de se obter uma elipse de confiança de 95% no mapa de preferência original. Os indivíduos significativamente ajustados pelo

modelo são aqueles que se encontram localizados fora da elipse, significando que a probabilidade desses indivíduos terem atribuído seus valores ao acaso é menor que 5%. Os mapas de preferência foram gerados através de um programa em linguagem SAS (SAS INSTITUTE INC., CARY, NC, 2000).

O desempenho das três escalas na obtenção dos Mapas Internos de Preferência foi avaliado através dos seguintes parâmetros: i) porcentagem da variação total explicada pelas dimensões de preferência significativas ($p=5\%$), ii) grau de segmentação das amostras de vinho, iii) número de consumidores ajustados significativamente pelo modelo de preferência ($p=5\%$) e, iv) grau de segmentação dos consumidores no espaço de preferências.

4. RESULTADOS E DISCUSSÃO

4.1 Caracterização dos consumidores segundo seus hábitos de consumo de vinho tinto seco

O perfil dos 112 consumidores recrutados que participaram desta pesquisa, reflete a grande diversidade de indivíduos existentes na Universidade Estadual de Campinas e no Instituto de Tecnologia de Alimentos, ambos considerando centros de referência na produção científica e tecnológica do Brasil. Nestes centros, pessoas procedentes do estado de São Paulo e demais estados brasileiros, como também estrangeiros procedentes de diversos países, realizam diariamente suas atividades. Assim, os resultados dos questionários preenchidos pelos consumidores amostrados, revelaram grandes diferenças em relação à ocupação, procedência, faixa etária, nível de escolaridade e hábitos de consumo, entre outras características. Desta forma, dos 112 consumidores que participaram do teste sensorial, 57,1% pertenciam ao sexo masculino e 42,9% ao sexo feminino. A maioria dos indivíduos pertencia à faixa de idade entre 25 e 34 anos, representando 44,6% do total de consumidores (Tabela 3). Em relação ao lugar de procedência, 59,8% dos consumidores procediam do estado de São Paulo, 28,6% procediam de outros estados brasileiros e 11,6% eram estrangeiros, na sua maioria latino-americanos (Tabela 3). No que se refere ao nível

de escolaridade, 8% dos indivíduos possuíam segundo grau completo; 21,5% superior incompleto; 22,3% superior completo e 48,2% cursavam ou já haviam finalizado a pós-graduação -mestrado ou doutorado- (Tabela 3). Quanto à frequência de consumo durante os meses de inverno, 3,6% dos entrevistados declararam consumir vinho tinto seco quase sempre, 13,4% consumiam de 3 a 4 copos por semana, 38,4% de 1 a 2 copos por semana, 36,6% até um copo por semana e 8%, menos de um copo por mês.

Tabela 3. Distribuição dos consumidores por faixa etária segundo sexo, local de procedência, nível de escolaridade e frequência de consumo (n=112 consumidores).

Variáveis sócio demográficas	Faixa etária (%)				Total (%)
	18-24 anos	25-34 anos	35-44 anos	45-55 anos	
Sexo					
M	13,4	22,3	12,5	8,9	57,1
F	11,6	22,3	6,2	2,7	42,9
Local de procedência					
SP	18,8	27,7	8,0	5,4	59,8
Outros estados	4,5	13,4	6,2	4,5	28,6
Estrangeiros	1,8	3,6	4,5	1,8	11,6
Nível de escolaridade					
2º grau completo	2,7	1,8	2,7	0,9	8,0
Superior incompleto	17,0	3,6	0	0,9	21,5
Superior completo	5,4	13,4	1,8	1,8	22,3
Pós-graduação	0	25,9	14,3	8,0	48,2
Frequência de consumo					
Quase sempre	0,9	0,9	0,9	0,9	3,6
3 a 4 copos/semana	2,7	6,2	4,5	0	13,4
1 a 2 copos/semana	8,0	21,4	7,1	1,8	38,4
Até 1copo/semana	9,8	16,1	4,5	6,2	36,6
Menos de 1 copo/mês	3,6	0	1,8	2,7	8,0

A Tabela 4 apresenta dados referentes à forma de obtenção de conhecimentos sobre vinhos, hábitos e local de consumo dos 112 consumidores que participaram do teste. Quanto à forma como os consumidores adquirem conhecimentos sobre a bebida, tanto homens como mulheres apontaram em primeiro lugar a obtenção de informações através de amigos (62,5% dos homens e

79,2% das mulheres), em segundo lugar, experimentando novos vinhos (67,2% dos homens e 54,2% das mulheres), seguido de informações oferecidas em restaurantes e revistas especializadas. Degustação em adegas e empórios, foi um dos itens menos indicados pelos consumidores como meio de obter informações ou novo conhecimento sobre vinhos, notadamente junto ao segmento feminino.

Quanto à forma de consumo (Tabela 4), tanto homens como mulheres revelaram alta preferência de consumo de vinho tinto como acompanhamento para queijos (68,8% dos homens e 62,5% das mulheres), seguido como acompanhamento para massas, pizzas e suflês (56,2% dos homens e 60,4% das mulheres). Em terceiro lugar foi relatado o consumo de vinho tinto como aperitivo, notadamente pelos homens. O acompanhamento desta bebida para carnes e aves entre outras formas de consumo, também foi relatado.

Quanto aos locais de consumo de vinho tinto, a Tabela 4 mostra que tanto homens como mulheres revelaram em primeiro lugar o consumo domiciliar (84,4% dos homens e 87,5% das mulheres), em segundo lugar, o consumo em restaurantes (35,9% dos homens e 50,0% das mulheres), seguido do consumo em festas, aniversários e reuniões sociais (37,5% dos homens e 43,8% das mulheres). Bares, empórios e casas noturnas obtiveram pouca indicação por parte dos entrevistados.

Tabela 4. Dados⁽¹⁾ dos consumidores entrevistados relacionados à forma de adquirir conhecimento sobre vinhos, hábitos e local de consumo segundo sexo, em %.

	Homens (%)	Mulheres (%)
Conhecimento sobre vinhos		
Revistas especializadas	28,1	25,0
Nos restaurantes	37,5	20,8
Informações com amigos	62,5	79,2
Adegas e empórios	34,4	12,5
Experimentando novos vinhos	67,2	54,2
Outros	31,2	27,1
Hábitos de consumo		
Sozinho como aperitivo	57,8	43,8
Com massas	56,2	60,4
Com carnes	42,2	54,2
Com queijos	68,8	62,5
Outros	34,4	31,2
Local de consumo		
Em casa	84,4	87,5
Em restaurantes	35,9	50,0
Em bares, empórios e casas noturnas	23,4	18,8
Festas, aniversários e reuniões sociais	37,5	43,8

(1) Dados segmentados por sexo e expressos como % = $\frac{n^\circ \text{ de respostas homens (mulheres)}}{n^\circ \text{ homens (mulheres)}} \times 100$; cada indivíduo pode marcar várias categorias de resposta.

4.2 ANOVA e Mapa Interno de Preferência

Estatísticas básicas referentes aos dados coletados através das escalas hedônica, autoajustável e híbrida são apresentadas na Tabela 5, onde se observa que as médias de aceitação das amostras mostraram, em geral, uma mesma tendência nas três escalas: a amostra 6 foi a mais aceita e as amostras 1 e 7 as menos aceitas.

Tabela 5. Estatísticas básicas dos dados de aceitação global de vinho tinto seco coletados através das escalas hedônica estruturada, autoajustável e híbrida (n= 112 consumidores).

Escala	Amostra	Média ⁽¹⁾		Desvio Padrão
Hedônica de 9 pontos⁽²⁾	1	5,1	ed	2,06
	2	6,0	abc	1,75
	3	6,2	ab	1,71
	4	5,5	cde	2,25
	5	5,6	bcd	2,05
	6	6,5	a	1,68
	7	4,8	e	2,20
	8	5,6	bcd	2,18
Autoajustável⁽³⁾	1	3,8	d	3,55
	2	6,0	ab	3,08
	3	5,5	abc	2,94
	4	4,4	cd	3,55
	5	4,4	cd	3,71
	6	7,0	a	3,15
	7	3,1	d	3,47
	8	5,4	bc	3,72
Hedônica híbrida⁽⁴⁾	1	4,4	c	2,60
	2	6,0	ab	2,38
	3	5,8	ab	2,17
	4	5,2	b	2,88
	5	5,4	b	2,34
	6	6,6	a	2,01
	7	4,2	c	2,76
	8	6,0	ab	2,43

(1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ao nível de 5% de significância. Esses resultados foram obtidos através do teste de REGWQ.

(2) 1= desgostei extremamente, 9= gostei extremamente; (3) 0= menos gostoso, 10= mais gostoso.

(4) 0=desgostei extremamente, 10= gostei extremamente.

Os valores dos desvios padrões (Tabela 5), mostram claramente que houve maior variabilidade nos dados produzidos pela escala autoajustável. Esses resultados confirmam aqueles já observados por VILLANUEVA *et al.* (2000) e VILLANUEVA *et al.* (2001). Assim, uma vez mais se verifica que na escala autoajustável, a eliminação das diferenças na expansibilidade entre os provadores, traz como decorrência, o aumento da variabilidade das respostas hedônicas.

No teste de REGWQ para comparação de médias (Tabela 5), a escala hedônica híbrida detectou 14 diferenças significativas (p= 5%), enquanto que a escala autoajustável detectou 13 diferenças e a escala hedônica de 9 pontos, 12

diferenças. Neste sentido, o poder discriminativo da escala hedônica de 9 pontos foi ligeiramente inferior ao da escala autoajustável, e esta por sua vez inferior à escala híbrida. Esses resultados confirmam aqueles já observados por VILLANUEVA *et al.* (2001) e sugerem que os dados coletados através da escala hedônica híbrida apresentam maior capacidade de discriminação de diferenças entre as amostras.

O fato da escala autoajustável ter apresentado poder discriminativo ligeiramente superior ao da escala hedônica de 9 pontos, a despeito de mostrar maior variabilidade dos dados, pode ser atribuído à própria natureza cognitiva desta escala, a qual exige que o provador primeiramente compare as amostras, ordenando-as em termos hedônicos, e somente então atribua seus valores hedônicos a cada produto. Desta forma, a escala autoajustável obriga o consumidor a realizar uma avaliação mais criteriosa das amostras. Em complementação, na escala autoajustável, o indivíduo é também forçado a utilizar o comprimento total da escala, e assim as distâncias entre as amostras no contínuo psicológico são maximizadas, tornando a escala mais discriminativa. Em decorrência desses fatos, a escala autoajustável produziu uma separação maior entre as amostras, apresentando uma diferença de 3,9 entre a média da amostra mais aceita (amostra 6) e a menos aceita (amostra 7), enquanto que na escala hedônica de 9 pontos esta diferença foi de 1,7 e na escala híbrida de 2,4. Porém, na escala autoajustável, os valores absolutos das médias de cada amostra não medem o grau de aceitação dos produtos e não permitem prever o quanto o consumidor gostou ou desgostou de cada produto, como ocorre nas outras duas escalas. Esta representa uma limitação da escala auto-ajustável que deve ser avaliada com cuidado pelo experimentador antes da aplicação do teste sensorial.

