A “magic number” of descriptive panelists does not exist

Edgar Chambers IV, Siim Koppel

Kansas State University; Center for Sensory Analysis and Consumer Behavior

Center for Sensory Analysis and Consumer Behavior

Abstract

Authors discussing trained descriptive panelists have recommended their selection be based on their ability to give reliable data. Although authors state the number depends on the quality of the panel, the size of the differences present, and the degree to which the researcher is willing to err, many people still believe a 'magic' number of 10-12 panelists is needed. Our objective was to determine the number of panelists needed of different quality to determine varying degrees of difference among products. We simulated data from 220,400 descriptive panels representing 3-25 members, qualities of 5, 10, 15, 20, and 25% standard deviations (SD) among panelists, and product tests where the mean differences were 5, 10, 25, and 40% of the scale. 400 generated data sets were analyzed repeatedly to determine how often significant differences were found with 25 or fewer panelists. Panel quality had an enormous impact on the necessary panel size. A 5 member panel with a 5% SD found differences in products with small 5% mean differences approximately 90% of the time while a panel with a 25% SD found differences in products with 25% differences 90% of the time. This suggests that small panels of high quality can find small differences and even small panels of intermediate quality can find key differences. With panels of lower quality (e.g. 10% SD) it took almost 20 panelists to be able to find 5% differences in products 90% of the time and when even lower quality panels were considered (i.e., panels with 15, 20, 25 SD) it was impossible to find 5% differences repeatedly and took 10-25+ members to find differences of 10%. Deterioration in panel quality caused a rapid deterioration in ability to find smaller differences repeatedly and showed that even large panels of mediocre quality do not work as well as small high quality panels.

Objectives

1. To explore the effect of panelist variation and product variation on the detection on significant differences between products
2. To generate a guideline and recommendation on how many panelists are needed for descriptive analysis based on the variability of panelists and products

Methods

Datasets:

Parameters used for generating the datasets were:

Number of panelists to start was 25.

Standard deviation (SD) for the panel was chosen to be 5%, 10%, 151%, 20% and 25% of the 15 point scale

Product deviation was set at 5%, 10%, 25% and 40% of the scale.

Distribution of the scores was assumed to be normal.

Simulation 1:

1. Generating random values for 5 products using normal distribution, mean of 7 and SD specified previously
2. Generating random values for each panelist (mean of 0 and SD specified previously)
3. Each panelist was assigned 3 random numbers from step 2 for each product
4. The 3 numbers generated for each panelist were added to the product score for each product from step 1.
5. The outcome was a table of scores for 5 products as assessed by 25 panelists repeated 3 times.
6. Each combination of product and panel deviations was repeated 20 times.

Overall number of generated datasets: 400

Data Analysis:

The generated data was analyzed using Analysis of Variance (ANOVA) with repeated measured (panelists). The Mean Squared Error (MSE) values for Product and interaction between Product and Panelist were recorded. This removes the number of panelists as separate source of variation.

The F value for hypothesis testing was calculated by :

$$F = \frac{MSE(Product)}{MSE(Product * Repetition)}$$

Next p value was calculated on the probability of calculated F value being larger than an F value with df1 = 4 and df2 = 8

If p <0.05, the decision was made that a difference between products was detected.

Simulation 2:

For each dataset:

1. Remove each panelist one by one to generate 25 datasets with 24 panelists in each
2. Run data analysis on each new dataset generating 25 p values
3. Remove random 2 panelists from the dataset to generate 25 datasets with 23 panelists in each
4. Run data analysis
5. Continue until 3 panelists were left to each dataset.

For each dataset 551 ANOVAs and p-values were generated.

Overall 220,400 “panel results” were generated.



Conclusion

The number of panelists needed to find significant differences among products is dependent on both the product and the panel standard deviation.

