Starry right? The real story behind online ratings

From Amazon to Airbnb, customer ratings are meant to help us understand the quality we can expect. But stars can be misleading, particularly across different rating systems.



January 27, 2021

- **Tapping the wisdom of the crowd in the form of star ratings is prone to systemic biases.**
- **Overall ratings of products differ based on whether or not people are asked to rate individual dimensions — and which dimensions are then included.**

©IESE Business School

- **Overall ratings are not comparable across platforms.**

In the ever-expanding world of online offerings, users are frequently asked to leave feedback on a purchase, often in the form of star ratings. This system is seen as tapping into the wisdom of the crowd — offering a more objective assessment than the sellers' pitch, safeguarding clients against potential bad purchases and giving the best products their due.

In reality — as Christoph Schneider examines in a paper with Markus Weinmann, Peter N.C. Mohr and Jan vom Brocke — star ratings are prone to systemic biases. Understanding those biases can help inform both consumers and companies as stakeholders in the rating game.

Dimensional rating bias

The focus of the authors' paper is how the rating system itself influences outcomes — that is, how the "choice architecture" influences assessments.

In it, they examine "single-dimensional" and "multidimensional" rating systems that are used to evaluate restaurants, movies and a university on online platforms.

While a single-dimensional system asks for an overall score only, its multidimensional counterpart breaks it down: how was the restaurant's food? Service? Atmosphere?

Over seven experiments, Schneider and his coauthors looked at how these different rating systems affect perceived customer satisfaction of the various products or services.

Through the series of experiments, they learned that people's overall ratings are influenced systematically by dimensional ratings, and that just being asked to think in terms of dimensions affects the overall rating.

Specifically, where individual dimensions are rated highly, the overall rating tends to be higher than it would be on a platform that asked for a single, all-encompassing rating. The reverse is also true: lower scores for individual dimensions pull down the overall score, whether or not the dimensions discussed were truly relevant to overall customer satisfaction.

Practical implications

The authors identify a range of stakeholders who can benefit from knowing about rating biases. These include the review-platform providers that want to provide a high-quality

©IESE Business School

platform. Platform providers should be aware that multidimensional ratings may give a more complete picture, yet they can unduly push ratings in one direction or the other. Instead of asking the crowd of reviewers for both dimensional and overall ratings, platform providers might create their own aggregates for less biased results.

Product and service providers also have vested interests when the rating systems are applied to their offerings. For multidimensional ratings, product providers should note the kinds of questions asked (and not asked). A rating system will be more advantageous if its dimensions match the product or service well.

Finally, some consumers will want to know whether or not they would enjoy a restaurant, movie or other product rated. The study suggests that consumers should place more trust in transparent rating systems.

Moreover, consumers should critically evaluate which dimensions of a product are being evaluated to determine whether those dimensions are the most helpful for them. For example, a restaurant rating that asks about decor may be less relevant to many of its hungry would-be patrons and yet still show up in an overall score.

In many ways, online ratings are still in their infancy. From one platform to another, a rating of 1 to 5 stars may seem a simple metric, but hidden biases suggest it is not.

Methodology

The coauthors began with the question: How do multidimensional rating systems influence overall ratings? They tested the idea over a series of seven studies containing a total of 17 conditions, both in highly controlled study settings and later as replications in a realistic setting. The studies took place between 2016 and 2019 with participants from Australia, Canada, England, Liechtenstein, New Zealand and the United States.

©IESE Business School