4.2.1 Efeito da ordem de apresentação das amostras e efeito residual

Os resultados da ANOVA dos dados obtidos através das escalas hedônica de 9 pontos, autoajustável e híbrida revelaram que nas três escalas os efeitos de ordem de apresentação foram significativos, com valores de pF_{ordem} de 0,026; 0,022 e 0,021 respectivamente. Estes efeitos podem ser claramente visualizados através da Tabela

6, que mostra que para as três escalas, os valores hedônicos médios atribuídos às amostras tendem a diminuir na medida em que a ordem de apresentação das amostras aumenta, ou seja, amostras apresentadas por último tendem a receber valores hedônicos menores. Observa-se na Tabela 6, que todas as escalas apresentaram além de efeitos de primeira ordem, também efeitos de outras ordens, as quais variaram de escala para escala. Assim, a despeito do tipo de escala a ser utilizada em testes afetivos, verifica-se a importância do controle dos efeitos de ordem de apresentação das amostras no incremento da sensibilidade do experimento após a eliminação da porção do erro experimental devido às ordens de apresentação.

Não foram detectadas efeitos residuais na análise da variância dos dados para nenhuma escala ($p \geq 0,23$).

Tabela 6. Médias⁽¹⁾ de aceitação global para cada ordem de apresentação das amostras dos dados coletados através da escala hedônica de 9 pontos, escala autoajustável e escala híbrida (n=112 consumidores).

Ordem de apresentação das amostras	Escala hedônica de 9 pontos ⁽²⁾	Escala autoajustável ⁽³⁾	Escala hedônica híbrida ⁽⁴⁾
1	6,2 a	6,3 a	6,4 a
2	5,6 ab	4,2 b	5,6 ab
3	5,6 ab	5,0 ab	5,4 b
4	5,8 ab	4,9 ab	5,8 ab
5	5,6 ab	4,5 b	5,0 b
6	5,5 b	4,9 ab	5,0 b
7	5,7 ab	5,0 ab	5,3 b
8	5,3 b	5,0 ab	5,2 b

(1) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ao nível de 5% de significância. Esses resultados foram obtidos através do teste de REGWQ.

(2) 1= desgostei extremamente, 9= gostei extremamente; (3) 0= menos gostoso, 10= mais gostoso.

(4) 0=desgostei extremamente, 10= gostei extremamente.

4.2.2 Análise dos Mapas Internos de Preferência

As Figuras de 6 a 13 mostram os Mapas Internos de Preferência (MDPREF) gerados através da escala hedônica de 9 pontos, escala hedônica híbrida e escala autoajustável, respectivamente. Esses mapas foram construídos de forma a representar graficamente as diferenças na aceitação das amostras e as preferências individuais de cada consumidor. Assim, as 112 respostas individuais dos

consumidores com relação a cada amostra de vinho avaliada, geraram para cada escala, um espaço sensorial multidimensional representado por dimensões de preferência que explicam a variação total das respostas sensoriais.

Nos Mapas de Preferência desenvolvidos neste estudo, cada amostra é representada no espaço sensorial afetivo dentro de uma elipse que delimita a região de confiança de 95% associada à aceitação da mesma. Desta forma, as elipses próximas entre si, indicam amostras que obtiveram níveis de aceitação similares; enquanto elipses ocupando posições opostas no gráfico, sugerem amostras com grandes diferenças de aceitação entre si. Em complementação, em cada mapa, cada consumidor encontra-se representado por um número entre 1 e 112, sendo que cada indivíduo foi alocado no gráfico, próximo às amostras de sua preferência. Nas Figuras de 6 a 13, os consumidores e amostras foram alocados em gráficos distintos para melhor visualização dos resultados.

De um modo geral, pode-se dizer que os Mapas Internos de Preferência consistem na representação gráfica tanto da variabilidade existente entre as amostras com relação à aceitação das mesmas entre os consumidores, como da variabilidade mostrada pelos indivíduos com relação às suas preferências individuais. Assim, antes que se faça a interpretação de cada mapa, é fundamental verificar a porcentagem da variância explicada por cada dimensão de preferência.

Na Tabela 7 apresentam-se para cada escala, as porcentagens da variância explicada pelas três primeiras dimensões de preferência dos mapas. Observa-se que a primeira e segunda dimensões de preferência produziram porcentagens de variância explicada muito similares entre as três escalas. Porém, na escala híbrida, a terceira dimensão de preferência produziu porcentagens maiores de explicação, acumulando nas suas três primeiras dimensões 66% de variância explicada. Já a escala hedônica de 9 pontos produziu menores porcentagens de explicação nas três dimensões de preferência, acumulando 61,73% de variância explicada. Esses resultados sugerem um melhor desempenho da escala híbrida em relação à escala hedônica tradicional.

Tabela 7. Porcentagens da variância explicada em cada escala pelas três primeiras dimensões de preferência.

Escala	Dimensão 1 (%)	Dimensão 2 (%)	Dimensão 3 (%)	Total acumulado (%)
Hedônica de 9 pontos	30,23*	18,69*	12,81	61,73
Hedônica híbrida	30,40*	19,52*	16,07*	66,00
Autoajustável	31,60*	18,84*	14,18	64,62

(*) Dimensões significativas ao nível de 5% de significância detectada pelo teste de permutação.

No teste de permutação para o estabelecimento da dimensionalidade do espaço de preferências (Tabela 7), foram detectadas duas dimensões significativas ($p= 5\%$) para as escalas hedônica de 9 pontos e autoajustável, reduzindo desta forma, a dimensionalidade das mesmas a duas. No caso da escala híbrida, foram detectadas três dimensões significativas ($p= 5\%$), reduzindo à dimensionalidade desta escala a três. Isto confirma que a escala híbrida produziu dados com maior porcentagem de explicação da variância total dos dados hedônicos do que as escalas autoajustável e hedônica de 9 pontos. Em decorrência, espera-se que o MDPREF da escala híbrida produza uma segmentação mais fina das amostras e dos consumidores do que os mapas correspondentes às demais escalas.

Na Tabela 8, apresenta-se o número total de consumidores significativamente ajustados ($p= 5\%$) pelas dimensões significativas dos mapas de cada escala. Estes foram 58 na escala hedônica estruturada, 89 na escala híbrida e 61 na escala autoajustável. Assim, a escala híbrida ajustou um número consideravelmente maior de consumidores do que a escala hedônica estruturada e autoajustável, indicando neste aspecto uma melhor performance desta escala. A Tabela 8, mostra também para cada escala, o número de consumidores significativamente ajustados em cada quadrante ($p=5\%$) do plano formado pelas duas primeiras dimensões de preferência dos mapas. No caso da escala híbrida, esse número foi determinado pelas três primeiras dimensões de preferência. Os consumidores significativos são aqueles que se localizaram fora do limite crítico da região de confiança de 95% (Figuras 7, 9, 11 e 13) tendo sido aqueles que mostraram correlações significativas ($p=5\%$) de seus

valores hedônicos com as duas ou três primeiras dimensões de preferência. Pode-se considerar que esses consumidores realmente revelaram possuir preferências com relação aos vinhos avaliados. Como se observa na Tabela 8, nas três escalas, os quadrantes à direita do mapa reúnem a maior proporção de consumidores. Entretanto, a segmentação desses consumidores mostrou-se similar para as escalas hedônica de 9 pontos e híbrida e distinta para a escala autoajustável. Nas escalas hedônica de 9 pontos e híbrida a maior proporção de consumidores localizou-se no quadrante superior direito, enquanto que na escala autoajustável, os consumidores dividem-se quase igualmente entre os dois quadrantes à direita do mapa.

Tabela 8. Número de consumidores significativamente ajustados ($p=5\%$) pelas dimensões significativas de preferência e número de consumidores significativamente ajustados ($p=5\%$) em cada quadrante do plano cartesiano dos MDPREF correspondentes às escalas hedônica de 9 pontos, híbrida e autoajustável.

Escala	Total	Superior direito	Inferior direito	Superior esquerdo	Inferior esquerdo
Hedônica de 9 pontos	58	30	16	11	1
Hedônica híbrida	89	37	27	20	5
Autoajustável	61	25	23	12	1

Para construção dos Mapas de Preferência mostrados nas Figuras de 5 e 6 (escala hedônica estruturada) e nas Figuras 8 e 9 (escala auto-ajustável), utilizou-se as duas primeiras dimensões geradas pelo espaço de preferências, uma vez que somente elas mostraram significância estatística ($p=5\%$). No caso da escala híbrida (Figuras 9, 10, 11 e 12), utilizou-se as 3 primeiras dimensões de preferência, dada a significância estatística das mesmas.

Na Figura 5, correspondente à escala hedônica de 9 pontos, o espaço de preferências definido pelas duas primeiras dimensões, mostra os vinhos divididos em três segmentos distintos: o primeiro segmento, composto pelas amostras 2, 3, 5 e 6, situadas nos dois quadrantes à direita do mapa; um segundo segmento, composto pelas amostras 4 e 8, situadas no lado esquerdo superior do mapa e um terceiro

segmento, composto pelas amostras 1 e 7, situado no quadrante esquerdo inferior do mapa. Os três segmentos revelados na Figura 5 distinguem-se claramente entre si ($p=5\%$), com relação aos níveis de aceitação entre os consumidores.

A Figura 6, que mostra a localização dos consumidores em função de suas preferências individuais com relação aos vinhos, indica que a maioria dos indivíduos significativamente ajustados pelo modelo de preferência, está alocada nos mesmos quadrantes onde se encontram as elipses correspondentes às amostras 2, 3, 5 e 6. Portanto, pode-se concluir que esses vinhos foram preferidos por esses indivíduos, os quais representam a maioria dos consumidores que participaram do teste. Duas dessas amostras (vinhos 5 e 6), correspondem aos dois vinhos nacionais de menor preço e menor definição varietal dentro do conjunto de amostras testadas (Tabela 1). As outras amostras (vinhos 2 e 3) representam outros dois vinhos nacionais, de preços intermediários, porém com varietal bem definido -Cabernet Sauvignon-. Aproximadamente 79% dos consumidores que mostraram preferência significativa ($p=5\%$) preferiram esses vinhos (Tabela 8).

