THE POWER OF PRODUCT TESTING WITH SYNTHETIC DATA

Humanizing AI series, part two

Ipsos' Research for Product Testing with Synthetic Data

Ipsos carried out two research waves to find ways to uncover effective strategies for making product testing more agile.

Wave 1

01. THE QUESTION

Will smaller[1] sample sizes yield results similar to those from larger[2] sample sizes?

02. THE APPROACH

Using data from Ipsos' extensive product testing database, we took a *three-step process* to correlate overall liking scores with sample sizes:

01 Identified the best-worst product gap

02 Estimated this gap across different sample sizes

03 Calculated the correlation

MARKETS	CATEGORIES	CONSUMERS
84	184	40K+
Geographical diversity from 84 markets	Products selected	Consumers sourced

03. THE DISCOVERY

A small human sample of 50 respondents can yield results like a larger human sample.

Larger product differences lead to more consistent results between sample groups.

>8%

When the top product differs from the worst by at least 8%, 50 respondents suffice, with *some limitations* (correlation coefficient = 0.8).

04. THE TAKEAWAY

Groups of 50 can mimic samples of 200 but have *limitations*:

Limited insights into subgroups

Lower statistical power for detecting product differences

1. Small Human Samples (n = 50)
2. Large Human Samples (n = 200)
3. Small Human Samples (n = 50) Augmented with Synthetic Samples (n = 150)

Wave 2

01. THE QUESTION

Will small human samples, augmented with synthetic data[3], mirror results from larger human samples?

02. THE APPROACH

To ensure robust validation, we used data from various product categories, markets, and testing formats (including blind and branded studies), and considered several factors:

- BUSINESS DECISIONS AND ACTION STANDARDS
- PRODUCT RANKINGS AND DIFFERENCES (1st, 2nd, 3rd)
- DATA DISTRIBUTION
- CORRELATION

03. THE DISCOVERY

R=+0.9

Results were consistent across all-human and human-synthetic groups, with an average correlation of +0.9.

In 60% of subgroup cases, adding synthetic data made existing product differences statistically significant.

For instance, when the human-only sample group showed unplausable results between the penalty analyses on flavor strength and overall liking of flavor, the augmented sample showed a much stronger and plausable response, aligning better with the penalty analyses.

04. THE TAKEAWAY

Small human samples, when combined with synthetic data, can predict outcomes of larger samples more efficiently, with greater statistical power, and deeper insights into subgroups. The quality and structure of the small human sample defines the quality of the synthetic data.

CONTACT US

Dr. Nikolai Reynolds,
Global Head of Product Testing
Nikolai.Reynolds@ipsos.com

Ipsos