Reprinted from IFPress (2011) 14(1) 3,4

How to Set Identicality Norms for *No Preference* Data

Daniel M. Ennis and John M. Ennis

Background: Expectations, or *norms*, are of crucial importance in science as they allow researchers to interpret results within a larger theoretical framework. For example, in 1866 Gregor Mendel published his now famous Mendelian ratio, which provided expected results that were instrumental in the early development of genetics. The importance of norms is just as great in product research where norms are essential for meaningful data interpretation. In particular, without an expectation regarding the level of *no preference* one has no metric with which to gauge the meaningfulness of *no preference* responses. For example, the ASTM advertising claims guide[1] currently recommends that unqualified advertising claims should not be made if more than 20% of the responses are *no preference*. While this recommendation correctly recognizes the importance of having an expected level of *no preference* responses, one could argue that the figure of 20% has been chosen somewhat arbitrarily. A greater level of awareness as to the meaning of *no preference* responses is required, and in this report we will discuss methods to obtain such an improved level of awareness before paired testing tools are applied. Specifically, we discuss the establishment of a metric that we call an identicality norm[2] that is based on the expected level of *no preference* votes for identical products and we show how such a norm can be established from historical data through the use of Thurstonian tools.

Scenario: You work on taste preferences for the salty seasoned snack category. Results of a recent experiment involving your main brand and a low-fat alternative as part of a company-wide wellness initiative are shown in Table 1. In this experiment 400 general population users of salty snacks appeared to prefer the two products similarly. In fact a chi-square test among those who expressed a preference is not significant ($\chi_1^2 = 0.14$, $p = 0.711$), so you cannot reject a null hypothesis of 50% preference for one or the other of the two products. Based on descriptive data, however, you see that these products are quite different so this result is somewhat counter intuitive, especially when you consider that the attributes on which they differ, mouthfeel and flavor strength, are typically drivers of preference in this category. You wonder if these products had been identical what the preference results would have been, as you would like to compare such results with Table 1.

Prefer Regular Brand	Prefer Low Fat Prototype	No Preference
175	182	43

Table 1. Results from paired testing.

Establishing an Identicality Norm: Leaving our scenario for the moment, suppose that two products, A and B, were tested in a paired preference study with a *no preference* option. To understand why we need an expectation regarding the level of *no preferences*, note that a result such as 45% (prefer A): 45% (prefer B): 10% (no preference), which we call 45:45:10, does not necessarily mean that the items were preferred equally by the test population. For example, this result could also occur if one product was preferred by one segment while the second product was preferred by a different segment. Such a scenario is depicted in Figure 1. In this example, we assume that there is a single variable driving preference, such as sweetness, and that the products differ on this attribute. Within a segment that prefers sweeter products, 7.5% prefer A, 82.5% prefer B and 10% have *no preference*. On the other hand, the remainder of the population comprises a second segment that prefers less sweet products and in this segment the preferences are reversed. If the two segments were of equal size then the total outcome would average 45:45:10. Unfortunately, implicit in the various statistical approaches currently used to test a null hypothesis of identicality is the assumption of an homogenous consumer group[4], which in this case is not satisfied. In order to correctly interpret a result such as 45:45:10 we need additional information that allows us to consider the possibility that the consumer group is not homogenous.

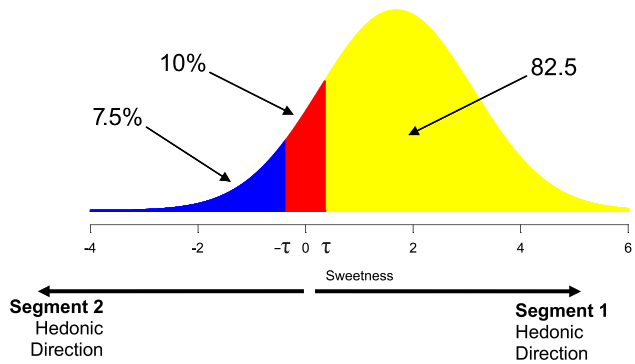

Figure 1. Two segments for which sweetness drives preference exclusively. Consumers in one segment prefer sweeter products; consumers in the other segment prefer less sweet products. There are equal numbers of consumers in each segment. The overall preference result will be 45%:45%:10% for the total sample. In this figure, $\tau = 0.36$ and $\delta = 1.68$.

An example of research seeking such information was reported by Ennis and Collins[5] for a series of choice experiments with a *no difference/preference* option. In these experiments, paired testing of four blind labeled brands of a major consumer products company was conducted to establish difference and preference testing norms for identical products. Products were manufactured at the company's main manufacturing plant, and samples from the same production run for each brand were divided into two groups for paired testing. The results of that research on identical products showed a narrow range of expected preference results with a mean of 40:40:20 based on a total sample of 1,787 consumers. In the Ennis and Collins[5] study, it was found that some highly analytical (non-hedonic) attributes induced a 20:20:60 norm when identical products were tested, suggesting that there may be a continuum of *no preference/difference* probabilities that differ depending on degree of hedonicity. Continuing in this vein, a recent study[6] involving diary products tested on a university community population found a higher *no preference* outcome than that reported by Ennis and Collins[5], while other norms have been found for other categories[7]. Importantly, these results indicate that identicality norms must be established in each particular category and that a single norm cannot be prescribed across categories and populations independently of testing. In particular, a number of factors contribute to the level of *no preference* counts. These include the nature of the products tested, the extent of product variability, the gender and age of the consumers, the test format and possibly even the personalities of the target population.