A Figura 6, mostra também uma pequena proporção de consumidores que foram alocados próximos às elipses correspondentes aos vinhos 4 e 8. Essas amostras correspondem aos vinhos importados de varietal Cabernet Sauvignon, tendo sido as amostras preferidas pelos consumidores situados nesse quadrante. Cerca de 19% dos consumidores que mostraram preferência significativa ($p=5\%$) preferiram esses vinhos (Tabela 8).

Finalmente, a Figura 6 mostra também, duas amostras -1 e 7- que se situam distantes da grande maioria dos consumidores que participaram do teste. Estes foram os vinhos menos aceitos pelos consumidores e neste caso, tratam-se justamente de dois vinhos nacionais varietais Cabernet Sauvignon de alto preço. Esses vinhos não atenderam às expectativas de qualidade dos consumidores nem aos padrões de qualidade estabelecidos pelos avaliadores especializados de vinhos (Tabela 1). Apenas 2% dos consumidores que revelaram preferência significativa preferiram esses vinhos ($p=5\%$) (Tabela 8).

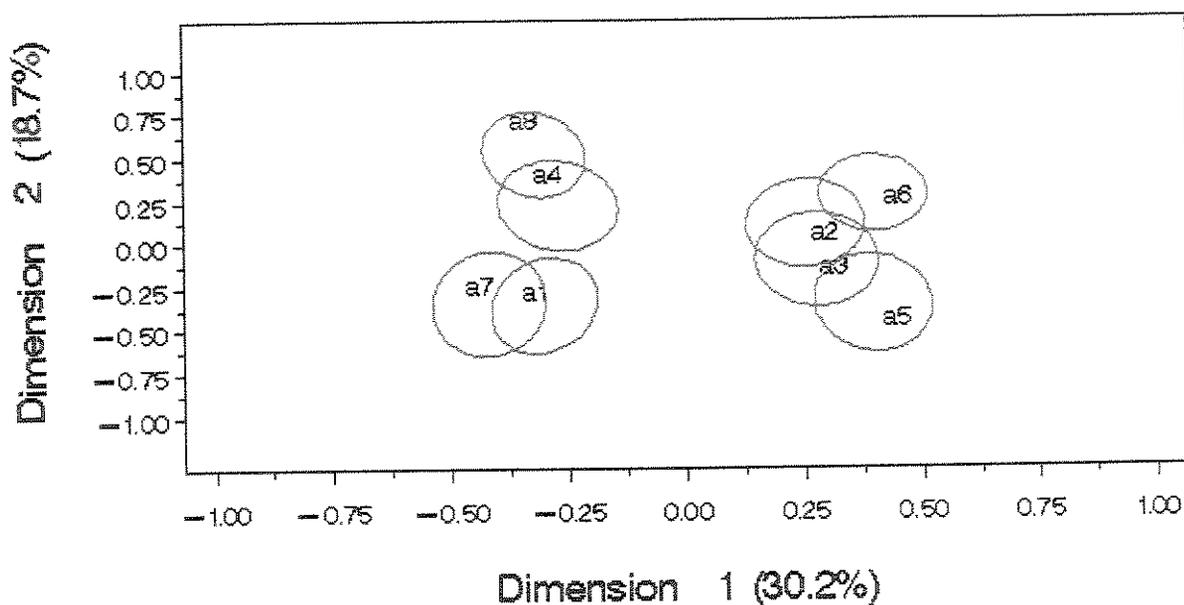


Figura 5. Mapa Interno de Preferência dos dados de aceitação global gerados pela escala hedônica de 9 pontos, que mostra a configuração dos vinhos nas dimensões de preferência 1 e 2.

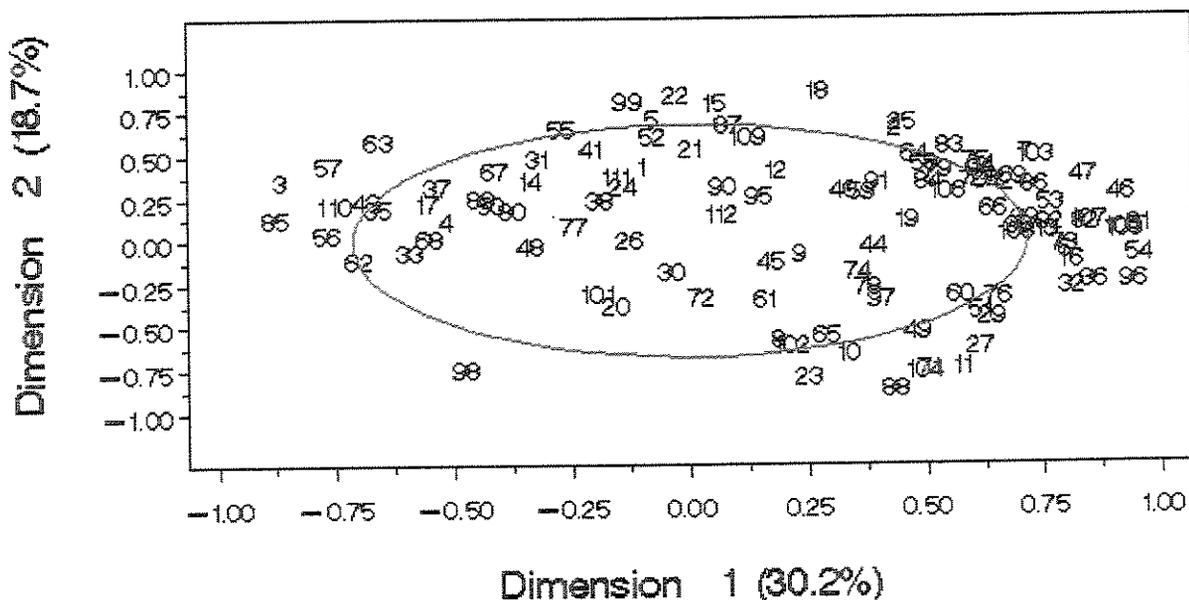


Figura 6. Mapa Interno de Preferência dos dados de aceitação global gerados pela escala hedônica de 9 pontos, que mostra a configuração dos consumidores nas dimensões de preferência 1 e 2.

A Figura 7, mostra a configuração das amostras de vinho no espaço sensorial gerado pelos dados da escala autoajustável. De forma similar ao mapa gerado pelos dados da escala hedônica estruturada, na Figura 7, o espaço de preferências definidos pelas dimensões 1 e 2 mostram os vinhos divididos em três segmentos distintos. O primeiro segmento, composto pelas amostras 2, 3, 5 e 6, situadas nos quadrantes à direita do mapa (Figuras 7 e 8), correspondem aos vinhos preferidos pelos consumidores alocados nesses quadrantes. Ainda na Figura 7, observa-se que a amostra 5, vinho nacional "Assemblage", elaborado com uvas Cabernet Franc e Merlot, encontra-se quase separada das amostras 2 (Cabernet Sauvignon), 3 (Cabernet Sauvignon) e 6 (Cabernet), indicando que a mesma distingue-se ligeiramente das demais em termos de preferência. A Tabela 8 revela que 79% dos consumidores significativos preferiram essas amostras ($p=5\%$).

Observa-se também na Figura 7, um segundo segmento composto pelas amostras 4 e 8, situadas no lado esquerdo superior do mapa. Essas amostras correspondem aos vinhos importados varietal Cabernet Sauvignon, e são as amostras preferidas dos consumidores que se situam no referido quadrante (Figura 8), os quais correspondem a cerca de 20% dos consumidores significativos ($p=5\%$).

Um terceiro segmento, composto pelas amostras 1 e 7, vinhos nacionais de alto preço, encontram-se situados no quadrante esquerdo inferior do mapa (Fig. 7) e correspondem às amostras de menor preferência, sendo preferidas por apenas 2% dos consumidores significativos ($p=5\%$).

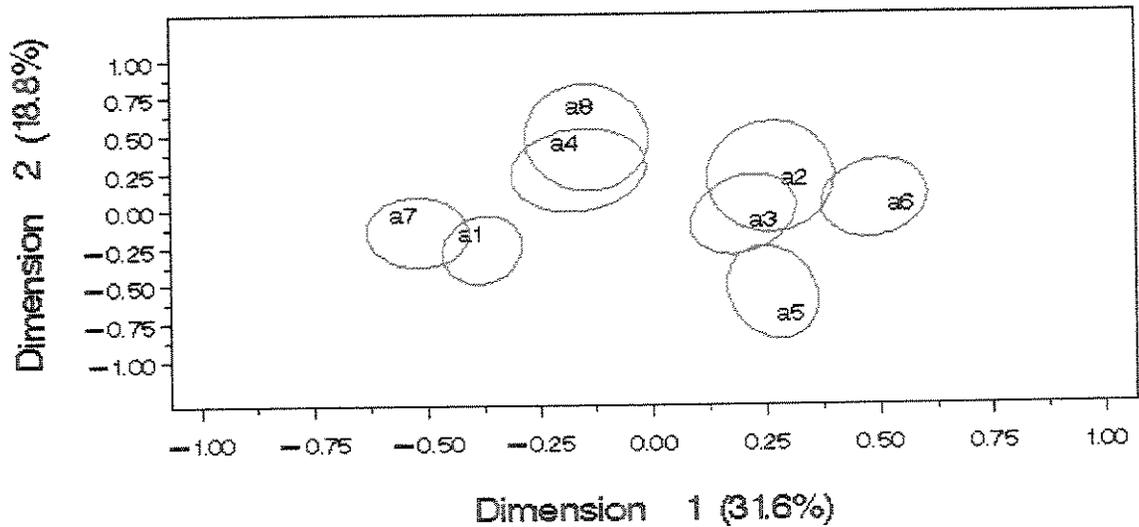


Figura 7. Mapa Interno de Preferência dos dados de aceitação global gerados pela escala autoajustável, que mostra a configuração dos vinhos nas dimensões de preferência 1 e 2.

