The Institute for Perception

For technical reports from The Institute for Perception or for information about short courses, please visit www.ifpress.com or email us at mail@ifpress.com

Reprinted from IFPress (2015) 18(1) 3,4

Identifying and Removing Sources of Bias in Product Tests and Surveys

Daniel M. Ennis and Benoît Rousseau

Background: In product tests and surveys, bias occurs when a measure from a sample systematically differs from the population measure of interest. To put this into statistical terms, a statistic is biased if it systematically deviates from a population parameter, irrespective of the sample size. The existence of bias determines whether one needs a control product or item in a product test or survey.

There are numerous sources of bias that have been identified. These include sampling bias, where a sample of participants or the items to be tested do not represent either the target population or the real test items. Bias also may occur when participation or non-response in a survey is not random, so that the opinions expressed do not represent the target population. Leading questions, interviewer effects, and uncontrolled individual differences can all contribute to bias. Position bias and code bias are two sources that will be discussed in this report. These two sources are relatively easy to control but code bias, in particular, is often ignored in practice. In the case of code bias, the codes themselves may contribute to the responses selected. An extreme example that we have observed involved data from a Chinese research supplier. The Chinese ideogram for the number 4 is close in appearance to that for "death" and therefore the number "4" is often avoided in practical situations such as a floor number in buildings and hotels. When one of the products in their study was coded as "444", it was not surprising that it received a poor hedonic rating. There are many other less dramatic sources of code bias that, if not controlled, may lead to inaccurate parameter estimates.

Scenario: You are a product development manager working for a food processing company and often submit samples for testing to the sensory and consumer evaluation group. Some of your submissions involve relatively minor blend and flavor modifications which are conducted when lower cost ingredients become available or to qualify a new supplier of a standard ingredient. It surprises you to find that in many instances the submitted samples are found to be significantly different. Sometimes even the same pair retested may give opposite results. When this occurs, there is often a delay in making a decision about a planned change, and sometimes it may lead to new testing. You want to know if these counter-intuitive results are just due to random variation or to an identifiable source of bias.

Evidence for Code Bias: Figure 1 shows the results of two real product tests on beverages conducted in the USA. In both studies there is little or no evidence that either product is preferred because in Study 1 the choice proportions were 44%:44%:11% (Prefer A, Prefer B, No Preference) and in Study 2 they were 45%:45%:10%. In this research, which anticipated the possibility of code bias, both high and low three-digit random code numbers were used for the same product. These particular tests were conducted in eight cities. In four of the cities one product was coded 457 and in the other four cities it was coded 892. Product B was coded the reverse so that both products appeared under both codes. The same design was used in Study 2. Figure 2 shows that when the results are presented by code number, irrespective of product, the higher codes are chosen more often in both studies, and they averaged 54%:46% over both studies in favor of the higher code. This result is significant at the 95% level.

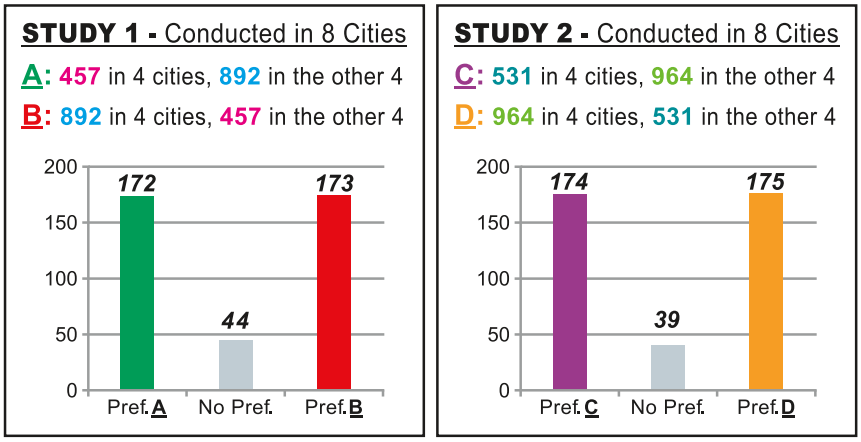

Figure 1. Results of two preference tests on beverages. As shown, both studies were conducted in eight cities and, to control code bias, each product was tested under a high code and a low code. The results are preference counts for A/B or C/D over both codes. There is no demonstrable preference for either product in each pair.

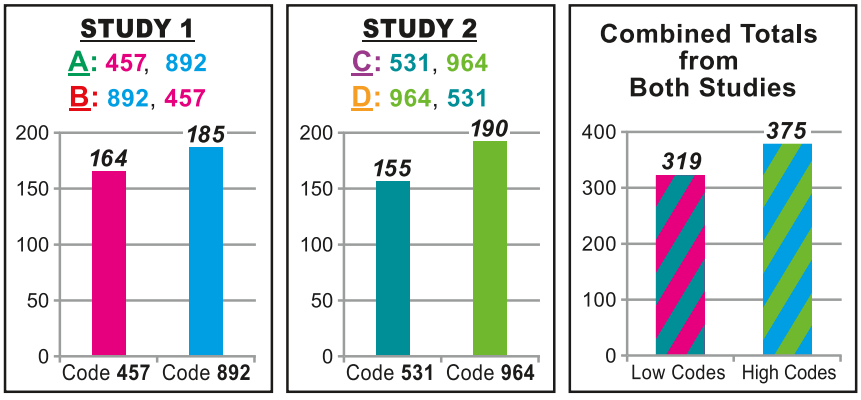

Figure 2. Preference counts for A/B and C/D product code numbers. The higher code numbers received the higher counts in both studies and these results are statistically significant at the 95% level.