Predicting an Identicality Norm: As we just saw, it is possible to test for identicality norms explicitly. Even so, such testing can be costly, time consuming and difficult to justify to management. Fortunately, to obtain identicality norms without incurring the costs of additional testing, it is instead possible to use historical data but there are three criteria that need to be met[8]. The first is that the historical data must come from the same category as that for which the norm will be applied (a norm for snack foods might not correspond to a norm for beer); the second is that the group of test subjects should be as similar as possible demographically to the current group of test subjects; and the third is that there needs to be no effective segmentation in the historical data. This last criterion will be satisfied if either: *a)* the tested group is homogenous regarding preferences or *b)* the products are essentially identical in which case preference segmentation does not matter. Examples of good candidate datasets from which to establish identicality norms include quality control data on existing products for which differences between products tested can be assumed to be trivial and data from an essentially homogenous consumer group, such as heavy users. Having selected appropriate data we then apply a Thurstonian 2-AC analysis to determine at what point differences between products become large enough that respondents begin to have preferences. We call this point τ and we estimate τ from the chosen data. This idea is expressed in Figure 2. Table 2 shows data from five of your past experiments where products from the salty seasoned snack category have been tested by high volume, loyal users of your company's main brand together with estimates of τ in each case obtained using *IFPrograms*.™ The τ values are quite similar and average to 0.435. From the Thurstonian theory[9], or using *IFPrograms*, you then compute the identicality norm to be 38:38:24 as shown in Figure 2.

Prefer Regular Brand	Prefer Prototype	No Preference	τ
195	48	57	0.431
159	75	66	0.424
114	113	73	0.438
111	117	72	0.432
114	111	75	0.450

Table 2. Historical data and results of Thurstonian analysis.

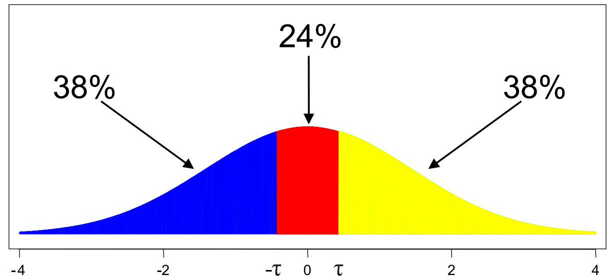

Figure 2. The distribution of product differences when products are identical ($\delta = 0$) and $\tau = 0.435$ with the identicality norm of 38:38:24 shown.

Interpretation of Table 1: Comparing the results from Table 1 with this norm leads to a highly significant chi-square value with two degrees of freedom ($\chi_2^2 = 38.66$, $p < 0.001$). Thus you now know: *a)* the products are different and *b)* the apparent similarity is due to opposing segments. Following up on this knowledge you consider the behavioral and demographic characteristics of the consumer sample and determine how segmentation is driven by differential preferences for low-fat snacks.

Conclusion: Expected theoretical outcomes are valuable in any scientific endeavor. In particular, identicality norms are essential for the meaningful analysis of paired experiments with a *no preference* option as they provide the basis for distinguishing two scenarios that often occur in paired testing: an equal split due to product equivalence and an equal split due to consumer segmentation. Classical methods of analysis, such as tests among those who express a preference, require an assumption of homogeneity within the consumer population for their correct interpretation. If this assumption is not met, then the two scenarios described above will be confounded.

References and Notes

1. ASTM International. (2007). E1958-07 *Standard Guide for Sensory Claim Substantiation*.
2. The term *identicality* is a theoretical construct wherein the products are identical and should not be confused with *equivalence*, in which differences may exist within defined bounds[3].
3. Ennis, D.M. and Ennis, J.M. (2009). Hypothesis testing for equivalence based on symmetric open intervals. *Communications in Statistics*, **38**(11):1792-1803.
4. Ennis, D.M. and Ennis, J.M. (2010). How to account for "no difference/preference" counts. *IFPress*, **13**(3):2-3. Available at www.ifpress.com.
5. Ennis, D.M. and Collins, J. (1980). The distinction between discrimination and splitting in paired testing. *Philip Morris Technical Report*, **80:** 233. Available at http://www.pmdocs.com/pdf/1000386683_6734.pdf.
6. Chapman, K.W. and Lawless, H.T. (2005). Sources of error and the no-preference option in dairy product testing. *Journal of Sensory Studies*, **20**(5):454-468.
7. Alfaro-Rodriguez, H. and Angulo, O. and O'Mahony, M. (2007). Be your own placebo: A double paired preference test approach for establishing expected frequencies. *Food Quality and Preference*, **18**(2):353-361.
8. Ennis, D.M. (2005). Relative scales and difference testing norms. *IFPress*, **8**(3):2-3. Available at www.ifpress.com.
9. $P(\textit{No Preference} \mid \textit{Identical Products}) = 1 - 2\Phi\left(\frac{-\tau}{\sqrt{2}}\right)$