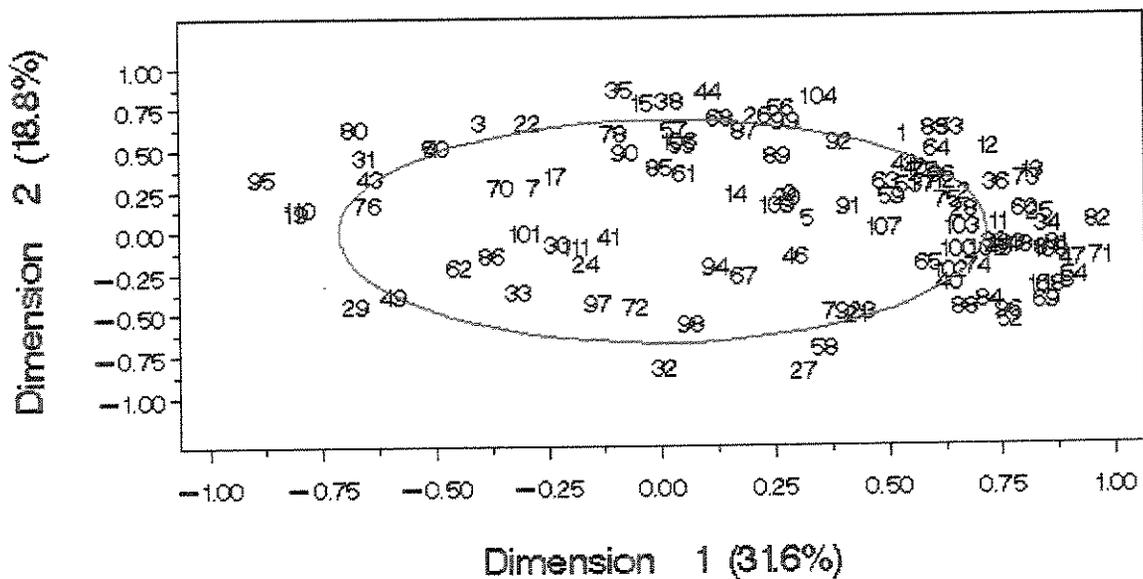


Figura 8. Mapa Interno de Preferência dos dados de aceitação global gerados pela escala autoajustável, que mostra a configuração dos consumidores nas dimensões de preferência 1 e 2.

A Figura 9, mostra a configuração das amostras de vinho no espaço sensorial gerado pelos dados da escala híbrida. Nesta figura, o espaço de preferências definidos pelas dimensões 1 e 2 mostram os vinhos divididos em cinco segmentos distintos. O primeiro segmento, composto por três vinhos varietais (amostras 2, 3, e 6), situadas no quadrante superior direito e área limítrofe do quadrante inferior, correspondem às amostras preferidas pelos consumidores situados nessa região, os quais, conforme mostra a Figura 10, representam uma grande proporção dos consumidores ajustados pelo modelo de preferência. Os vinhos 2 e 3 correspondem a vinhos varietais Cabernet Sauvignon e o vinho 6 ao varietal Cabernet, todos nacionais. Uma segunda segmentação pode também ser observada no quadrante direito inferior do mapa (Figura 10), a qual corresponde à amostra 5, vinho nacional "Assemblage" elaborado com uvas Cabernet Franc e Merlot, que obteve maior preferência junto a um grupo de consumidores distintos daqueles que preferiram as amostras 2, 3 e 6.

A Figura 10 revela ainda um terceiro grupo de consumidores, correspondentes àqueles situados mais à direita do mapa, os quais preferiram igualmente os dois segmentos de amostras. Esses resultados confirmam aqueles obtidos através do MDPREF da escala hedônica de 9 pontos (Figuras 5 e 6), porém no MDPREF construído com os dados coletados pela escala hedônica híbrida (Figuras 9 e 10), tanto as amostras como os consumidores foram mais bem segmentados: a amostra 5 se localizou numa região de 95% de confiança distinta das amostras 2, 3 e 6. Da mesma forma que ocorreu com a escala hedônica estruturada onde a maioria dos consumidores preferiram as amostras 2, 3, 5 e 6, no MDPREF da escala híbrida, 72% dos consumidores que mostraram preferências significativas, também preferiram essas amostras ($p=5\%$).

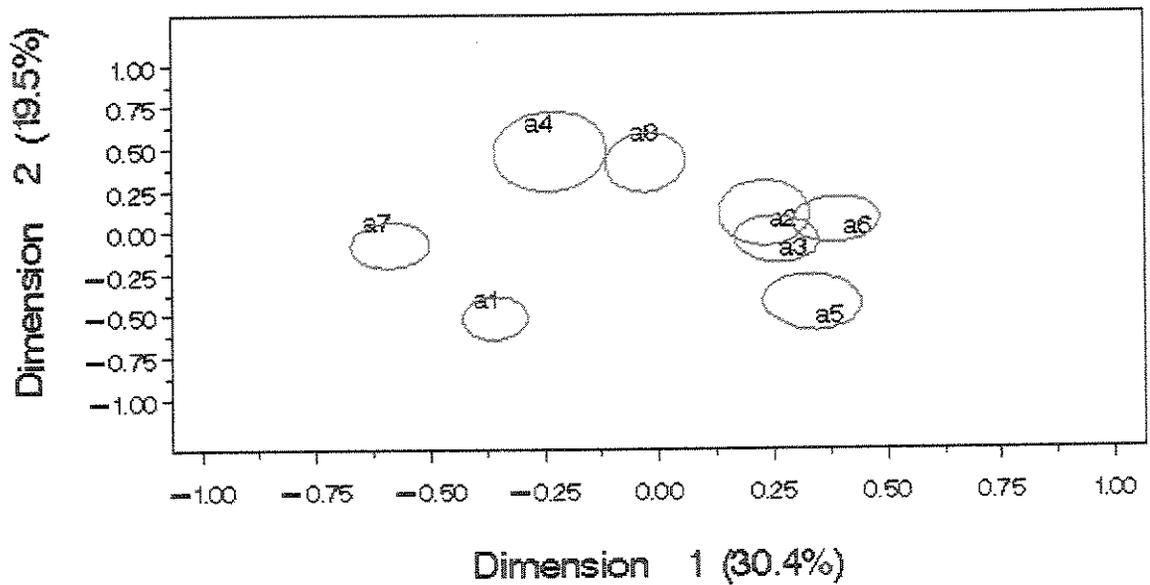


Figura 9. Mapa Interno de Preferência dos dados de aceitação global gerados pela escala hedônica híbrida, que mostra a configuração dos vinhos nas dimensões de preferência 1 e 2.

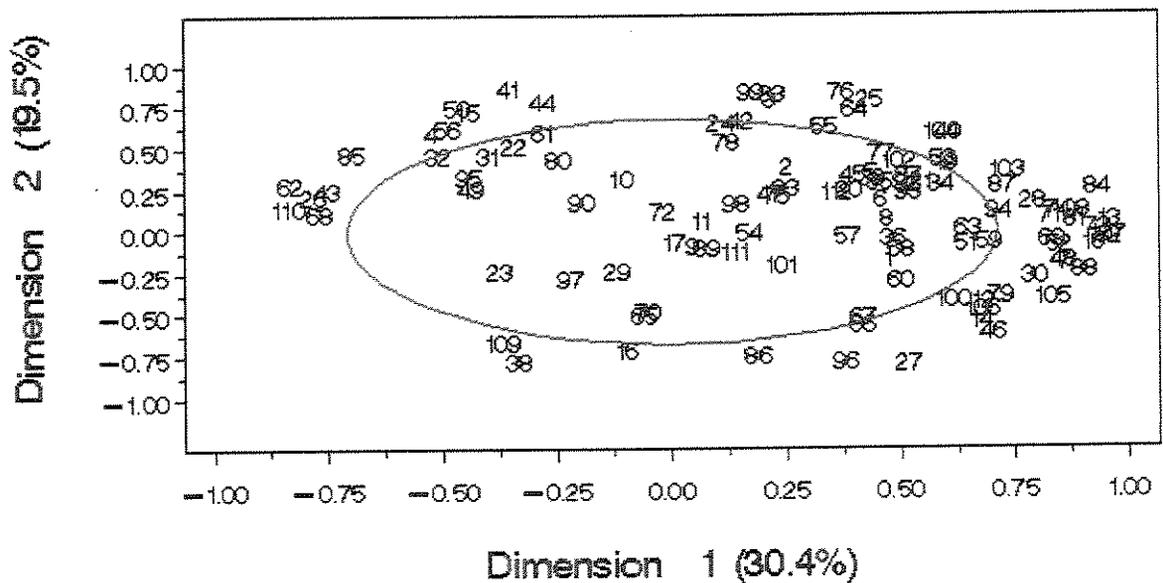


Figura 10. Mapa Interno de Preferência dos dados de aceitação global gerados pela escala hedônica híbrida, que mostra a configuração dos consumidores nas dimensões de preferência 1 e 2.

A Figura 9 mostra também, um terceiro segmento de vinhos, composto pelas amostras 4 e 8, situadas no lado esquerdo superior do mapa, correspondentes aos vinhos importados de varietal Cabernet Sauvignon. Esses resultados confirmam aqueles obtidos através do MDPREF das escalas hedônica de 9 pontos e autoajustável (Figuras 5 e 7). Entretanto, no MDPREF da escala hedônica híbrida (Figuras 9 e 10), tanto as amostras como os consumidores foram mais bem segmentados: a amostra 8, se localizou em uma região de 95% de confiança distinta da amostra 4, sendo que os consumidores à direita da amostra 8, dividem sua preferência entre esta amostra e os vinhos 2, 3 e 6. Da mesma forma, os consumidores situados à esquerda da amostra 4, preferem-na à amostra 8, dividindo sua preferência com o vinho 7. Os consumidores que preferiram as amostras 4 e 8 representam 22,5% dos consumidores significativos ($p=5\%$).

Um quarto segmento, foi também observado, correspondendo à amostra 7, situada em área limítrofe entre os quadrantes superior e inferior à esquerda do mapa. Esta amostra corresponde a um vinho nacional Cabernet Sauvignon de alto preço, porém, com baixa preferência entre os consumidores. Finalmente, observa-se na Figura 9, um quinto segmento referente à amostra 1 (Cabernet Sauvignon nacional de alto preço), que obteve a menor preferência entre os consumidores e encontra-se situada no quadrante esquerdo inferior do mapa. Esses resultados confirmam aqueles obtidos através dos mapas das escalas hedônica de 9 pontos e autoajustável (Figuras 5, 6, 7 e 8). Porém, no MDPREF dos dados gerados pela escala híbrida (Figuras 9 e 10), tanto as amostras como os consumidores foram melhor segmentados: as amostras 1 e 7, correspondentes a vinhos nacionais de alto preço, se localizaram em regiões de aceitação de 95% de confiança bastante distintas entre si. Também, no mapa da escala híbrida (Figuras 9 e 10), a amostra 7 revelou ser mais bem aceita por uma maior proporção de consumidores que a amostra 1. Este resultado é interessante porque somente o mapa desta escala revelou que um segmento de consumidores considerou o vinho 7 de melhor qualidade que o vinho 1, coincidindo com a opinião dos especialistas (Tabela 1). Os consumidores que preferiram as amostras 1 e 7 às demais, representaram 5,6% dos consumidores significativos ($p=5\%$).

