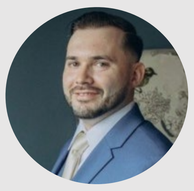
A statistical exploration of going "beyond liking" in category reviews by changing the response metric



Michael Nestrud, PhD
VP, Strategic Accounts
Curion, Chicago, IL
Correspondence:
mnestrud@curioninsights.com



Tony Gualtieri
Sr. Statistician
Curion, Old Bridge NJ,
Correspondence:
tgualtieri@curioninsights.com



Roman Igudin
Statistician
Curion, Old Bridge NJ
Correspondence:
rigudin@curioninsights.com



Rebecca Maine
Client Services Director
Curion, Deerfield IL
Correspondence:
rmain@curioninsights.com



Emily Kimmins
R&D Lead - Sensory and Consumer Science
The Kraft Heinz Company, Chicago, IL
Correspondence: emily.kimmins@kraftheinz.com

Background

The traditional sensory KPI of Overall Liking, developed and used by the industry since the 1940s, is not necessarily predictive of sales. What makes a product well liked and what makes it sell well is not always the same. Curion has developed a calculated metric that integrates the concepts of liking and optimization together to create a comprehensive way to evaluate product performance. This metric (commercially known as the CurionScore™) is currently being used to provide category benchmarking on product performance and to assess alienation risk based on category slippage. It provides a unique perspective of category performance that is independent of sales and marketing metrics, and provides the ability to differentiate performance and size up opportunities in a broader context with a greater degree of sensitivity than liking [1].

Because this metric is uniquely calculated at the individual product * respondent level, using a series of normalization calculations we have the ability substitute CurionScore™ into most analytical methods where we previously have used liking, thus shifting the interpretation from "liking" to "product performance." Both internal preference maps and clustering analyses were performed on a large retail CLT data set using both traditional Liking hedonics and the CurionScore™ product performance metric as the dependent variable and results are compared.

Results indicate that the CurionScore™ metric differentiates performance at the higher levels better than liking.

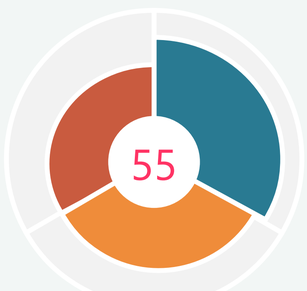
Methodology

Consumer Testing:

- In-person evaluation of 30 in-market frozen entrees
- Recruit: Category users
- N=402 adults aged 18 to 54
- Provides Ratings of Overall Liking on 9-point scale
- Date: Fall 2022
- Locations: Chicago, Dallas, Boston, NJ, San Francisco and Atlanta

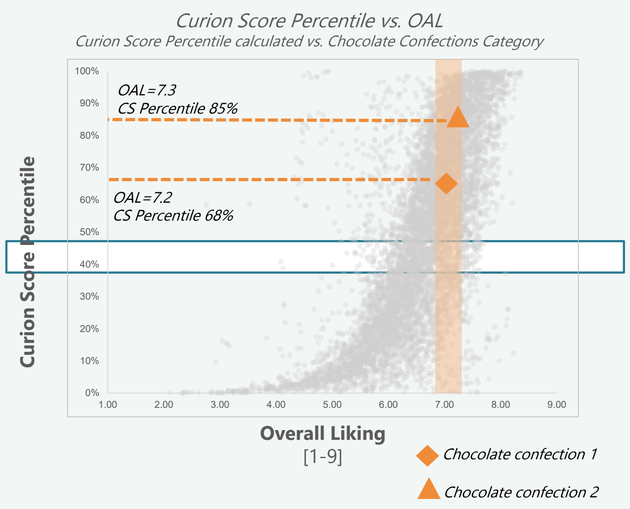
CURION SCORE
Product Performance Measurement.

- Calculated at the individual * product level
- Normalized and controlled for context
- Provides basis for Curion's product performance benchmarking
- Eligible on ~90% of research studies without modification of measures
- Retail Food & Bev and QSR supported
- Hundreds of product categories scored



Overall Liking has a range compression that makes it poor at differentiation at the upper end of the scale.