Evidence for Position Bias: Figure 3 shows the results for the two product tests discussed in the previous section averaged over high and low codes for each position in which the products appeared. It is clear from Figure 3 that the well-known preference for the first item tested was demonstrated in the two studies. A simple solution to overcome position bias is to balance the order of testing by product, a practice commonly adhered to by product testing suppliers.

Blend and Flavor Modifications: An investigation into possible sources of bias in your product test submissions revealed that the typical practice used by the sensory and consumer evaluation group is to assign single three-digit randomly chosen codes to each product tested. This procedure may have led to the unexpected results you obtained, especially in comparisons where the differences between products are small. The testing participants may then resort to any strategy that would justify expressing a preference. You recommend to the product testing group that multiple product codes be used either across different sessions or within the same session so that each product is presented under high and low codes. There was no evidence for position bias as the testing group routinely rotates the samples in a balanced fashion.

Recommended Coding Practice: In previous research[1,2] we showed that large numbers of consumers, often as high as 80%, express preferences for putatively non-existent differences when identical products are tested. Under these circumstances code bias may emerge strongly. When small changes in products are of interest, it is most important to control code bias. In addition to code bias created by large and small code numbers, other biases can occur due to area codes, emergency numbers, airplane codes (747, 757), sequential runs up or down (234, 654), and numbers repeated in a three-digit triad (222). A set of three-digit numbers that are free from these effects can be downloaded at The Institute for Perception's website, www.ifpress.com. From this set, codes for product tests can be selected. It may be expected that the effect of code bias would decrease as the product differences or preferences are detectable and the evaluators then have an item-specific basis for their choices.

Conclusion: Code and position bias can occur in any product test or survey context, but in particular when direct comparisons are made between items that are very similar. In these situations we recommend the use of high and low codes for each item but the adoption of this practice in all comparative testing is the best strategy to ensure control over code bias. If multiple pairs of codes are used within a session, we recommend that the difference between the high and low codes for a given pair should be similar to the difference in codes for a different pair. Position bias is easily controlled using balanced rotations[3].

References

1. Ennis, D. M. and Collins, J. (1980). The distinction between discrimination and splitting in paired testing. *Philip Morris Technical Report,* 80:233.
2. Ennis, D. M. and Ennis, J. M. (2012). Accounting for no difference/preference responses or ties in choice experiments. *Food Quality and Preference,* **23**(1), 13-17.
3. Ennis, D. M., Rousseau, B., and Ennis, J. M. (2014). *Tools and Applications of Sensory and Consumer Science.* (pp. 118-119). Richmond, VA: The Institute for Perception.

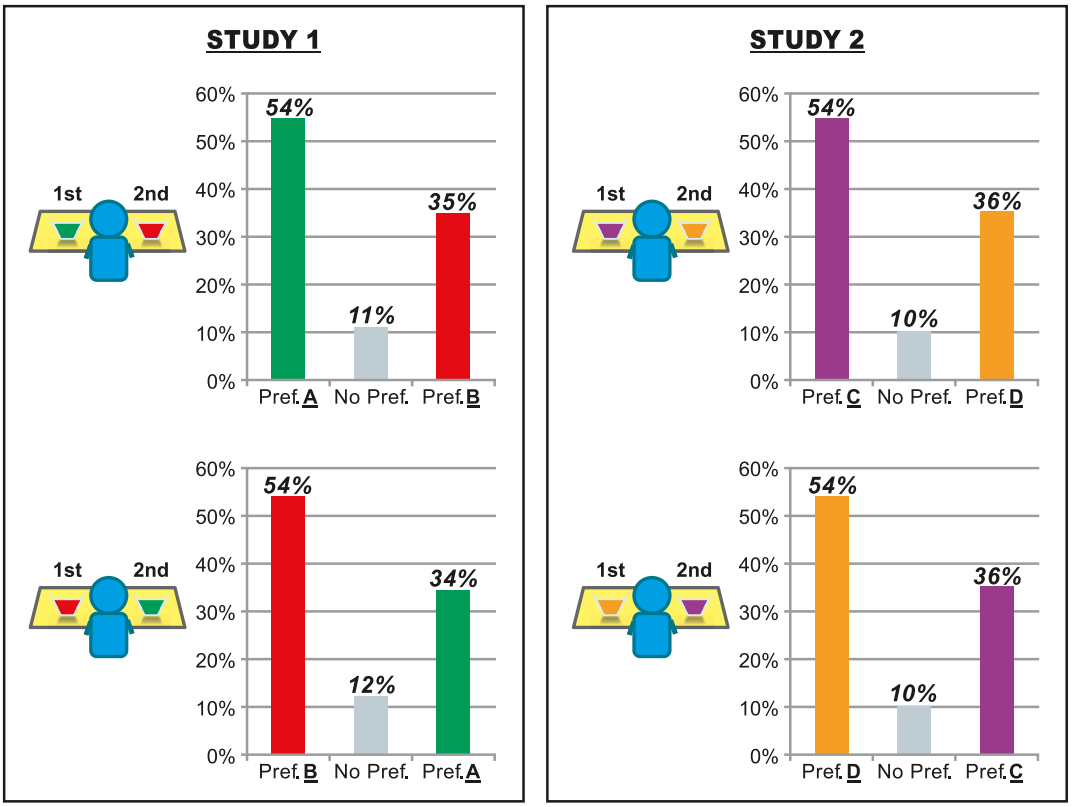

Figure 3. Preference counts for the A/B and C/D products. The product tested first received the higher counts in both studies.