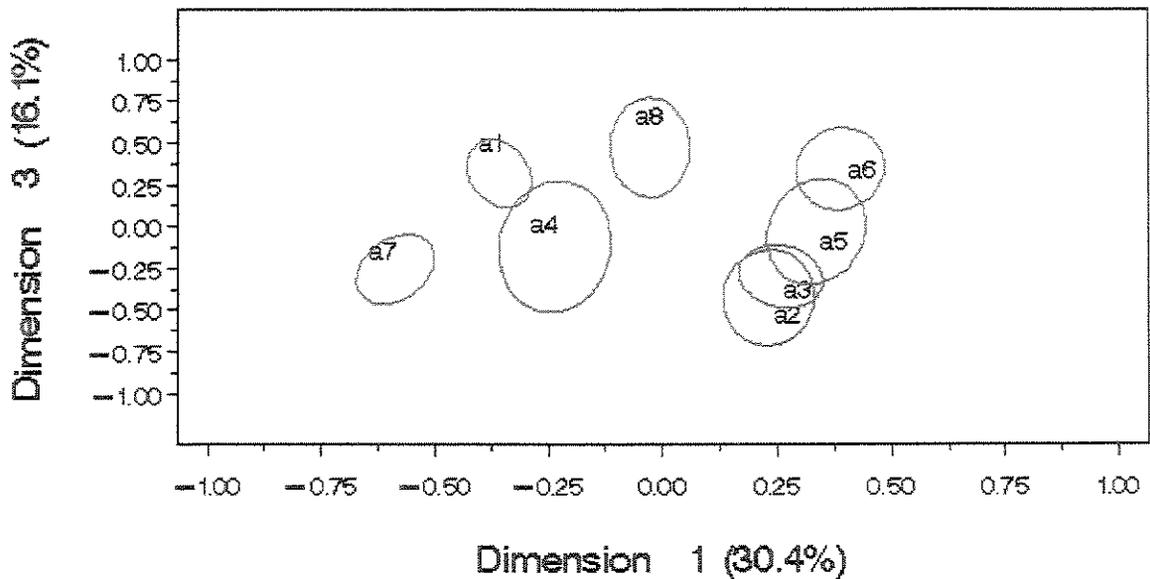


Figura 11. Mapa Interno de Preferência dos dados de aceitação global gerados pela escala hedônica híbrida, que mostra a configuração dos vinhos nas dimensões de preferência 1 e 3.

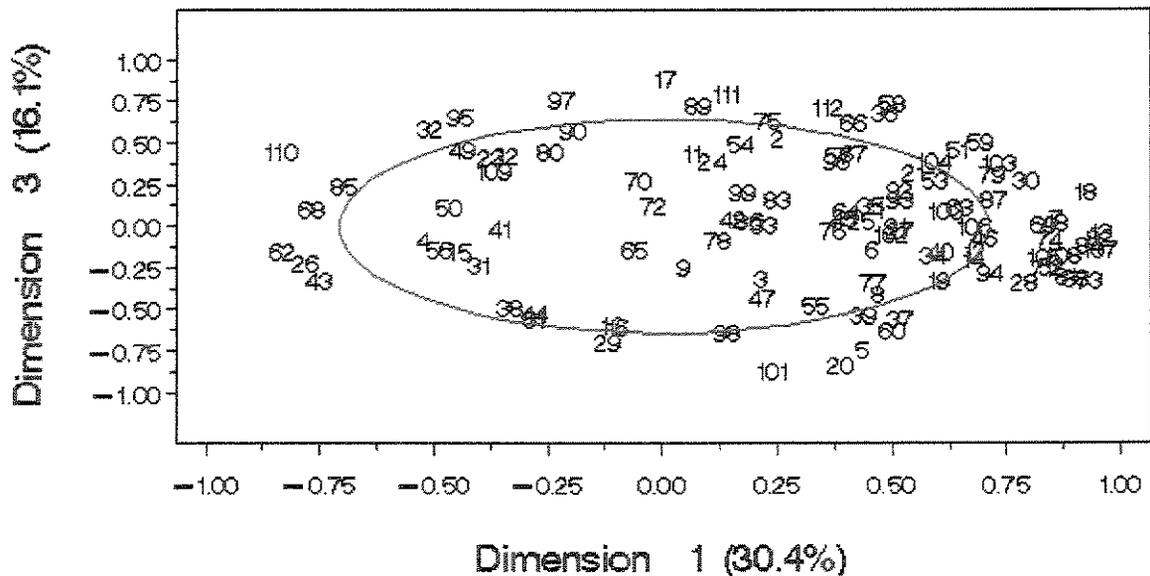


Figura 12. Mapa Interno de Preferência dos dados de aceitação global gerados pela escala hedônica híbrida, que mostra a configuração dos consumidores nas dimensões de preferência 1 e 3.

A Figura 11 mostra o efeito da terceira dimensão de preferência no Mapa gerado pela escala híbrida. Observa-se que a terceira dimensão separou claramente a amostra 6 (vinho nacional Cabernet de baixo preço) das amostras 2 e 3 (vinhos nacionais Cabernet Sauvignon de preços intermediários) e a amostra 8 (vinho chileno) da amostra 4 (vinho argentino). A Figura 12 mostra a configuração dos consumidores nas dimensões de preferência 1 e 3.

Pelo exposto, de um modo geral, pode-se concluir que: i) o Mapa de Preferência gerado com os dados da escala hedônica híbrida produziu uma segmentação mais fina de amostras e consumidores, do que o mapa gerado com os dados coletados com as escalas hedônica estruturada de 9 pontos e autoajustável, ii) as preferências dos consumidores que participaram do presente estudo, em geral não seguiram os critérios de qualidade estabelecidos pelo sistema internacional de avaliação de vinhos tintos, uma vez que as amostras consideradas de maior qualidade através deste sistema (amostras 4 e 8), foram com muita clareza as amostras preferidas por um segmento menor de consumidores, independente da escala por eles/elas utilizadas na avaliação dos vinhos, iii) os resultados desta pesquisa mostram que vinhos brasileiros de alto preço, não necessariamente atendem às expectativas de qualidade dos consumidores e nem aos padrões de qualidade estabelecidos por avaliadores especializados de vinhos.

4.2.3 Perfil de preferência e caracterização sócio-demográfica dos principais segmentos de consumidores

Com o objetivo de aprofundar a compreensão sobre os padrões de preferência dos principais segmentos de consumidores identificados pelo MDPREF de cada escala, foram realizadas para cada segmento, análises da variância dos dados de aceitação, testes de REGWQ para comparação de médias e tabelas de frequência de algumas variáveis sócio-demográficas.

A Tabela 9, mostra os resultados do teste REGWQ dos dois principais segmentos de consumidores identificados pelo Mapa de Preferência da escala

hedônica de 9 pontos: o primeiro segmento constituído por 46 consumidores significativos ($p=5\%$) localizados nos quadrantes direito superior e inferior do mapa (Figura 6), e o segundo segmento composto por 11 consumidores significativos localizados no quadrante esquerdo superior (Figura 6). A Tabela 9 mostra claramente que os dois segmentos identificados pelo MDPREF representam preferências marcadamente diferentes. Os consumidores do primeiro segmento, os quais constituem-se a maioria, mostraram preferência pelas amostras 2, 3, 5 e 6, tendo rejeitado as amostras 1, 4, 7 e 8, cujas médias de aceitação neste grupo foram inferiores a 5, que corresponde ao termo “nem gostei/nem desgostei” da escala. Já os consumidores do segundo segmento indicaram claramente sua preferência pelos vinhos importados (amostras 4 e 8) e sua rejeição pela amostra 5, vinho nacional elaborado com uma mistura de uvas Cabernet Franc e Merlot, a qual obteve média de aceitação próxima a 4, valor correspondente a “desgostei ligeiramente” na escala hedônica. Estes resultados exemplificam a importância da utilização do Mapa de Preferência como ferramenta complementar de análise estatística dos dados hedônicos, pois permite identificar segmentos pré-existentes de consumidores com padrões de preferências completamente opostas, os quais no caso de segmento 2 (Tabela 9) representam aproximadamente 10% dos consumidores que participaram do teste e 19,6% dos consumidores significativamente ajustados.

Tabela 9. Médias de aceitação global⁽¹⁾ dos dados gerados através da escala hedônica de 9 pontos, considerando-se: i) todos os consumidores ii) consumidores significativamente ajustados ($p=5\%$) situados no quadrante superior e inferior direito do MDPREF (Tabela 8) e iii) consumidores significativamente ajustados ($p=5\%$) situados no quadrante superior esquerdo do MDPREF (Tabela 8).

Amostra	Todos os⁽²⁾ consumidores (n=112 consumidores)	Segmento 1⁽²⁾ (n= 46 consumidores)	Segmento 2⁽²⁾ (n= 11 consumidores)
1	5,1 ed	4,2 bc	5,6 cd
2	6,0 abc	6,4 a	5,5 cd
3	6,2 ab	6,5 a	5,4 cd
4	5,5 cd	4,8 b	7,7 ab
5	5,6 bcd	6,4 a	4,3 d
6	6,5 a	6,9 a	6,1 bc
7	4,8 e	3,8 c	6,4 abc
8	5,6 bcd	4,4 bc	7,9 a

(1) 1= desgostei extremamente, 9= gostei extremamente; (2) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ao nível de 5% de significância. Esses resultados foram obtidos através do teste de REGWQ.

A Tabela 10, apresenta os resultados correspondentes à escala autoajustável. Também, nesta escala foram detectados dois segmentos majoritários: o primeiro segmento constituído por 48 consumidores significativamente ajustados ($p=5\%$) localizados nos quadrantes direito superior e inferior do MDPREF (Tabela 8), e o segundo segmento composto por 12 consumidores ($p=5\%$), localizados no quadrante esquerdo superior (Tabela 8).

A Tabela 10 confirma observações anteriores de que os dados coletados através da escala autoajustável segmentaram bem melhor as amostras junto aos consumidores do primeiro segmento que os dados gerados através da escala hedônica estruturada (Tabela 9). Através da Tabela 10, observa-se que este grupo, composto pela maioria de consumidores ajustados, realmente teve menor preferência pelas amostras 1 e 7, ambas vinhos Cabernet Sauvignon brasileiros de alto preço. Estas duas amostras diferiram significativamente ($p=5\%$) do grupo de amostras que obtiveram maior preferência entre esses consumidores, quais sejam: amostra 6 (Cabernet nacional), 2 (Cabernet Sauvignon nacional), 3 (Cabernet Sauvignon nacional) e 5 (Cabernet Franc e Merlot, nacional). As amostras 4 (Cabernet Sauvignon, Argentina) e 8 (Cabernet Sauvignon, Chile), ambas julgadas pelos

especialistas como os vinhos mais competitivos em termos de qualidade sensorial dentro do conjunto de vinhos avaliados, formaram o segundo grupo de amostras de menor preferência entre os consumidores do grupo 1. O vinho 6 (Cabernet nacional) foi o preferido por esse grupo de consumidores, diferindo significativamente ($p=5\%$) de todos os demais vinhos com exceção da amostra 2 (Cabernet Sauvignon nacional).

