Considerations and Implications of Including Auxiliary Products in Design CLTs for Benchmarking Purposes: Case-Studies from Industry

Authors: Catherine Lee, Carol Uy

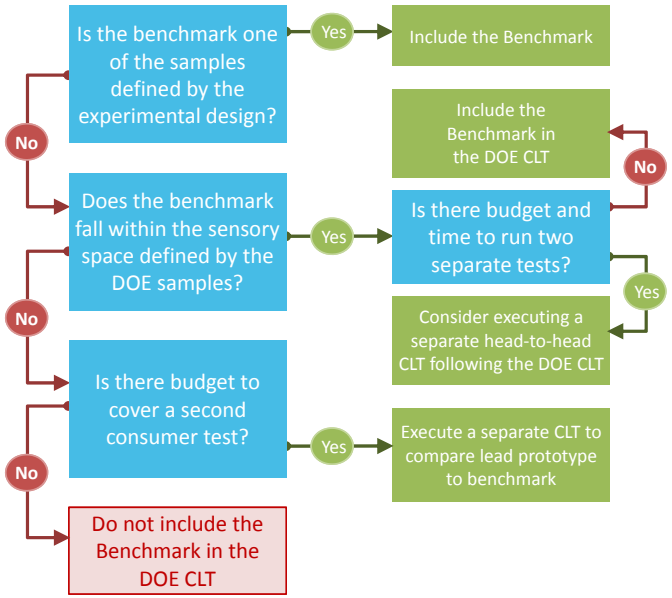
Introduction

Design of Experiment (DOE) CLTs are powerful and efficient tests capable of providing rich learning, yet a secondary objective often arises for comparison of product scores to a benchmark, such as a control or competitive product.

When the benchmark is part of the design, the power of the test to detect differences between the benchmark and individual prototypes is maximized. This situation is win-win: foundational learning is obtained on attributes driving liking within the set of prototypes, and the prototypes can each be benchmarked to the control.

However, the benchmark doesn't often fit neatly into the experimental design, and sensory-related context effects may bias the data, such that no statistical correction can be made. The effect on the total body of research can be substantial, particularly when the benchmark is clearly different.

In these instances, two general circumstances can arise. Risks, best practices, rationale, and case studies for these two situations are presented.



Questions? Contact: catherine.lee@genmills.com

Benchmark is NOT part of design but SIMILAR to other products

Rationale

- Context effects are mitigated due to limited sensory differences
- Expand learning from effects controlled within DOE to include relative performance against a benchmark

Risks

- Compared to DOE analysis, pairwise means-comparison of each individual sample against benchmark will be less precise, increasing risk of Type 2 error
- Inclusion of benchmark is subject to team's discretion; context effects may be larger than anticipated

Recommendations

- Ensure minimal sensory differences exist between benchmark and DOE samples; maintain full rotation of all products within CLT

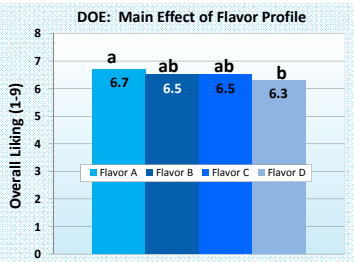
CASE STUDY: Including a benchmark among DOE samples augments product learning

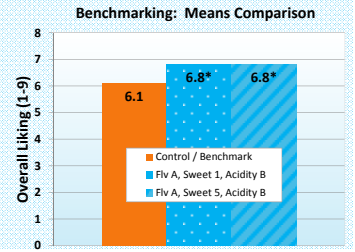
Background

- Team sought to improve existing product through a flavor system change, maintaining texture and appearance, using a complex DOE
- Benchmark = in-market control, added within full rotation of DOE samples

Results

- DOE delivered insights on impact of product levers on consumer liking
- Dunnett's analysis supported DOE learning, with some samples performing significantly better than control
- One option was pursued for further optimization and launch. DOE learning was applied to other SKU's within portfolio





Benchmark is NOT part of design and markedly DIFFERENT from other products

Rationale

- Context effect cannot be mitigated due to large sensory differences between benchmark and DOE samples, causing unpredictable response bias
- Unknown impact caused by large response bias poses too great a business risk

Recommendations

- Identify lead prototype from DOE CLT
- Execute separate follow-up CLT that includes benchmark and lead CLT
- Never append benchmark onto DOE CLT in fixed/last position

CASE STUDY: Including benchmarks at the end of a CLT results in misleading scores

Background

- Team was tasked with creating dry dinner mixes to combat successful line of competitive products
- Benchmark = competitive product, served in last position in an effort to protect main objective of CLT and understand performance against competitor

Results

- Contextual expectations generated by prototypes skewed competitive CLT liking scores, creating false confidence
- Contradictory data from subsequent IHUT demonstrated that CLT liking scores for competitive products were misleading

