Reprinted from IFPress (2013) 16(2) 3,4

# Transitioning from Proportion of Discriminators to Thurstonian δ

*Benoît Rousseau, Daniel M. Ennis, and John M. Ennis*

**Background:** In a previous technical report[1] we discussed a commonly used measure of sensory difference, the proportion of discriminators ($P_d$), and explained that this measure is method specific. We concluded that:

> *"The concept of the proportion of discriminators is intuitive and appealing. Unfortunately, it is also method-specific and one can achieve very different values using different methods, which reduces its appeal as a management criterion. The use of a method–independent index, such as the Thurstonian d', can provide more stable and accurate information that may be valuable in a decision that involves the consideration of consumers' ability to discriminate between alternate product formulations."*

Many companies have implemented sensory discrimination programs using $P_d$ as a parameter for power calculations. Since this measure has shortcomings, there is a need to consider a unit that is method-independent such as the Thurstonian δ. Fortunately, tools exist that permit the translation of a method-specific $P_d$ into the corresponding δ value. In this report we discuss how these tools can be used in practice.

**Scenario:** You manage the sensory and consumer insights group of a global consumer products company. Among your duties you oversee the daily operations in your carbonated lemonade business. When investigating potential ingredient modifications, your team has established a specific experimental protocol using the triangle test and associated test specifications. These specifications are: α level (Type I error) at 1%, power at 90% (Type II error at 10%), and level of acceptable difference measured as $P_d$ of 20%. Based on published tables[2], you need a sample size of 176. You reach this sample using internal employees with replication.

Following budget cuts, you are strongly encouraged to improve the efficiency of your internal testing. Reducing the sample size would lower your power and the precision of your measurements. You decide to look into alternative testing protocols. Recent research suggests the feasibility of using the tetrad test, as it is theoretically more powerful than the triangle test, and thus would require a lower sample size[3].

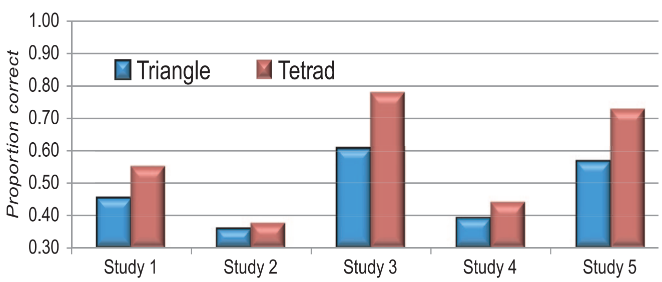


**Figure 1.** Proportion correct in 5 studies involving the triangle and tetrad tests.

**Triangle vs. Tetrad:** You study this new methodology according to a recommended experimental approach[4] using five studies in which the panelists evaluated the samples using both the triangle and tetrad tests. As expected, you obtain larger proportions of tests correct with the tetrad methodology as shown in Figure 1.

These first results are encouraging. The tetrad test indeed appears more powerful and thus will potentially permit a reduction in the needed sample size, freeing resources and enabling faster decision-making. But since your usual protocol uses $P_d$, you calculate the proportion of discriminators in each of the five studies. The results are shown in Figure 2.

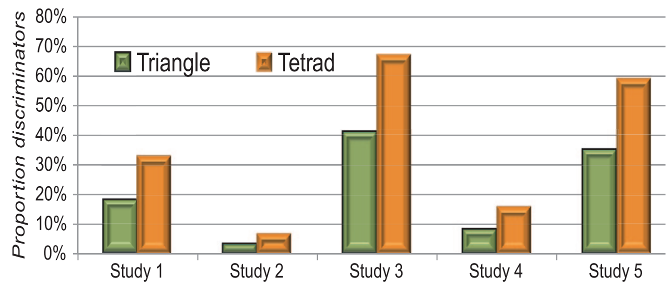


**Figure 2.** Proportion discriminators in the 5 studies.

The proportions of discriminators differ between the two methodologies. How is this possible since both tests were measuring the same underlying sensory difference and the two methods have the same guessing probability? Setting the power specification for the tetrad test using a $P_d$ of 20%, as you did with the triangle test, will not provide the same accept/reject probabilities. Since the two methods have the same guessing probability, using your original tables[2] will again provide a recommended sample size of 176. Since the tetrad method seems to always yield a larger proportion of tests correct, you will thus reject a change more often than with the triangle test. You decide to investigate this issue further, knowing that you will need to account for the difference in methodology.

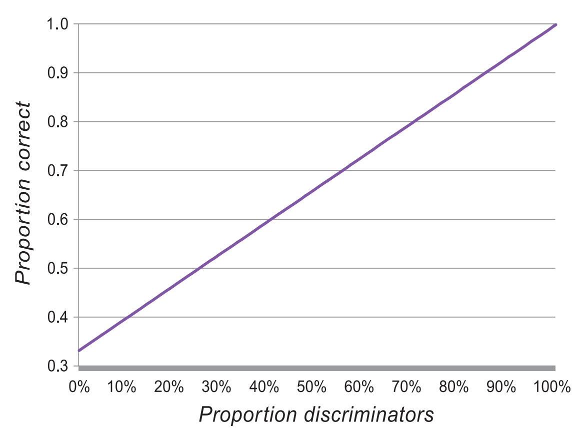


**Figure 3.** Relationship between proportion of discriminators and proportion of correct responses for the tetrad and triangle tests.

**Translating $P_d$ into δ:** A basic problem with $P_d$ is that the measure is method-specific, and thus the same underlying sensory difference will correspond to different proportions of discriminators depending on the method used. For instance, a δ of 1 will correspond to a $P_d$ of 13% for a triangle test and 24% for a tetrad test[5]. The method-dependency of this measure suggests the need to make sample size estimates with a measure that avoids this problem, such as δ.

