An Efficient Way to Determine the Difference Threshold from Consumer Test Data

E. Koehn, Hamburg University of Applied Sciences*

D. Minkner, British American Tobacco (Hamburg International) GmbH

Keywords: psychophysics, difference threshold, just-about-right

Introduction

Knowing the psychophysical function of key sensory properties amongst the target population is important in guiding intended/required product modifications, new product development and quality assurance. The classic psychophysical methods are laborious, expensive and vulnerable to a number of experimental biases as e.g. outlined by H. Heymann and H.L. Lawless. From a practitioner's point of view, difference thresholds or Just Noticeable Differences (JND) should be measured amongst target consumers and with marketable products, a requirement which can be difficult to achieve with classic approaches at acceptable costs.

Consideration has been given to whether "just-right" data from aggregated consumer tests could be used as an alternative method. "Just-right" judgements are understood as a perceived difference rating between test product's and respondent's "ideal" sensation. To use "just-right"-judgements would however require estimating the intensity of the "ideal" sensation. Other requirements include measurements of the causing ingredient or alternatively using an indicator variable which is highly correlated with the sensation of interest.

This approach was investigated using a data base which contained information regarding consumer test results of cigarettes as well as analytical figures of the test products and the respondent's most frequently used brand (regular brand) at time of testing. The sensation of interest was strength, which is known to be highly correlated with tar and nicotine delivery of a cigarette.

The study detailed below is an exploratory data analysis aimed at encouraging sensory scientists to further explore this approach. Further exploration and research would be needed to confirm the benefits but if deemed successful it could become a useful further addition to our sensory methods repertoire.

Objective

To evaluate whether JAR data in connection with analytical results of the test products can be used to establish an empirical, target group averaged psychophysical function in order to allow determining "no-difference ranges" and estimating sensorial based product class boarders.

Materials and Methods

A consumer test data base was provided containing judgements of cigarettes, collected over 4 years in Germany. Additional to "liking" ratings and a range of attribute and "just-right" ratings, the data base also contained analytical measurements of the tested products as well as each respondent's regular brand at time of the test. The number of cases was approximately 90,000. Test products were brands and development products of the same tobacco blend style.

- ✓ Test method: monadic (single product) home-use tests with independent matched consumer samples, self-completion questionnaire.
- ✓ JAR-Scale: End-labelled 7-point Degree-of-Change scale, labels were "should have more" and "should have less" of the attribute.
- ✓ Analysed sensation: Strength, a term commonly used by consumers. According to ASTM Manual, MNL 63, complex terms are acceptable JAR-attributes if consumer use of the term is confirmed. Only the JAR-proportions (at scale midpoint) are used for the analysis.
- ✓ Independent variable: smoke nicotine [mg/cigarette] difference between test product and reference (respondent's regular brand at time of testing). This measurement is understood as an indicator variable, not necessarily as cause for the strength sensation.

Assumptions:

The following assumptions constitute the basis for the proposed use of JAR rating to approximate the psychophysical function:

- ✓ In a blind consumer product test the "just right"-scale measures the perceived deviation of a test product's from the respondent's ideal sensation.
- ✓ Consumer prefer (and choose for regular consumption) such products which are close to their "relative ideal" with regard to key sensory properties.
- ✓ The proportion of respondents which rate a test product "just-about-right" (JAR=midpoint of the "just right"-scale) will reach its maximum if the intensity difference between test product and respondents "ideal" becomes zero.
- ✓ Therefore, JAR-proportion from tests in which respondents judged their most frequently consumed product (regular brand) would constitute an estimate of the maximum JAR-proportion, the proportion of consumer for which regular brand is at the ideal point in the respective sensation.
- ✓ With increasing stimulus intensity difference between test product and "ideal", the proportion of JAR's will drop.
- ✓ If the JAR-proportions are plotted against the logarithm of the absolute stimulus difference one would expect an inverse dose response function.

Steps in the Analysis:

1. Calculate the "sensorial active compound"-difference between test product and each respondent's regular brand, delivering the "stimulus difference" variable.
2. Group cases by equal "stimulus difference" (binning).
3. Determine the maximum JAR-proportion (Rmax) of all bins.
4. Divide the JAR-proportion of each "stimulus differences"-bin (R) by Rmax, to derive with the "normalised response" variable (R/Rmax).
5. Calculate a non-linear (dose-response) regression with R/Rmax as dependent and the logarithm of the absolute "stimulus difference" as independent variable.

Software:

Data handling, transformation, and aggregation was performed with SPSS 16, nonlinear regression with XLSTAT(2008).

Results and Discussion

Since respondents most frequently used (regular) brand will be the reference one first needs to know the JAR-proportion for those cases when test product is regular brand (blind). As reported by J.-F. Meullenet, a 70% figure seems to be standard for many product categories. This initial check of the data revealed, that the maximum JAR-proportion was lower than 70%, but differs between consumer strength classes (Figure 1). The analysis was therefore performed separately for the high and medium strength classes. The numbers of cases for the low strength class were too low to be analysed and were therefore excluded from the analysis.

Regression model used was a 4-parameters dose response function (Ratkowski's model, provided with the XLSTAT software). Figure 2 is a graphical representation of the regression, separately for 2 consumer classes (high and medium strength), Table 1 summarise the goodness of fit statistics and Table 2 the coefficients of the two regression equations. The regression functions for both classes are nearly identical; indicating that class belonging does not affect the psychophysical function.