Tabela 10. Médias de aceitação global⁽¹⁾ dos dados gerados através da escala autoajustável, considerando-se: i) todos os consumidores ii) consumidores significativamente ajustados ($p=5\%$) situados no quadrante superior e inferior direito do MDPREF (Tabela 8) e iii) consumidores significativamente ajustados ($p=5\%$) situados no quadrante superior esquerdo do MDPREF (Tabela 8).

Amostra	Todos os⁽²⁾ consumidores	Segmento 1⁽²⁾ (n= 48 consumidores)	Segmento 2⁽²⁾ (n= 12 consumidores)
1	3,8 d	1,8 f	6,6 ab
2	6,0 ab	7,2 ab	5,3 bc
3	5,5 abc	6,2 bc	3,7 c
4	4,4 cd	3,7 e	7,2 ab
5	4,4 cd	5,3 cd	0,2 d
6	7,0 a	8,3 a	3,7 c
7	3,1 d	1,0 f	7,2 ab
8	5,4 bc	4,5 de	9,2 a

(1) 0= menos gostoso, 10= mais gostoso; (2) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ao nível de 5% de significância. Esses resultados foram obtidos pelo teste de REGWQ.

Características bastante peculiares associadas às preferências dos consumidores do segundo segmento podem ser visualizadas na Tabela 10. Neste segmento houve consenso claro entre os provadores ao considerarem a amostra 5 a menos preferida, discriminando-a significativamente dos demais vinhos ($p=5\%$). Por tratar-se de um vinho elaborado com uma mistura de uvas Cabernet Franc e Merlot, este vinho, foi de fato, considerado pelos especialistas, como um produto de menor qualidade sensorial (Tabela 1), razão pela qual, esta amostra é a que apresenta o menor preço de mercado em relação às demais (Tabela 1). Por sua vez, as amostras 1, 4, 7 e 8, todas vinhos varietal Cabernet Sauvignon de alto valor comercial e melhor qualidade sensorial à exceção da amostra 1, segundo avaliadores especializados, foram as amostras preferidas desse grupo de consumidores, os quais mostraram

preferência clara pela amostra 8 (Cabernet Sauvignon, Chile), que obteve a máxima nota na avaliação dos especialistas, discriminando-a significativamente ($p=5\%$) das amostras 2, 3, 5 e 6, que foram os vinhos preferidos pelos consumidores do primeiro segmento. Em complementação, a escala autoajustável segmentou com mais nitidez que a escala hedônica estruturada: as duas outras amostras menos preferidas, quais sejam, amostras 3 e 6 (vinhos nacionais Cabernet Sauvignon e Cabernet, respectivamente), as quais foram significativamente discriminadas da amostra 5 ($p=5\%$), a amostra menos preferida por esses consumidores. Assim, através da escala autoajustável, foi possível maximizar-se a segmentação dos consumidores que participaram do estudo, expondo claramente as diferenças entre eles com relação às suas preferências: enquanto o primeiro grupo, majoritário, não mostrou concordância com os critérios de qualidade estabelecidos pelos avaliadores especializados, o segundo grupo correlaciona-se claramente com eles. Portanto, apesar da escala autoajustável ter a desvantagem de não medir o grau de aceitação das amostras, de um modo geral pode-se concluir no presente estudo, que esta escala mostrou clara superioridade sobre a escala hedônica de 9 pontos com relação à segmentação de amostras e consumidores através de MDPREF. Portanto, em testes cujo objetivo for a segmentação e identificação do perfil de preferência do mercado consumidor com relação às várias marcas que competem entre si, a escala autoajustável parece ser um instrumento de melhor desempenho que a escala estruturada de 9 pontos.

A Tabela 11 apresenta os resultados correspondentes à escala hedônica híbrida. Nesta escala foram detectados 4 segmentos de consumidores (Tabela 8), dos quais três podem ser considerados majoritários. O primeiro segmento constituído por 37 consumidores significativos localizados no quadrante direito superior do MDPREF mostrado na Figura 10, o segundo segmento constituído por 20 consumidores significativos localizados no quadrante esquerdo superior da citada figura, e o terceiro segmento composto por 27 consumidores significativos localizados no quadrante direito inferior do MDPREF.

Tabela 11. Médias de aceitação global⁽¹⁾ dos dados gerados através da escala hedônica híbrida⁽²⁾, considerando-se: i) todos os consumidores, ii) consumidores significativamente ajustados (p=5%) situados no quadrante superior direito do MDPREF (Tabela 8), iii) consumidores significativamente ajustados (p=5%) situados no quadrante superior esquerdo do MDPREF e, iv) consumidores significativamente ajustados (p=5%) situados no quadrante inferior direito do MDPREF (Tabela 8).

Amostra	Todos os⁽²⁾ consumidores (n=112 consumidores)	Segmento 1⁽²⁾ (n= 37 consumidores)	Segmento 2⁽²⁾ (n= 20 consumidores)	Segmento 3⁽²⁾ (n= 27 consumidores)
1	4,4 c	2,5 c	5,9 bc	5,6 ab
2	6,0 ab	6,9 ab	5,6 bc	5,6 ab
3	5,8 ab	6,3 ab	5,0 c	6,3 a
4	5,2 b	5,7 b	8,2 a	3,3 cd
5	5,4 b	5,7 b	3,4 d	6,9 a
6	6,6 a	7,2 a	5,4 c	7,0 a
7	4,2 c	3,1 c	7,2 ab	2,5 d
8	6,0 ab	6,6 ab	7,2 ab	4,5 bc

(1) 0=desgostei extremamente, 5= nem gostei/nem desgostei, 10= gostei extremamente; (2) Médias com letras em comum numa mesma coluna, indicam amostras que não diferem significativamente entre si ao nível de 5% de significância. Esses resultados foram obtidos através do teste de REGWQ.

A Tabela 11 mostra que os dados coletados pela escala hedônica híbrida também produziram boa segmentação das amostras junto aos consumidores do primeiro grupo, revelando que as amostras 1 e 7 (vinhos nacionais Cabernet Sauvignon) tiveram a menor aceitação/preferência, diferindo significativamente de todas as demais amostras (p=5%). O grau de aceitação dessas amostras situou-se próximo a 3, valor associado à região de rejeição da escala híbrida, a despeito de trataram-se de vinhos varietais nacionais de alto preço. Diferentemente do que ocorreu com as demais escalas, o segundo grupo de amostras de menor aceitação junto a esses consumidores diz respeito às amostras 4 (Cabernet Sauvignon, Argentina) e 5 (Cabernet e Merlot, nacional), os quais foram discriminados significativamente (p=5%) da amostra de maior preferência entre os consumidores, ou seja, a amostra 6, vinho nacional varietal Cabernet. Finalmente, a amostra 6, os vinhos nacionais 2 e 3 (varietais Cabernet Sauvignon nacionais) e o vinho 8 (varietal Cabernet Sauvignon, Chile), compuseram o grupo de amostras de maior preferência entre os consumidores do segmento 1. Uma singularidade dos resultados obtidos através da escala hedônica híbrida é que, por ter permitido uma segmentação mais

final dos consumidores que as demais escalas, ela identificou o presente grupo de consumidores os quais mostraram boa aceitação da amostra 8, que obteve baixa aceitação/preferência entre os indivíduos alocados no primeiro segmento das escalas hedônica estruturada e autoajustável. Este resultado evidencia o grau de impacto e a importância do tipo de escala utilizada para coletar dados afetivos, sobre as informações geradas pelo Mapa Interno de Preferência.

Os consumidores agrupados no segundo segmento gerado pelo MDPREF da escala híbrida (Figuras 9 e 10), mostraram preferência pelos vinhos Cabernet Sauvignon importados -amostras 4 e 8- além do vinho 7 (Cabernet Sauvignon, nacional de boa qualidade), tendo rejeitado a amostra 5, vinho "Assemblage" de menor preço e menor qualidade, elaborado com uvas Cabernet Franc e Merlot. Observa-se que as preferências dos consumidores deste segmento mostraram clara concordância com os critérios de qualidade estabelecidos pelos especialistas, pois as amostras 4, 7 e 8 receberam as maiores notas na avaliação cega e, a amostra 5, a menor nota (Tabela 1). Esses resultados assemelham-se àqueles gerados pelos indivíduos que compõem o segmento 2 dos Mapas de Preferência das escalas hedônica estruturada (Figuras 5 e 6) e autoajustável (Figuras 7 e 8).

Finalmente, o MDPREF dos dados gerados pela escala hedônica híbrida produziu um terceiro segmento de consumidores, composto pela segunda maior proporção de indivíduos, os quais caracterizaram-se por preferirem os vinhos de menor preço (amostras 3, 5 e 6) entre as amostras avaliadas, e alta rejeição pelas amostras 4 e 7, ambos Cabernet Sauvignon de boa qualidade sensorial, sendo o primeiro nacional e o segundo Argentino. Este grupo assemelhou-se aos indivíduos do segmento 1 por demonstrar alta aceitação e preferência pela amostra 6 (vinho Cabernet nacional) e baixa preferência e aceitação pela amostra 7. No entanto, os consumidores do segmento 3 distinguiram-se do segmento 1 por apresentarem menor aceitação das amostras 4 e 8, vinhos varietais importados, e melhor aceitação da amostra 1, vinho varietal nacional de qualidade sensorial intermediária, segundo opinião dos especialistas.

Uma das grandes vantagens do Mapa Interno de Preferência consiste na possibilidade de caracterizar cada segmento de consumidores de acordo com variáveis sócio demográficas como: faixa etária, sexo, lugar de procedência, grau de instrução e frequência de consumo. Estas informações são de grande importância e utilidade para a área de marketing das empresas vinícolas, as quais podem conhecer com maior profundidade as preferências e características dos diversos segmentos do mercado consumidor de vinhos tintos para melhor atendê-los.