- For highly trained panels with low deviation (e.g. 5% SD), the number of panelists required, even to find differences in products with small differences is low. **For example, such a well-trained panel of only 4 panelists would find differences more than 80% of the time even for products that are barely different.**
- **Even panels of 5 members with a deviation of 25% would find differences in products with moderate differences more than 90% of the time.**
- **Panels with less training having higher deviation (e.g. 25%) need 15 or more members to find smaller differences in products (<10% SD) 80% of the time and will rarely find small differences even with 25 members**.

Results and Discussion

For all graphs: blue line – 5% product deviation, yellow line- 10% product deviation, dark blue line 25% product deviation and brown line – 40% product deviation

Figure 1. The results of the simulations for the panel deviation of 5%.

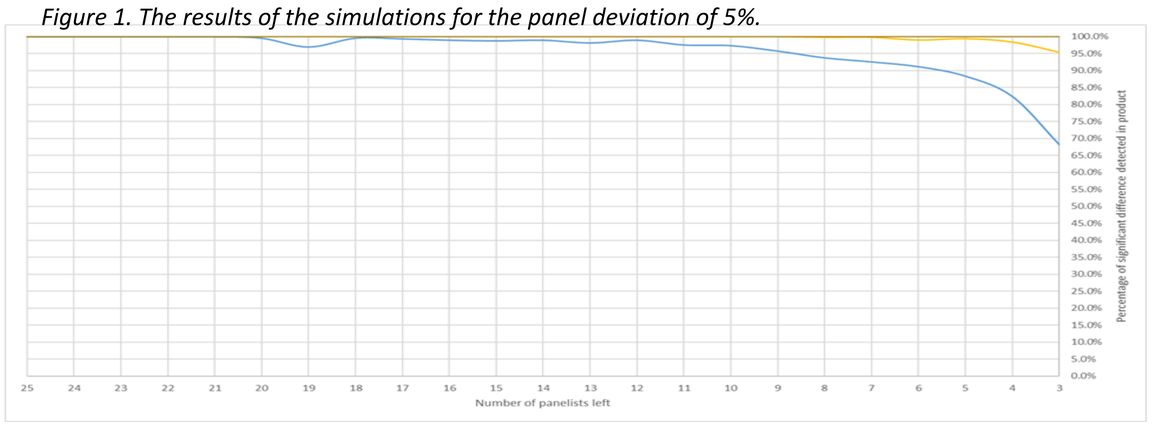

Figure 2. The results of the simulations for the panel deviation of 10%.

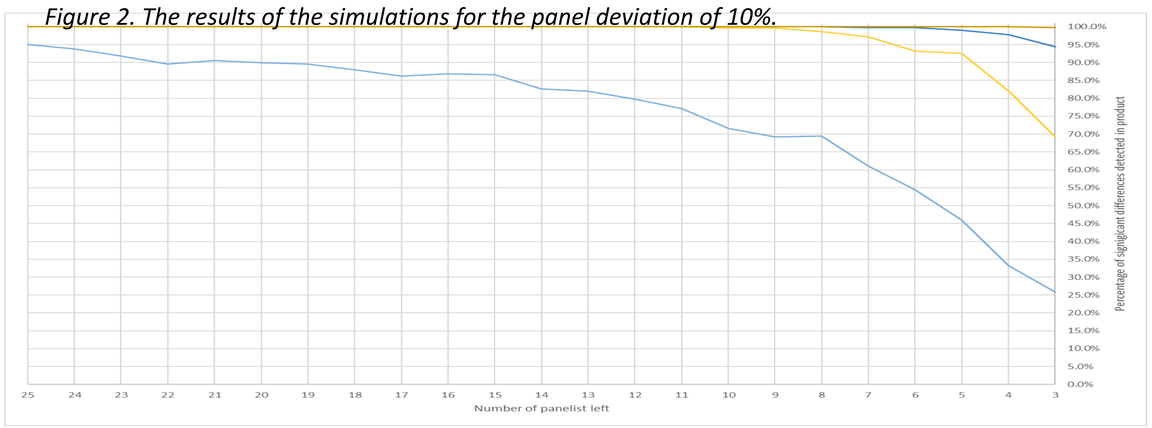

Figure 3. Results of the simulations for the panel deviation of 15%.