When using different methodologies, it is thus necessary to translate $P_d$ into δ and then conduct the power and sample size calculation at this new level. Such translation is actually fairly straightforward. As shown in Figure 3, the reference $P_d$ is transformed into the corresponding proportion of correct answers ($P_c$):

$$P_c = P_d + P_{chance}\,(1\text{-}P_d).$$

Thus, for 25% $P_d$ in a triangle test, $P_c$ = 50%. Published tables[6] can be used to translate $P_c$ into δ. A criterion of 25% $P_d$ in the triangle test corresponds to a δ of 1.47. If study specifications for a triangle test were set at an α of 0.05, a power of 80%, and a δ of 1.47, then a sample size of 60 is required[3]. If the same study specifications must be used with the tetrad test, the required sample size is 22. Tables allowing the direct translation of $P_d$ into δ are also available[7] and an excerpt is shown in Table 1. For values not in the tables, interpolation can be used.

| $P_d$ | 2-AFC | Duo-Trio | Triangle | Tetrad |
|---|---|---|---|---|
| … | … | … | … | … |
| 10% | 0.18 | 0.76 | 0.88 | 0.62 |
| 15% | 0.27 | 0.95 | 1.10 | 0.77 |
| 20% | 0.36 | 1.12 | 1.29 | 0.90 |
| 25% | 0.45 | 1.27 | 1.47 | 1.02 |
| 30% | 0.55 | 1.42 | 1.64 | 1.14 |
| … | … | … | … | … |

**Table 1.** δ as a function of proportion discriminators for the 2-AFC, duo-trio, triangle and tetrad tests.

Figure 4 summarizes the concepts just discussed. If the same reference of 25% discriminators was used for the tetrad test as it was with the triangle test, the criterion would become more stringent as the size of the corresponding underlying difference would be a δ of 1.02.

**Translating Specifications from the Triangle to the Tetrad Test:** Using the referenced tables, you translate your reference $P_d$ of 20% and find the value of δ to be 1.29. You now have a measure of difference that is method-independent. You use it to find the necessary sample size for the tetrad test based on the same risks you had specified with the triangle test [α level at 1%, power at 90%, and level of acceptable difference measured as $P_d$ of 20% (δ of 1.29)]. The corresponding sample size is 56 for the tetrad method[3]. Therefore, by switching to the tetrad test, you reduced your necessary subject resource requirements by a third, without compromising the rigor of your future investigations compared to your historical research.

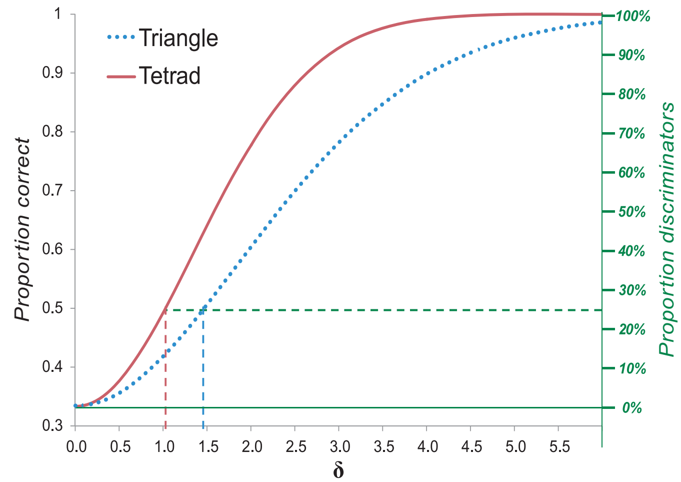


**Figure 4.** Relationship of δ to proportion correct and proportion discriminators for the triangle and tetrad tests; example of a $P_d$ of 25%.

**Conclusion:** $P_d$ is an attractive but ultimately misleading measure of sensory difference. It depends on the methodology used and does not carry its implied meaning. In fact, even if a single methodology is only ever used, one should still investigate the consumer relevance of sensory differences to establish the equivalence criterion instead of accepting $P_d$ at face value[8]. Fortunately, if a company has established a reference value of $P_d$ that has been shown historically to be consumer relevant, translating this reference value into a Thurstonian δ eliminates the inherent problems associated with the proportion of discriminators, and facilitates more informed investigations as new developments in sensory testing arise.

## References and Notes

1. Rousseau, B. and Ennis, D. M. (2007). Why proportion of discriminators is method-specific. *IFPress,* **10**(3), 2-3.
2. Schlich, P. (1993). Risk tables for sensory discrimination tests. *Food Quality and Preference,* **4,** 141-151.
3. Ennis, J. M. and Jesionka, V. (2011). The power of sensory discrimination methods revisited. *Journal of Sensory Studies,* **26,** 371-382.
4. Ennis, J. M. (2012). Guiding the switch from triangle testing to tetrad testing. *Journal of Sensory Studies,* **27,** 223–231.
5. δ is the Thurstonian measure of underlying sensory difference.
6. Ennis, D. M., Rousseau, B., and Ennis, J. M. (2012). *Short Stories in Sensory and Consumer Science*. Richmond, Virginia: The Institute for Perception.
7. Jesionka, V., Rousseau, B., and Ennis, J. M. *(In press).* Transitioning from proportion of discriminators to a more meaningful measure of sensory difference. *Food Quality and Preference.*
8. Rousseau, B. and Ennis, D. M. (2013). When are two products close enough to be equivalent? *IFPress,* **16**(1) 3-4.