No presente estudo, esses dados são mostrados na Tabela 12, e revelaram algumas diferenças claras no perfil dos consumidores dos três segmentos identificados pelo MDPREF da escala hedônica híbrida. Assim, no primeiro segmento formado pelos indivíduos que mostraram alta rejeição pelos vinhos varietais 1 e 7, observou-se consumidores pertencentes a todas as faixas etárias; 62,2% deles pertenciam ao estado de São Paulo; 43,3% dos consumidores declararam consumir vinho tinto entre 1 a 2 copos por semana e 37,8%, até um copo por semana nos meses de inverno. Já no segundo segmento, formado pelos indivíduos que preferiram os vinhos varietais importados e um nacional de boa qualidade sensorial, 75% dos consumidores pertenciam ao sexo masculino e, 80% encontrava-se na faixa etária entre 25 e 44 anos. Este segmento também distinguiu-se dos demais por apresentar maior porcentagem de consumidores estrangeiros (20%), além de apresentar uma maior frequência de consumo de vinho, pois 25% destes consumidores declararam consumir vinho tinto no mínimo 3 a 4 copos/semana nos meses de inverno. Finalmente, o terceiro segmento, cujos indivíduos preferiram os vinhos de menor preço e definição varietal, caracterizou-se por apresentar consumidores mais jovens, pois 77,8% desses consumidores possuía idade variando entre 18 e 34 anos. Este grupo de consumidores apresentou os menores níveis de consumo de vinho tinto, sendo que 55,6% deles declararam consumir menos de 1 copo/mês de vinho tinto ou no máximo, 1 copo/semana, durante os meses de inverno.

Tabela 12. Características sócio demográficas expressas em porcentagens dos consumidores pertencentes a cada segmento, referentes aos subgrupos de preferência detectados pelo MDPREF correspondente à escala hedônica híbrida.

Variáveis sócio demográficas	Segmento 1	Segmento 2	Segmento 3
	(n=37 consumidores) %	(n=20 consumidores) %	(n=27 consumidores) %
Sexo			
M	48,7	75,0	63,0
F	51,4	25,0	37,0
Faixa etária			
18-24 anos	27,0	15,0	29,6
25-34 anos	37,8	45,0	48,2
35-44 anos	16,2	35,0	11,1
45-44 anos	19,0	5,0	11,1
Local de Procedência			
SP	62,2	45,0	59,3
Outros estados	29,7	35,0	29,6
Estrangeiros	8,1	20,0	11,1
Nível de Escolaridade			
2º grau completo	2,7	0,0	14,8
Superior incompleto	21,6	20,0	14,8
Superior completo	29,7	30,0	25,9
Pós-graduação	46,0	50,0	44,5
Freqüência de consumo de vinho tinto			
Quase todo dia	2,7	10,0	0
3 a 4 copos/semana	8,1	15,0	14,8
1 a 2 copos/semana	43,3	40,0	29,6
Menos de 1 copo/semana	37,8	30,0	40,8
Menos de 1 copo/mês	8,1	5,0	14,8

5. CONCLUSÕES

De um modo geral, os resultados obtidos através da ANOVA e dos Mapas Internos de Preferência, evidenciaram um desempenho mais pobre da escala hedônica de 9 pontos em relação às escalas hedônica híbrida e autoajustável no que se refere à capacidade de segmentação de amostras e consumidores através da técnica de Mapa Interno de Preferência.

Observou-se também uma superioridade da escala híbrida sobre as demais com relação à: 1) segmentação de amostras, 2) segmentação de consumidores, 3) número de dimensões de preferência significativas ($p \leq 0,05$) extraído pelo MDPREF e, 4) porcentagem de consumidores ajustados significativamente ($p \leq 0,05$). Como consequência, o MDPREF da escala híbrida produziu um número maior de segmentos de amostras e consumidores que as demais escalas, possibilitando uma melhor caracterização do perfil de preferências de cada segmento.

A despeito da escala autoajustável apresentar a limitação de não medir o grau de aceitação dos produtos e não permitir prever o quanto o consumidor pode gostar ou não de um produto, o presente estudo evidenciou uma clara superioridade desta escala sobre a escala hedônica de 9 pontos em relação à segmentação de amostras e consumidores. Assim, em testes associados à identificação de segmentos pré-existent de consumidores e à determinação de seus padrões de preferência no mercado consumidor em relação às várias marcas que competem por um mesmo grupo de consumidores, a escala autoajustável parece ser um instrumento de melhor desempenho que a escala estruturada de 9 pontos.

Pelo exposto, de um modo geral, os resultados do presente estudo sugerem fortemente que a escala hedônica híbrida apresenta melhor desempenho que as demais escalas, tanto quando os dados por ela gerados são avaliados por ANOVA e teste de REGWQ, como quando os mesmos são avaliados pela técnica Mapa Interno de Preferência. Isto evidencia o grau de impacto e a importância do tipo de escala utilizada para coletar dados afetivos, sobre as informações geradas pelo MDPREF.

Em relação aos vinhos, os resultados obtidos no presente estudo revelaram informações importantes para as vinícolas nacionais. Através da escala hedônica híbrida -a mais discriminativa das escalas- o presente estudo revelou que, no grupo de consumidores que participaram do teste, aproximadamente 80% dos indivíduos mostraram preferências significativas ($p=5\%$) por um grupo de amostras de vinhos discriminando-as das demais. Considerando-se o perfil socioeconômico desses

consumidores, e o fato deles/elas consumirem no mínimo entre 1 a 2 copos de vinho tinto/por semana durante o inverno, esses resultados são importantes porque indicam que consumidores de vinhos tintos no Brasil distinguem, de fato, os vinhos comerciais disponíveis no mercado, em função de suas qualidades sensoriais. Entretanto, os resultados do presente estudo sugerem também, que a preferência de uma proporção considerável desses consumidores não se correlaciona com a classificação utilizada por avaliadores especializados de vinhos. De fato, a grande maioria de consumidores preferiu significativamente ($p=5\%$) os vinhos com preços e qualidade varietal classificados entre baixos e intermediários. Uma menor proporção de consumidores preferiu vinhos considerados mais nobres pelos avaliadores especializados. Esses resultados são indicativos que, no futuro para ganhar competitividade dentro dos seus mercados, as vinícolas devem introduzir em complementação aos testes de degustação conduzidos por especialistas, testes sensoriais afetivos, realizados através de métodos validados pela área de análise sensorial e aplicados a reais consumidores do produto.

AGRADECIMENTO

Os autores agradecem à Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro que fez possível a realização deste projeto de pesquisa.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ALTMAN, D. G. Plotting probability ellipses. *Applied Statistics*, v. 37, p. 347-349, 1978.
- AMERINE, M. A.; ROESSLER, E. B. *Wines – Their Sensory Evaluation*. San Francisco: W. H. Freeman and Company, 1983, 432p.
- ARDITTI, S. Preference mapping: a case study. *Food Quality and Preference*, v. 8, p. 323-326, 1997.
- CHANG, J. J.; CARROL, J. D. How to use MDPREF, a computer program for multidimensional analysis of preference data. Unpublished report, Bell Telephone Laboratories, 1968.

- DAILLANT-SPINLER, B.; MACFIE, H. J. H.; BEYTS, P. K.; HEDDERLY, D. Relationships between perceived sensory properties and major preference directions of 12 varieties of apples from the southern hemisphere. *Food Quality and Preference*, v. 7, p. 113-126, 1996.
- EDGINGTON, E. S. *Randomization Test*, 2nd. Edn. Marcel Dekker, New York, 1987.
- EFRON, B.; TIBSHIRANI, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science*, v. 1, p. 54-77, 1986.
- FALAHEE, M.; MACRAE, A. W. Consumer appraisal of drinking water: multidimensional scaling analysis. *Food Quality and Preference*, v. 6, p. 327-332, 1995.
- GAY, C.; MEAD, R. A statistical appraisal of the problem of sensory measurement. *Journal of Sensory Studies*, v. 7, p. 205-228, 1992.
- GIOVANNI, M. E.; PANGBORN, R. M. Measurement of taste intensity and degree of liking of beverages by graphic scales and magnitude estimation. *Journal of Food Science*, v. 48, p. 1175-1182, 1983.
- GREENHOFF, K; MACFIE, H. J. K. Preference mapping in practice. *Measurement of food preferences*. Eds. H. J. H. MacFie and D. M. H. Thomson. Blackie Academic and Professional London, p. 137-166, 1994.
- GUINARD, J.-X.; UOTANI, B.; SCHLICH, P. Internal and external mapping of preferences for commercial lager beers: comparison of hedonic ratings by consumers blind versus with knowledge of brand and price. *Food Quality and Preference*, v. 12, p. 243-255, 2001.
- HEYD, B.; DANZART, M. Modeling consumers' preferences of coffees: Evaluation of different methods. *Lebensm.-Wiss. u.-Technol.*, v. 31, p. 607-611, 1998.
- HELGESEN, H.; SOLHEIM, R.; NAES, T. Consumer preference mapping of dry fermented lamb sausages. *Food Quality and Preference*, v. 8, p. 97-109, 1997.
- MACFIE, H. J.; WAKELING, I. N. Designing consumer trials balanced first and higher orders of carry-over effects when only a subset of k samples from t may be tested. *Food Quality and Preference*, v. 6, p. 299-308, 1995.
- MACPHERSON, R. S; RANDALL, E. Line length measurement as a tool for food preference research. *Ecology of food and Nutrition*, v. 17, p. 149-156, 1985.
- MONTELONE, E.; FREWER, L.; WAKELING, I.; MELA, D. J. Individual differences in starchy food consumption: The application of preference mapping. *Food Quality and Preference*, v.9, p. 211-219, 1998.
- MURRAY, J. M.; DELAHUNTY, C .M. Mapping consumer preference for the sensory and packaging attributes of Cheddar cheese. *Food Quality and Preference*, v.11, p. 419-435, 2000.
- O'MAHONY, M. Some assumptions and difficulties with common statistics for sensory analysis. *Food Technology*, p. 75-82. 1982.

- PERYAM, D. R.; PILGRIM, F. J. Hedonic scale method of measuring food preferences. *Food Technology*, 11(9), Suppl. 9-14, 1957.
- PIGGOT, J. R. Design questions in sensory and consumer science *Food Quality and Preference*, v.6, p. 217-220, 1995.
- SAS INSTITUTE INC. SAS/IML[®] Software: Usage and Reference, Version 8.0, SAS Institute Inc., Cary, NC, USA, 2000.
- SIDEL J. L.; STONE H., The role of sensory evaluation in the food industry. *Food Quality and Preference*, v.4, p. 65-73, 1993.
- VILLANUEVA, N. D. M.; PETENATE A. J.; DA SILVA, M. A. A. P. Performance of three affective methods and diagnosis of the ANOVA model. *Food Quality and Preference*. Vol. 11, p. 363-370, 2000.
- VILLANUEVA, N. D. M.; PETENATE A. J.; DA SILVA, M. A. A. P. *Self-adjusting and hybrid hedonic scales: Advantages and disadvantages with respect to traditional affective methods*. In: The 4th Pangborn Sensory Science Symposium. Dijon, 2001. Conference Program Abstracts, p. 136.
- YACKINOUS, C.; WEE, C.; GUINARD, J-X. Internal preference mapping of hedonic ratings for Ranch salad dressings varying in fat and garlic flavor. *Food Quality and Preference*, v.10, p. 401-409, 1999.