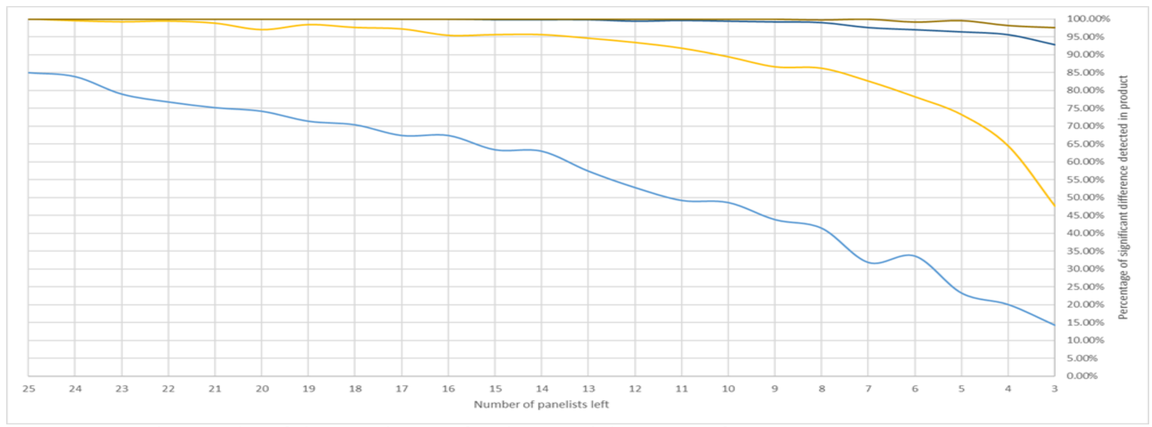

Figure 4. The Results of the simulations for the panel deviation of 20%.

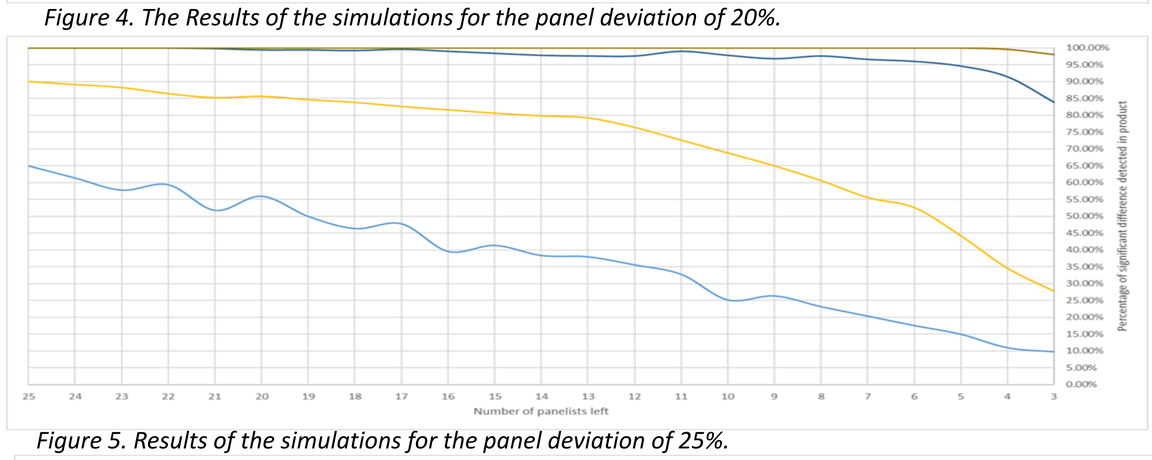

Figure 5. Results of the simulations for the panel deviation of 25%.