The resulting sigmoid relationship delivers a very accurate approximation for small product differences and less accurate prediction for greater product differences (>0.6mg nicotine) due to lack of empirical data. The difference threshold range information can be used for determining tolerance limits for quality assurance which are based on consumer perception rather than technical feasibility. The JND50 information (or other ranges) is a helpful marketing tool for classifying brands according to perceivable sensorial differences and for setting sensory goals for line extensions.

For the data set analysed, the "no-difference range" is determined with ±0.06mg/cigarette. The JND50 was determined with 0.4mg nicotine/cigarette for both target populations (high and medium strength classes).

Conclusions

The method offers a cost-efficient way to establish an empirical, target group averaged psychophysical function in order to allow calculating difference threshold ranges. It is expected that the described approach can also be applied to data from multiple product sequential monadic tests.

Acknowledgements

The authors wish to thank British American Tobacco for providing the data for this exploratory study and also giving their permission to publish these results.

References

L. Rothman, M.J. Parker, Editors (2009) *Just-About-Right (JAR) Scales: Design, Usage, Benefits and Risks*, ASTM international manual series, MNL 63

J.-F. Meullenet et al (2007), *Multivariate and Probabilistic Analyses of Sensory Science Problems*, Blackwell Publishing

H. Heymann, H.L. Lawless (1999) *Sensory Evaluation of Food, Principles and Practise*, Kluwer Academic/Plenum Publisher

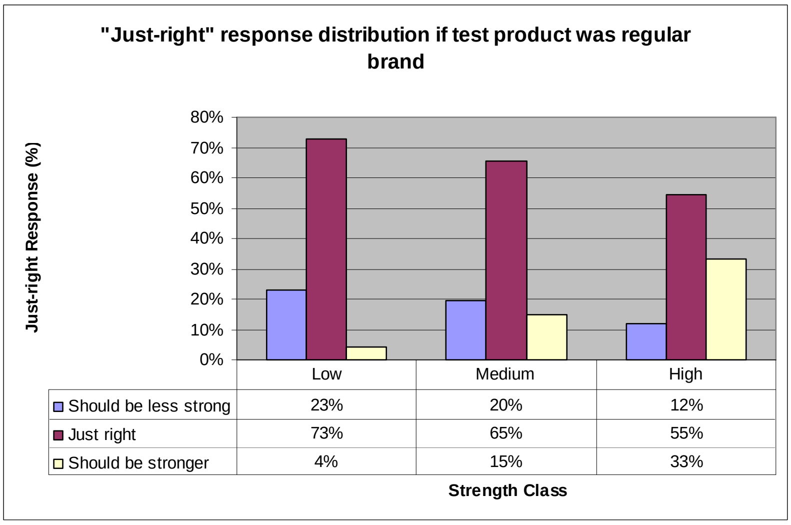

Figure 1: Maximum JAR response differs between consumers of different strength classes.

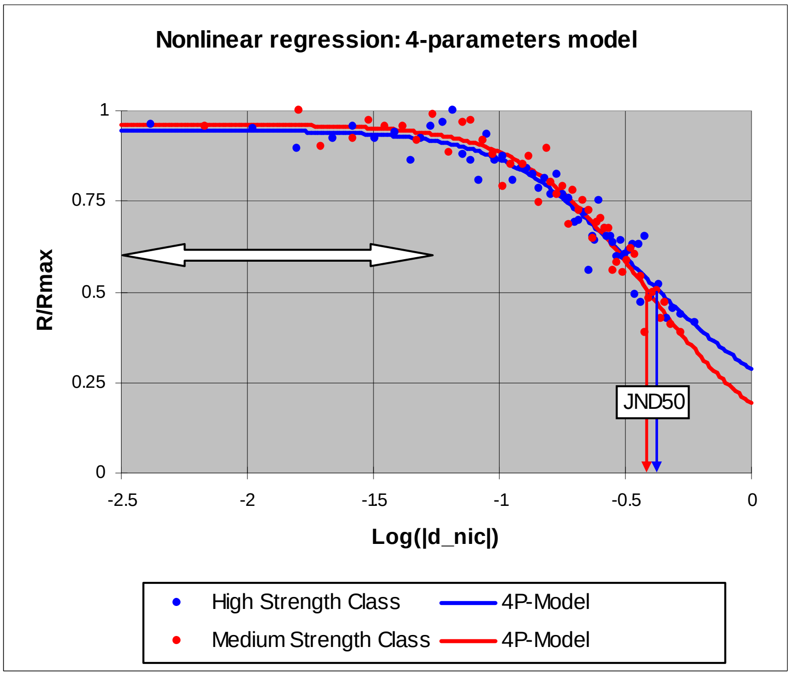

Figure 2: R/Rmax in relation to |d_nic|, the absolute nicotine difference (test product – regular brand) for both Strength Classes. The dual arrow indicates the “no-difference” range.

	Medium Strength Class	High Strength Class
Observations	48	55
DF	44	51
R²	0.944	0.927
SSE	0.089	0.108
MSE	0.002	0.002
RMSE	0.045	0.046
Iterations	29	25

Table 1: Goodness of fit statistic

Parameter	Medium Strength Class	High Strength Class
p_1	-2.096	-2.991
p_2	-4.216	-4.374
p_3	0.962	0.945
p_4	1.369	2.545

Table 2: Coefficients of the does-response function

$$y = \frac{p_3}{(1 + e^{(-p_1 - p_2 * x)})^{1/p_4}}$$