CONCLUSÕES GERAIS

O desempenho da escala autoajustável em testes afetivos com várias amostras mostrou-se superior ao da tradicional escala hedônica tanto quando os dados foram analisados através de modelos normais de análise da variância como quando foram analisados através da técnica de Mapa Interno de Preferência, o que torna o uso desta escala promissor. No MDPREF a escala autoajustável apresentou melhor segmentação de amostras e consumidores comparativamente à escala hedônica tradicional. Porém, quando a mesma é utilizada para discriminar apenas três amostras por provador, como no caso do primeiro experimento, esta escala apresentou baixo desempenho. Uma limitação da escala autoajustável é o fato de ela não ser genuinamente uma escala hedônica devido a que a mesma não avalia o grau de aceitação das amostras e não permite predizer o quanto o consumidor gostou ou desgostou de cada produto. Finalmente, outra limitação da escala autoajustável identificada no presente estudo é que os dados por ela gerados apresentam sérios desvios da normalidade, tornando os níveis de significância estabelecidos para os testes de hipóteses associados à ANOVA, tal vez, de pouco confiáveis.

Os resultados da presente pesquisa sugeriram uma larga superioridade da escala hedônica híbrida sobre a tradicional escala hedônica estruturada e escala autoajustável, tanto em função do poder discriminativo como em relação à adequação dos dados coletados aos pressupostos estatísticos associados ao uso dos modelos normais aditivos para análise da variância. O poder discriminativo da escala hedônica híbrida foi superior ao das demais escalas tanto na análise dos dados de aceitação realizadas através de ANOVA e do teste não paramétrico de Friedman, como através da técnica multivariada de Mapa Interno de Preferência (MDPREF). Na análise MDPREF, os dados coletados através da escala hedônica híbrida apresentaram um maior número de dimensões de preferência significativas ($p=5\%$), uma maior porcentagem de consumidores significativamente ajustados

($p=5\%$), e uma segmentação mais fina de amostras e de consumidores comparativamente às demais escalas.

Os resultados desta pesquisa mostraram fortes evidências de que a distribuição dos dados de aceitação depende da natureza psicométrica da escala com a qual esses dados foram coletados. No caso da escala hedônica híbrida os resultados deste trabalho mostram fortes indícios de que a distribuição dos dados gerados por esta escala apresentam formas mais regulares e próximas à distribuição normal. A escala autoajustável produziu distribuições de dados bastante assimétricas ou bimodais, mostrando modas em um ou ambos os extremos das distribuições e apresentando desvios da normalidade significativos ($p=5\%$). Já no caso da escala hedônica tradicional, nossos estudos confirmaram problemas de heteroscedasticidade nas variâncias das amostras, problemas estes já apontados em estudos anteriores por diversos cientistas sensoriais.

Finalmente, os resultados dos experimentos realizados nesta pesquisa evidenciaram o grau de impacto que o tipo de escalonamento utilizado na coleta de dados afetivos teve sobre as informações geradas nos experimentos. Tanto na análise dos dados realizados através de modelos ANOVA, como através da técnica de Mapa de Preferência, pôde-se observar que a distribuição dos dados, o poder discriminativo e o grau de segmentação de amostras e consumidores dependeram do tipo de escala utilizado. Por outro lado, os resultados experimentais revelaram que não somente o número ou o tipo de amostras testadas teve influência nos efeitos de ordem de apresentação, nos efeitos residuais e na variação da expansibilidade entre provadores, mas também, o tipo de escala utilizada no teste. Isto significa que um aspecto fundamental na geração de dados afetivos confiáveis é a escolha da escala a ser utilizada na realização dos testes. Esta escolha deve fundamentar-se em conhecimentos prévios dos aspectos psicométricos da escala e ainda em suas vantagens e limitações.

APÊNDICE

AVALIAÇÃO DAS ESCALAS

Nome: _____ Data: _____

Idade: _____ Curso: _____

1. Apresentamos abaixo três métodos de escala por você utilizados para avaliação das amostras de suco de laranja.

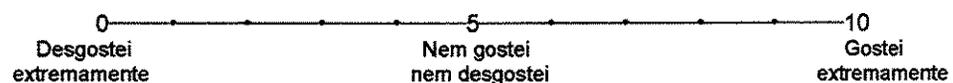
ESCALA HEDÔNICA ESTRUTURADA

Você esta recebendo— amostras codificadas de —. PROVE por favor da esquerda para direita cada uma das amostras e utilizando a escala abaixo indique o quanto você gostou ou desgostou das mesmas.

- 9 Gostei extremamente
- 8 Gostei muito
- 7 Gostei moderadamente
- 6 Gostei ligeiramente
- 5 Nem gostei / nem desgostei
- 4 Desgostei ligeiramente
- 3 Desgostei moderadamente
- 2 Desgostei muito.
- 1 Desgostei extremamente

ESCALA HEDÔNICA HÍBRIDA

Você esta recebendo — amostras codificadas de —. Por favor PROVE as amostras de esquerda para a direita e marque, na escala abaixo um “x” no lugar que você achar que melhor represente o quanto você gostou ou desgostou de cada amostra.



ESCALA AUTOAJUSTÁVEL

Você esta recebendo — amostras codificadas de —. PROVE agora as amostras da esquerda para a direita. Escolha primeiro a amostra que você mais gostou e depois a que você menos gostou e anote os códigos dessas amostras nos extremos direito e esquerdo da escala abaixo. Avalie agora as outras amostras de forma comparativa com aquelas alocadas nos extremos, marcando um “x” e o código da amostra no ponto da escala que você achar que melhor represente o quanto você gostou ou desgostou cada amostra após degustá-las.

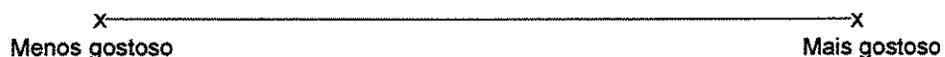


Figura 1. Questionário utilizado para avaliação das escalas pelos consumidores.

2. Por favor, leia com atenção cada item do questionário abaixo e responda:

2.1 Você já participou antes deste estudo, de algum teste sensorial utilizando alguma das escalas mostradas acima? Sim Não

2.2 Qual? _____

2.3 Em relação à facilidade de uso dessas escalas, avalie cada escala marcando um "x" no lugar que você ache mais conveniente:

Escala hedônica estruturada:

- extremamente fácil
- muito fácil
- moderadamente fácil
- ligeiramente fácil
- Nem fácil / nem difícil
- ligeiramente difícil
- moderadamente difícil
- muito difícil
- extremamente difícil

Escala hedônica híbrida:

- extremamente fácil
- muito fácil
- moderadamente fácil
- ligeiramente fácil
- Nem fácil / nem difícil
- ligeiramente difícil
- moderadamente difícil
- muito difícil
- extremamente difícil

Escala autoajustável:

- extremamente fácil
- muito fácil
- moderadamente fácil
- ligeiramente fácil
- Nem fácil /nem difícil
- ligeiramente difícil
- moderadamente difícil
- muito difícil
- extremamente difícil

2.4 Avalie cada escala em relação à liberdade que lhe deu para expressar suas percepções sensoriais:

Escala hedônica estruturada:

- total liberdade
- muita liberdade
- moderada liberdade
- ligeira liberdade
- pouca liberdade
- nenhuma liberdade

Escala hedônica híbrida:

- total liberdade
- muita liberdade
- moderada liberdade
- ligeira liberdade
- pouca liberdade
- nenhuma liberdade

Escala autoajustável:

- total liberdade
- muita liberdade
- moderada liberdade
- ligeira liberdade
- pouca liberdade
- nenhuma liberdade

2.5 Indique em ordem decrescente de preferência o método que você escolheria para avaliar em futuros testes de consumidores:

1°) escala _____ + preferido

2°) escala _____

3°) escala _____ - preferido

2.6 Justifique sua preferência _____

Figura 1. Questionário de avaliação das escalas pelos consumidores (continuação).

4. Que tipo de vinho você consome com maior frequência?

- Tinto Suave
- Tinto Demi Sec
- Tinto Seco
- Branco Suave
- Branco Demi Sec
- Branco Seco
- Rosé
- Champagne
- Espumantes
- Outros

5. Indique três vinícolas ou marcas de vinho que você costuma comprar com maior frequência

1. _____
2. _____
3. _____

6. Indique três vinhos varietais* que você costuma consumir com maior frequência ?

1. _____
2. _____
3. _____

*Vinhos varietais são aqueles produzidos a partir de uvas de cultivares europeus, para portarem esta definição devem trazer no rótulo do produto o nome do cultivar de uva a partir do qual foram elaborados.

7. Indique o(s) lugar(es) onde você mais costuma consumir vinho:

- Em casa
- Em restaurantes
- Em bares, empórios, casas noturnas
- Em festas, aniversários, reuniões sociais

8. Indique utilizando a escala abaixo, com que frequência você costuma comprar cada tipo de vinho

Tinto	Branco	Outro
<input type="checkbox"/> Compro muito (+ de 4 garrafas/ano)	<input type="checkbox"/> Compro muito (+ de 4 garrafas/ano)	<input type="checkbox"/> Compro muito (+ de 4 garrafas/ano)
<input type="checkbox"/> Compro moderadamente (até 4 garrafas/ano)	<input type="checkbox"/> Compro moderadamente (até 4 garrafas/ano)	<input type="checkbox"/> Compro moderadamente (até 4 garrafas/ano)
<input type="checkbox"/> Compro ocasionalmente (1 garrafa/ano)	<input type="checkbox"/> Compro ocasionalmente (1 garrafa/ano)	<input type="checkbox"/> Compro ocasionalmente (1 garrafa/ano)
<input type="checkbox"/> Compro raramente (menos de 1 garrafa/ano)	<input type="checkbox"/> Compro raramente (menos de 1 garrafa/ano)	<input type="checkbox"/> Compro raramente (menos de 1 garrafa/ano)

Assinatura

Figura 2. Questionário utilizado para levantamento do perfil dos consumidores. (continuação).