This chart shows the general effect using two chocolate confections. In a consumer study, they were differentiated by 0.1 OAL unit. However, when compared using Curion Score to the category, the difference in performance **is 68% vs. 85%,** 17% of the entire chocolate confections category, a fact obscured by the similar OAL scores.



CurionScore™ corrects the non-normality of Overall Liking, making it better suited as a dependent variable.

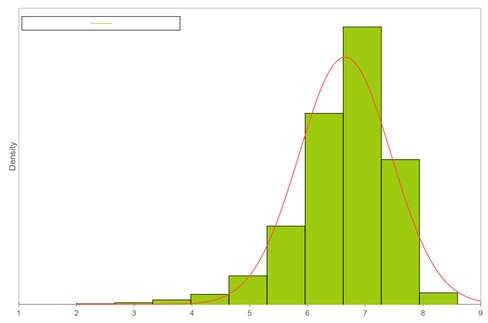
Overall Liking has well known properties of range compression and left truncation.

This causes scores to be compressed at the top end of the scale.

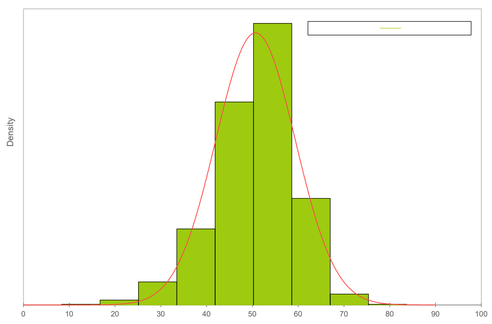
Analysis such as preference maps, clustering, and other methods often struggle differentiate performance.

CurionScore™ may be better suited as a dependent variable when understanding consumer differentiation against competition or oneself is a research objective.

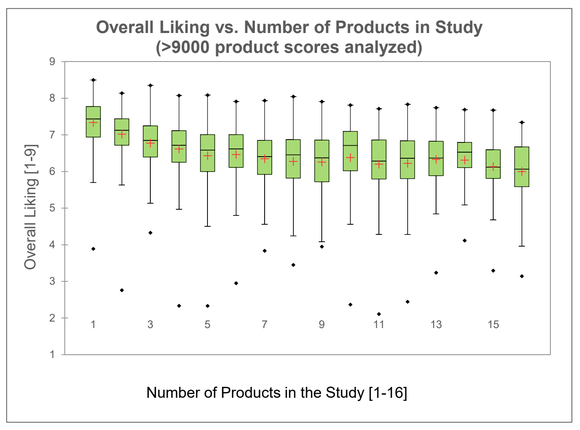
9-point OAL Distribution
> 8500 products across Curion's Research Database.

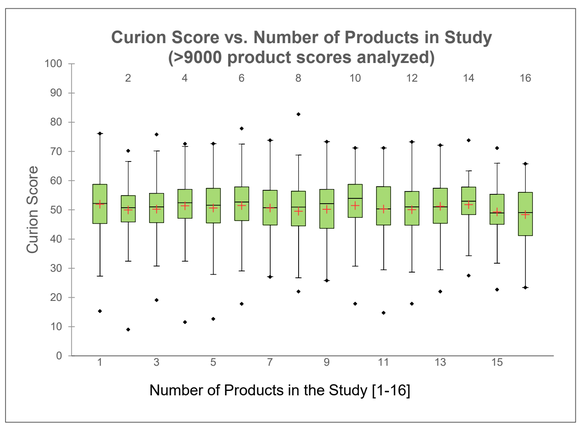


CurionScore™ Distribution
>8500 products across Curion's Research Database.



CurionScore™ also corrects for study design variables that impact Overall Liking, especially when comparing between studies.

