

Misfit of rating scales

Before launching their innovation contest, the members of the innovation division of **AwesomeTech** held a meeting to define the details of the challenge. One very important topic was the definition of the selection criteria. After a short brainstorming session, the members collect about 17 criteria. While discussing the criteria, **Bob** affirms, *“it doesn’t make sense to evaluate 300 ideas with 17 criteria, that’s too much at this point”*. Everyone agrees and they proceed with a short deliberation, after which they define the final five criteria: Innovativeness, feasibility, customer value, market potential and elaborateness. The meeting ends and a few days later, the contest is launched. At the end of the submission period, the expert team starts evaluating the ideas on a 5-point Likert scale. Soon, **Bob** reviews the first assessments and realizes that the experts ratings vary significantly. The same idea was given an overall score of 2.5 by **Lisa**, 4 by **John** and 5 by **Mike**. **Bob** calls for another meeting to solve this issue. At the update meeting, **Brad** affirms: *“rating with those criteria is not realistic, customer value is relatively difficult to assess, market potential is super difficult to assess... so if you ask three people, you’ll probably get different assessments. That is very complex.”* **Mike** complements: *“it’s probably not working out because the evaluation criteria are softly described. There are people who evaluate very positively, and there are people who evaluate very negatively”*. **Bob** decides to average the ratings and define the 30 finalist ideas based on the highest scores. However, when seeing the final ranking, **Bob** notices that an idea he finds incredibly promising did not make it. Thinking the idea has a chance to win, **Bob** changes the ranking and includes it in the shortlist. At the shortlisting meeting, **Bob** presents the ideas to the experts and they proceed to nominate their favorites. After processing all votes, **Bob** presents the ranking to the evaluators. To his surprise, no one seems truly satisfied with the result, because all evaluators feel that it does not really reflect their opinion. Reflecting on the dissatisfaction of the shortlisting phase, **Bob** starts preparing the materials for the winner determination with the senior-level managers of **AwesomeTech**. **Bob** adds more holistic, coarse-grained criteria, such as a simple “would you invest?” question, in order to incorporate the intuitive assessments. At the winner determination meeting every submission is discussed, and since there is no ranking included, the jury discusses whether there should be just one winner or more, given the high level of submissions.