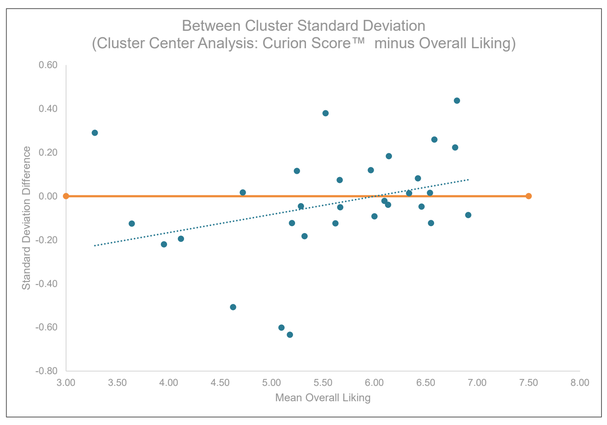




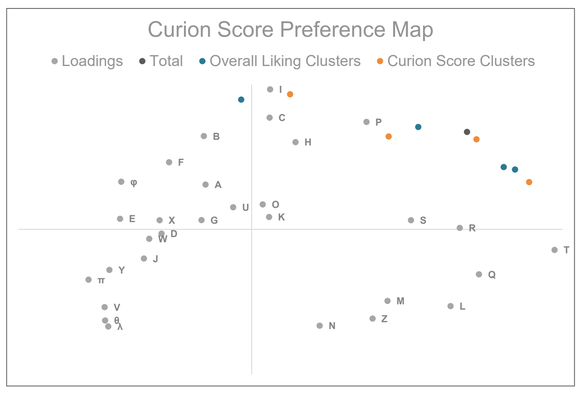
Overall Liking varies greatly based on study context – the chart on the left shows that between 1 and 6 product study designs, OAL will drop by up to a point on average. This means that fixed OAL benchmarks are not reliable, which leads to expensive 1st position or monadic designs. **Curion Score (right) is designed to be a better benchmark** by accounting for the largest design effects, showing no position effects.

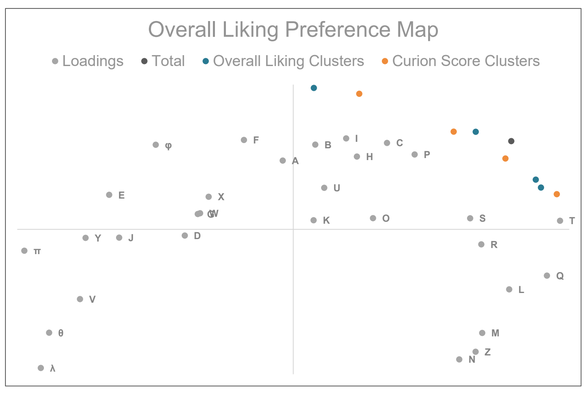
CurionScore™ Clusters Better Differentiate High Performing Products than OAL Clusters

- Four-Cluster Solutions were developed on CurionScore™ and OAL (not shown).
- OAL means for each product were calculated for both sets of cluster solutions and the difference was taken and plotted (right).
- The results show that CurionScore clusters capture a wider variation in OAL than Liking, thus creating clusters that better show product performance differences.



Internal Preference Maps using OAL and CS show expected similarities.





RV=0.87, NRV=17.8, p-value < 0.0001
The NRV coefficient indicates maps are significantly similar.

Discussion and Conclusions

- Overall Liking has been the de-facto primary performance optimization metric for product testing since it was developed by the U.S. Army in the 1950s.
- Its statistical design is well suited for comparing products within a study. However, limitations prevent it from being an ideal benchmark for external validity and for comparing across studies.
- Modern demands of business are requiring more from our research, including externally valid measures and approaches permitting benchmarking and comparison across studies. This can lead decreased reliability in performance measurement due to the reasons shown in this poster.
- CurionScore™ is designed from the ground up as an externally valid product performance metric that makes it more robust and discriminating as a benchmark and as a dependent variable in statistical analyses.

[1] Nestrud et al. 2022. An Exploration of Sensory Performance Benchmarking. Society of Sensory Professionals Annual Conference, Savannah, Georgia.

PANG2023_0702
S002 Poster session 1, 21st Aug 23, 14:00 - 15:15 CEST, Great Hall